DYNA

20nm

UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

Publication Information

Indexing and Databases

DYNA

CONTENTS

Our cover
Image alluding to Article:
Facile one-pot synthesis of uniform silver nanoparticles and growth mechanism

Authors:
Daniel Ramirez & Franklin Jaramillo

CONTENIDO

Nuestra carátula
Imagen alusiva al artículo:
Fácil síntesis en un paso y mecanismo de formación de nanopartículas de plata

Autores:
Daniel Ramírez & Franklin Jaramillo

Resistance correlation with compression uniaxial effective humidity and porous rocks

Manolo Galván-Ceballos & Iván A. Restrepo

Escuela de Ingeniería Civil y Geomática, Universidad del Valle, Cali, Colombia, manolo.galvan@correounivalle.edu.co, ivan.restrepo@correounivalle.edu.co

Abstract
This research consists of the determination of the mechanical parameter of resistance to uniaxial compression in the rock and its correlation with moisture and effective porosity, physical properties of the rock. By its goal of study, the project comes within the framework of the mechanics of rocks. Rocks selected for research were taken from a quarry in South West Colombia. The investigation led to the determination of equations that allow to estimate the resistance to uniaxial compression of rock from its effective porosity and moisture content. It is important to highlight that such correlations are representative for rocks benefiting from lithological characteristics similar to those used in the experimentation. Finally the results are presented with their analysis and conclusions of the research.

Key words: *Rock mechanics; Rock properties; Unconfined compressive strength; Quarry rock resistance.*

Correlación de la resistencia a compresión uniaxial con la humedad y porosidad eficaz en rocas

Resumen
Esta investigación, consiste en la determinación del parámetro mecánico de resistencia a compresión Uniaxial en la roca y su correlación con la humedad y porosidad eficaz, propiedades físicas de la roca. Por su objetivo de estudio, el proyecto se engloba dentro del marco de la mecánica de rocas. Las rocas seleccionadas para la investigación fueron tomadas de una cantera del Sur Occidente Colombiano. La investigación condujo, a la determinación de ecuaciones que permiten estimar la resistencia a compresión Uniaxial de la roca a partir de su porosidad eficaz y del contenido de humedad. Es importante resaltar que este tipo de correlaciones son representativas para rocas que gocen de características litológicas similares a las utilizadas en la experimentación. Finalmente se presentan los resultados con sus análisis y conclusiones de la investigación.

Palabras clave: Mecánica de rocas; Propiedades de la roca; Resistencia a compresión inconfinada; Resistencia en rocas de cantera.

1. Introducción

Una definición ampliamente aceptada de la mecánica de rocas es la presentada por el Comité Nacional de Mecánica de Rocas de la Academia Nacional de Ciencias, Washington D.C. en 196, y complementada en 1974 [1],"la mecánica de rocas es la ciencia teórica y aplicada del comportamiento mecánico de la roca y macizos rocosos. Esta es la rama de la mecánica concerniente con la respuesta de la roca y macizos rocosos, a los campos de fuerza de su entorno físico".

La mecánica de rocas hace parte del campo de la geotécnica y diferente de la Mecánica de Suelo, la cual se define como "la aplicación de la ciencias de la mecánica de suelos y rocas, ingeniería geológica y otras disciplinas relacionadas, en la construcción civil, la industria minera y la conservación del medioambiente" [1].

Los requerimientos para realizar el ensayo a resistencia a compresión simple de las probetas son [2]:
- La relación altura/diámetro de 2.5 a 3.0
- Diámetro mayor a 10 veces el tamaño máximo de grano de la roca.
  - Diámetro no inferior a 50 mm.
  - Tolerancia de las caras de la probeta (ver Tabla 1).

Tabla 1.
Tolerancia de las probetas

Tolerancia respecto a:	Deformabilidad de la roca:		
	Poca	Media	Alta
Desviación de la generatriz respecto a la dirección axial	±0.3 mm	±0.4 mm	±0.5 mm
Planitud de la base	±0.02 mm	±0.5 mm	±0.1 mm
Desviación, respecto al ángulo recto, del ángulo del eje de la probeta con la base.	10'	20'	30'

Fuente: Tomada de [2]

Dichos requerimientos son de importante consideración, debido a que en algunos casos sean muy difíciles de alcanzar o hacen que el ensayo sea muy costoso. Algunas veces se ven obligados al incumplir con las recomendaciones de las normas, por ejemplo:

- Las discontinuidades del macizo rocoso pueden impedir la obtención de cilindros de roca de longitudes iguales o superiores a 125 mm, que corresponde a la probeta de menor longitud, para el menor diámetro de 50 mm.
- En el caso de algunas rocas cuyos granos o clastos que superen el centímetro (algunos granitos o pegmatita), las dimensiones de la probeta correspondiente, son imposibles de obtener y en caso de lograrlo, la probeta no podría romperse con una prensa convencional.

Para mitigar estos hechos, investigadores como [3,4], han correlacionado experimentalmente la resistencia a compresión de la roca con los resultados de ensayos indirectos o las características físicas de la roca estudiada, brindando procedimientos económicos y rápidos para la estimación indirecta de la resistencia de la roca.

Los parámetros mecánicos se obtiene a partir del ensayo de resistencia a compresión uniaxial y los parámetros físicos (Porosidad eficaz, contenido de humedad, gravedad especifica) se determinan acorde a los procedimientos indicados en las normas [5,6]. Con los resultados de los ensayos, se procedió a la formulación de las ecuaciones que relacionan los parámetros, empleando métodos estadísticos de correlación y regresión lineal.

2. Geología de la zona de estudio

2.1. Características de la zona

A nivel nacional, se han adelantado investigaciones sobre el subsuelo por el instituto Ingeominas [7]. (A partir del 2012, Ingeominas cambia su nombre por el Servicio Geológico Colombiano), dando como resultado la construcción de los mapas geológicos en el año 2007, que brindan información de la distribución de distintos tipos de roca en Colombia.

De los mapas Geológicos, la plancha número 5-13, brinda la distribución de las unidades cronoestratigrafícas y litológicas del Valle del Cauca. Es de especial interés indagar sobre las características de las rocas de Santiago de Cali y sus alrededores, lugares entre los que se encuentra ubicada la cantera. El análisis del mapa brinda la siguiente información del subsuelo, edad, litología y provincia:
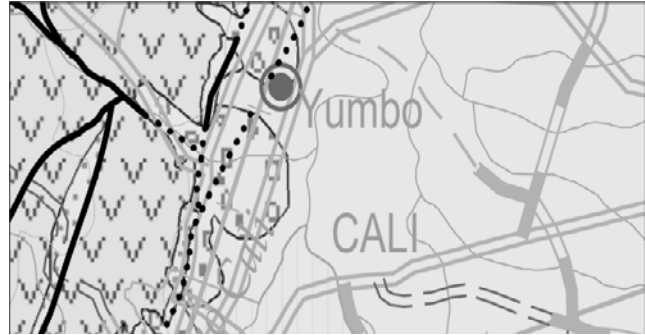
Figura 1. Detalle de la unidad crono-estratigráfica de Cali y Yumbo, en el Valle del Cauca
Fuente: Tomado de [7].

- Santiago de Cali: Suelo de la época Pleistoceno-holoceno, depósito aluvión de la provincia Litosferica Continental Mezo-proterozoica Grenvilliana (Q1Q2 -Q al, PLCMG).
- Yumbo: Suelo de la época Pleistoceno-holoceno, depósito Abanico de la provincia Litosferica Continental Mezo-proterozoica Grenvilliana (Q1Q2-Q ca, PLCMG).

3. Estado del arte

Para estimar la resistencia a compresión simple de una roca, se pueden usar métodos y/o ensayos que se realizan en campo y/o laboratorio. Su obtención va desde una estimación meramente subjetiva hasta una medición indirecta. Uno de estos métodos es la obtención de la RCS a través de las propiedades índice de la roca.

El investigador [8], presenta correlaciones obtenidas por diferentes investigaciones en rocas del golfo de México: areniscas, esquistos, caliza y dolomita. Estas correlaciones, son el resultado de múltiples ensayos y análisis de los núcleos de rocas, lo que permite caracterizar y correlacionar los parámetros para formaciones específicas de roca.

A excepción de la primera ecuación, todas las ecuaciones para la arenisca tienen una mala predicción de los datos de resistencia para tiempos altos de viaje $\Delta t > 100$ μs/ft, e inferiores a $\Delta t < 3000$ μs/ft, como informa [8].

En relación con esta investigación se puede observar, que a mayor porosidad de la roca se reduce su resistencia.

4. Trabajos experimentales

4.1. Propiedades Índice y Alterabilidad.

4.1.1. Porosidad (n)

Es definida como la razón entre el volumen de los espacios vacíos entre la roca y el total volumen aparente de la roca, también se refiere a la probabilidad de encontrar vacíos en el volumen total:

Tabla 2.
Correlación de la resistencia a compresión (MPa) con las propiedades físicas de la roca [8].

Tipo de roca	Autor	Correlación (MPa)	Nomenclatura
Arenisca de Turingia, Alemania.	Freyburg, 1972	$UCS = 0,035v_p - 31,5$	Velocidad ultrasónica (P) v_p (m/s)
Arenisca grano fino, consolidada y no consolidada en la Cuenca de Bowen de Australia	McNally, 1987	$UCS = 1200exp(-0,036\Delta t)$	Δt (μs/ft)= $1/v_p$
Areniscas débiles y no consolidadas en la costa del Golfo de los EE.UU.	Chang et al. 2006	$UCS = 1,4138 \times 10^7 \Delta t^{-3}$	-
Arenisca en el Golfo de México.	Chang et al. 2006	$UCS = 3,87exp(1,14x10^{-10}\rho v_p^2)$	densidad - ρ (g/cm³):
Pizarra	Lal, 1999	$UCS = 10(304,8/\Delta t - 1)$	v_p(km/s)
Esquistos del Mar del Norte	Horsrud, 2001	$UCS = 0,77v_p^{2,93}$ $UCS = 243,6 \phi^{-0,96}$	Φ: Porosidad en porcentaje.
		$UCS = 243,6\rho_0^{-0,96}$	ρ_0: Porosidad en porcentaje
Caliza y dolomita	Milizer and Stoll, 1973	$UCS = (7682/\Delta t)^{1,82}/145$	-
Caliza y dolomita	Golubev and Rabinovich, 1976	$UCS = 10^{(2,44+109,14/\Delta t)}/145$	-
Caliza y dolomita, en el Oriente medio.	Chang et al. 2006	$UCS = 143,8$ ex(-6,95Φ)	Φ: Porosidad en fracción.

Fuente: Tomada de [8]

$$n = \frac{V_{poros}}{V_{muestra}} * 100 \qquad (1)$$

Las dos descripciones más comunes de la porosidad, son la porosidad total definida anteriormente que considera todos los poros presentes en la roca y la porosidad efectiva o interconectada, que representa la razón entre el espacio de los poros interconectados y el volumen aparente de la roca descartando los poros aislados y cerrados. [5]

La porosidad se relaciona en proporción directa con la deformación de la roca y es inversamente proporcional con la resistencia y densidad. El aumento de la porosidad afecta ampliamente las características mecánicas de las rocas ya que la presencia de poros dan lugar a zonas de debilidad en la matriz rocosa.

Como lo indican los autores [8], la porosidad es controlada por la forma, tamaño y arreglo de los granos de la roca. Esto, a su vez depende de los procesos mecánicos (compactación, deformación y evaluación de fractura) y químicos (disolución, precipitación, cambios en la mineralogía).

4.1.2. Contenido de Humedad (w%)

Es definida [5], como la razón en porcentajes entre, la masa de agua contenida en los espacios o poros y la masa de la las partículas sólidas que conforman la roca, una temperatura estándar de 110 +/- 5 ºC es usada para determinar de manera individual esta última masa:

$$\%w = \frac{M_w}{M_s} * 100 \qquad (2)$$

Se consideran dos condiciones extremas de contenido de humedad: Roca libre de humedad en la cual la masa de agua en la roca es cero y roca saturada para la cual la masa de agua en la roca es máxima, ocupando todos los espacios de vacío o poros.

4.2. Ensayo de resistencia a compresión simple (RCS)

La resistencia a compresión simple se determina mediante el "ensayo de resistencia a compresión uniaxial" (ECU), de una forma directa. El ensayo consiste en aplicar una fuerza axial F a una probeta cilíndrica de área A, llevándola hasta la rotura mediante una prensa. La resistencia viene dada por la Ec. (3).

$$\sigma_c = \frac{F_c}{A} \qquad (3)$$

Los requisitos según la norma UNE 1990 [2], para las probetas a ensayar son:
- Forma cilíndrica, con dimensiones:
  - Relación altura / diámetro de 2,5 a 3,0.
  - Diámetro, mayor de 10 veces el tamaño máximo de grano de la roca.
  - Diámetro no inferior a 50 mm.
- Las probetas se elaborarán mediante perforación, corte torneado y pulido o cualquier otro método apropiado, de modo que:
  - La superficie lateral de la probeta debe ser lisa y estar libre de irregularidades.
  - Las bases deben ser planas y formar un ángulo recto con el eje de la probeta.
  - Se debe evitar el empleo de materiales de recubrimiento como igualadores para conseguir el paralelismo requerido de las superficies de las bases de la probeta de ensayo. Si fuera necesario desviarse de esta regla debido a las características propias del material a ensayar, habrá que indicarlo en el protocolo de ensayo.

Por lo tanto los requerimientos, en cuanto a la muestra a ensayar, no son fáciles de cumplir y en algunas ocasiones resultan imposibles.

5. Resultados y discusión

5.1. Correlación y regresión simple

Una vez concluida la experimentación sobre toda las muestras y procesado los datos, se presenta la Tabla 3, en la cual se indica el valor de varios parámetros para cada núcleo, como son: el contenido de humedad, porcentaje de porosidad interconectada, resistencia última a compresión simple, densidad real y módulo de Young. Los colores en la gráfica indican la magnitud de cada valor comparado con los demás de su grupo (o variable).

La Fig. 2, permite visualizar la relación ente el contenido de humedad y la RCU de los datos presentados en la Tabla 3.

Se observa en la Fig. 2, la dispersión de los datos, y se observa un índice de coeficiente de determinación cerca al 50%.

El análisis estadístico de los valores ha permitido obtener las ecuaciones que describen la variación de la resistencia a compresión de la roca según los diferentes estados de contenido de humedad:

Tabla 3.
Propiedades índices y esfuerzo último de los núcleos de roca.

Id	Núcleos	Contenido de Humedad (%)	Porcentaje de Porosidad	Esfuerzo ultimo (Kf/cm²)	Esfuerzo ultimo (MPa)	Densidad real ρ_r (g/cm³)	Módulo de Young, E (MPa)
1	M3-3	0.53%	1.53%	564.0	55.3	2.89	12.0
2	M4-1	0.60%	1.72%	401.9	39.4	2.96	16.5
3	M4-2	0.34%	1.33%	785.1	77.0	2.96	15.0
4	M5-1	0.33%	1.29%	446.0	43.7	2.93	8.0
5	M5-5	0.28%	1.50%	366.4	35.9	2.93	13.9
6	M5-6	0.25%	1.43%	747.1	73.2	2.93	22.8
7	M7-1	0.15%	1.00%	1321.5	129.6	2.92	35.4
8	M7-2	0.12%	0.76%	1628.0	159.6	2.92	36.9
9	M7-3	0.20%	1.37%	1191.9	116.9	2.92	32.8
10	M8-1	0.20%	1.37%	1072.3	105.1	2.89	26.0
11	M8-2	0.15%	1.20%	1326.8	130.1	2.89	24.7
12	M8-3	0.16%	1.28%	972.3	95.3	2.89	31.5
13	M8-4	0.17%	1.38%	1342.8	131.6	2.89	25.2
14	M8-5	0.13%	1.00%	1225.6	120.2	2.89	25.5
15	M8-6	0.11%	0.97%	1373.8	134.7	2.89	34.2
16	M8-7	0.15%	1.32%	1305.2	128.0	2.89	27.1
17	M8-8	0.12%	1.30%	1187.5	116.4	2.89	28.5
18	M8-9	0,11%	1,38%	800,3	78,5	2,89	18.4
19	M8-10	0,08%	1,02%	764	74,9	2,89	23.8
20	M9-1	0,13%	1,21%	1273,3	124,8	2,9	28.1
21	M9-2	0,15%	1,29%	1378,8	135,2	2,9	25.0
22	M9-3	0,15%	1,31%	1057,8	103,7	2,9	29.0
23	M10-1	0,09%	1,30%	1049,5	102,9	2,94	26.4
24	M10-2	0,09%	1.26%	816.8	80.1	2,94	44.6
Valores máximos		0.60%	1.72%	1628.0	159.6	3.0	44.6
Valores mínimos		0.08%	0.76%	366.4	35.9	2.9	8.0
Media aritmética		0.20%	1.27%	1016.6	99.7	2.9	25.5
Desviación típica (σ)		0.131%	0.206 %	341.36	33.19	0.024	8.51
Varianza (S^2)		1.715 E-06	4.07 E-06	116525.67	1120.01	0.0004255	69.42

Fuente: Propia.
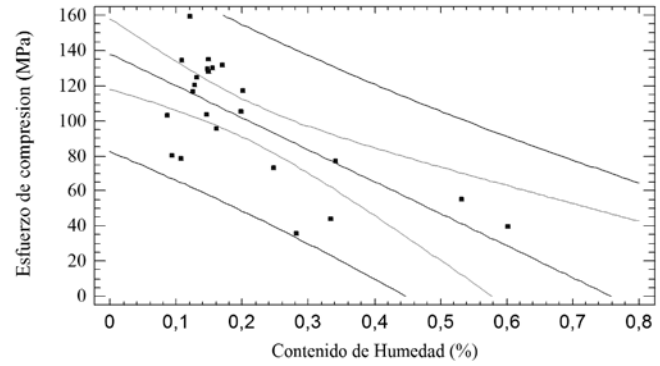
Figura 2. Variación de la resistencia a compresión uniaxial según el contenido de humedad.
Fuente: Propia.

Tabla 4.
Estadísticas de la regresión, RCU vs Contenido de Humedad

Coeficiente de correlación R	- 0,71
Coeficiente de determinación R^2	50,05%
R^2 ajustado	47,67%
Error típico	19,67
Observaciones	23

Fuente: Propia.

Tabla 5.
Coeficientes de la regresión, RCU vs Contenido de Humedad

Parámetro	Coeficientes	Error típico	Estadístico T	Valor-P
Intercepto	137,993	9,6475	14,3035	0,0000
Pendiente	-181,868	49,6481	-4,5871	0,0002

Fuente: Propia.

$$RCU[MPa] = 137,993 - 181,868 * w(\%) \qquad (4)$$

$$R^2 = 50.05\% \qquad (5)$$

Del contenido de humedad se puede evidenciar que su aumento genera una disminución considerable en la resistencia a compresión de la roca, el coeficiente de correlación es de -0,71, lo que indica que hay una correlación negativa media, el estadístico R^2 indica que en el modelo el 50,05% de la variabilidad de la RCU se ve explicada por la variación del contenido de humedad.

5.2. Correlación entre la RCU y la porosidad (n).

La Figura permite visualizar la relación entre la RCU y el porcentaje de porosidad eficaz en las muestras de roca analizadas, descartando una muestra por su elevado valor de dispersión.

Se observa en la Fig. 3, la dispersión de los datos y hay una correlación del 57%.

El análisis estadístico de los valores nos ha permitido obtener las ecuaciones que describen la variación de la resistencia a compresión de la roca según los diferentes estados de porosidad interconectada.
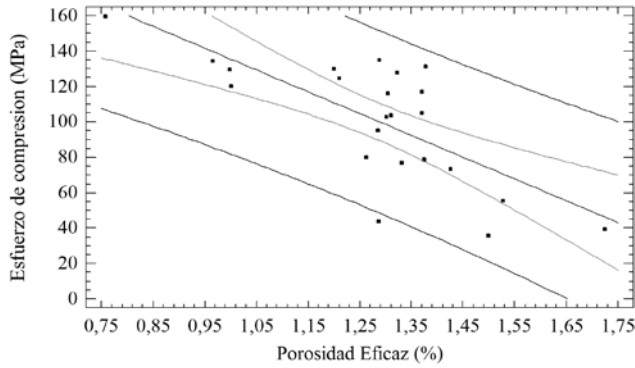
Figura 3. Variación de la resistencia a compresión uniaxial según la porosidad interconectada.
Fuente: Propia.

Tabla 6.
Estadísticas de la regresión, RCU vs %Porosidad

Coeficiente de correlación R	-0,7264
Coeficiente de determinación R^2	57,76%
R^2 ajustado	50,51%
Error típico	24,30
Observaciones	23

Fuente: Propia.

Tabla 7.
Coeficientes de la regresión, RCU vs %Porosidad

Parámetro	Coeficientes	Error típico	Estadístico T	Valor-P
Intercepto	258,997	33,07	7,833	0,0000
Pendiente	-123,401	25,48	-4,843	0,0001

Fuente: Propia.

$$RCU[MPa] = 258,997 - 123,401 * n(\%) \qquad (6)$$

$$R^2 = 57,76\% \qquad (7)$$

Respecto al porcentaje de porosidad interconectado también se visualiza en la Figura 3. La correlación lineal inversa, a la que le corresponde un coeficiente de Pearson -0.73 que indica una correlación negativa media (Ver Tabla 7).A mayor porosidad en la roca se puede esperar una menor resistencia a compresión, desarrollando un análisis simple se puede notar en la Tabla 6 , el impacto que tiene la porosidad. El coeficiente de determinación R^2 indica que en el modelo el 57,76% de las variaciones de la resistencia en la roca se debe a la variación en el volumen de los poros interconectados.

5.3. Correlación entre la RCU y el contenido de humedad y la porosidad eficaz.

Con el objetivo de explicar y predecir el efecto que tiene la porosidad eficaz y el contenido de humedad de la roca sobre su capacidad de resistencia a compresión simple, se emplean nuevamente los datos de la La Fig. 2, permite visualizar la relación ente el contenido de humedad y la RCU de los datos presentados en la Tabla 3.

Se observa en la Fig. 2, la dispersión de los datos, y se observa un índice de coeficiente de determinación cerca al 50%.

El análisis estadístico de los valores ha permitido obtener las ecuaciones que describen la variación de la resistencia a compresión de la roca según los diferentes estados de contenido de humedad:

Tabla Las Figs. 4 y 5 permiten apreciar la dispersión de las observaciones y la tendencia del modelo.
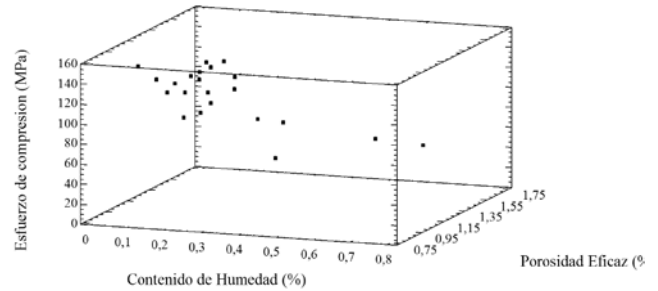
Figura 4. Efecto del Contenido de humedad y el porcentaje de Porosidad en la Resistencia a Compresión Uniaxial.
Fuente: Propia.
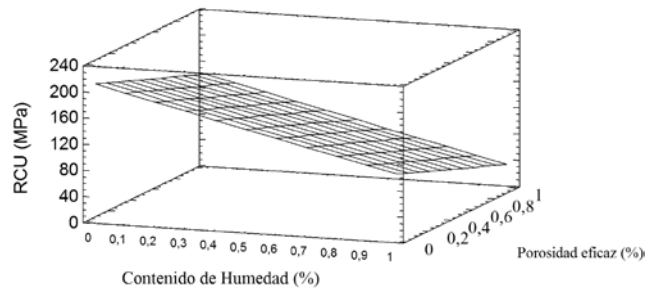
Figura 5. Modelo de regresión lineal múltiple, RCU vs contenido de humedad y porosidad.
Fuente: Propia.
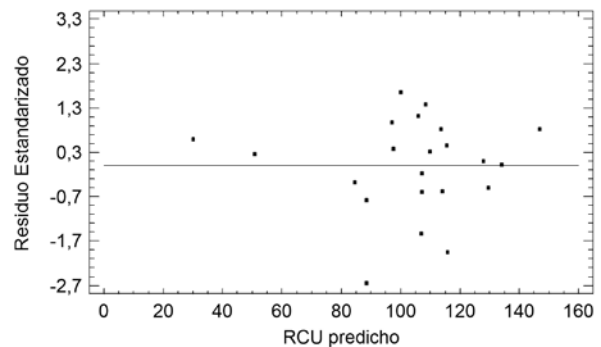
Figura 6. Gráfico de residuos estandarizados RCS vs porosidad eficaz y contenido de humedad.
Fuente: Propia.
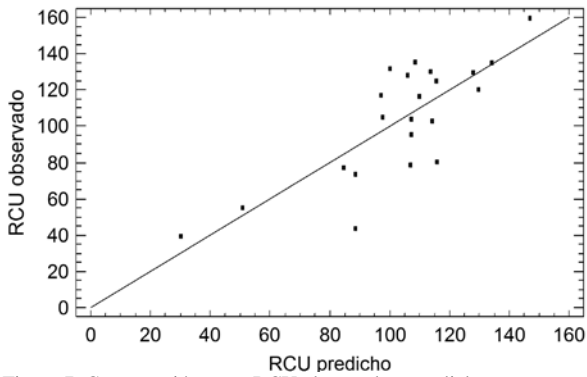
Figura 7. Comparación entre RCU observado y predicho.
Fuente: Propia.

Tabla 8.
Estadísticas de la regresión, RCU vs Contenido de Humedad y Porosidad

Coeficiente de correlación múltiple R	-79,37,
Coeficiente de determinación R^2	62,99%
R^2 ajustado	59,10%
Error típico	20,631
Observaciones	22

Fuente: Propia.

Tabla 9.
Coeficientes de la regresión, RCU vs Contenido de Humedad y % Porosidad

Parámetro	Coeficientes	Error típico	Estadístico T	Valor-P
Intercepto	210,442	32,27	6,5202	0,0000
Variable – Humedad (w, %)	-108,171	42,88	-2,5223	0,0207
Variable – Porosidad (n)	-66,7724	28,91	-2,3094	0,0323

Fuente: Propia.

Tabla 10.
Residuos estandarizados de la regresión, RCU vs Contenido de Humedad y % de Porosidad.

Fila	RCU	Predicción RCU	Residuos	Residuo estandarizado
4	43,7	88,55	-44,85	-2,65

Fuente: Propia.

Del modelo multi-variado se puede concluir que la información entregada por el software Statgraphics Centurion, es la ecuación que describe la recta que mejor se ajusta a las observaciones es:

$$RCU[MPa] = 210,442 - 108,171 * \%w - 66,77 * n(\%) \quad (8)$$

$$R^2 = 62,99\% \quad (9)$$

Entre las variables hay una asociación inversa, a medida que la porosidad eficaz y el contenido de humedad aumentan la RCU disminuye, esto es evidente por el signo negativo de los coeficientes humedad y porosidad, por el signo del coeficiente de correlación.

El coeficiente de determinación ajustado R^2 es 59,097%, Es decir que para el modelo logra explicar el 59,097% de la variación de la RCU en la roca.

6. Conclusiones

Los resultados obtenidos en esta investigación son un avance importante en la mecánica de rocas del departamento del Valle. Es importante recordar, que el tipo de modelos de regresión presentados en este documento, son aplicables y representativos para rocas que gocen de características similares a las utilizadas en la presente investigación.

Como conclusiones generales de la investigación se resaltan las siguientes consideraciones:

Se le determinaron parámetros físicos – mecánicos a las muestras de rocas consideradas en la investigación. A continuación, se presentan un resumen de los rangos de estos valores.

Tabla 11.
Rangos de magnitud de los parámetros físicos – mecánicos.

Parámetro	Rango	Valor medio	Desviación estándar	Unidades
Resistencia a Compresión Uniaxial	$35,9 < \sigma_c < 159,6$	99,67	34,19	MPa
Módulo de Young	$8,0 < E < 44,6$	25,47	8,51	GPa
Porosidad eficaz	$0,76 < n < 1,72$	1,27	0,21	%
Densidad Real	$2,89 < \rho_r < 2,96$	2,91	0,024	g/cm^3
Contenido de humedad	$0,08 < w\% < 0,60$	0,20	0,13	%

Fuente: Propia.

6.1. Relación entre el resistencia a compresión uniaxial (σ_c) y el contenido de humedad (%w)

De esta relación se puede concluir que al aumentar el contenido de humedad (pasar de seco a parcialmente saturado) su resistencia a compresión uniaxial se ve reducida, aproximadamente entre un 48% (con los datos experimentales) y en un 79% (con los datos analíticos). Se planteó un modelo matemático lineal, uno exponencial y otro inverso; para describir la relación entre las variables, la Tabla 12 presenta los modelos. Para el modelo lineal el coeficiente de correlación es de -0,71, indicando que hay una correlación negativa media entre las variables. El coeficiente de determinación R^2 indica que en el 50% de la variación de la resistencia a compresión uniaxial de la roca se debe a la variación de la porosidad eficaz.

6.2. Relación entre la resistencia a compresión uniaxial (σ_c) y la porosidad (n)

Las variables presentan una relación inversa de magnitud media (R = -0,73), por tanto a mayor porosidad eficaz se espera una disminución en la resistencia a compresión. Para la muestra de rocas estudiadas, esta pérdida de resistencia varía entre un 15% y un 59% para las rocas de porosidad media y alta, respectivamente, comparadas con las de baja porosidad. El coeficiente de determinación R^2 indica que en el modelo el 57,76% de las variaciones de la resistencia en la roca se deben a la variación en la porosidad eficaz.

6.3. Relación entre la RCU, el contenido de humedad y la porosidad eficaz

Del análisis de esta relación se generó un modelo matemático lineal multivariado para caracterizar la relación entre las variables. Este modelo aumentó la capacidad predictiva en comparación con los modelos lineales. Ambas variables continúan mostrando una relación lineal inversa, de modo que, al aumentar el contenido de humedad o la porosidad eficaz la resistencia a compresión disminuye. Los resultados del modelo indican que el 59,09% de las variaciones de la resistencia se deben a la variación del contenido de humedad y porosidad eficaz de la roca.

Tabla 12.
Resumen de ecuaciones desarrolladas en investigación.

	Modelo (RCU[MPa])	CD	CC
RCU vs. Contenido de humedad	$RCU = 137.993 - 181.868 * w$	$r^2 = 50,05\%$	$r = -71,0\%$
	$RCU = \dfrac{1}{0.0019 + 0.034 * w}$	$r^2 = 55,76\%$	$r = -74,67\%$
	$RCU = e^{5,02 - 2.35 * w}$	$r^2 = 55,74\%$	$r = -74,66\%$
RCU vs. Porosidad eficaz	$RCU = 258,997 - 123,401 * n$	$r^2 = 57,76\%$	$r = -72,64\%$
	$RCU = (18,244 - 6,531 * n)^2$	$r^2 = 50,20\%$	$r = -70,85\%$
	$RCU = \sqrt{42072,2 - 24002,6 * n}$	$r^2 = 56,30\%$	$r = -75,03\%$
RCU vs. Porosidad y humedad	$RCU = 210,442 - 108,171 * w - 66 * n$	$R^2 = 59,09\%$	
RCU v. porosidad y densidad	$RCU = 1439,54 - 410,245 * \rho_r - 111,05 * n$	$R^2 = 65,51\%$	

Fuente: Propia.

6.4. Relación entre la RCU, la porosidad eficaz y la densidad real

En este análisis se determinó que hay una relación lineal inversa considerable (R=-0,83). El coeficiente R^2 ajustado indica que el 65,51% de la variación de la RCU se debe a la variación de la porosidad eficaz y la densidad real de la roca. Se resalta en el modelo que contrario a lo esperado, al aumentar la densidad se ve disminuida la resistencia a compresión, esto puede ser debido a que hay un pequeño rango de densidades entre las rocas (2.9-3.0).

En la Tabla 12, se presenta de forma resumida las ecuaciones planteadas a lo largo de la investigación y presenta su poder predictivo.

Donde CD y CC son los coeficientes de determinación y correlación, respectivamente. Como se observa en la Tabla 12, se recomienda utilizar los modelos con el valor de R^2 mayor, aunque por no encontrarse muy alejados y por facilidad de utilización para ingenieros e investigadores se podría utilizar los modelos lineales para el contenido de humedad.

Agradecimientos

Referencias

[1] Brady, B.H.G. and Brown, E.T., Rock Mechanics and Mining Engineering. Springer, Netherlands. 2006

[2] Española, Una Norma. Ensayos para la determinación del módulo de elasticidad (Young) y del coeficiente de Poisson. 1990. UNE 22-950-90.

[3] Galván, M., Correlación entre la resistencia a compresión simple y el índice a carga puntual Is (50) en calizas de la Comunidad Valenciana (España). Tesis Doctoral, Universidad Politécnica de Valencia, España, 2011.

[4] Hudson, J.A. and Harrison, J.P., Introduction. Engineering Rock Mechanics. Oxford: Pergamon. 1997.

[5] ASTM 2010b. Standard test methods for laboratory determination of water (Moisture) content of soil and rock by Mass1. D2216-10.

[6] ASTM 2010a. Standard test method for compressive strength and elastic moduli of intact rock core specimens under varying states of stress and temperatures. D7012-10.

[7] Gómez, J., Nivia, A., Montes, N.E., Jiménez, D.M., Tejada, M.L., Sepúlveda, M.J., Osorio, J.A., Gaona, T., Diederix, H., Uribe, H. y Mora, M., Compiladores. Mapa Geológico de Colombia. Escala 1:1'000.000. INGEOMINAS, Bogotá. 2007.

[8] Peng, S. and Zhang, J., Rock properties and mechanical behaviors. Engineering Geology for Underground Rocks. Springer Berlin Heidelberg, 2007.

M. Galván-Ceballos, es docente desde el año 2011 de la Escuela de Ingeniería Civil y Geomática de la Universidad del Valle, Cali, Colombia. En el año 2002 se graduó como Ing. Civil de la Universidad del Quindío, Armenia, Colombia, es MSc. en consultoría de empresas de ingeniería civil de la Universidad Politécnica de Valencia y Dr. en la misma universidad en el 2011 con una tesis doctoral sobre las propiedades de las rocas calizas de la comunidad Valenciana. Fue ingeniero residente de la empresa Seven S.L en Valencia, España y docente de la Universidad Politécnica de Valencia (España) en la asignatura Geotecnia y Cimientos I.
ORCID: 0000-0002-5503-2689

I. Restrepo, es Ing. Civil en 2012 de la Universidad del Valle, Cali, Colombia. Actualmente trabaja en una empresa consultora de ingeniería civil en la ciudad de Cali.
ORCID: 0000-0002-7997-6975
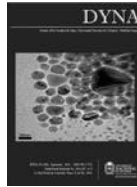
Configurable parameters of physical and MAC layer of a new DTT standard wireless return channel to provide Internet connectivity in rural areas of Latin America

Ciro Diego Radicelli-García [a] & Narcis Cardona-Marcet [b]

[a] Facultad de Ciencias de la Educación, Humanas y Tecnologías, Universidad Nacional de Chimborazo, Riobamba, Ecuador. cradicelli@unach.edu.ec
[b] Instituto de Telecomunicaciones y Aplicaciones Multimedia, Universidad Politécnica de Valencia, Valencia, España. ncardona@iteam.upv.es

Abstract
This paper proposes the elaboration of a guide to configure the parameters of a physical and MAC layer of a new second-generation digital terrestrial television (TDT) with a wireless return channel called DVB-RCT2, particularly on the physical layer; (i) transmission modes, (ii) the burst structures (BS1, BS2, BS3), and (iii) the medium access schemes (MAS) for different FFT sizes; and as for the MAC layer (i) the access modes by considering an internet band of 6 MHz. To obtain the information, the conditions of the RCT2 channel have been emulated in a broadcasting laboratory, where the transport stream (TS) has been generated to get different values on the operation from both the physical layer and MAC which are represented in different charts that demonstrate the DVB-RCT2 capacity over its predecessor DVB-RCT which makes this proposal an interesting alternative to provide internet access to rural areas.

Keywords: TDT; physical-layer; MAC-layer; DVB-RCT2.

Parámetros configurables de capa física y MAC de un nuevo estándar de TDT con canal de retorno inalámbrico para proporcionar conectividad a Internet en zonas rurales de Latinoamérica

Resumen
Este artículo propone una guía para la configuración de los parámetros de capa física y MAC de un nuevo estándar de televisión digital terrestre (TDT) de segunda generación con canal de retorno inalámbrico denominado DVB-RCT2, particularizando en la capa física; (i) los modos de transmisión, (ii) las estructuras de ráfagas (BS1, BS2, BS3), y (iii) los esquemas de acceso al medio (MAS) para diferentes tamaños de FFT; y en la capa MAC (i) los modos de acceso, considerando un ancho de banda de 6 MHz. Para obtener los datos de esta investigación se ha emulado las condiciones del canal RCT2 en un laboratorio de radiodifusión, donde se ha generado el flujo de transporte de RCT2, obteniendo valores sobre el funcionamiento tanto de capa física como de MAC, que están representados en tablas, constatando así las capacidades de DVB-RCT2 sobre DVB-RCT, lo que lo convierte en una interesante alternativa para proporcionar acceso a Internet a zonas rurales.

Palabras clave: TDT; capa-física, capa-MAC; DVB-RCT2.

1. Introducción

Dado que Internet amplía el acceso a la información, tiene un gran potencial para mejorar la educación, la ciencia, la cultura, la comunicación, y la información; incluidos también los principios de los derechos humanos como la democracia, la libertad de expresión y el acceso al saber [1]. En este sentido la Cumbre Mundial sobre la Sociedad de la Información (CMSI), celebrada en Ginebra (2003) y Túnez (2005), trató una amplia gama de temas relacionados con las Tecnologías de la Información y

Comunicación (TICs) para el desarrollo, definiendo diez objetivos que se enuncian en el Informe sobre el Desarrollo Mundial de las Telecomunicaciones [2] además de diversas recomendaciones para el 2015, destinados a fomentar la creación de una sociedad de la información inclusiva.

Sin embargo, el acceso a Internet en zonas rurales de América Latina es deficiente, debido a que la población rural no tiene acceso a esta tecnología o lo hace desde lugares públicos a baja velocidad de conexión. Situación que no sucede por ejemplo con la tecnología móvil, que cuenta en América Latina con una acogida del 130% y específicamente en Ecuador del 105% y va en aumento [3,4]. Es decir que el número de usuarios que acceden a esta tecnología es alto. Por otro lado la televisión (TV), suele ser más popular y accesible que el Internet, la telefonía móvil, e incluso que la radio, teniendo valores de penetración del 90%, según el informe de la UIT [2].

Latinoamérica está migrando de la TV analógica a la Televisión Digital Terrestre (TDT). Debido a lo cual para las emisiones de TDT los países de la región han adoptado en su mayoría el estándar japonés-brasileño ISDB-Tb (Integrated Services Digital Broadcasting – Terrestrial Brasil), a excepción de Panamá y Colombia que utilizan las normas europeas DVB-T (Digital Video Broadcasting – Terrestrial) y DVB-T2 (Terrestrial 2nd Generation) respectivamente, y los países de América Central que usan la norteamericana ATSC (Advanced Television System Committee).

La digitalización de la televisión conlleva la liberación de parte del espectro radioeléctrico cuyo uso está previsto para servicios de comunicaciones de cuarta generación 4G. Este proceso conocido como dividendo digital (DD), corresponde en América (Región 1 de la UIT) a la banda de 700 MHz (canales 52 al 69 de TV) [5]. Es así por ejemplo que en Colombia ya se ha producido este dividendo y la subasta del espectro 4G está pendiente [6], mientras que en Ecuador se dispuso que el Consejo Nacional de Telecomunicaciones (CONATEL), sea el organismo que defina las bandas para la implementación de la TDT y que la Secretaría Nacional de Telecomunicaciones (SENATEL), efectúe un análisis para ocupar el rango de 482-512 MHz [7]. Además la situación con respecto al manejo del espectro radioeléctrico se detalla en [8,9].

2. Tecnologías para proporcionar conectividad a Internet en zonas rurales

Con el despliegue de redes 4G en la banda de 700 MHz se espera mejorar el acceso a Internet en zonas rurales, gracias a las mejores condiciones de propagación y penetración que presentan. Sin embargo, es probable que dicho despliegue no sea suficiente para garantizar el acceso en zonas con baja densidad poblacional, por lo que la única posibilidad sería utilizar conectividad vía satélite, cuyo elevado costo no permitiría universalizar el servicio.

Bajo este contexto, una posible solución para lograr acceso universal a Internet en zonas rurales de Latinoamérica sería la utilización de un canal TDT con retorno inalámbrico, con lo que se reutilizaría la infraestructura de televisión digital para aumentar la penetración a Internet y disminuir los costos de implementación, puesto que llegaría a todos los hogares que tengan acceso a la TDT.

Existe una opción en el estándar europeo, DVB-RCT (Return Channel Terrestrial) que proporciona un canal de retorno inalámbrico para la tecnología de TDT DVB-T [10]. Sin embargo esta nunca llegó a desplegarse comercialmente debido al deficiente apoyo de la industria, a la oposición de los operadores de telefonía móvil, y a que se necesitan cambios en la regulación del uso del espectro de radiodifusión.

En este artículo se propone mejorar el canal de retorno UL (uplink), a nivel de capa física y MAC desarrollando una segunda-generación del estándar DVB-RCT, llamada DVB-RCT2 (Return Channel Terrestrial 2nd Generation), para lo cual se utilizará los principios de DVB-RCT y se aprovechará los avances tecnológicos de DVB-T2 para aplicarlos en el UL. En este sentido DVB-RCT2 podrá ser utilizado en DL (downlink) con cualquier tecnología de TDT presente en América Latina (ISDB-Tb, DVB-T o DVB-T2) si se usan el transmisor y decodificador adecuados, mientras que para el UL se utilizaría la propia tecnología DVB-RCT2. En la Fig. 1 se ilustra la arquitectura RCT2.

La tecnología DVB-RCT2 es definida como una red inalámbrica de difusión bidireccional, puesto que para la comunicación desde el proveedor del servicio hasta el usuario, se utiliza un canal de difusión (forward interaction path), pudiendo emplear en el DL, cualquiera de las tecnologías antes descritas, que en el caso de Ecuador sería ISDB-Tb, utilizando modulación OFDM normal, mientras que para la comunicación inversa se emplearía un canal de retorno (return interaction path), que utiliza propiamente la arquitectura DVB-RCT2 utilizando modulación OFDMA (Orthogonal Frequency Division Multiple Access). Específicamente en el canal de retorno se utiliza la capa de control de acceso al medio MAC (Medium Access Control), que está constituida sobre la capa física para proveer seguridad y un canal de retorno inalámbrico compartido, entre el terminal de usuario (RCTT - Return Channel Terrestrial), y la estación base. El protocolo MAC ofrece un uso transparente de la capa física a los protocolos de capas superiores, habilitando en la estación base el adaptador de red interactivo (INA – Interactive Network Adapter), quien es el responsable de asignar los recursos de subida (ranuras de frecuencia y tiempo) a los RCTTs, cuando las capas superiores lo necesiten. Por su parte el RCTT, es responsable de solicitar más recursos, si la estación base no se los ha provisto, o le ha dado muy pocos. Dichos mensajes MAC de subida (interacción) pueden ser enviados como se detallará posteriormente como acceso de tasa fija, contención, y reserva en el canal de subida, y pueden ser encapsulados en células ATM o mapeados directamente dentro de una ráfaga física, en cambio que los datos de bajada (difusión), si se utilizaría ISDB-Tb, serían encapsulados en paquetes MPEG-4.

3. Escenario de simulación

Para la tecnología DVB-RCT2, tal como se muestra en la Fig. 2, se ha emulado en el laboratorio de TDT de la Universidad Nacional de Chimborazo (UNACH) las características del canal RCT2, utilizando la plataforma VillageFlow trabajando con una tarjeta DTA115, la misma que permite generar flujos de transporte (TS) multiestándar, para el caso específico de esta investigación se ha generado el TS para DVB-RCT2, mediante el cual se parametrizó tanto la capa física como la MAC del estándar propuesto.
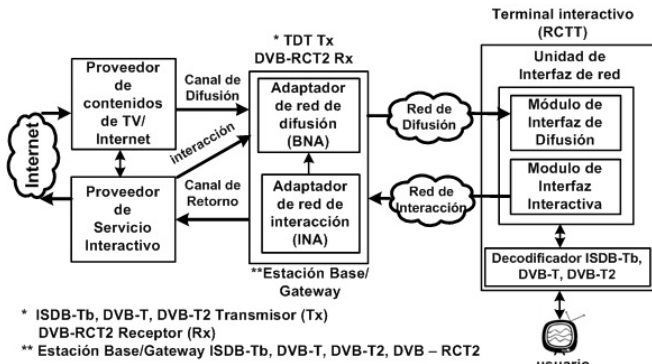
Figura 1. Arquitectura propuesta de DVB-RCT2.
Fuente: Elaboración propia con base en [11].

Figura 2. Escenario de pruebas.
Fuente: Elaboración propia

4. Capa física de DVB-RCT

4.1. Generalidades

Para proveer un canal de retorno inalámbrico compartido para DVB-T, DVB-RCT usa un canal de radiofrecuencia dedicado y lo organiza para permitir el acceso concurrente desde muchos terminales de usuario individuales (RCTTs). El método usado para organizar dicho canal de retorno es dividiéndolo en dominios de tiempo y frecuencia, obteniendo de esta manera una malla de ranuras tiempo/frecuencia, en donde cada ranura es asignable a un RCTT.

Para la organización del canal de retorno a nivel de capa física, el estándar provee dos conformados de señal; con el fin de dotar de inmunidad contra la interferencia entre portadora (ICI) y la interferencia entre símbolo (ISI) [11].

- **Conformado de Nyquist (Nyquist shaping).-** usa filtrado de Nyquist en tiempo sobre cada portadora, sirve también para brindar inmunidad contra jammers.
- **Conformado Rectangular (Rectangular shaping).-** usa una disposición ortogonal de las portadoras y un intervalo de guarda (GI) con valores de 1/4, 1/8, 1/16 y 1/32 entre los símbolos modulados, sirve además para combatir los efectos de la propagación multicamino.

El uso de los conformados es estrictamente exclusivo, es decir que estos no pueden ser utilizados al mismo tiempo en el canal de retorno.

DVB-RCT define dos tipos de tramas de transmisión (Transmission Frame - TF), las mismas que proveen las características necesarias para permitir la demodulación en la estación base.

- TF1.- se compone de un conjunto de símbolos OFDM, que contienen varios Subcanales de Datos, un Símbolo Nulo y una serie de Símbolos de Sincronización/Aleatorio.
- TF2.- se compone de un conjunto de símbolos OFDM de propósito general, que contienen ya sea Datos o Subcanales de Sincronización/Aleatoriedad.

Cada RCTT transmite ráfagas de datos, mediante las estructuras de ráfagas de datos (Burst Structure - BS), y lo hace en base a un número entero de celdas ATM (una celda ATM es el contenedor común usado para llevar el control MAC o los mensajes de datos MAC). Sea cual sea la tasa de codificación y la modulación física, las BS tienen una capacidad de ráfaga constante de 144 símbolos modulados.

DVB-RCT define también tres estructuras de ráfagas BS1, BS2 y BS3 que tienen sus propias características con respecto a la división de las ráfagas de datos y las portadoras piloto a través de las ranuras tiempo/frecuencia.

La asignación de las BS en las TF, es controlado por el proceso MAC que se ejecuta en la estación base, para esto existen tres métodos llamados Esquemas de Acceso al Medio (Medium Access Scheme - MAS), que a su vez están relacionados con los modos de acceso que tendrán los RCTTs para solicitar recursos a la estación base.

El TF1 es adecuado para los Esquemas de Acceso al Medio 1 y 2 (MAS1 y MAS2), que describen el método de asignación para BS1 y BS2 respectivamente. El TF2 en cambio es usado solamente para MAS3, y provee el método de asignación a ser usado en BS3 y BS2. En la Tabla 1 se definen las 4 combinaciones existentes para crear los MAS, y se indican los modos de acceso correspondientes.

4.2. Modos de transmisión DVB-RCT

DVB-RCT especifica 6 modos de operación definidos por el máximo número de portadoras usadas, dado por el tamaño de la FFT (1K y 2K), y su distancia entre-portadora (Carrier Spacing – CS). Solamente un modo de transmisión podrá ser implementado en un canal de radiofrecuencia. En la Tabla 2 se indican las distancias entre-portadora aproximadas para un canal de 6 MHz.

Tabla 1.
Esquemas de acceso al medio del estándar DVB-RCT.

Esquema de acceso	Trama de transmisión	Estructura de ráfaga	Modo de acceso
N/A	N/A	N/A	Acceso de sincronización
MAS1	TF1	BS1	Acceso de tasa fija
MAS2	TF1	BS2	Acceso de contención
MAS3	TF2	BS3	Acceso de reserva
MAS3	TF2	BS2 (opcional)	

Fuente: Elaboración propia.

Tabla 2.
Distancias entre-portadora para un canal de 6MHz.

CS1	0,84 kHz
CS2	1,67 kHz
CS3	3,35 kHz

Fuente: Elaboración propia.

Tabla 3.
Modos de transmisión de DVB-RCT para un canal de 6MHz.

Portadoras totales del Sistema (Tsc)	1024 (1K)	2048 (2K)
Portadoras útiles (Cu)	**842**	**1712**
Reloj del sistema RCT (T)	1,167 us	0,583 us
Duración del símbolo útil (Tu)	1195 us	1195 us
Distancia entre-portadora1 (Cs1)	837 Hz	837 Hz
Ancho de banda del canal RCT (Bu)	0,705 MHz	1,433 MHz
Reloj del sistema RCT (T)	0,583 us	0,292 us
Duración del símbolo útil (Tu)	597 us	597 us
Distancia entre-portadora2 (Cs2)	1674 Hz	1674 Hz
Ancho de banda del canal RCT (Bu)	1,410 MHz	2,866 MHz
Reloj del sistema RCT (T)	0,292 us	0,146 us
Duración del símbolo útil (Tu)	299 us	299 us
Distancia entre-portadora3 (Cs3)	3348 Hz	3348 Hz
Ancho de banda del canal RCT (Bu)	2,819MHz	5,732 MHz

Fuente: Elaboración propia con base en [11].

La distancia entre-portadora controla la robustez del sistema en lo que respecta a la posible desincronización de cualquier RCTT. Cada valor implica un tamaño de celda de transmisión máxima y brinda una resistencia al efecto Doppler cuando el RCTT está en movimiento.

En la Tabla 3, se indican los parámetros de los modos de transmisión aplicables para el estándar DVB-RCT utilizando un canal de radiofrecuencia de 6 MHz.

Cada combinación tiene una compensación específica entre la diversidad en frecuencia y la diversidad en tiempo, y además entre el rango de cobertura y la capacidad de portabilidad.

4.3. Tramas de transmisión (TF) de DVB-RCT

Las tramas de transmisión proveen al canal de radiofrecuencia de una estructura repetitiva, compuesta de un conjunto de ranuras tiempo/frecuencia, en las que Símbolos Nulos, Símbolos Sincronización/Aleatorios, Símbolos de Datos y Símbolos Pilotos son embebidos para proveer recursos para la sincronización y transmisión de datos.

El proceso MAC que se ejecuta en la estación base maneja los recursos provistos por las tramas de transmisión.

TF1 organiza el canal en el dominio del tiempo, usando conjuntos específicos de símbolos OFDM para transmitir símbolos nulos, símbolos aleatorios/sincronización y símbolos de usuario. Mientras que TF2 organiza el canal en el dominio de la frecuencia, donde un conjunto de 5 subcanales consecutivos (llamados subcanales aleatorios), son usados para sincronización, mientras que el resto de subcanales son usados para transmisión de datos.

Para organizar TF1 en el dominio del tiempo, la trama de transmisión puede llevar las siguientes tres categorías de símbolos.

- **Símbolo Nulo.**- ninguna transmisión ocurrirá en el primer símbolo OFDM de la trama de transmisión. Este símbolo provee la detección del jammer de la estación base receptora.
- **Símbolos Aleatorios.**- muchos símbolos consecutivos OFDM (6, 12, 24 o 48) son provistos para permitir características de aleatoriedad al RCTT.

- **Símbolos de usuario.**- parte de la trama de transmisión permite la transmisión de estructuras de ráfagas de datos (BS1 o BS2), las cuales incluyen datos de usuario y portadoras piloto. Tanto BS1 como BS2 hacen uso de un conjunto de portadoras, llamadas subcanales que se encuentran repartidos en todo el canal.

Además los símbolos de usuario pueden llevar una BS1 o cuatro BS2, y tanto BS1 como BS2 no podrán ser utilizadas al mismo tiempo.

Si los símbolos de usuario usan el conformado rectangular, el valor del GI será idéntico para todos los símbolos OFDM embebidos en la trama TF1. Si en cambio usa el conformado de Nyquist el valor del GI se aplicará sobre el Símbolo Nulo y los Símbolos Aleatorios y el mismo será de ¼.

TF2 por su parte, podrá llevar de las siguientes categorías de símbolos en el mismo símbolo OFDM.

- **Símbolos aleatorios.**- 8 intervalos aleatorios (hechos de 6 símbolos consecutivos), que brindan funciones de aleatoriedad.
- **Símbolos de usuario.**- llevan las estructuras de ráfagas de datos (BS3 o BS2), que incluyen portadoras de datos y pilotos. Tanto BS2 como BS3 hacen uso de un conjunto de portadoras, llamadas subcanales que se encuentran repartidos en todo el canal.

Los símbolos de usuario de la TF2 permite el uso de la BS3 u opcionalmente de la BS2. Cuando una BS2 es transmitida, esta deberá ser completada por un conjunto de 4 símbolos nulos modulados con el objetivo de alcanzar una duración igual a la duración de ocho BS3.

4.4. Estructuras de ráfagas de datos (BS) de DVB-RCT

El canal de retorno es dividido en una malla de ranuras de tiempo/frecuencia, en donde en cada ranura de tiempo se pueden usar diferentes frecuencias simultáneamente para transmitir. Esto se conoce como estructura de ráfaga de datos (BS) y es la ranura de transmisión básica asignada a un RCTT, en la Fig. 3 se ilustra el concepto.

El estándar define tres estructuras de ráfaga que brindan varias combinaciones diversas en tiempo y frecuencia, proporcionado de esta manera diferentes grados de robustez, duración de ráfaga y un amplio rango de capacidad tasa de bits al sistema.
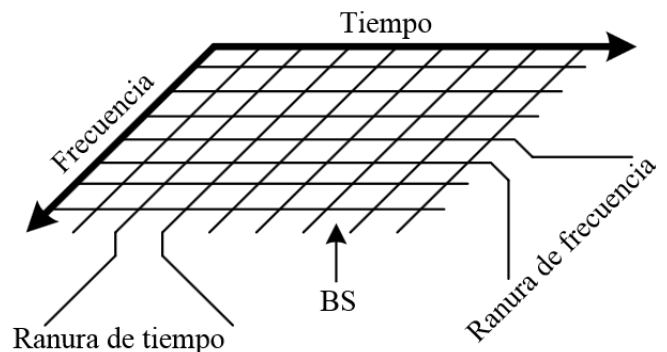
Figura 3. Ranuras de transmisión básicas de un RCTT.
Fuente: Elaboración propia.

Tabla 4.
Asignación de parámetros para BS1, opción sin salto de frecuencia.

	Rectangular	Nyquist
Símbolos de datos en una mini-ráfaga	36	36
Símbolos piloto en una mini-ráfaga	9	9
Símbolos de datos en una ráfaga	144	144
Símbolos piloto en una ráfaga	36	36
Símbolos nulos pre/post preámbulo	0	8
Total de símbolos OFDM en la portadora (tiempo de transmisión)	180	188

Fuente: Elaboración propia con base en [11].

Cada BS hace uso de un conjunto de portadoras llamadas subcanales. Uno o muchos subcanales pueden ser usados simultáneamente por un RCTT dependiendo de la asignación realizada por el proceso MAC.

BS1.- Está formada por una única portadora (1 subcanal) para llevar la ráfaga de datos total (144 símbolos de datos modulados) sobre la componente del tiempo, con un salto de frecuencia opcional aplicado dentro de la duración de la ráfaga.

Para la asignación de datos BS1 provee dos opciones, BS1 sin salto de frecuencia y BS1 con salto de frecuencia (opcional).

BS1 sin salto de frecuencia.- La asignación de datos será hecha secuencialmente, después de la decodificación, el entrelazo y la conversión del flujo binario a las constelaciones, sobre las 4 mini-ráfagas consecutivas de una ráfaga BS1. Cada una de las 4 mini-ráfagas, se asignarán a la misma portadora, en una secuencia de tiempo. La asignación de una portadora al subcanal, para esta opción, será conforme a los valores de la Tabla 4.

BS1 con salto de frecuencia (opcional).- para proveer diversidad de frecuencia para BS1, la opción de salto de frecuencia puede ser implementada para la transmisión de las mini-ráfagas. La asignación de datos será hecha secuencialmente, como en la opción sin salto de frecuencia.

Cada una de las 4 mini-ráfagas, serán asignadas a 4 portadoras diferentes. La asignación de una portadora a un subcanal, con la opción de salto de frecuencia, será conforme los valores de la Tabla 5.

BS2.- usará simultáneamente 4 portadoras (4 subcanales) llevando cada una, la cuarta parte del total de ráfaga de datos (36 símbolos de datos modulados por portadora).
BS2 llevará 144 símbolos de datos, organizados en 4 ráfagas que serán transmitidas simultáneamente en una portadora por separado. Las 4 portadoras usadas forman un subcanal. El subcanal será definido de acuerdo al uso de BS2 ya sea en el MAS2 o en el MAS3 (opcional), mientras que la asignación de datos será hecha secuencialmente como en BS1, así como la inserción de pilotos.

Tabla 5.
Asignación de parámetros para BS1, opción con salto de frecuencia.

	Rectangular	Nyquist
Símbolos de datos en una mini-ráfaga	36	36
Símbolos piloto en una mini-ráfaga	9	9
Símbolos de datos en una ráfaga	144	144
Símbolos piloto en una ráfaga	36	36
Símbolos nulos pre/post preámbulo	0	32
Total de símbolos OFDM en la portadora (tiempo de transmisión)	180	212

Fuente: Elaboración propia con base en [11].

La asignación de portadoras, será conforme los valores de la Tabla 6.

Cuando se usa MAS2, la asignación de las 4 portadoras al subcanal seguirá los valores descritos en la Tabla 7.

Cuando se usa MAS3, la asignación de las 4 portadoras al subcanal quedará según lo definido en la Tabla 8.

BS3.- usará simultáneamente 29 portadoras (29 subcanales) llevando cada una, la vigésima novena parte del total de ráfaga de datos (aproximadamente 4,97 símbolos de datos modulados por portadora). Es decir BS3 llevará 144 símbolos de datos transmitidas sobre 29 portadoras, mismas que forman un subcanal. Esta estructura de ráfaga puede usarse solamente con conformado de señal Rectangular.

Para la asignación de portadoras, BS3 usará los parámetros del dominio del tiempo y del dominio de la frecuencia descritos en las Tablas 9 y 10 respectivamente.

Tabla 6.
Asignación de parámetros para conformados de señal Rectangular y Nyquist, subcanal BS2.

	Rectangular	Nyquist
Símbolos de datos en la ráfaga	144	144
Símbolos piloto en la ráfaga	36	36
Símbolos nulos pre/post preámbulo	0	32
Número total de símbolos de usuario	176	208
Total de símbolos OFDM por portadora (tiempo de transmisión)	45	53

Fuente: Elaboración propia con base en [11].

Tabla 7.
Asignación de parámetros cuando se MAS2, subcanal BS2 de DVB-RCT.

Modos	1K	2K
Número de portadoras en un símbolo OFDM	1024	2048
Portadoras totales	842	1712
Número máximo de portadoras útiles (4 portadoras)	840	1708
Máximo número de portadoras útiles por cada portadora (número subcanales BS2)	210	427
Numeración de las portadoras	de 0 a 841	de 0 a 1711
Banda de guarda en la izquierda	91	168
Banda de guarda en la derecha	91	168
Número de portadoras excluidas	0, 421	0, 1, 856,1711
Numeración del subcanal: SCN	1 al 210	2 al 428
X	210	427
Y	420	855
Z	630	1282

Fuente: Elaboración propia con base en [11].

Tabla 8.
Asignación de parámetros cuando se MAS3, subcanal BS2 de DVB-RCT (conformado Rectangular).

	Modo 1K	Modo 2K
Número de portadoras en un símbolo OFDM	1024	2048
Numeración del subcanal: SubcanalBS2	1..7, 11..17, 21..27,..281..287	1..7, 11..17, 21..27,..581..587

Fuente: Elaboración propia con base en [11].

Tabla 9.
Asignación de parámetros dominio del tiempo, conformado de señal Rectangular, subcanal BS3 de DVB-RCT.

	Rectangular	Nyquist
Símbolos de datos en la ráfaga	144	No aplica
Símbolos piloto en la ráfaga	30	No aplica
Número total de símbolos de usuario	174	No aplica
Total de símbolos OFDM por portadora (tiempo de transmisión)	6	No aplica

Fuente: Elaboración propia con base en [11].

Tabla 10.
Asignación de parámetros dominio de la frecuencia, conformado de señal Rectangular, subcanal BS3 de DVB-RCT.

	Modo 1K	Modo 2K
Número de portadoras en un símbolo OFDM	1024	2048
Portadoras totales	842	1712
Numeración de las portadoras	de 0 a 841	de 0 a 1711
Banda de guarda en la izquierda	91	168
Banda de guarda en la derecha	91	168
Número de portadoras excluidas	421	856
Número de subcanales BS3	29	59
Numeración del subcanal: SCN	0, 10, 20, 30, ..280	0, 10, 20, 30, ..580

Fuente: Elaboración propia con base en [11].

Tabla 11.
Parámetros de TF1, BS1 para conformado de señal Rectangular y Nyquist (MAS1).

Conformados	Rectangular con y sin salto de frecuencia	Nyquist sin salto de frecuencia	Nyquist con salto de frecuencia
Símbolos OFDM de usuario	180	188	212
Símbolos OFDM nulos	1	1	1
Símbolos OFDM aleatorios/sincronización	6, 12, 24, o 48	6, 12, 24, o 48	6, 12, 24, o 48
Número total de símbolos en una TF	187, 193, 205, o 229	195, 201, 213, o 237	219, 225, 237, o 261

Fuente: Elaboración propia con base en [11].

En BS2 y BS3 las portadoras no son consecutivas, pero están esparcidas a través de todo el canal de radiofrecuencia.

Las 3 estructuras de ráfaga proveen un esquema de modulación asistida de piloto que permite la detección coherente en la estación base. La relación de inserción de pilotos es de aproximadamente 1/6, lo que significa que un patrón de portadoras piloto (PPP) es insertado cada 5 portadoras de datos.

4.5. Esquemas de acceso al medio (MAS) de DVB-RCT

El estándar DVB-RCT ofrece 2 TF y 3 BS, los MAS1, MAS2 y MAS3 proveen la asignación de una o más BS sobre una TF. El número de símbolos usados, símbolos de datos y símbolos de sincronización que constituyen la TF para MAS1, MAS2 y MAS3 deberá cumplir con los valores dados en las Tablas 11, 12 y 13 respectivamente.

5. Capa MAC de DVB-RCT

5.1. Modos de acceso

El usuario envía las peticiones a través del Módulo de Interfaz Interactiva hacia el adaptador de red de interacción INA (Interactive Network Adapter) en la estación base, y esta responde a las peticiones desde el BNA (Broadcast Network Adapter), hacia el módulo de Interfaz de Difusión del RCTT.

La capa MAC en el UL proporciona seguridad y compartición del canal entre los RCTTs y la estación base. El INA ejecuta el proceso controlador de la capa MAC para asignar los recursos de subida (ranuras de frecuencia y tiempo) a los RCTTs. Dicho controlador realiza tres procesos fundamentales: sincronización inicial, mantenimiento de la sincronización, y asignación de recursos.

La sincronización inicial para cada RCTT se realiza en tiempo, potencia y frecuencia. Para sincronizarse en tiempo y potencia cada RCTT demodula la señal de difusión DL que provee la estación base obteniendo información sobre la operación de la red, luego de lo cual transmite un código aleatorio de los 32 definidos para tareas de inicialización, mantenimiento o solicitud de ancho de banda en el subcanal de servicio (subcanal de subida predefinido por el controlador MAC mediante el cual entra en la red, y que luego sirve para para mantener la sincronización durante la conexión y para solicitar ancho de banda). Dichos códigos son CDMA (Code Division Multiple Access) por lo que la estación base es capaz de identificar varios códigos recibidos

Tabla 12.
Parámetros de TF1, BS2 para conformado de señal Rectangular y Nyquist (MAS2).

Conformados	Rectangular	Nyquist
Símbolos OFDM de usuario	176	208
Símbolos OFDM nulos	1	1
Símbolos OFDM aleatorios/sincronización	6, 12, 24, o 48	6, 12, 24, o 48
Número total de símbolos en una TF	187, 193, 205, o 229	219, 225, 237, o 261

Fuente: Elaboración propia con base en [11].

Tabla 13.
Parámetros de TF2, con BS2 y BS3 para conformado de señal Rectangular (MAS3).

Conformados	Rectangular con BS2 (opcional)	Rectangular con BS3
Símbolos OFDM de usuario	44	48
Símbolos OFDM nulos	1	8
Símbolos OFDM aleatorios/sincronización	0	0
Número total de símbolos en una TF	48	48

Fuente: Elaboración propia con base en [11].

al mismo tiempo. La sincronización en frecuencia se realiza al final y se deriva de la señal emitida por la estación base. Además el controlador MAC define cuatro modos de acceso en el UL; (i) Acceso de sincronización, para sincronizar inicialmente al RCTT y mantener la sincronización durante toda la conexión; (ii) Acceso de tasa fija, para asignar al RCTT una serie de ranuras cuando se establece la conexión; (iii) Acceso de contención, para enviar mensajes MAC en las ranuras no reservadas que el INA identifica en el DL; y (iv) Acceso de reserva, para conceder a un RCTT ranuras mediante mensajes MAC o por motivos de sincronización.

El controlador MAC continuamente monitorea el UL, asignando el MODCOD que deberán tener los RCTTs por conexión y verificando que estén dentro de los umbrales de sincronización predefinidos; si se detecta algún equipo fuera de estos parámetros, empieza un proceso de recalibración.

Una vez efectuada la sincronización inicial, se establece una conexión entre el RCTT y la estación base, la misma que envía en DL un reconocimiento al RCTT para identificarlo.

Para solicitar ancho de banda (ranuras) el RCTT envía un mensaje MAC utilizando los modos de acceso de tasa fija, contención o reserva, según la configuración de la estación base. Dichas ranuras le son asignadas en DL mediante los subcanales de datos. Si la asignación no es obtenida se vuelve a transmitir la solicitud.

6. Capa física de DVB-RCT2

6.1. Generalidades

El nuevo estándar DVB-RCT2 está pensado para proveer un canal de retorno inalámbrico para las tecnologías de DL ISDB-Tb, DVB-T y DVB-T2, siendo una propuesta innovadora que podría ser utilizada en toda Latinoamérica.

Para la organización del canal de retorno a nivel de capa física, DVB-RCT2 al igual que su predecesor DVB-RCT tendrá dos conformados de señal; con el fin de dotar de inmunidad contra ICI e ISI.

- **Conformado de Nyquist (Nyquist shaping).-** usa filtrado de Nyquist en tiempo sobre cada portadora, sirve también para brindar inmunidad contra jammers.
- **Conformado Rectangular (Rectangular shaping).-** usa una disposición ortogonal de las portadoras y un GI con valores de 1/4, 19/128, 1/8, 19/256, 1/16, 1/32, 1/128 entre los símbolos modulados.

Dependiendo del modo de transmisión usado, la señal total en el aire se compondrá de un conjunto de portadoras, basadas en DVB-T2 (1K, 2K, **4K, 8K, 8K extendido (ext), 16K, 16K extendido, 32K, 32K extendido**) adyacentes moduladas sincrónicamente por el RCTT activo. Cabe destacar que los valores en negrita son tomados del estándar DVB-T2 y originalmente no existen en el estándar DVB-RCT. Además al igual que DVB-RCT, DVB-RCT2 utilizará dos tramas de transmisión (TF), tres estructuras de ráfagas de datos (BS1, BS2, BS3), y tres esquemas de acceso al medio (MAS1, MAS2, MAS3).

6.2. Modos de transmisión DVB-RCT2

DVB-RCT2 especificará 27 modos de operación definidos por el máximo número de portadoras usadas, dado por el tamaño de la FFT (1K, 2K, 4K, 8K, 8K ext, 16K, 16K ext, 32K, 32K ext), y su distancia entre-portadora (Carrier Spacing – CS). Se debe considerar como en DVB-RCT que solo un modo de transmisión podrá ser implementado en un canal de radiofrecuencia. En la Tabla 14, se indican los parámetros de los modos de transmisión aplicables para el estándar DVB-RCT2 utilizando un canal de radiofrecuencia de 6 MHz, en donde:

- **Portadoras totales del sistema (Tsc):** es el número total de portadoras manejadas por el estándar DVB-RCT2.
- **Portadoras útiles (Cu):** es el máximo número de portadoras efectivamente usadas por el RCTT.
- **Reloj del sistema RCT2 (T):** se deriva del estándar DVB-T2, en donde el reloj del sistema definido para un canal de 6MHz es de 48/7 MHz o 7/48 us. Por consiguiente el reloj del sistema para DVB-RCT2 será:
  - En el caso de CS1, 4 veces el período del reloj del sistema de DVB-T2. Para CS2, 2 veces el período del reloj del sistema de DVB-T2. Y para CS3, 1 vez el período del reloj del sistema de DVB-T2;
- **Duración del símbolo útil (Tu):** es el período útil del símbolo, expresado como: $Tu = Tsc * T$
- **Distancia entre-portadora 1,2,3 (Cs1, Cs2, Cs3):** viene expresada como: $Cs= 1/Tu$
- **Ancho de banda del canal RCT2 (Bu):** es usado en el canal RCT2, viene expresado como: $Bu = Cs_{1,2,3}*Cu.$

A modo de ejemplo, en las tablas a definirse a continuación se representarán los valores para los modos 4, 8, 16, y 32K.

6.3. Estructuras de ráfagas de datos (BS) DVB-RCT2

Al igual que en DVB-RCT, para DVB-RCT2 el subcanal de BS1 está compuesto de 4 mini-ráfagas que a su vez tienen portadoras de datos y portadoras piloto, estas últimas con un patrón de distribución específico. La asignación de una portadora a la primera mini-ráfaga del subcanal será conforme con el valor dado en la Tabla 15.

En lo referente a BS2, en DVB-RCT2 el subcanal será definido de acuerdo al uso de esta estructura de ráfaga ya sea en el MAS2 o en el MAS3 (opcional). Cuando se usa MAS2, la asignación de las 4 portadoras al subcanal seguirá los valores de a Tabla 16.

Tabla 14.
Modos de transmisión DVB-RCT2, canal 6 MHz.

Tsc	4096(4K)	8192 (8K)	16384 (16K)	32768 (32K)
Cu	3409	6817	13633	27265
T	0,291 us	0,145 us	0,072 us	0,036 us
Tu	1195 us	1195 us	1195 us	1195 us
Cs1	837 Hz	837 Hz	837 Hz	837 Hz
Bu	2,853 MHz	5,705 MHz	11,410 MHz	22,820 MHz
T	0,145 us	0,072 us	0,036 us	0,018 us
Tu	597 us	597 us	597 us	597 us
Cs2	1674 Hz	1674 Hz	1674 Hz	1674 Hz
Bu	5,706 MHz	11,411 MHz	22,821 MHz	45,641 MHz
T	0,072 us	0,036 us	0,018 us	0,009 us
Tu	298 us	298 us	298 us	298 us
Cs3	3355 Hz	3355 Hz	3355 Hz	3355 Hz
Bu	11,437 MHz	22,871 MHz	45,738 MHz	91,474 MHz

Fuente: Elaboración propia.

Tabla 15.
Parámetros conformado Rectangular y Nyquist, subcanal BS1 DVB-RCT2.

Modos	4K	8K	16K	32K
Número virtual de portadoras en un símbolo OFDM	4096	8192	16384	32768
Portadoras totales	3409	6817	13633	27265
Banda de guarda en cada borde	346	687	1375	2751
Número máximo de portadoras útiles (número subcanales BS1)	3403	6810	13625	27256
Número de portadoras excluidas	0, 1, 426, 852, 1704, 3408	0, 1, 426, 852, 1704, 3408, 6816	0, 1, 426, 852, 1704, 3408, 6816, 13632	0, 1, 426, 852, 1704, 3408, 6816, 13632, 27264
Numeración del subcanal	2 al 425; 427 al 851; 853 al 1703; 1705 al 3407	2 al 425; 427 al 851; 853 al 1703; 1705 al 3407; 3409 al 6815	2 al 425; 427 al 851; 853 al 1703; 1705 al 3407; 3409 al 6815; 6817 al 13631	2 al 425; 427 al 851; 853 al 1703; 1705 al 3407; 3409 al 6815; 6817 al 13631; 13633 al 27263

Fuente: Elaboración propia.

Tabla 16.
Parámetros MAS2, subcanal BS2 de DVB-RCT2.

Modos	4K	8K	16K	32K
Número de portadoras en un símbolo OFDM	4096	8192	16384	32768
Portadoras totales	3409	6817	13633	27265
Número máximo de portadoras útiles (4 portadoras)	3403	6810	13625	27256
Máximo número de portadoras útiles por cada portadora (número subcanales BS2)	850	1702	3406	6814
Numeración de las portadoras	de 0 a 3408	de 0 a 6816	de 0 a 13632	de 0 a 27264
Banda de guarda en la izquierda	346	687	1375	2751
Banda de guarda en la derecha	346	687	1375	2751
Número de portadoras excluidas	0, 1, 426, 852, 1704, 3408	0, 1, 426, 852, 1704, 3408, 6816	0, 1, 426, 852, 1704, 3408, 6816, 13632	0, 1, 426, 852, 1704, 3408, 6816, 13632, 27264
Numeración del subcanal: SCN	1 al 850	1 al 1702	1 al 3406	1 al 6814
X	850	1702	3406	6814
Y	1703	3407	6815	13631
Z	2553	5109	10221	20445

Fuente: Elaboración propia.

Las siguientes fórmulas definen la indexación de la portadora dentro de una asignación de un subcanal para los modos desde 1K hasta 32K ext.

- Portadora 0 = SCN
- Portadora 1 = SCN + X
- Portadora 2 = SCN + Y
- Portadora 3 = SCN + Z

Por su parte la asignación de portadoras para MAS3, sigue los mismos principios que en DVB-RCT, en este sentido la asignación de portadoras a los subconjuntos de BS2 es obtenida por la ec. (1):

$$\text{Subconjunto\# (SCN,}n) = \text{Portadora\#}(n\text{-}1), \text{Portadora\#}(n+6), \text{Portadora\#}(n+13), \text{Portadora\#}(n+20) \quad (1)$$

Donde n indica el número de subconjunto para la asignación del subcanal BS2 dentro de los subconjuntos del subcanal BS3 (7 en total), mediante la ec. (2):

$$\text{Subcanal BS2\#} = \text{SCN}+n; \text{ donde } 1<=n<=7 \quad (2)$$

El número de subcanales en BS3 está hecho en múltiplos de 10 para permitir la numeración de los subcanales de BS2 dentro de los subcanales de BS3 para el MAS3. Los parámetros de dicha asignación se definen en la Tabla 17.

Tabla 17.
Parámetros subcanal BS2, MAS3 en DVB-RCT2 (conformado Rectangular).

Modos	4K	8K	16K	32K
Número de portadoras en un símbolo OFDM	4096	8192	16384	32768
Numeración del subcanal: Subcanal BS2	1..7..11..17, 1161..1167	1..7..11..17, 2341..2347	1..7..11..17, 4691..4697	1..7..11..17, 9391..9397

Fuente: Elaboración propia.

Tabla 18.
Parámetros dominio de la frecuencia, conformado Rectangular, subcanal BS3 de DVB-RCT2.

Modos	4K	8K	16K	32K
Número de portadoras en un símbolo OFDM	4096	8192	16384	32768
Portadoras totales	3409	6817	13633	27265
Numeración de las portadoras	de 0 a 3408	de 0 a 6816	de 0 a 13632	de 0 a 27264
Banda de guarda en la izquierda	346	687	1375	2751
Banda de guarda en la derecha	346	687	1375	2751
Número de portadoras excluidas	1704	3408	6816	13632
Número de subcanales BS3	117	235	470	940
Numeración del subcanal: SCN	0,10, 20..1160	0,10, 20..2340	0,10, 20..4690	0,10, 20..9390

Fuente: Elaboración propia.

Por otro lado, la asignación de portadoras de BS3 en el dominio del tiempo para DVB-RCT2 será prácticamente la misma que para DVB-RCT, considerando como en los casos anteriores los tamaños de FFT del nuevo estándar propuesto; en lo referente al dominio de la frecuencia, los parámetros están descritos en la Tabla 18.

6.4. Esquemas de acceso al medio (MAS) DVB-RCT2

Al igual que en DVB-RCT, DVB-RCT2, utilizará las dos TF, las tres BS, y los tres MAS, en donde estos últimos serán usados para proveer la asignación de una o más BS sobre una TF. Para todas las BS, la estación base controla el número de subcanales simultáneamente asignados a un RCTT, con las siguientes limitaciones:

- Para conformado de señal Nyquist no más de 64 subcanales podrán ser asignados a un RCTT. Además, cuando muchos subcanales son asignados, la numeración de sus subcanales deberá ser consecutiva.
- Para conformado de señal Rectangular no más de 128 subcanales usando modulación 64QAM, o 192 subcanales usando 16QAM o 384 subcanales usando QPSK, podrán ser asignados a un RCTT.

7. Capa MAC de DVB-RCT2

7.1. Modos de acceso DVB-RCT2

Aunque DVB-RCT2 tenga los mismos modos de acceso que el estándar DVB-RCT: tasa fija, contención, reserva, y sincronización. Para proveer de Internet a zonas rurales mediante DVB-RCT2 se utilizará el modo de acceso de sincronización, puesto que tiene un rendimiento excelente para servicios con altas tasas de datos, además de un rendimiento considerable para servicios con bajas tasas de datos, con respecto al modo de acceso de contención [10], reduciendo además las posibles colisiones que podrían existir en el subcanal de sincronización (más no en el de datos). Además dado que los códigos aleatorios que se envían son códigos CDMA, la estación base es capaz de identificar muchos códigos sin ambigüedad al mismo tiempo. Esta característica no es posible en el modo de acceso de contención, donde un considerable número de subcanales de datos deben ser dedicados únicamente para solicitudes de ancho de banda, así los nuevos usuarios que llegan no encontrarán a todos los subcanales con cargas de tráfico elevadas. Otra ventaja de este modo es la baja capacidad que se utilizaría en el DL para implementar el protocolo MAC, debido a que la estación base no necesita informar sobre las ranuras en UL no usadas, ni sobre las posibles colisiones. Cabe señalar que la estación base en lugar de esto, realiza reconocimientos de los códigos de sincronización recibidos correctamente para asignar ancho de banda a los RCTTs.

8. Resultados

8.1. Prestaciones DVB-RCT2 con respecto a DVB-RCT

En la Tabla 19 se presentan los valores de CNR y capacidades de transmisión para DVB-RCT y DVB-RCT2,

Tabla 19.
Valores de CNR y capacidades de transmisión en Mbps para DVB-RCT y DVB-RCT2.

MODCOD	DVB-RCT		DVB-RCT2	
	CNR	Capacidad	CNR	Capacidad
QPSK ½	4,6	4,26	2,67	5,64
QPSK ¾	7,8	6,40	5,87	8,49
QPSK 5/6	--	--	7,07	9,46
16QAM ½	10,6	8,53	7,87	11,34
16QAM ¾	14,0	12,80	11,87	17,03
16QAM 5/6	--	--	13,17	18,91
64QAM ½	15,7	12,80	12,27	16,98
64QAM ¾	19,6	19,21	16,91	25,52
64QAM 5/6	--	--	18,68	28,43
256QAM ½	--	--	16,31	22,67
256QAM ¾	--	--	22,06	34.07
256QAM 5/6	--	--	24,38	37,88

Fuente: Elaboración propia.

considerando diferentes MODCODs [12],, ancho de banda de 6MHz, método de recepción fija (Perfil de Canal Rice) y los siguientes modos de transmisión. Para DVB-RCT; FFT 2K, GI 1/32, para DVB-RCT2; FFT 32K extendido, GI 1/128, PP7, y en común para las dos tecnologías; conformado de tráfico rectangular, BS3, CS1 y esquema de acceso de sincronización.

En los modos de transmisión descritos anteriormente se consideró la FFT de mayor tamaño permitida en cada tecnología junto con el GI más pequeño, esto debido a que dicho GI sólo se lo utilizará para proteger la señal de ecos que pudieran existir en la transmisión del mismo usuario.

Además en los dos estándares se utilizó el modo de acceso de sincronización, que reduce las colisiones durante la transmisión. La distancia entre portadora CS1, para obtener un tamaño máximo de celda de transmisión. BS3, debido a que el protocolo MAC del estándar 802.16a (Wimax), está basado en OFDMA y usa una capa física similar a la del estándar DVB-RCT. Es importante recalcar que BS3 cuenta con 59 subcanales, de los cuales en este estudio se utilizaron 5 para formar el subcanal de sincronización y los 54 restantes fueron utilizados para transmisión de datos. Con lo cual se logró mejorar el rendimiento del sistema DVB-RCT2 al utilizar 5 y no 4 subcanales en el subcanal de sincronización como lo hacía DVB-RCT. Además se utilizó el conformado de tráfico rectangular para combatir los efectos de la multipropagación.

Con las consideraciones anteriormente mencionadas, se observa que DVB-RCT2 proporciona mayor capacidad de transmisión (37,88 Mbps en 6 MHz), mientras que la capacidad máxima de DVB-RCT es de 19,21 Mbps. Así mismo DVB-RCT2 Presenta una mínima CNR (2,67 dB para QPSK 1/2), mientras que la CNR mínima de DVB-RCT es 4,6 dB (QPSK ½).

9. Conclusiones

Al estar basado en DVB-RCT, el estándar de TDT propuesto DVB-RCT2, seguirá los mismos principios que su predecesor, sin embargo al utilizar características de DVB-T2 y combinarlas con DVB-RCT, se logra que este nuevo estándar tenga capacidades superiores en cuanto a transmisión utilizando una CNR pequeña, así como a nivel

de capa física y MAC, contando por ejemplo con 27 modos de operación definidos por el máximo número de portadoras usadas, en comparación con los 6 modos con lo que trabaja DVB-RCT.

Para el envío de las peticiones de usuario hacia la estación base (UL), RCT2 utilizará un canal de retorno adaptado de la arquitectura DVB-RCT, mientras que para el (DL) se podría usar las arquitecturas ISDB-Tb, DVB-T o DVB-T2, lo que permitirá que RCT2 pueda ser utilizado en toda Latinoamérica, lo cual constituye una alternativa interesante para brindar acceso a Internet a zonas rurales, puesto que ofrecería un alto porcentaje de penetración y abarataría los costos de despliegue con respecto a otras soluciones, puesto que podría reutilizar infraestructura de la TV analógica y/o digital.

Referencias

[1] UNESCO., Conferencia General 36ª reunión, 36C/54, Reflexión y análisis sobre Internet, París, 2011, 13 P.

[2] UIT., Informe sobre el desarrollo mundial de las telecomunicaciones/TIC de 2010, Verificación de los objetivos de la CMSI, Examen Intermedio, Resumen Ejecutivo, Suiza, 2010, 28 P.

[3] Rojas, P., La Televisión Digital Terrestre y su incidencia en las estaciones televisivas ecuatorianas. Advicom Cia. Ltda. [Online]. 4, 2012. [fecha de referencia 30 de mayo de 2015]. Disponible en: http://www.advicom.ec/userFiles/files/Publicaciones/La%20TDT%20y%20su%20incidencia%20en%20las%20estaciones%20televisivas%20ecuatorianas-ADVICOM-PRV-FINAL%20%2826-Marzo-2012%29.pdf

[4] Martínez, F., Uribe, G. y Mosquera, F., OneWeb: Plataforma de adaptación de contenidos Web basada en las recomendaciones del W3C Mobile Web Initiative. Ingeniería e Investigación. [Online]. 13(1), 2011. [fecha de referencia 10 de junio de 2015]. Disponible en: http://www.bdigital.unal.edu.co/23588/1/20571-69500-1-PB.pdf

[5] Kholod, A. y Lewis, J., El dividendo digital: Oportunidades y retos. Actualidades de la UIT [Online]. 2010. [fecha de referencia 17 de junio de 2015]. Disponible en: http://www.itu.int/net/itunews/issues/2010/01/27-es.aspx

[6] Martinez-Pinzón, G., Primera fase de optimización de la red de televisión digital terrestre en Colombia con el nuevo estándar DVB-T2. Tesis de Maestría, Instituto de Telecomunicaciones y Aplicaciones Multimedia. Universidad Politécnica de Valencia, Valencia, España, 2013.

[7] CONATEL., Resolución RTV-38-02-CONATEL-2012, Ecuador, Consejo Nacional de Telecomunicaciones, Ecuador, 2012, 3 P.

[8] CONARTEL., Norma Técnica para el Servicio de Televisión Analógica y Plan de Distribución de Canales (Resolución No.1779-CONARTEL-01), Ecuador, Consejo Nacional de Radiodifusión y Televisión , 2001, 11P.

[9] Cruz-Herrera, D., Análisis de disponibilidad de espectro radioeléctrico para la transición de Televisión Analógica a Televisión Digital Terrestre en el Ecuador. Tesis de Ingeniería, Departamento de Eléctrica y Electrónica. Escuela Superior Politécnica del Ejército, Sangolqui, Ecuador, 2011.

[10] Fraile-Gil F. y Gómez-Barquero, D., On the Capability of DVB-RCT to provide Interactive Services. Tesis de Maestría, Departamento de Tecnología. Universidad de Gävle, Gävle, Suecia, 2004.

[11] ETSI., EN 301 958 v1.1.1, Digital Video Broadcasting (DVB); Interaction Channel for Digital Terrestrial Television (RCT) incorporating Multiple Access OFDM, Francia, 2002, 164P.

[12] Castellanos, G., Perspectivas de utilización de los métodos de modulación digital con fase continua en los sistemas de telecomunicaciones modernos. Ingeniería e Investigación [Online]. 29,(993). [fecha de referencia 2 de junio de 2015]. Disponible en: http://www.bdigital.unal.edu.co/23755/1/20799-70339-1-PB.pdf

C.D. Radicelli-García, se graduó de Ing. en Sistemas Informáticos en 2006 y de MSc. en Interconectividad de Redes en 2010 en la Escuela Superior Politécnica de Chimborazo (ESPOCH), Ecuador. Obtuvo el título de MSc.en Tecnologías, Sistemas y Redes de Comunicación en 2013 en la Universidad Politécnica de Valencia (UPV), España, y actualmente es estudiante de Doctorado (PhD) en Telecomunicación en la UPV. Es además profesor de la Facultad de Ciencias de la Educación, Humanas y Tecnologías de la Universidad Nacional de Chimborazo (UNACH), Ecuador. Sus intereses de investigación están centrados en TDT.
ORCID: 0000-0001-9188-0514

N. Cardona-Marcet, recibió en 1990 el título de MSc. en Telecomunicaciones de la Universidad Politécnica de Cataluña, España y en 1995 el título de Dr. en Telecomunicaciones de la UPV. Desde 1990 está vinculado con la UPV donde es profesor catedrático y director del grupo de comunicaciones móviles (GCM) del Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM). El profesor Cardona ha sido vicepresidente del COST273, ha estado a cargo del WG3 del COST2100 en el área redes de acceso vía radio, y actualmente preside el EU COST IC1004.
ORCID: 0000-0001-5697-1453.
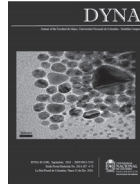
High performance control of a three-phase PWM rectifier using odd harmonic high order repetitive control

Germán Ramos [a], Iván Dario Melo-Lagos [b] & Jenny Cifuentes [c]

[a] Departamento de Ingeniería Eléctrica y Electrónica, Universidad Nacional de Colombia, Bogotá, Colombia. garamosf@unal.edu.co
[b] Departamento de Ingeniería Eléctrica y Electrónica, Universidad Nacional de Colombia, Bogotá, Colombia. idmelol@unal.edu.co
[c] Programa de Ingeniería Eléctrica, Universidad de La Salle, Bogotá, Colombia. jacifuentesq@unal.edu.co

Abstract
The control goal for three-phase pulse width-modulated rectifiers focuses on generating sinusoidal input currents and regulating the DC output voltage. Despite the fact that control strategies such as resonant and repetitive control have been proposed in recent works, with many notable results on the area, they have significant performance decay when the frequency changes in the exogenous signal. In this paper, it is shown that the use of an Odd Harmonic High Order Repetitive Controller can be used to control the three-phase rectifier current loops with a performance that is considerably superior to traditional alternatives developed in this field. This compensator's Odd Harmonic property keeps a computational complexity similar to that of the conventional repetitive controllers but it has the advantage of increasing the robustness when the signal frequency varies. Simulation and experimental results show the high performance that was obtained even in the case of deviation of network frequency from its nominal value.

Keywords: current control; Odd Harmonic High Order Repetitive Controller; Three- Phase PWM rectifier.

Control de alto desempeño de un rectificador PWM trifásico usando control repetitivo de alto orden

Resumen
El objetivo de control en rectificadores de potencia trifásicos se basa en generar corrientes de entrada sinusoidales y regular el voltaje de salida DC. Aunque el Control Repetitivo y Resonante son enfoques de control que han presentado excelentes resultados, su principal desventaja se basa en la pérdida considerable de desempeño cuando la frecuencia de la red se desvía de su valor nominal. En este artículo, se presenta el uso de un Controlador Repetitivo Impar de Alto Orden para controlar los lazos de corrientes de un rectificador trifásico con un desempeño considerablemente superior a otras alternativas tradicionalmente implementadas en este campo. Este controlador permite el rechazo de los armónicos impares introducidos en el sistema, lo que mantiene una complejidad computacional similar a la obtenida con los controladores repetitivos convencionales con la ventaja de incrementar la robustez cuando la frecuencia de la señal varíe. Las simulaciones y los resultados experimentales muestran un alto desempeño aún cuando se presenten desviaciones de la frecuencia de la red respecto a su valor nominal.

Palabras clave: Control de Corriente; Controlador Repetitivo de Alto Orden Impar; Rectificador PWM Trifásico.

1. Introduction

More recently, reactive power and harmonic currents generated by power converters have become a crucial issue in many electric power systems [1]. In this case, diode bridge and phase controlled bridge rectifiers have been considered as one of the most important causes of electrical pollution [2]. Although these solutions have been widely implemented in industrial applications because of their economic advantages, they begin to look unsustainable due to the impositions of stricter harmonic standards.

Pulse width modulation rectifiers (PWM) have arisen as

a solution to overcome these problems, reducing harmonic pollution [3,4] and adapting in accordance to the highest requirements for energy quality [5,6]. Using these devices, it is possible to control the output voltage and to obtain sinusoidal ac currents. In order to assess the performance of a three-phase PWM rectifier, two parameters are taken into account: the Total Harmonic Distortion (THD), which is a measurement of the harmonic content of the current signal; and the Power Factor (PF), which accounts for phase differences between the current and the voltage. With this thread in mind, PWM rectifiers seek a unity PF and a small THD.

Several strategies have been proposed to control three-phase PWM rectifiers. In the reported research, multiple Proportional Integral controllers [7] have been implemented in the current loops, mainly due to the ease of implementing them as well as their satisfactory performance. However, the main limitation of this approach lies in an inherent tracking (amplitude and phase) error [8].

Over the last few years, interesting emerging control techniques such as resonant control [9,10] and repetitive control (RC) [11-13] have been developed to track/reject periodic signals. Good response characteristics are achieved by using these strategies, ensuring near unit power factor and constant output dc voltage under parameter uncertainties, as well as load disturbances. The main disadvantage is related to the lost of performance caused by variations of the network frequency from its nominal value [14].

In this paper, an odd harmonic HORC Controller is proposed for the current loops of a three-phase PWM rectifier. Hence, taking advantage of its inherent properties, only the odd harmonics are removed. By using this strategy, two aims are achieved: 1) Computational complexity is substantially reduced by the odd property of this compensator, which is very similar to that obtained with the conventional repetitive controllers, and 2) Robustness is increased when the signal frequency varies. Experimental results show a high performance at nominal frequency, even when harmonics are present in the voltage source waveforms and have the additional characteristic of preserving low THD and unitary PF, despite network frequency variations.

The paper is organized as follows: section 2 describes the three-phase PWM rectifier model, introduces the control objectives, and presents the controller structure. In Section 3, the structure of an Odd HORC Controller for the current loops is explained in detail; the control implemented for the voltage loop is described in Section 4. Section 5 reports experimental results and demonstrates a comparison to a classical HORC Controller. Finally, in Section 6 we present conclusions and future directions.

2. System description

2.1. Three-phase PWM rectifier

PWM techniques have been widely implemented to control the output of power converters as they allow shaping voltage and/or current waves based on specific applications [15].
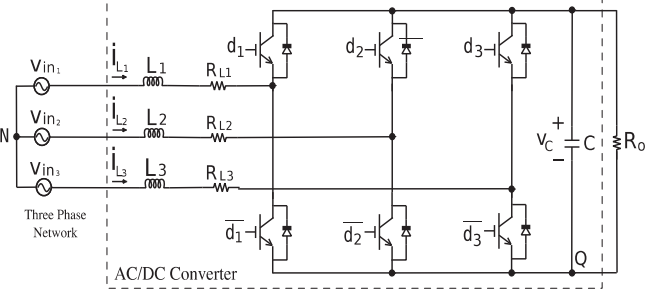
Figure 1. Three-phase pulse width-modulated (PWM) rectifier.
Source: The authors.

The three-phase PWM rectifier circuit, as shown in Fig. 1, has three-legs with IGBT transistors. It is known as a bi-directional boost rectifier (increasing the voltage) and works using fixed DC voltage polarity. Input inductors are an integral part of the rectifier and are selected based on their design [16]. We achieve harmonic compensation through the input inductors, and we ensure smooth voltage waveforms by using capacitors.

The dynamics of the three-phase PWM rectifier (as shown in Fig. 1) can be described as:

$$L_1 \frac{di_{L1}}{dt} = -R_{L1} i_{L1} - (V_c * d_1 + V_{QN}) + V_{in1}, \qquad (1)$$

$$L_2 \frac{di_{L2}}{dt} = -R_{L2} i_{L2} - (V_c * d_2 + V_{QN}) + V_{in2}, \qquad (2)$$

$$L_3 \frac{di_{L3}}{dt} = -R_{L3} i_{L3} - (V_c * d_3 + V_{QN}) + V_{in3}, \qquad (3)$$

$$C \frac{dV_c}{dt} = (i_{L1} * d_1) + (i_{L2} * d_2) + (i_{L3} * d_3) - \frac{V_c}{R_o}, \qquad (4)$$

where L_n, C and R_o are the nominal values of inductances, the capacitor and the load resistance, respectively; i_{L1}, i_{L2} and i_{L3} are the phase currents and V_c is the capacitor voltage; V_{in1}, V_{in2} and V_{in3} are the known three-phase sinusoidal voltages; $V_{in1}+V_{in2}+V_{in3}=0$; R_{Ln} is the parasitic inductor resistance for each phase; $V_{QN} = -\frac{V_d}{3}(d_1 + d_2 + d_3)$; and d_n is the control action used to trigger the IGBTs. Their values range within the interval [0,1].

2.2. Control system structure

Control objectives for three-phase PWM rectifier systems are quite similar to the objectives for the single-phase case. They are focused on achieving unit power factor, constant output dc voltage and a suitable low harmonics sinusoidal waveform for the input current. The difference is to be found in the fact that the control aim for the current loop should be achieved for every phase, taking into account the shifts associated with a three-phase current flow.

Fig. 2 depicts the general control system structure. It consists of a double-loop structure: two inner current loops that ensure a sinusoidal waveform in the current input, and an outer dc voltage loop designed to regulate the output voltage and proportionate the active power balance in the system. The voltage loop gives the amplitude, $I_{dn}(t)$, of the

reference currents that is to be used in the two inner current loops. The reason behind only using two current loops is that, under the assumption of a balanced system ($V_{in1}+V_{in2+}V_{in3}$=0), the compensation action applied in two phases will generate the action to be applied for the third phase. Thus, control action d_3 will be generated form d_1 and d_2 using $d_3 = -d_1 - d_2$, as shown in Fig. 2.

A PLL (Phase Locked Loop) system takes the phase voltages and produces sine waves with appropriate phases. These signals, together with the amplitude $I_{dn}(t)$, create the two current references for the current loops.

In this way, the control system should track sinusoidal signals with suitable I_{dn} magnitudes and phases and reject the system disturbances. Among these disturbances, two are recurrent in these kinds of systems: low-frequency signals and harmonics generated by distortions in the network voltage. Due to the fact that symmetrical waveforms only contain odd harmonics, distortions in the voltage waveform are regularly found between the 3th and 9th harmonic [17].

In addition, in order to keep a constant output dc voltage, the system should react to changes in the network voltage and in the rectifier load. Finally, it is worth drawing the reader's attention to the fact that the voltage control loop uses the average of the measured dc voltage to avoid introducing its ripple into the loops.

In this work, an Odd Harmonic High Order Repetitive Controller is adopted for the current loops in order to compensate the odd harmonics that are introduced to the system by the network voltage and increase system robustness to changes in the network frequency. A PI controller is integrated into the voltage loop for the purpose of achieving an active power balance and keeping the average value of the output voltage constant.

3. Current control loops

This section describes the Odd Harmonic High Order Repetitive Controller that is proposed for the current loops. Identical current control loops are designed for two of the system phases and can be seen in Fig. 2.
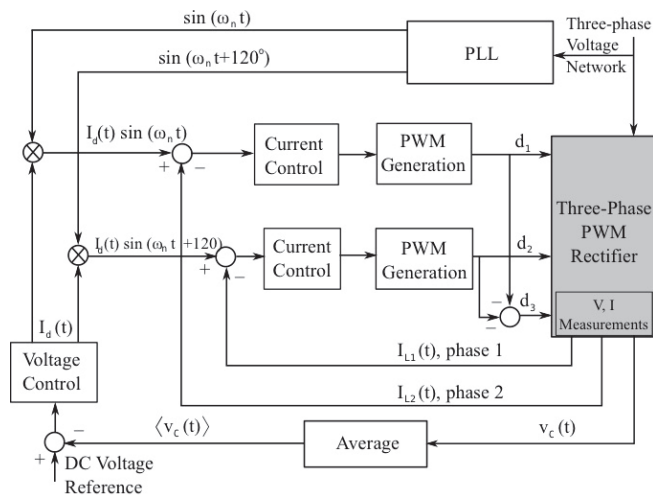
Figure 2. Control System Structure.
Source: The authors.

Primarily, the structure of conventional repetitive control scheme is explained in order to introduce the high-order controller design.

3.1. The odd harmonic high order repetitive controller

Repetitive control (RC) has proven to be a useful control strategy to track/reject periodic signals [18,19]. Using this control strategy, the sinusoidal signal references in the current loops are tracked and the harmonics generated by the network voltage are rejected. In order to accomplish this, the RC scheme incorporates a generator of periodic signals (also known as internal model) in the control structure [20]:

$$I(z) = \frac{\sigma W(z)H(z)}{1 - \sigma W(z)H(z)}, \qquad (5)$$

where $W(z)$ is a time delay function, $H(z)$ is a low-pass filter which provides robustness at high frequencies, and σ takes the values 1 or -1 and determines if the control is applied to all harmonic components; $\sigma = 1$ or $\sigma = -1$ only for odd harmonics. The way in which the internal model (eq. 5) is added to the control structure will be described in the next section.

The standard internal model, used in repetitive control [21], can be determined by using $H(z) = 1$, $\sigma = 1$, or $W(z) = z^{-N}$ with $N=T_p/T_s$, where T_p is the period of the signal being tracked/rejected and T_s is the sampling period of the control system. This model provides infinite gain at fundamental frequency and all its harmonics until the $(N/2-1)$-th. Having high gains at selected frequencies guarantees the tracking/rejecting capability in closed-loops for signals with this frequency content.

However, an internal model which only generates infinite gain at odd-harmonic components [22] can be obtained using $H(z) = 1$, $\sigma = -1$, $W(z) = z^{-N/2}$.

The magnitude frequency responses of the standard and odd internal models are shown in Fig. 3. The sampling period is selected to be $T_s=15$ kHz, which yields N=250, and the fundamental frequency is set to 60 Hz. As can be appreciated from the figure, these models generate infinite gain at fundamental and harmonic frequencies. This establishes a
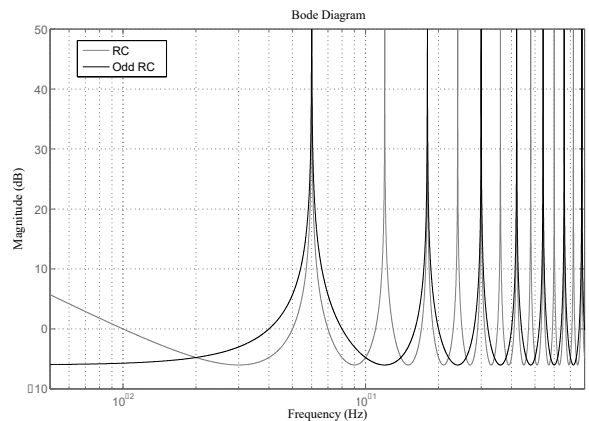
Figure 3. Frequency responses for the standard and odd internal models.
Source: The authors.

very limited frequency band that is centered around each harmonic component. This feature currently represents the main disadvantage of standard and odd internal models. As a result, the standard repetitive control exhibits a considerable loss of performance when the exogenous signal frequency varies from its nominal value, even with very small deviations [14].

In this regard, by modifying the time delay function $W(z)$, an internal model can provide high gain over a broader frequency interval around the fundamental and harmonic modes. This increases the robustness against changes in network frequency. Accordingly, $H(z) = 1, \sigma = -1$, and the time delay function is given by [5]:

$$W(z) = -1 + \left(1 + z^{-\frac{N}{2}}\right)^M, \tag{6}$$

where M denotes the Harmonic High Order Repetitive Controller (HORC). By using this definition, the internal model (eq. 5) is given by:

$$I_{odd}(z) = \frac{1 - \left(1 + z^{-\frac{N}{2}}\right)^M}{\left(1 + z^{-\frac{N}{2}}\right)^M} \tag{7}$$

Fig. 4 compares the magnitude response of the internal models used in odd-harmonic RC and odd-harmonic HORC for second (M=2) and third (M=3) order case; these have a nominal frequency of 60 Hz and a sampling period of 15 KHz. Fig. 4 also highlights the gain of the internal models designed for nominal frequency deviations of 59.5 and 60.5 Hz and their corresponding harmonics. It can be seen that the gain is higher for the odd-harmonic HORC (M=2 and M=3) compared with the odd-harmonic RC (M=1) for the same frequency variation. This improves the robustness against frequency changes and the performance under non-ideal operating conditions. However, internal models with high orders can decrease the robustness due to the elevated gains in the highest harmonics. Although this problem can be overcome by using a lower cut-off frequency in the low-pass filter H(z), this could cause performance decay because of the action reduction in the superior harmonics. For this reason, a tradeoff between the internal model order M and the filter bandwidth H(z) is required in order to achieve performance and stability robustness [18,19].

Based on the facts that frequency variations in the Colombian electric system are quite small (59.8-60.2) [23] and that it is desirable to keep the distortion as low as possible in the input current of the rectifier, the Odd Harmonic High Order internal model with M=2 (2ORC) has been experimentally validated as the controller with the best performance.

The odd internal model with the low-pass filter $H(z)$ and $M=2$ becomes:

Furthermore, it is important to note that the odd internal model (eq. 8) and the standard repetitive controller are mostly of the same order, which results in a similar computational complexity during their implementations. This implies that
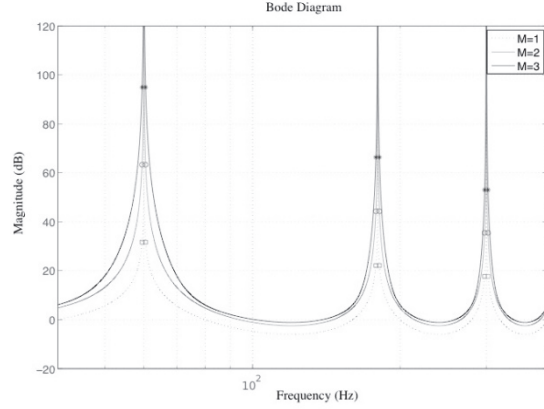
Figure 4. Odd-harmonic RC and odd-harmonic HORC internal models gain diagram
Source: The authors.

standard internal models and second order odd harmonic internal models with same N parameter would perform almost the same number of multiplications and additions, and also use the same amount of memory during an experimental implementation.

$$I_{2odd}(z) = \frac{-\left(2z^{-\frac{N}{2}} + z^{-N}\right)H(z)}{1 + \left(2z^{-\frac{N}{2}} + z^{-N}\right)H(z)} \tag{8}$$

3.2. Repetitive controller structure

Each current control block shown in Fig. 2 has a corresponding repetitive controller, which is as the one shown in Fig. 5. In order to include one of the internal models that were defined in the previous section, i.e. $I(z), I_{odd}(z),$ or $I_{2odd}(z)$ (equations (5), (7) and (8), respectively), repetitive control is usually implemented using the plug-in configuration presented in Fig. 5. As can be seen, the internal model $(I(z), I_{odd}(z)$ or $I_{2odd}(z))$ is added, as part of the existing controller $G_c(z)$, and the filter $G_x(z)$ is also incorporated. The compensator $G_c(z)$ is used to stabilize the plant $G_p(z)$ and to provide disturbance attenuation across the rectifier bandwidth. Additionally, compensator $G_c(z)$ should offer good robustness margins, and filter $G_x(z)$ is designed to ensure system stability in closed loop.

The transfer function of the plant is obtained from (eq. (1)-(3)), which performs the variable change $\alpha_n = d_n v_c$ and causes each phase to result in the following:
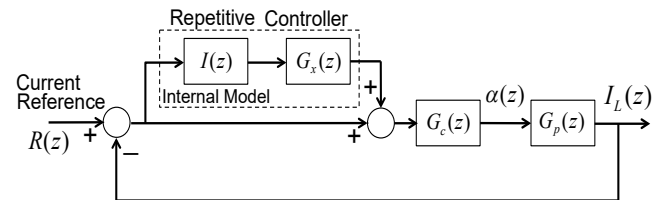
Figure 5. Repetitive control structure.
Source: The authors.

$$G_p(s) = \frac{1}{sL + R_L} = \frac{I_L(s)}{\alpha(s)} \qquad (9)$$

From the discretization of (9) the following can be obtained:

$$G_p(z) = Z\left[\frac{1}{sL + R_L}\frac{1 - e^{-T_s}}{s}\right]_{T_s} \qquad (10)$$

Control action in each phase results in:

$$\alpha(z) = (I(z)G_x(z) + 1)G_c(z)(R(z) - I_L(z)) \qquad (11)$$

with $R(z)$ being the z-transform of $r(k)$. $r_1(k) = i_d(t)\sin(\omega_n k)$ and $r_2(k) = i_d(t)\sin(\omega_n k + 120°)$, which are the reference signals for phase 1 and phase 2, respectively. Finally, dutty cycle d_n is calculated as $d_n = \alpha_n/v_c$ and sent to eah PWM generator, as seen in Fig. 2.

3.3. Stability

The closed-loop system of Fig. 5 is stable if the following conditions are fulfilled [24]:
The closed-loop system without the repetitive controller is stable, which means the transfer function $T_o(z)$ should be stable:

$$T_o(z) = \frac{G_c(z)G_p(z)}{1 + G_c(z)G_p(z)} \qquad (12)$$

1. $\|W(z)H(z)(1 - T_o(z)G_x(z))\|_\infty < 1$, where H(z) and G$_x$(z) should be chosen adequately.

Whole system dynamics, shown in Fig. 2, are divided into two parts: fast dynamics given by the current loop, and slow dynamics represented by the voltage loop. Therefore, to ensure the convergence of the whole system, two considerations are assumed: 1) the voltage loop provides very slow changes in the current reference, and 2) the current loop responds so fast that its dynamic behavior does not affect the evolution of the voltage loop. In order to obtain this behavior, the bandwidth of each loop should be properly selected in order to achieve a sufficient separation to differentiate between both dynamics. See reference [25] for a more detailed analysis on this topic.

3.4. Design of filter H(z)

Filter H(z) is used to limit the repetitive controller bandwidth, and, as such, it impacts performance, limiting the number of harmonics that will be taken into consideration. In general, the choice of filter bandwidth implies a trade-off between stability robustness and performance. Although numerous optimization techniques to define a conventional filter H(z) have been proposed, H(z) is chosen to be a null phase low-pass FIR filter [26].

3.5. Design of filter G$_x$(z)

In the event of a minimum phase systems, $G_x(z)$ is usually selected as $G_x(z) = k_r(T_o(z))^{-1}$, while for non minimum phase systems, it is possible to use the approach proposed in
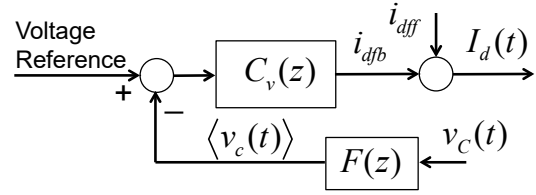
Figure 6. Voltage loop structure.
Source: The authors.

[21]. Finally, the value of the gain k_r requires a trade-off between stability robustness and steady state performance [27].

4. Voltage control loop

Voltage loop control is designed to achieve an active power balance in the rectifier and to provide the magnitude of the current reference for the current loops. Given that voltage at the rectifier output terminal has a waveform that is highly rippled, a low pass filter is usually applied to obtain a smoother signal. The purpose of this is to avoid the ripple propagation to the current reference $i_d(t)$, which would negatively impact the harmonic distortion of the rectifier input current $i_L(t)$. In this work, the low pass filter has been replaced by an averaging filter that has been tuned at 60 Hz:

$$F(z) = \frac{1}{N}\frac{(1 - z^{-N})}{(1 - z^{-1})}. \qquad (13)$$

By using this filter, the average voltage is calculated from the output voltage $V_{dc} = \langle v_c \rangle$.

Voltage loop structure is shown in Fig. 6. This structure is comprised of two basic components: a feedback control action, i_{dfb}, generated by a PI compensator, and a pre-feeding action i_{dff}. The PI controller in the feedback loop will regulate the average value of the output voltage to the desired value with zero error:

$$C_v(z) = k_i\frac{T_s}{2}\frac{(z + 1)}{(z - 1)} + k_p \qquad (14)$$

Moreover, the pre-feeding signal, i_{dff}, is designed in order to avoid sudden changes in voltage levels causing transients with high current peaks. This pre-feeding stage is calculated based on the rectifier power balance in steady state. As such, the balance between the AC active power and the DC power is given by:

$$\frac{1}{2}V_m I_m = V_{dc}I_{dc} + P_{losses} \qquad (15)$$

where $V_m = V_{in}\sqrt{2}$ is the peak value of the network voltage, $I_m = I_d$ is the peak value of the rectifier input current, $V_{dc} = \langle v_c \rangle$ is the average value of the capacitor voltage, $I_{dc} = \langle i_L \rangle$ is the average value of the load current, and P_{losses} are the rectifier losses. On that basis, the pre-feeding value can be computed by:

$$i_{dff} = \frac{1}{2}\frac{V_m I_m}{V_{dc}} + I_{losses} \qquad (16)$$

where I_{losses} is a factor that weighs the rectifier losses. As a result, the reference signal amplitude for the current loop is calculated by:

$$I_d = i_{dfb} + i_{dff} \qquad (17)$$

5. Experimental results

In this section the experimental results are presented and analyzed. Two different setups are defined for nominal and varying frequency conditions. A comparison of HORC and RC is provided in order to show the advantages of implementing the odd harmonic HORC.

5.1. Experimental setup

The utility network frequency is 60 Hz, thus defining a period of Tp=1/60 s. The sampling period is set to Ts=1/15000 s, which corresponds with the PMW switching frequency. As a result, the discrete period is N=Tp/Ts=250.

The transfer function of the plant is obtained from (eq. 10) with *L=300* mH and *R_L=0.1* ohms. The internal controller is *Gc(z)=(6,293z-6,283)/(z-0.998)*, which provides enough robustness margins. Filter *H(z)*, which has a good robustness-performance tradeoff, is comprised of $H(z) = -0,003871z^5 + 0,03209z^3 + 0,1167z^2 + 0,2207z + 0,2687 + 0,2207z^{-1} + 0,1167z^{-2} + 0,03209z^{-3} - 0,003871z^{-5}$. Concurrently $G_x(z) = k_r(T_o(z))^{-1}$ with $k_r=0.7$. A PI controller for the voltage loop defined in (9) takes the values $k_p=0,01$ and $k_i=0,7$.

For comparison purposes, a conventional RC with internal model, as in (5), has been designed with $H(z) = 0,175z + 0,65 + 0,175z^{-1}$, $\sigma = -1$, $W(z) = z^{-N/2}$, $N=250$, $G_c(z) = (6,293z - 6,283)/(z - 0.998)$, and $G_x(z) = k_r(T_o(z))^{-1}$ with $k_r=0,3$.

Results at nominal frequency are obtained with the rectifier connected to the grid, and results at different frequencies are acquired using an AC varying frequency voltage source from PowerSun. In both cases, a three phase autotransformer is used to provide $V_{in}=20$ Vrms to the rectifier. The three-phase rectifier is a power converter from Semikron, and control algorithms were implemented in an Intel PC with Matlab XPCTarget real time software. Bus dc voltage is $V_c=80$ V with a resistive load of 25 ohms.

5.2. Nominal frequency performance

Fig. 7 shows the response of the non-controlled rectifier. It can be seen that non-sinusoidal currents with a THD=34.4% are obtained. It is important to note that Fig. 7 also shows the distortion in the voltage network with 5th and 7th harmonics with a THD of 4.3 %. This distortion acts as a disturbance in the current loops.

Fig. 8 presents the results with the proposed odd harmonic HORC. It can be seen that currents now have a sinusoidal shape in the phase with the voltage source. The THD obtained is 2.2 %. Power factor is presented in Fig. 8 in which it can be seen that PF=1 was obtained.

These results show that odd harmonic HORC satisfies the control objectives, achieving a low THD and unitary power factor, even in the case of distortion in the voltage source. Based on its definition, conventional RC should also achieve the same performance. However, since HORC provides smaller robustness margins, the cut-off frequency of filter *H(z)* is usually smaller than the one selected for RC. This entails a slight performance reduction that appears as a small THD degradation. Fig. 9 shows the THD obtained with RC under the same conditions.

a)

b)

c)

d)
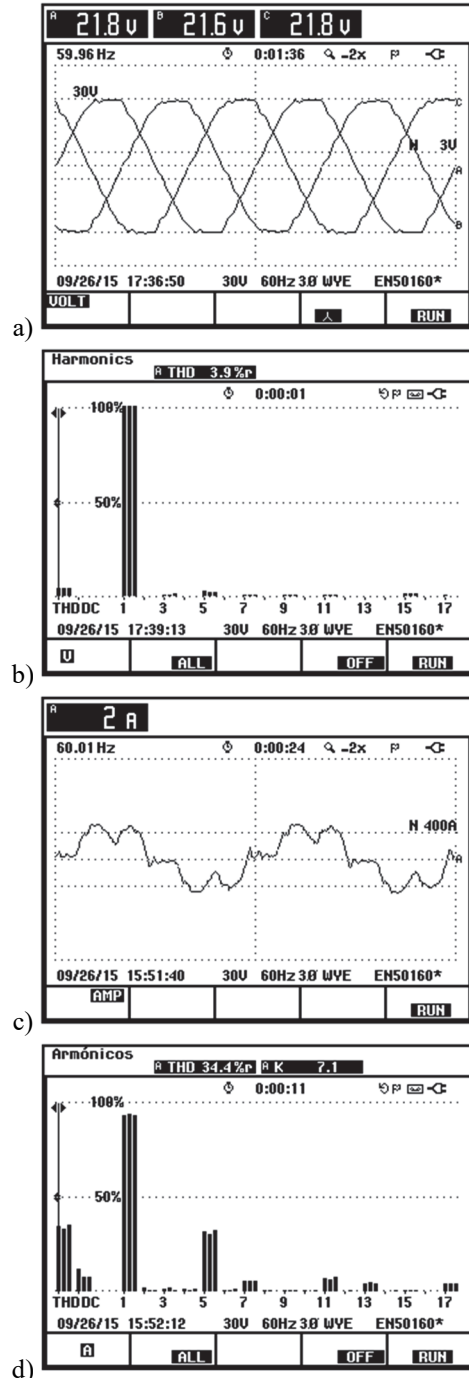
Figure 7. Non-controlled rectifier response. a) Voltage signals. b) Voltage harmonic content. c) Current signals. d) Current harmonic content.
Source: The authors.

a)

b)

HARMONICS TABLE

Volt	A	B	C	
THD%r	2.6	2.7	2.8	95.4
H3%r	0.3	0.3	0.4	19.3
H5%r	0.8	0.7	0.9	11.6
H7%r	0.3	0.6	0.6	10.6
Amp	A	B	C	
THD%r	1.8	1.7	2.1	98.6
THD%r	1.8	1.7	2.1	98.6
THD%r	1.8	1.7	2.1	98.6

09/26/15 17:11:48   30V 60Hz 3Ø WYE   EN50160*

c)

Power & Energy

	A	B	C	
kW	0.2	0.2	0.2	0.5
kVA	0.2	0.2	0.2	0.5
kVAR	0.0	0.0	0.0	0.0
PF	1.00	1.00	1.00	1.00
DPF	1.00	1.00	1.00	1.00
A rms	8	9	8	
	A	B	C	
V rms	20.90	20.32	20.54	

09/26/15 17:33:58   30V 60Hz 3Ø WYE   EN50160*

d)
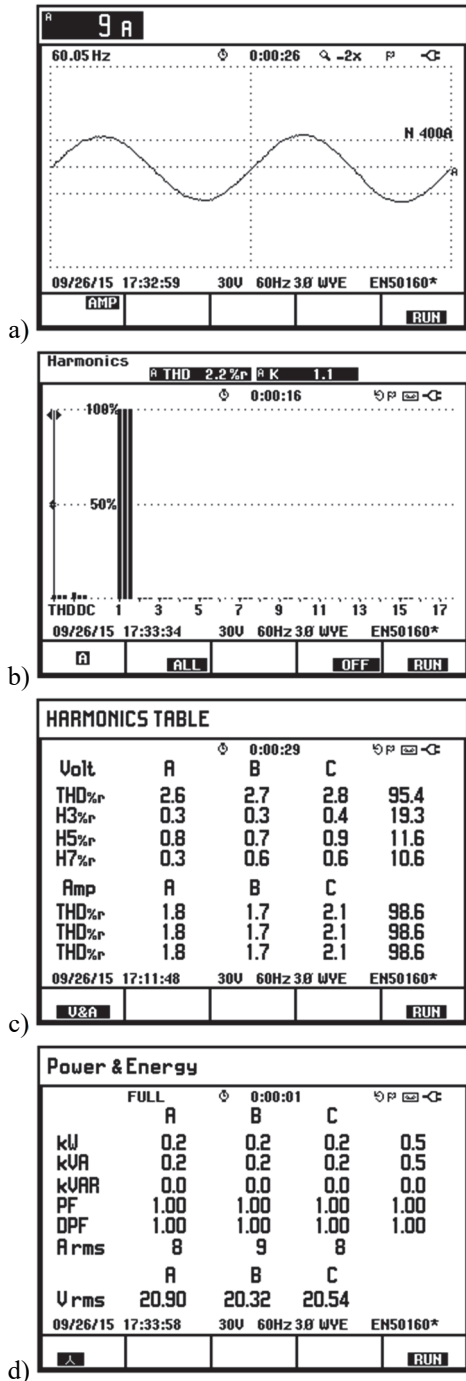
Figure 8. Controlled rectifier response (odd harmonic HORC). a) Current signal.
b) Current harmonic content. c) Harmonic table. d) Power and energy table.
Source: The authors

5.3. Varying frequency performance

These experiments use an AC voltage source that comes from PowerSun and has a configurable frequency. Fig. 10 depicts the current waveforms when the source frequency is 62 Hz and uses conventional RC. The distortion of the obtained currents can be seen with a THD=5 % with 5th and 7th harmonics. Also, the degradation of the power factor is now present with PF=0.98.

HARMONICS TABLE

Volt	A	B	C	
THD%r	3.9	3.9	3.9	91.8
H3%r	0.9	1.1	1.2	30.8
H5%r	2.6	2.4	2.4	13.7
H7%r	0.9	0.9	0.9	8.6
Amp	A	B	C	
THD%r	1.7	1.6	1.8	98.4
THD%r	1.7	1.6	1.8	98.4
THD%r	1.7	1.6	1.8	98.4

09/26/15 17:39:44   30V 60Hz 3Ø WYE   EN50160*
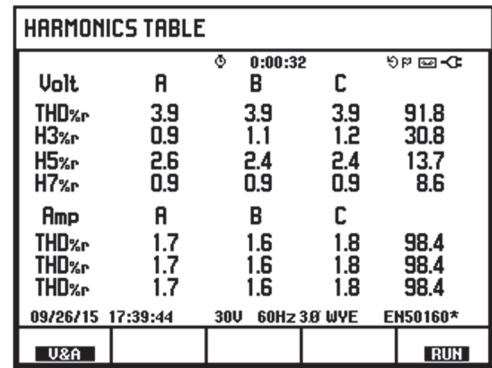
Figure 9. Harmonic content table for conventional RC.
Source: The authors

a)

b)

HARMONICS TABLE

Volt	A	B	C	
THD%r	3.5	3.3	3.8	95.8
H3%r	0.3	0.2	0.3	18.0
H5%r	1.8	1.7	2.0	10.6
H7%r	1.2	1.2	1.2	8.4
Amp	A	B	C	
THD%r	5.0	4.1	4.9	98.5
THD%r	5.0	4.1	4.9	98.5
THD%r	5.0	4.1	4.9	98.5

09/26/15 16:57:44   30V 60Hz 3Ø WYE   EN50160*

c)

Figure 10. RC response at 62 Hz. a) Current signals. b) Current harmonic content. c) Harmonic table. d) Power and energy table.
Source: The authors

d)

Figure 10. RC response at 62 Hz. a) Current signals. b) Current harmonic content. c) Harmonic table. d) Power and energy table.
Source: The authors

d)

Figure 11. Response of odd harmonic HORC at 62 Hz. a) Current signals b) current harmonic content c) harmonic table d) power and energy table
Source: The authors

Using odd harmonic HORC yields, a significantly better performance can be seen in Fig. 11. The obtained THD and PF are 3.8 % and 1, respectively.

a)

b)

c)

a)

b)

Figure 12. Response of RC at 59 Hz. a) Harmonic table d) Phasor diagram.
Source: The authors

Similar behavior is expected at lower frequencies. Fig. 12 and Fig. 13 present the THD and phasor diagram results for 59 Hz for RC and HORC, respectively. It can be seen that HORC performs better with a lower THD and a smaller current phase deviation.

6. Conclusion

This work has proposed an odd harmonic HORC has for the current control loops in a three phase-PMW power rectifier. Experimental validation has shown that the proposed controller achieves high performance with THD=1.7% and a unitary PF. Compared to conventional RC, the HORC has the advantage of providing higher robustness against network frequency variations. This characteristic allows the rectifier to accomplish the more demanding standards of power quality, even under varying frequency conditions.
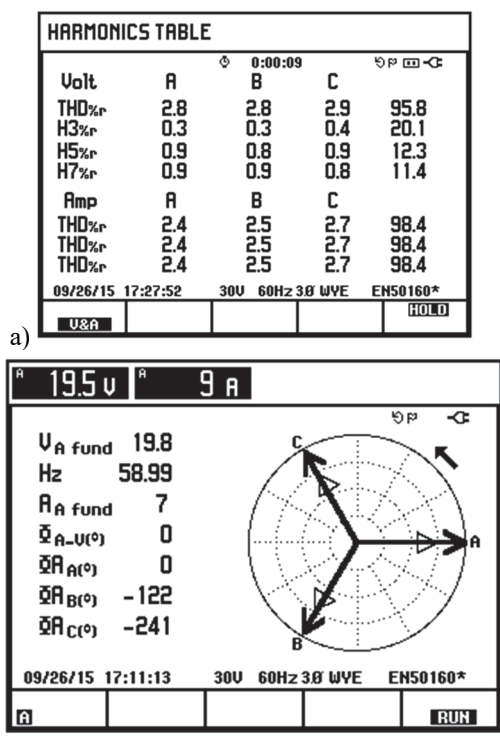
Figure 13. Response of the odd harmonic HORC at 59 Hz. a) Harmonic table. d) Phasor diagram.
Source: The authors

Despite being a high order controller, it was found that odd harmonic HORC has a similar computational burden when it was compared to a full- harmonic RC.

References

[1] Peña-Huaringa, O.J., Studying and simulating transformer configuration to improve power quality. Ingeniería e Investigación, 31, pp.125-130, 2011.

[2] Acha, E. and Madrigal, M., Eds., Power system harmonics, computer modeling and analysis, New York: Wiley, 2001.

[3] Etz, R., Patarau, T. and Petreus, D., Comparison between digital averafe current mode control and digital one cycle control for a bridgeless PFC boost converter, Design and Tehnology in Electronic Packaging (SIITME), IEEE 18th International Symposium for Design and Technology in Electronic Packaging (SIITME), pp. 211-215, 2012. DOI: 10.1109/SIITME.2012.6384378

[4] IEC 1000-3-2. Electromagnetic Compability. Part 3: Limits – Sect. 2: Limits for harmonic current emission (Equipment Input Current 16A per Phase), 1995.

[5] Rodriguez, J.R., Dixon, J.W., Espinoza, J.R., Pontt, J. and Lezana, P., PWM regenerative rectifiers: State of the art. Industrial Electronics, IEEE Transactions on, 52(1), pp. 5-22, 2005. DOI: 10.1109/TIE.2004.841149

[6] IEEE Recommended practices and requeriments for harmonic control in electrical power systems IEEE Std 519-1992, pp.1-112, 1993.

[7] Kazmierkowski, M.P. and Malesani, L., Current control techniques for three-phase voltage-source PWM converters: A survey. Industrial electronics, IEEE Transactions on, 45(5), pp. 691-703, 1998. DOI: 10.1109/41.720325

[8] Yuan, X., Merk, W., Stemmler, H. and Allmeling, J., Stationary-frame generalized integrators for current control of active power filters with zero steady-state error for current harmonics of concern under unbalanced and distorted operating conditions. Industry

[9] Liu, F., Maswood, A., Kang, Y., Zhang, Y. and Duan, S., Proportional-resonant current control for three-phase three-level rectifier. In Power Engineering Conference, IPEC 2007. International, pp. 1018-1022, 2007.

[10] Cheng, Q.M., Cheng, Y.M., Xue, Y., Hu, X.Q. and Wang, Y.F., A summary of current control methods for three-phase voltage-source PWM rectifiers. Power System Protection and Control, 40(3), pp. 145-155, 2012.

[11] Longman, R.W., Iterative learning control and repetitive control for engineering practice. International Journal of Control, 73(10), pp. 930-954, 2000. DOI: 10.1080/002071700405905

[12] Zhou, K. and Wang, D., Digital repetitive controlled three-phase PWM rectifier. Power Electronics, IEEE Transactions on, 18(1), pp. 309-316, 2003. DOI: 10.1109/TPEL.2002.807150

[13] Ramos-Fuentes, G.A., Olm J.M. and Costa-Castelló, R., A survey of repetitive control in varying frequency conditions, Ingeniería e Investigación, 31(2), pp.29-37, 2011.

[14] Steinbuch, M., Repetitive control for systems with uncertain period-time, Automatica, 38, pp. 2103-2109, 2002. DOI: 10.1016/S0005-1098(02)00134-6

[15] Rahim, N.A., Green, T.C. and Williams, B.W., PWM ASIC Design for the three phase bidirectional buck converter, INT. J. Electronics, 1996. DOI: 10.1080/002072196136517

[16] Malinowski, M. and Kazmierkowski, M.P., CHAPTER 11 - Control of Three-Phase PWM Rectifiers, In Academic Press Series in Engineering, edited by Kazmierkowskir, M.P. and Blaabjerd, K., Academic Press, Burlington, Control in Power Electronics, pp. 419-459, 2002, ISBN 9780124027725, DOI: 10.1016/B978-012402772-5/50012-0.

[17] Siemens Industry, Inc. Harmonics in power systems. Causes, effects and control, USA, 2013.

[18] Ramos, G.A. and Costa-Castelló R. and Olm J.M., Digital repetitive control under varying frequency conditions. Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin, Germany. 446, 2013, 159 P. DOI: 10.1007/978-3-642-37778-5

[19] Pipeleers, G., Demeulenaere, B. and Sewers, S., Robust high order repetitive control: Optimal performance trade offs, Automatica, 44, pp. 2628-2634, 2008. DOI: 10.1016/j.automatica.2008.02.028

[20] Inoue, T., Practical repetitive control system design. Proc. 29th IEEE Conf. on Decision and Control, pp. 1673-1678, 1990.

[21] Tomizuka, M., Tsao, T.C. and Chew, K.K., Analysis and synthesis of discrete-time repetitive controllers, Meas. Control J. Dyn. Syst., 111, pp. 353-358, 1989. DOI: 10.1115/1.3153060

[22] Ramos, G.A., Costa-Castelló, R. and Olm, J.M., Analysis and design of a robust odd-harmonic repetitive controller for an active filter under variable network frequency, Control Engineering Practice, 20(9), pp. 895-903, 2012. DOI: 10.1016/j.conengprac.2012.05.009

[23] Colombia. Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC). Norma Técnica Colombiana NTC 5001. Calidad de la Potencia Eléctrica. Límites y metodología de evaluación en puente de conexión común, 2008b.

[24] Griñó, R. and Costa-Castelló, R., Digital repetitive plug-in controller for odd-harmonic periodic references and disturbances, Automatica, 41(1), pp. 153-157, 2005. DOI: 10.1016/j.automatica.2004.08.006

[25] Costa-Castelló, R., Griñó, R., Parpal, R.C. and Fossas, E., High-performance control of a single-phase shunt active filter. Control Systems Technology, IEEE Transactions on, 17(6), 1318-1329, 2009. DOI: 10.1109/TCST.2008.2007494

[26] Tsao, T.C. and Tomizuka, M., Robust adaptive and repetitive digital tracking control and application to a hydraulic servo for noncircular machining, Journal of Dynamic Systems, Measurement, and Control, 116(1), pp. 24-32, 1994. DOI: 10.1115/1.2900676

[27] Hillerström, G. and Lee, R.C.H., Trade-offs in repetitive control, University of Cambridge, CUED/F-INFENG/TR 294, 1997.

G.A. Ramos, received his BSc in Electrical Engineering in 1999 and his MSc degree in Industrial Automation in 2006, both from the Universidad Nacional de Colombia, Bogotá, Colombia. He obtained his PhD in Automatics in 2012, from Universitat Politècnica de Catalunya, Barcelona,

Spain. He is currently is an Associate Professor in the Electrical and Electronic Department, Facultad de Ingeniería, Universidad Nacional de Colombia. His research interests include: control theory, control applied to power converters, repetitive control and active disturbance rejection control.
ORCID: 0000-0003-1393-6943

I.D. Melo Lagos, received his BSc in Telematics Engineering in 2008 from the University Distrital Francisco Jose de Caldas, Bogotá, Colombia, and his BSc in Electrical Engineering in 2010 from the Universidad Nacional de Colombia. In the years 2010-2015 he worked in the area of the design of power electronics equipment, in the CDP and Powersun groups. In 2015 he finished his MSc studies of in industrial automation at the Universidad Nacional de Colombia. His research interests include: control theory applied to power electronics.
ORCID: 0000-0002-5818-5808

J.A. Cifuentes, received her BSc in Mechatronics Engineering in 2008, her MSc. degree in Industrial Automation in 2010, both from the Universidad Nacional de Colombia, Bogotá, Colombia. She was awarded a PhD in Industrial Automation from the Institute National des Sciences Apliqueés de Lyon (INSA), France and a PhD in Mechanical and Mechatronics Engineering from the Universidad Nacional de Colombia, Bogotá, Colombia in 2015. She is currently an Associate Professor in the program of Electrical Engineering, Facultad de Ingeniería, Universidad de la Salle. Her research interests include: Modeling and analysis of dynamic systems, Signal Processing, Biomedical Signal Classification, and Gesture Acquisition.
ORCID: 0000-0001-7421-291X

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN

FACULTAD DE MINAS

Área Curricular de Ingeniería
Eléctrica e Ingeniería de Control

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Eléctrica

Mayor información:

E-mail: ingelcontro_med@unal.edu.co
Teléfono: (57-4) 425 52 64

FPGA implementation of the AES-128 algorithm in non-feedback modes of operation

Ian Carlo Guzmán, Rubén Darío Nieto & Álvaro Bernal

*Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia. ian.guzman@correounivalle.edu.co,
ruben.nieto@correounivalle.edu.co, alvaro.bernal@correounivalle.edu.co*

Abstract
In this paper, we present a hardware implementation of the pipelined AES-128 algorithm that works on non-feedback modes of operation (ECB and CTR). The architecture was implemented using the Xilinx Virtex 5 FPGA platform. We compared two modes of operation (ECB, CTR) for encryption and decryption according to device utilization, throughput, and security. A clock frequency of 272.59Mhz for the ECB encryption process was obtained, which is equivalent to a throughput of 34.89 Gb/s. Also, we obtained a clock frequency of 199.48Mhz for the decryption process, which is equivalent to a throughput of 25.5Gb/s. In CTR mode, we obtained a clock frequency of 272.59Mhz, which is equivalent to a throughput of 34.89Gb/s.

Keywords: AES; $G(2^8)(2^8)$; ECB; CTR; Pipelined; Throughput.

Implementación en FPGA del algoritmo AES-128 en modos de operación no realimentados

Resumen
En este artículo, presentamos una implementación hardware segmentada del algoritmo AES-128 en modos de operación no realimentados (ECB, CTR). La arquitectura fue implementada en la FPGA Virtex 5 de Xilinx. Dos modos de operación (ECB,CTR) para encriptación y desencriptación de acuerdo a uso de recursos, rendimiento y seguridad fueron comparados. Una frecuencia de reloj de 272.59Mhz para el proceso de encriptación ECB fue obtenida, la cual es equivalente a un rendimiento de 34.89 Gb/s. Además, una frecuencia de reloj de 199.48Mhz para el proceso de desencriptación, equivalente a un rendimiento de 25.5Gb/s fue obtenido. En el modo CTR, una frecuencia de reloj de 272.59Mhz. equivalente a un rendimiento de 34.89Gb/s fue obtenido.

Palabras clave: AES; $G(2^8)(2^8)$; ECB; CTR; Segmentado; Rendimiento.

1. Introduction

In 1997, the National Institute of Standards and Technology (NIST) initiated a public request for researchers to develop a new cryptographic algorithm that would be called the Advanced Encryption Standard (AES) and would replace its predecessor, the Data Encryption Standard (DES) [1]. Fifteen proposals were made. In October 2000, NIST announced that Rijndael, the algorithm proposed by the two Belgian cryptographers, Joan Daemen and Vincent Rijmen, had been selected as the Advanced Encryption Standard (AES) and was published as FIPS 197 [2] in 2001. Rinjdael can be implemented in both hardware and software, but hardware implementations are faster. Reprogrammable devices such as FPGA's are widely used for cryptographic algorithms' hardware implementations [3-4]. ASIC implementations offer optimized structure, a smaller area, and ahigher operation speed. However, ASIC implementations cannot be modified once they have been implemented, and the cost is higher than for that of reconfigurable devices.

This paper presents a pipelined architecture for the AES (Advanced Encryption Standard) in non-feedback modes of

operation (ECB, CTR). The non-feedback modes of operation allow data to be processed in parallel whereas the feedback modes of operation (CBC, CFB, OFB) do not. [5]

Therefore, ECB and CTR modes can be implemented in pipelining architectures, which are faster than iterative architectures [6].

The purpose of this paper is to show the hardware implementation results of the AES algorithm in nonfeedback modes of operation (ECB and CTR). There is an emphasis on explaining and describing our architecture for the AES-CTR as there are very few reports that explain in detail hardware architectures for the AES-CTR. Fu, Hao designed and explained in detail an architecture for the AES-CTR [6].

This paper is organized as follows: in section 2, we discuss the components of the AES algorithm and explain how it works. The design and implementation of the proposed hardware architectures are presented in section 3. Results such as, device utilization, throughput, and comparison with others pipelined implementations are presented in section 4. Finally, conclusions are stated in section 5.

2. Aes algorithm and modes of operation

2.1. Encryption and decryption process

The AES algorithm is a symmetric block cipher that can process data blocks of 128 bits, using cipher keys with lengths of 128, 192 and 256 bits. The encryption process consists of $N_r - 1$ rounds, where N_r depends on the key length and $N_r = 10$ on a 128 key-length. A round is made up of four basic operations: *SubBytes, ShiftRows, MixColumns and AddRoundKey*; the last round of the algorithm omits the *MixColumns* operation.

The *SubBytes* transformation is a non-linear byte substitution that operates independently on each state byte using a substitution table (*S-Box*). The *S-Box* is constructed by performing two transformations, the first one calculates the multiplicative inverse of the input bytes in the finite field $GF(2^8)$, the second one applies an affine transformation over $GF(2)$.

The *ShiftRows* transformation is a cyclic shift operation in each row of the state. The bytes in the last three rows of the state are cyclically shifted over a different number of bytes (offsets). The first row is not shifted.

The *MixColumns* transformation involves addition and multiplication over $GF(2^8)$ and can be expressed as a matrix multiplication for each column of the state.

The *AddRoundKey* transformation adds a round key to the state. Each round key is generated in the Key expansion process.

The decryption process consists of $N_r - 1$ rounds. A round is made up of four inverse operations: *InvSubBytes, InvShitRows, InvMixColumns and AddRoundKey*. The last round of the algorithm omits the *InvMixColumns* operation.

The *InvSubBytes* transformation is the inverse of the *SubBytes* transformation in which the inverse *S-box* is applied to each byte of the state. This is obtained by applying the inverse of the affine transformation and is followed by taking the multiplicative inverse in $GF(2^8)$.

The *InvShiftRows* transformation is the inverse of the *ShiftRows* transformation. The bytes in the last three rows of

the state are cyclically shifted over a different number of bytes. The first row is not shifted.

The *InvMixColumns* transformation is the inverse of the *MixColumns* transformation. *InvMixColumns* operates on the state column-by-column and treats each column as a four-
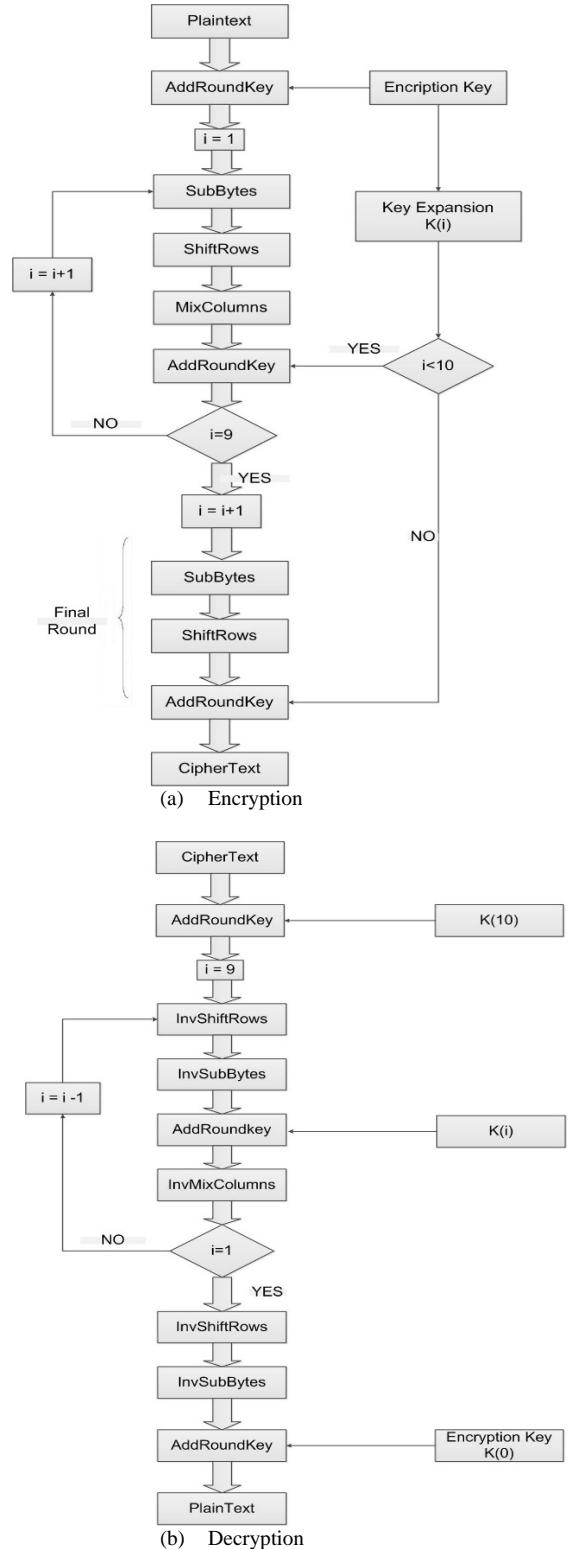
Figure 1. AES-128 Algorithm for: a) Encryption and b) Decryption
Source: The authors

term polynomial. It can be expressed as a matrix multiplication for each column of the state.

The *AddRoundKey* transformation is its own inverse due to the fact it only involves a XOR operation. [2]

Encryption and Decryption algorithms are shown in Fig. 1(a) and 1(b), respectively.

2.2. Key expansion process

The AES algorithm takes the cipher Key, **K**, as four 32-bit words and performs a key expansion. The key expansion generates a total of $N_b(N_r + 1)$ words. A word, $w(i)$, is equal to the XOR between the previous word, $w(i - 1)$, and the word $w(i - k)$ is located N_k in earlier positions,. For words in positions that are a multiple of N_k, a transformation is applied to $w(i - 1)$ prior to the XOR and followed by an XOR by a round constant, $Rcon(i)$. This transformation consists of a cyclic shift, followed by the application of the *SubBytes* transformation to all four bytes. The key expansion algorithm is shown in Fig. 2. [2]

2.3. Modes of operation

The Electronic Codebook (ECB) mode processes each block of the plaintext directly and independently and encrypts the same plain text block into the same ciphered text block. In ECB encryption and decryption, multiple cipher functions and inverse cipher functions can be computed in parallel.

The feedback modes of operation: Cipher block chaining mode (CBC), Output feedback mode (OFB), and Cipher feedback mode (CFB) offer better security properties than ECB. However, encryption of the blocks depends on the previous encrypted blocks, so the encryption cannot be performed in parallel; therefore, the speed of CBC, OFB, and CFB has a lower performance than the ECB. The counter mode (CTR) eliminates the security problem of ECB and it allows for encryption and decryption to be performed in parallel using only the cipher forward function. As shown in Fig. 3, the modes of operation can be classified as feedback and non-feedback. [5]

3. Hardware implementation

The AES algorithm may be implemented in hardware using different architectures such as iterative, inner round pipelining, loop unrolled, pipelining, and subpipelining or Mixed inner and outer round pipelining [7-8]. In this work we used a pipelined architecture in order to achieve a high speed to encrypt and decrypt the data.

3.1. Pipelined architecture

The pipelined architecture shown in Fig. 4 allows the speed of the encryption and decryption process to be increased by processing blocks of data in parallel. Pipelining is introduced by inserting registers between the rounds. This architecture allows a encrypted block every clock cycle to be obtained after the first block has been encrypted.

Figure 2. Key Expansion algorithm.
Source: The authors
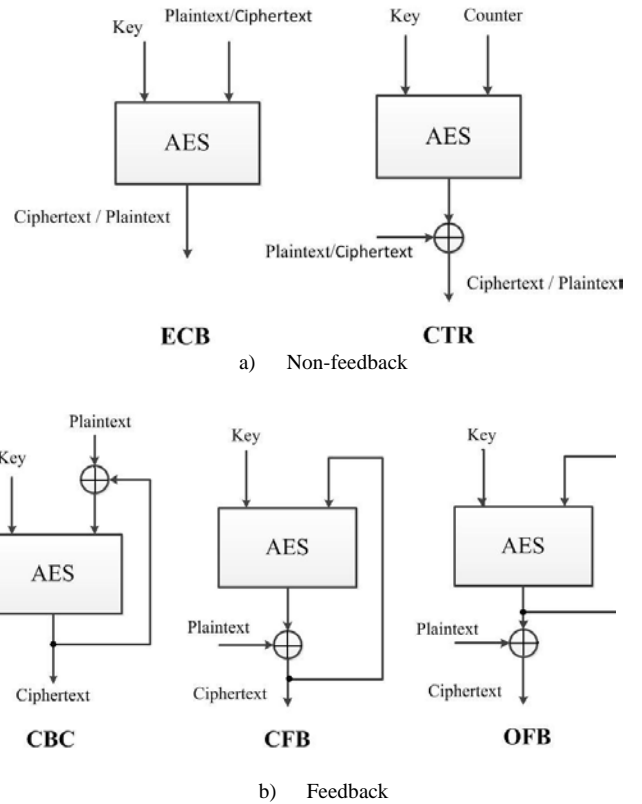
Figure 3. Modes of operation
Source: The authors
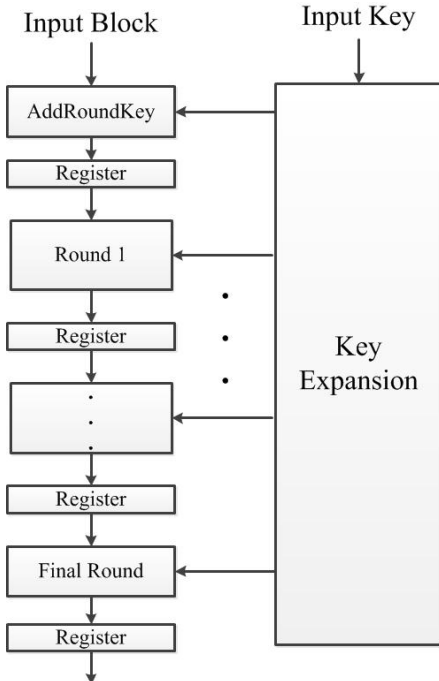
Figure 4. Pipelined Architecture
Source: The authors
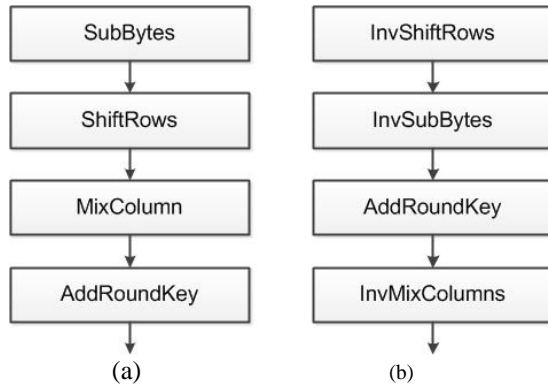
Figure 5. Round Module for: a) encryption and b) decryption.
Source: The authors

Figure 6. Implementation of SubBytes based on look-up tables.
Source: The authors

one is based on look-up tables. The design based on look-up tables was faster and required less hardware resources. The *S-Box* was stored in a 256x8 ROM memory.

In order to achieve parallelism and finish one round in less than one clock cycle, the same *S-box* was duplicated 16 times, as in Fig. 6.

3.1.2. ShiftRows/InvShiftRows

The *ShiftRows* and *InvShiftRows* transformation only changes the positions of the bytes; therefore, this transformation can be implemented by changing the order of the interconnection lines without using additional hardware components.

3.1.3. MixColums/InvMixColums

This transformation is based on the *Xtime* function [9], which performs a multiplication by 2 over the Galois domain $GF(2^8)$ Based on the *Xtime* function, it is possible to design a circuit that multiplies four bytes by a matrix over $GF(2^8)$, as is shown in Fig. 7. To achieve parallelism and finish one round in less than one clock cycle, the same circuit is duplicated 4 times in order to process each column of the state array in parallel.

The Round module is shown in Fig. 5(a), this module is composed of four circuits, which are *SubBytes, ShiftRows, MixColumns,* and *AddRoundKey*. Therefore, ten of these modules are needed to meet the rounds of the algorithm. The last module is slightly different due to the fact that the *MixColumns* transformation is not included. The round module for the decryption process is very similar but it uses the inverses of the transformations, as shown in Fig. 5(b). The last module is slightly different due to the fact that the *InvMixColumns* transformation is not included.

3.1.1. SubBytes/InvSubBytes

Two designs were made for this transformation, the first one consists of the mathematical operation, and the second

Figure 7. Implementation of MixColumns transformation.
Source: The authors

3.2. Key expansion module

This module is based on a round key expansion circuit, as is shown in Fig. 8. This circuit expands a round key; therefore, it performs the *subword, rotword* and *Rcon* transformations for the least significant input bytes $b_3 b_2 b_1 b_0$ (words in positions that are a multiple of 4), and then it starts to perform the XOR operations with the words that are four positions earlier.

The *SubWord* transformation was designed similarly to the *SubBytes* transformation. It required 4 256x8 ROM memories; each memory stores the S-Box.

The *RotWord* transformation was designed in the same way as the *ShiftRows* transformation since it only involves the order of the interconnection lines to be changed.

The *Rcon* transformation was designed using a combinational circuit based on NOT gates which is different depending on the number of the round, since each round has a different *Rcon* value.

The pipelined architecture of the key expansion module required ten round key expansion circuits, as shown in Fig. 9, in which each round key expansion circuit calculates a round key. In this architecture, all round keys are available at the same time for each one of the ten rounds. Moreover, a clock signal is not necessary due to the fact that there is no need for synchronization; therefore, all round keys expand very quickly.
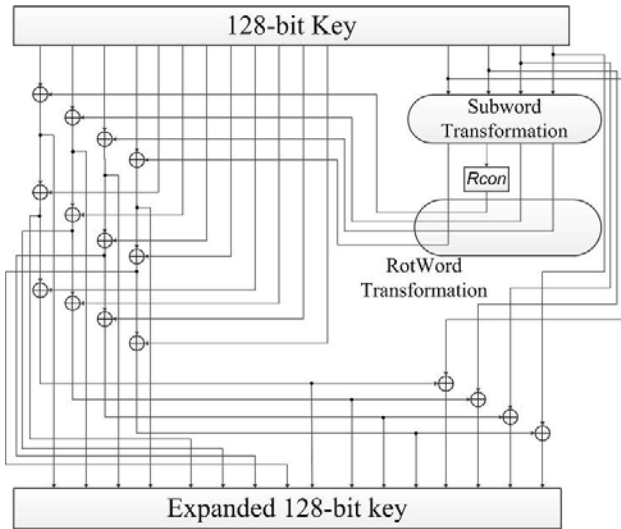
Figure 8. Round key expansion circuit.
Source: The authors
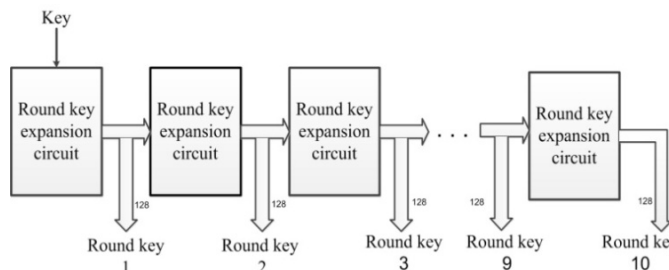
Figure 9. Key expansion module.
Source: The authors

3.3. AES in CTR mode

This mode encrypts counter values to produce a sequence of output blocks that are exclusive-ORed with plaintext to produce the ciphertext, and vice versa. For this mode to be properly operated, all counter values must be different for each plaintext block that is encrypted. Otherwise, if a counter is used more than once, then the confidentiality of all of the plaintext blocks corresponding to that counter value may be compromised [5]. To fulfill this condition, we designed a counter block module, as shown in Fig. 10, which provides a 128-bit word, and is, in turn, divided into two 64-bit values. The 64 most significant bits correspond to a message nonce, and the remaining 64 bits correspond to the count of the counter, which can start counting from any value. It increases by one for each clock cycle. According to [5], the number of plaintext blocks to be encrypted must satisfy $n < 2^{64}$ in order to counter values do not repeat. If this condition is not met, counter values can repeat themselves. However, using the maximum frequency of $272.59 Mhz$, it would take around 2145.86 years for a counter to repeat. Therefore, the condition $n < 2^{64}$ is satisfied.
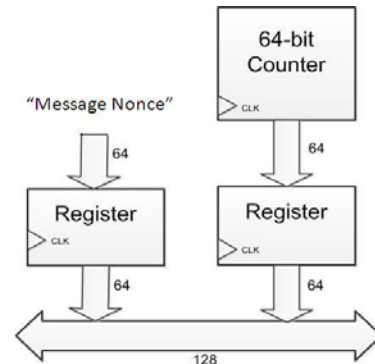
Figure 10. Counter block module.
Source: The authors
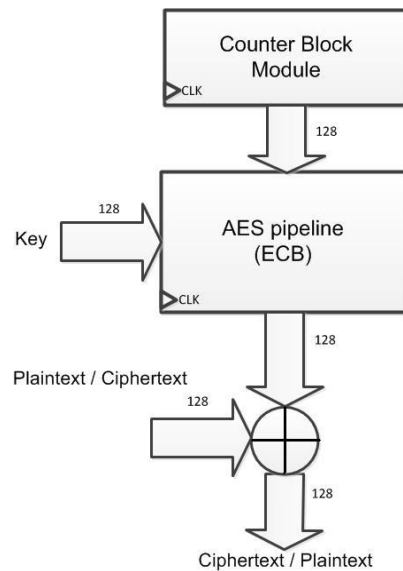
Figure 11. AES CTR mode operation.
Source: The authors

The counter mode provides better security properties than the ECB mode since the encrypted blocks are always different due to all the counter values are distinct. Only the cipher forward function is needed for the encryption and decryption process, as is shown in Fig. 11. However, the counter values must be the same for both encryption and decryption.

4. Comparison and performance results

All the modules described above were designed using VHDL and implemented using the Xilinx Virtex-5 XC5VLX110T FPGA [10], which is integrated into the XUPV5-LX110T Development System [11]. The proposed pipelined AES architecture provides a throughput of 34.89 Gbps and a clock frequency of 272.59 Mhz. Both features are higher than other AES designs on FPGA reports.

Table 1 shows the resources used in this implementation. From Table 1, we can see that for both designs the throughput is the same for both encryption on ECB and CTR mode. Although, on the CTR mode, a little more hardware resources are used. However, the throughput of the decryption process

Table 1.
Resources utilization of Virtex 5.

Mode	Slice LUT's	Slice Registers	Fmax (Mhz)	Throughput (Gbit/s)
ECB Enc.	11359	1289	272.6	34.89
ECB Dec.	13952	1289	199.4	25.53
CTR Enc / Dec	11677	1484	272.6	34.89

Source: The Authors

on ECB mode is less than it is on the CTR mode, and it uses more hardware resources.

The structure of slices between Virtex 5 and other series of Virtex before Virtex 5 are very different. Therefore, we cannot compare the amount of hardware resources used for other designs. However, we can compare the throughput to other AES pipelined approaches and we can see from Table 2 that our design has a high throughput and good resource efficiency.

5. Conclusions

The hardware implementation of the advanced encryption standard (AES) in non-feedback modes of operation as well as design details for the counter mode have been presented in this paper.

Our design achieves a high throughput of 34.9 Gbit/s and a good resource efficiency when compared to other AES pipelined designs.

Comparisons related to the implementation of the modes of operation have revealed that the CTR mode is an option that has considerable advantages over the ECB mode, such as the level of security. This is because the encrypted blocks on CTR mode are always different. Also, the ECB mode requires more hardware resources utilization than the CTR mode as the encryption and decryption hardware are different on ECB mode. Conversely, the hardware for encryption and decryption on CTR mode are the same.

Future research will consider carrying out ASIC implementations since these can offer higher speeds for data processing and an optimized structure.

Table 2.
Comparison of our results with other pipelined AES architecture designs.

Reference	Device	Fmax (Mhz)	Mode of operation	Function E/D	Slice LUT's	Throughput (Gbit/s)	T/S (Mbps Slice)
[12] Qu, Shou	XC5vlx85	576.07	CTR	E/D	22994	73.73	3.21
[13] Chih-Peng	Xc4vlx200	250.0	ECB	E	86806	32.00	0.37
[6] Nalini	XCV 2000	242.3	ECB	E/D	4626 (160 BRAM's)	30.88	6.67
	XCV 3200	241.3	ECB	E/D	4626 (160 BRAM's)	30.88	6.67
[14] M.R.M Rizk	Xc4vlx60	222.7	ECB	E	18855 (200 BRAM's)	28.51	1.51
	Xc4vlx60	180.4	ECB	D	20155 (200 BRAM's)	23.09	1.14
[5] Fu, Hao	Virtex 2	212.5	CTR	E/D	17887	27.1	1.51
[15] Hesham, Abd	XC5VLX50-3	294.8	ECB	E/D	6741	37.7	5.61
[16] Granado, Criado	XC2V6000-6	194.7	ECB	E	3576 (80 BRAM's)	24.9	1.8
[17] Liu, Xu	XC7VX690T	516.8	ECB	E	3436	66.10	19.20
Our Work	Xc5vlx110T	272.59	ECB	E	11359	34.89	3.07
Our Work	Xc5vlx110T	199.4	ECB	D	13952	25.53	1.83
Our Work	Xc5vlx110T	272.59	CTR	E/D	11677	34.89	2.99

Source: The Authors

References

[1] National Institute of Standard and Technology, Data Encryption Standard, Federal Information Processing Standards 46, November 1977.

[2] National Institute of Standards and Technology (NIST), Federal Information Processing Standards Publication 197. Advanced Encryption Standard (AES), [Online]. 2001. Available at: http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

[3] Bolaños, F. y Bernal, A., Una implementación hardware optimizada para el operador exponenciación modular, DYNA, 75(156), pp. 55-63, 2008.

[4] Naidu, A.P.A. and Joshi, P.K., FPGA implementation of fully pipelined advanced encryption standard. International Conference on Communications and Signal Processing (ICCSP), pp. 0649-0653, 2015. DOI: 10.1109/ICCSP.2015.7322568

[5] Dworkin, M., Recommendation for block cipher modes of operation, methods and Techniques, NIST special publication 800-38A. [Online].   2001.   Available   at: http://csrc.nist.gov/publications/nistpubs/800-38a/sp800-38a.pdf

[6] Fu, Y.F.Y., Hao, L.H.L., Zhang, X.Z.X. and Yang, R.Y.R. Design of an extremely high performance counter mode AES reconfigurable processor. Second International Conference on Embedded Software and Systems (ICESS'05). 2005.

[7] Nalini, C., Anandmohan, P., Poornaiah, D. and Kulkarni, V.D., An FPGA Based performance analysis of pipelining and unrolling of AES Algorithm. International Conference on Advanced Computing and Communications. 2006.

[8] Nieto, R., Diseño e implementación de un cripto procesador asincrono de bajo consumo basado en el algoritmo de Rinjdael, PhD Thesis, Escuela de Ingeniería Electrica y Electrónica, Universidad del Valle, Cali, Colombia, 2009.

[9] Shao, F., Chang, Z., Zhang, Yi., AES encryption algorithm based on the high performance computing of GPU, Second International Conference on Communication and Networks, IEEExplore, pp 598-590, 2010.

[10] XILINX, Virtex 5 FPGA User Guide, UG190(v5.4), [Online]. 2012. Available   at: http://www.xilinx.com/support/documentation/user_guides/ug190.pdf

[11] XILINX, ML505/ML506/ML507 User Guide, UG347(V 3.1.2), [Online].   2011.   Available   at: http://www.xilinx.com/support/documentation/boards_and_kits/ug347.pdf

[12] Qu, S., Shou, G., Hu, Y., Guo, Z. and Qian, Z., High throughput, pipelined implementation of AES on FPGA. 2009 International Symposium on Information Engineering and Electronic Commerce, (x), pp. 542-545. 2009. DOI: 10.1109/IEEC.2009.120

[13] Fan, C. and Hwang, J., Implementations of high throughput sequential and fully pipelined AES processors on FPGA. 2007 International Symposium on Intelligent Signal Processing and Communication Systems, pp. 353-356. 2007. DOI: 10.1109/ISPACS.2007.4445896

[14] Rizk, M.R.M., Member, S. and Morsy, M., Optimized area and optimized speed hardware implementations of AES on FPGA. 2007.

[15] Hesham, S., Abd-El Ghany, M.A. and Hofmann, K., High throughput architecture for the advanced encryption standard algorithm. 17th International Symposium on Design and Diagnostics of Electronic Circuits   &   Systems,   pp.   167-170.   2014.   DOI: 10.1109/DDECS.2014.6868783

[16] Granado-Criado, J.M., Vega-Rodriguez, M.A., Sanchez-Perez, J.M. and Gómez-Pulido, J.A., A new methodology to implement the AES algorithm using partial and dynamic reconfiguration, Integr. VLSI J., 43, pp. 72-80, 2010.

[17] Liu, Q., Xu, Z. and Yuan, Y., A 66.1 Gbps single-pipeline AES on FPGA. 2013 International Conference on Field-Programmable Technology   (FPT),   pp.   378-381,   2013.   DOI: 10.1109/FPT.2013.6718392

I.C. Guzmán-Velásquez, received his BSc. in Electronic Engineering from the Universidad del Valle, Cali, Colombia in 2013. Currently, he is pursuing a MSc. degree in electrical engineering at the Universidad del Valle, Cali, Colombia. His research interests include: digital circuit design, computer architecture and signal processing.
ORCID: 0000-0002-4532-8814

R.D. Nieto-Londoño, received his BSc. in Electrical Engineering from the Universidad del Valle, Cali, Colombia in 1995. He received his MSc. degree with an emphasis in Automatic Engineering from the Universidad del Valle, Cali, Colombia in 2001 and his PhD. from the Universidad del Valle, Cali, Colombia in 2009. Currently, he is a professor at the Universidad del Valle in the School of Electronic and Electrical Engineering. His research interests include: digital circuits design, low-power digital design and computer architecture.
ORCID: 0000-0002-1113-3269

A. Bernal-Noreña, received his BSc. in Electrical Engineering from the Universidad del Valle, Colombia, and his MSc. in Electrical Engineering, majoring in VLSI circuit design, from Sao Paulo University, Brazil in 1992. In 1999 he received his PhD. with emphasis in microelectronic engineering from the Institute National Polytechnique, Grenoble, France. In 1993, he joined the School of Electronic and Electrical Engineering at the Universidad del Valle where he teaches CMOS VLSI Design, Physics of Semiconductor and Electronic Devices. His research interests include: digital circuits design, low-power digital CMOS and embedded systems. Currently, he is the director of the Digital Architectures and Microelectronic research Group.
ORCID: 0000-0003-4766-8086

UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería
Eléctrica e Ingeniería de Control

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Eléctrica

Mayor información:

E-mail: ingelcontro_med@unal.edu.co
Teléfono: (57-4) 425 52 64

Determination of the domain dimensions in embankment numerical modeling

Milena Mesa-Lavista [a], José Álvarez-Pérez [a] Eduardo Tejeda-Piusseaut [a] & Carlos A. Recarey-Morfa [b]

[a] School of Civil Engineering, Polytechnic University José Antonio Echeverría (CUJAE), Havana, Cuba. milenaml@civil.cujae.edu.cu, joseap@civil.cujae.edu.cu, etejeda@civil.cujae.edu.cu
[b] School of Constructions, Center for Mechanical Computational Investigations and Numerical Methods in Engineering. Central University of Las Villas (UCLV), Villa Clara, Cuba. recarey@uclv.edu.cu

Abstract
This article presents a study of the influence of subdomain dimensions on the stress-strain state of road embankments. Its purpose is to determine the most advisable distance from the base of the slope to the boundary condition using computational numerical modeling as the chief tool, and performing an experimental design 3^3 in order to achieve the stated goal. Recommendations are given for using different boundary conditions when varying the embankment slopes in problems related to road embankment modeling.

Keywords: Boundary conditions, Road embankment modeling, Numerical Modeling.

Determinación de las dimensiones del dominio en la modelación numérica de terraplenes

Resumen
El presente artículo presenta un estudio de la influencia de las dimensiones del subdominio en el estado tenso-deformacional de terraplenes viales. Su objetivo principal es determinar la distancia más aconsejable desde el pie de talud hasta la condición de frontera empleando como herramienta fundamental la modelación numérica y realizando un diseño de experimento 3^3 para cumplir con el mismo. Este trabajo ofrece recomendaciones para el empleo de diferentes condiciones de contorno al variar las pendientes de los taludes en problemas relacionados con la modelación de terraplenes de obras viales.

Palabras claves: Condiciones de frontera; Modelación de terraplenes viales, Modelación numérica.

1. Introduction

In order to determine the active depth of the soil foundations, it is necessary to calculate the stress that emerges due to the load imposed by the embankment weight; that is, the ratio result of the stresses produced by the weight of the soil foundation and the load imposed by the embankment, which should be between 0.1 and 0.2 [1,2]. As the depth of the soil foundation is infinite, the smaller the mathematical model of the problem, the smaller the computational resources [3]. Therefore, to optimize the computational cost, we have to decrease the size of the model without affecting the solutions in the working area.

2. Model conceptualization

The modeling of a real problem is described through a physical and mathematical model in which there is an interrelation between the characteristics each of them possesses: geometric model, constitutive model, load model, boundary condition model, governing equations, initial conditions, and definition of the work domain (Fig. 1). It is precisely the latter that is the most essential feature dealt with in this paper. Since soil is considered an infinite area, it is necessary to carry out a study of the adjacent domain of the geometric model that characterizes it as well as of the influence of boundary conditions.
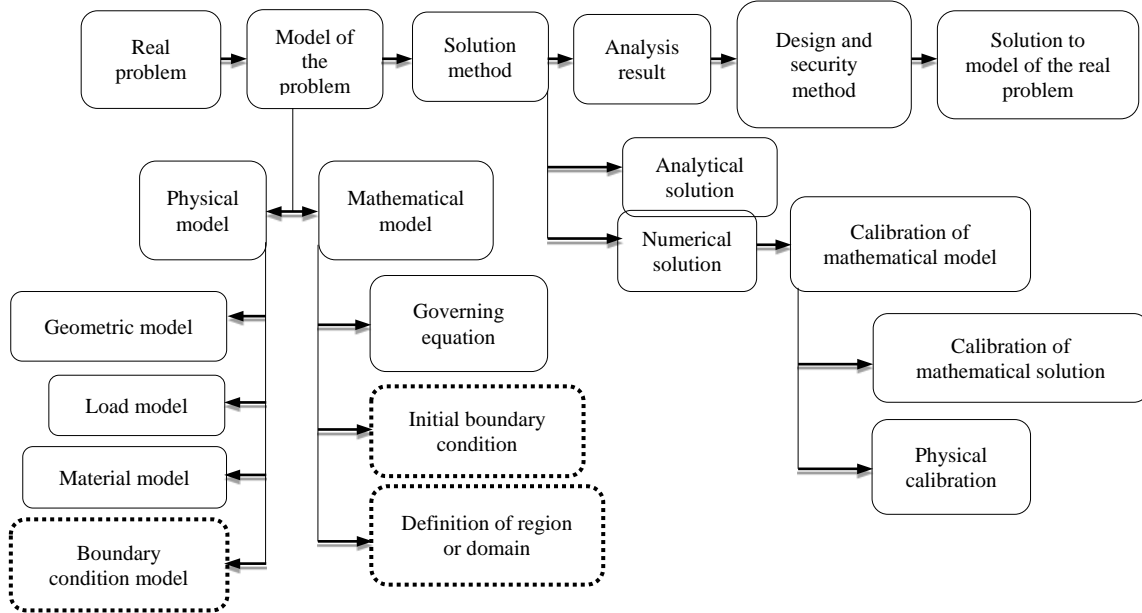
Figure 1 Methodological diagram for the solution to a real problem
Source: [4]

2.1. Construction of the numerical model

In this study, the numerical model was implemented based on the Finite Element Method (FEM); a plane strain was applied to the numerical model, and the slopes and the height of the embankments varied (Fig. 2). The letter C denotes the distance of the subdomain which is defined according to the relationship B/H, (ratio between the slope base and its height) (Table1). The constitutive model and the load model were used, taking into consideration the specialized literature and standard regulations [1,5-7] (Table 2).

In previous studies, the analysis was carried out [4,8,9] using the physical and mathematical parameters of the road embankment problems calibrated simultaneously after a homogenization process. As a result, the four-node quadrilateral Finite Element Type (FET) and 50cm of the domain fragmentation were chosen as a solution to this problem.

2.2. Boundary condition modeling

In this study, the boundary conditions applied were vertical and horizontal lineal supports, infinite elastic supports (Winkler´s Model), and supports with infinite elements.
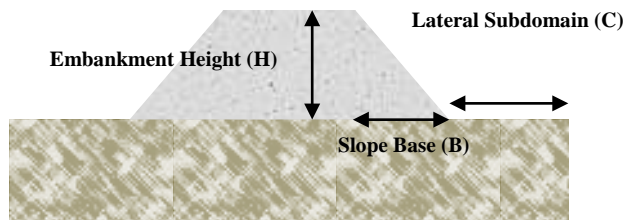
Figure 2 Geometric model
Source: The Authors

Table1.
Lateral Subdomain for the slope base and embankment height

C	Slope		
	(2:1)	(3:1)	(4:1)
4B/H	8m	12m	16m
6B/H	12m	18m	24m
8B/H	16m	24m	32m
10B/H	20m	30m	40m
12B/H	24m	36m	48m
14B/H	28m	42m	56m
16B/H	32m	48m	64m

Source: The authors

Table 2.
Mechanical properties of soil

Parameters	Embankment Crown and embankment foundation	Nucleus and first soil foundation	Second soil foundation	Third soil foundation
E (kPa)	30000	22000	32000	33000
μ	0.30	0.32	0.33	0.28
Ɣ (kN/ m3)	19	18	17	18
Ø (º)	41	36	38	35
C (kPa)	0	20	0	0
Thickness (m)	0.75	(12,14,16) and 6	22	35

Source: [9]

- Vertical and horizontal lineal supports are used to inhibit the displacements in the X and Y axes as stated in the literature [10], by restricting the lateral and depth displacements, or imposing the null displacements to the ends of the model of the problem (Fig. 3, a). When these elements are used, the constitutive material model is not restricted.

a) Vertical and horizontal lineal supports

b) Infinite elastic supports

c) Supports with infinite elements
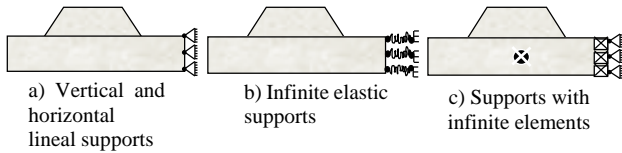
Figure 3 Representation of boundary conditions
Source: The Authors

a) Key area

b) Node in front of the support

Figure 4 Ways for determining the subdomain dimensions
Source: The Authors

- Infinite elastic supports (Fig. 3, b) are equivalent to an infinite number of elastic springs whose rigidity is the relationship between the pressure applied to the soil and the resultant displacement. For the numerical representation, the ballast coefficient of Winkler´s theory [11-14] is considered. Winkler´s Model supports are also employed in solving the problems concerning soil-structure interaction. Thus, an elastic constitutive material model is preferable when using this model.

- Supports with infinite elements (Fig. 3, c) are special elements which use a polynomial interpolation in such a way that the interpolation function describes an asymptote. When the working area decreases, the function makes the working space a continuous and infinite soil area. Consequently if it is necessary to design a numerical model of large-scale geometry, the infinite elements will guarantee the continuity of the soil foundation to reduce the geometric model to a smaller size. In this case, the constitutive model has to be lineal and elastic.

Table 3.
Experimental design

Boundary conditions	Height	Slope
Vertical and horizontal lineal supports	12m	2:1
Infinite elastic supports	14m	3:1
Supports with Infinite Elements	16m	4:1

Source: The Authors

3. Determination of the subdomain dimensions

Different methodologies can be followed to determine the dimension of a subdomain in an infinite soil space:

1.1. The displacements in the nodes are analyzed in front of the supports, which should tend to zero. Otherwise, the approximation to the numerical solution of the boundary will affect the answer to the problem in the working area (Fig. 4, a).

   1.1.1. If vertical and horizontal lineal supports are used, the null displacements are imposed.

   1.1.2. If infinite elastic supports are employed, the solution will be found according to the ballast coefficient.

   1.1.3. If the supports are infinite elements, the continuity of the soil foundation area will be guaranteed.

1.2. Stress, strain and displacement can also be analyzed at a point inside a key area, and the order of these values is compared to the order of the corresponding variables values near the supports, to determine the distance where the variables acquired the approximation desired (Fig. 4, b).

1.3. Another criterion to obtain the value of the subdomain dimensions is the analysis of several nodes in front of the supports or in the key area by applying the Standards of Error at several points [3].

Node1551

Figure 5 Embankment model
Source: The Authors

To obtain the subdomain dimensions of the numerical model in a semi-space of soil foundation corresponding to an embankment, an experimental design 3^3 was performed. The three following factors: boundary conditions, height, and slope varied by three levels each (Table 3). In the experimental design presented, the dependent variables were obtained in a node located within a chosen key area (Fig. 5).

4. Results and discussion

Twenty-seven models were applied to the embankment design to determine the subdomain dimensions which varied from 4B/H up to 16B/H each (Table 1). Throughout the analysis of the post-process, the results for the horizontal strain values were studied in terms of the influence of the boundaries.

The relative errors (Eq. 1) were obtained taking into account the subdomain maximum values. The maximum value for parameter "C" was 16B/H, which was considered the standard error value. Errors were analyzed when the results were lower than 5%.

$$E_R = \frac{V_i - V_{Patron}}{V_{Patron}}$$

Equation 1 Standard of Relative Error
Source: [3,9]

Relative errors in the problem solution were calculated for the three specific heights analyzed by using slope 2:1 with vertical and horizontal lineal supports as well as infinite elastic supports. Errors were lower than 5% in the subdomain values ranking between 8B/H and 16B/H whereas relative errors increased for the smallest values of the subdomains 6B/H and 4B/H (Figs. 6-8).

Since the relative error values were in the same numerical order and the results were lower than 5%, the recommended value for the subdomain dimension is 4B/H (Fig. 9) if supports with infinite elements are used. On the other hand, when the results of slopes 3:1 and 4:1 were analyzed with vertical and horizontal lineal supports and infinite elastic supports then, the subdomain values were 4B/H. However, slopes 3:1 and 4:1 with infinite elements provided a relative error of lower than 5% for each value of the sub-domain dimensions including B/H. Hence any value for the subdomain dimension can be used. Each case analyzed is shown in Table 4.
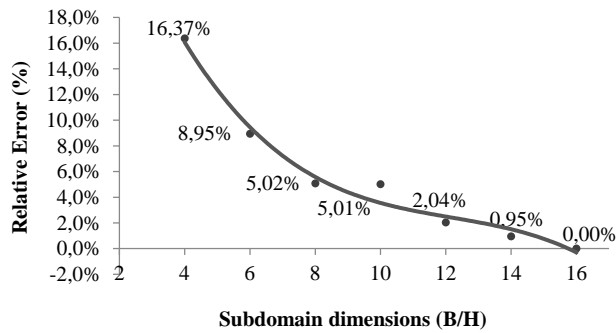
Figure 6 Relative errors using vertical and horizontal lineal supports in the 12 meter-high model with slope 2:1
Source: The Authors
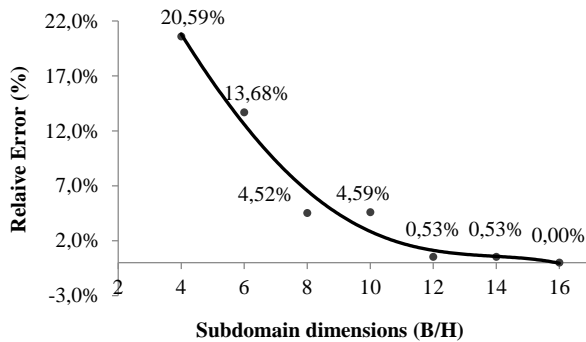
Figure 7 Relative errors using vertical and horizontal lineal supports in the 14 meter-high model with slope 2:1
Source: The Authors
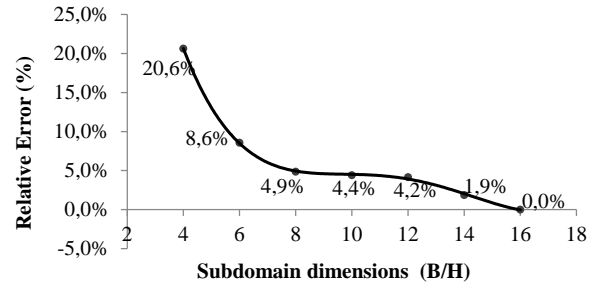
Figure 8 Relative errors using vertical and horizontal lineal supports in the 16 meter-high model with slope 2:1
Source: The Authors
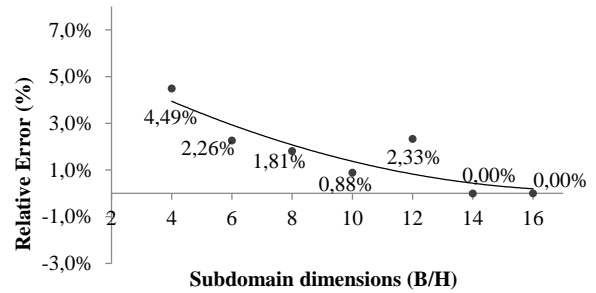
Figure 9 Relative errors using supports with infinite elements in the 14 meter-high model with slope 2:1
Source: The Authors

Table 4.
Summary of subdomain dimension values and boundary conditions for road embankment numeric model

Boundary conditions	Slope	Subdomain dimension values	Constitutive model
Vertical and horizontal lineal supports	2:1	$8\frac{B_T}{H}$	Linear – elastic Elastic – plastic
	3:1	$4\frac{B_T}{H}$	
	4:1		
Infinite elastic supports	2:1	$8\frac{B_T}{H}$	Elastic
	3:1	$4\frac{B_T}{H}$	
	4:1		
Supports with infinite elements	2:1	$4\frac{B_T}{H}$	Linear – elastic
	3:1	$\frac{B_T}{H}$	
	4:1		

Source: The Authors

5. Conclusions

1. The C=4B/H value is suggested for the sub-domain dimension when slope 2:1 is employed with infinite elements.
2. The recommended value of the sub-domain dimension is C=4B/H when slopes 3:1 and 4:1 are used.

3. With slope 2:1 and vertical and horizontal lineal supports, the appropriate proportion for the geometrical model is C=8B/H.
4. When using supports with infinite elements the constitutive model should be elastic and lineal, and the sub-domain dimension can vary indistinctly for slopes 3:1 and 4:1.
5. If infinite elastic supports are employed the constitutive model should be elastic according to the ballast coefficient.

Acknowledgments

We would like to thank professor Odalys Morales Chacon for reviewing the English in this text.

References

[1] Jiménez-Salas, J.A., De-Justo, J.L. y Serrano, A., Geotecnia y Cimientos III, 2a ed. vol. 1, Madrid, España, 1981.
[2] Juarez-Badillo, E. and Rico, A., Mecánica de Suelos vol. 1. Ciudad Habana, Cuba: Pueblo y Educación, 1972, pp. 20-443.
[3] Recarey, C.A., Modelación del terreno y las estructuras en el dominio del tiempo, Tesis Dr., Departamento de Ingeniería Civil, Facultad de Construcciones, Universidad Central "Marta Abreu" de las Villas. (UCLV), Villa Clara, Cuba, 1999, 140 P.
[4] Mesa, M., Álvarez, J. y Recarey, C., Proceso de algoritmo para calibración matemática, Memorias de la 16 Convención Científica de Ingeniería y Arquitectura, 2012, 13 P.
[5] Escario, V., Los materiales para terraplenes. Construcción de terraplenes, in Kraemer, C., Uriel, R.S., ed., Explanaciones y drenaje I, E.T.S. de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Madrid, España, 1993, pp. 43-74.
[6] Torres-Vila, J.A., Introducción a la construcción de las explanaciones, in Diseño y construcción de explanaciones. Vol. 1, ed La Habana, Cuba: Avenida del Bosque, 1986, pp. 227-242.
[7] NC-63-153, Empuje de tierra, procedimiento de calculo,"ed: Norma cubana de la construcción, 1985.
[8] Mesa, M. y Álvarez, J., Técnicas de calibración numérica para problemas de ingeniería, Memorias de la XXXIII Convención Panamericana de ingenierías (UPADI), 2012, 12 P.
[9] Mesa, M. y Álvarez, J., Calibración numérica de un problema de ingeniería vial, Revista de la Construcción, vol. 10, pp. 52-63, 2011. DOI: 10.4067/S0718-915X2011000300006
[10] Zienkiewicz, O.C. y Taylor, R.L., El método de los elementos finitos, 5ta ed. vol. 1, 2 y 3, Barcelona, España: CIMNE, 2004.
[11] Muzás, F., Consideraciones sobre la elección de coeficientes de balasto, Revista de Obras Públicas, [En línea]. 3(425), pp. 45-51, 2002. [Fecha de consulta julio de 2011]. Disponible en: http://worldcat.org/issn/00348619
[12] Leoni, A.J., Apunte de coeficiente de balsto, Facultad de Ingenieria U.N.L.P., [En líena]. 2009, 30 P, [fecha de consulta Marzo de 2011]. Disponible en: http://materias.fi.uba.ar/7411/curso/teoria/balasto/leoni.pdf
[13] López-Menchero, J. y Lozano, J., Presión admisible por el terreno para losa., 2009, 11 P.
[14] Aragón, J., El coeficiente de balasto, [En línea]. 2006, [fecha de consulta marzo de 2011]. Dsiponible en: http://radioamericahn.net

M. Mesa-Lavista, has a degree in Civil Engineering from the Polytechnic University José Antonio Echeverría (CUJAE), Cuba. At present, she is researching Mechanical Computational and Numerical Methods in Engineering for her PhD. She has worked as a researcher at the Center for Research and Development of Structures and Materials (CIDEM) and at the National Company for Applied Investigations of Cuba (ENIA). As well as teaching, she currently works as a researcher at the Road Department of the School of Civil Engineering, at the CUJAE. Her research topics include simulation, embankment numeric modeling, mathematical calibration problems and road modeling problems.
ORCID:0000-0001-5310-3255

J. Álvarez-Pérez, is graduated from the School of Civil Engineering at the Polytechnic University José Antonio Echeverría (CUJAE), Cuba in 2009. He completed his PhD at the Central University of Las Villas (UCLV), Cuba, in 2014 with a specialization in "Laminar Structures". In 2010, he worked at the National Project Engineering Company (EMPROI). Besides teaching, he also carries out research at the Department of Structures at the School of Civil Engineering, CUJAE. His research interests include simulation, numerical modeling, mathematical calibration problems, laminar structures, and calculus of laminar structures in relative coordinates.
ORCID: 0000-0002-6210-8477

E. Tejeda-Piusseaut, is a full professor at the Roads Department at the School of Civil Engineering at the Polytechnic University José Antonio Echeverría (CUJAE), Cuba. He has a PhD from the University of Madrid, Spain, specializing in pavement recycling. He has carried out several studies on recycled materials for pavement construction, flexible and rigid pavements, and embankment modeling.
ORCID: 0000-0002-0589-9989

C.A. Recarey-Morfa, completed his PhD at the Central University of Las Villas (UCLV) in 1999, specializing in Simulation of Rheology Modeling. In addition to being the director of the CIMCNI (Research Center on Computational and Numerical Method in Engineering), he is also a professor and researcher at the School of Construction at the UCLV. He is a member of the International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain. His research areas comprise particles simulation, numerical simulation, rheological and structural modeling, and geotechnical and foundation modeling.
ORCID: 0000-0002-0408-4032
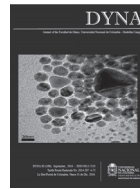
Modeling and behavior of the simulation of electric propagation during deep brain stimulation

Pablo A. Alvarado [a], Cristian A. Torres-Valencia [a], Álvaro A. Orozco-Gutiérrez [a], Mauricio A. Álvarez [a], Genaro Daza-Santacoloma [b] & Hans Carmona-Villada [b]

[a] Universidad Tecnológica de Pereira, Pereira, Colombia. dapa@utp.edu.co
[b] Instituto de Epilepsia y Parkinson del Eje Cafetero – Neurocentro, Pereira, Colombia

Abstract
Deep brain stimulation (DBS) is an effective treatment for Parkinson's disease. In the literature, there are a wide variety of mathematical and computational models to describe electric propagation during DBS; however unfortunately, there is no clarity about the reasons that justify the use of a specific model. In this work, we present a detailed mathematical formulation of the DBS electric propagation that supports the use of a model based on the Laplace Equation. Moreover, we performed DBS simulations for several geometrical models of the brain in order to determine whether geometry size, shape and ground location influence electric stimulation prediction by using the Finite Element Method (FEM). Theoretical and experimental analysis show, firstly, that under the correct assumptions, the Laplace equation is a suitable alternative to describe the electric propagation, and secondly, that geometrical structure, size and grounding of the head volume affect the magnitude of the electric potential, particularly for monopolar stimulation. Results show that, for monopolar stimulation, basic and more realistic models can differ more than 2900%.

Keywords: DBS; Parkinson disease; electric brain propagation; Laplace equation; FEM.

Modelado y comportamiento de la simulación de propagación eléctrica durante la estimulación cerebral profunda

Resumen
La Estimulación Cerebral Profunda (DBS) es un tratamiento efectivo para la enfermedad de Parkinson. Gran variedad de modelos matemáticos y computacionales para describir la propagación eléctrica debido a la DBS han sido propuestos, desafortunadamente, no existe claridad sobre las razones que justifican el uso de un modelo específico. En el presente trabajo se presenta una formulación matemática detallada de la propagación eléctrica debido a DBS que soporta un modelo basado en la ecuación de Laplace. Se realizan simulaciones para diferentes modelos geométricos del cerebro para determinar si la geometría, el tamaño y la ubicación de la tierra del modelo afectan la predicción de la estimulación eléctrica mediante el uso del Método de Elementos Finitos (FEM). Los análisis teórico y experimental muestran en primera instancia que la ecuación de Laplace es adecuada para describir la propagación eléctrica en el cerebro, y en segunda instancia que la estructura geométrica, tamaño y ubicación de la tierra afectan la magnitud del potencial eléctrico, particularmente para modos de estimulación monopolar. Los resultados muestran que para modelos básicos y más realistas pueden existir diferencias en la propagación de hasta un 2900%.

Palabras Clave: Estimulación Cerebral Profunda; Ecuación de Laplace; Enfermedad de Parkinson; FEM.

1. Introduction

Parkinson's disease (PD) is a degenerative disorder of the central nervous system that results in impaired motor skills and speech. Its most prevalent symptoms are tremor and rigidity [1]. PD is the second most common neurodegenerative disorder after Alzheimer's disease, often affecting the elderly population [2].

Deep brain stimulation (DBS) is a clinically effective treatment for medically intractable PD [3]. To improve all PD

symptoms, it is best for DBS to target the Subthalamic Nucleus (STN) [4], the brain structure related to sensorimotor, cognitive, and limbic functions [5]. The fundamental purpose of DBS is to modulate neural activity with applied electric fields [6]. However, the mechanisms by which DBS works are not yet well understood [7]. In this sense, DBS's therapeutic action seems to depend on the electrical excitation of neural elements [8]. Moreover, there are also studies that support neuronal inhibition [9]. Other studies suggest that DBS reduces the PD symptoms through the excitation of axons and the inhibition of the dendritic activity [10, 11].

To achieve successful stimulation, it is necessary to excite the intended brain areas while preventing the unintended excitation of other zones: the spread of current to non-motor areas of the STN or adjacent structures is implicated in cognitive and cognitive-motor declines [12-14]. The stimulation of the dorsolateral STN and the bottom (ventral) part of the thalamus could reduce parkinsonian tremor and trigger dyskinesias, whereas stimulation outside the STN could induce adverse effects [15].

A suitable stimulation protocol involves not only the accurate placement of the electrode inside the brain, but also the proper configuration of some electrical and geometrical parameters for the DBS device [4]. The electrical parameters for DBS are pulse width, frequency and the voltage amplitude. Additionally, each of the lead's electrodes can be designated as anode or cathode [4], To facilitate the configuration of the DBS device it is propitious to employ computational models, this allows the electric propagation of the stimulation to be predicted as a function of the previously mentioned electrical and geometrical parameters.

These computational simulations help to visualize the electric behavior of the stimulus in the brain. In this sense, several works [16 − 18] have developed simulators of the electric activity for DBS.

The mathematical and computational models found in the literature [18 − 21] require information such as the conductivity and permittivity of brain tissue, geometrical description of the head volume, the physical laws that govern the system, and the associated equation constraints. Most of the simulation approaches are specifically based on electrostatic models. The electric potential is often computed using the Laplace [17, 22] or Poisson's [23, 24] equation. Unfortunately, there are no major justifications about the use of this mathematical background, which is essential to define the scope, realism and accuracy of the simulation. The core of these simulations is the Finite Element Method (FEM) that has been widely used in DBS problems and other engineering fields (see [25] and [26]).

Previous research undertaken by authors such as [27] and [28] address some of the effects of the DBS that show some simulations from schemes different to the one proposed in this work. In [27], a latent force model was developed in order to include the dynamics of the electric propagation in the brain, unlike several state-of the-art works that only focus on the quasi-static or static approach. In [28], some propagation models following the quasi-static approach were developed using an open source library of finite element methods with no deep analysis of the physical laws that govern the DBS problem. Additionally the results are difficult to compare against the state-of the-art works due to the difference in the simulation tool used.

It is unusual to find academic discussions about the physical laws that support the behavior of the deep brain electric fields induced by an external source. In fact, there is no interpretation or explanation about the consequences of most of the mathematical simplifications carried out by the basic equations that describe the phenomenon. Moreover, in order to establish which kind of representations are appropriate to describe the electric propagation inside the human brain's behavior, it is convenient to make a quantitative comparison of several geometrical head models, taking into account the ground positioning that is assumed by the computational algorithms.

In this work, we present a mathematical formulation of the electric propagation during DBS. Indeed, we offer an argument that sustains the use of an electrostatic propagation model based on the Laplace equation. The theoretical framework is corroborated by a set of computer simulations of the electric potential generated by DBS. Furthermore, the simulation analysis indicates that, for monopolar stimulation, the geometrical structure, size and grounding of the conducting head volume alter the magnitude of the electric field. In fact, a voltage comparison between basic and more realistic models can differ by more than 2900%.

2. Deep brain stimulation considerations

An accurate treatment of Parkinson's disease using DBS should analyze the different effects of potential propagation around the objective structure, that is the STN [15]. Adverse effects could be produced from undesired potential propagation to non-motor regions of the brain, as is presented in Fig. 1. In order to improve the Parkinsonian motor symptoms, the electrode must be placed at the motor section of the STN, as presented in Fig. 2 [15].

Given a specific electrode, e.g. the Medtronic DBS lead model 3389 that has four configurable electrodes, there are several geometrical possible arrangements to configure the stimulation parameters. In clinical practice, usually one or two stimulation contacts are used at most. Fig. 3 shows three different monopolar (Fig. 3(a)) and bipolar (Fig. 3(b)-(c)) configurations and their corresponding electric potential [8].
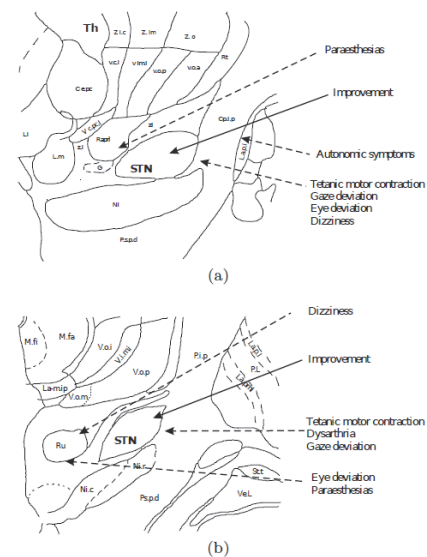
Figure 1. Sites of stimulation-induced effects in the STN region.(a) Sagittal view. (b) Coronal view.
Source: [15]
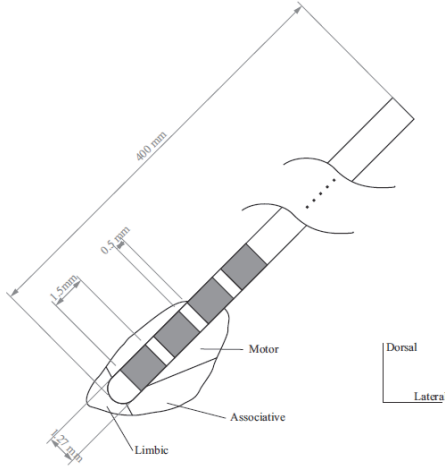
Figure 2 Positioning of the electrode at STN (coronal view) (Medtronic DBS lead model 3389).
Source: [15]
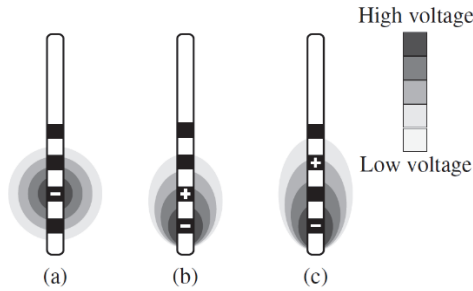
Figure 3 Examples of commonly used electrode configurations. (a) Monopolar. (b), (c) Bipolar.
Source: [8]

3. Electric stimulation modeling

Electromagnetic fields generated by DBS are dynamic since the source field or electric stimulation is time-varying and has a fundamental frequency range from 130Hz to 185Hz [7, 20, 29, 30] (the frequency commonly used is around 140Hz). Moreover, the electric potential induced throughout the brain tissue close to the stimulating electrode is commonly modeled using the Laplace equation, which assumes a quasi-static or static field [17 − 21].

It is worthwhile mentioning that the quasi-static approximation is only valid when the electrodynamic system analyzed is *a low frequency time-varying* field [31 − 33]. In this section, we provide a detailed explanation of how to derive the quasi-static model in order to support a DBS propagation model based on the Laplace equation. This explanation involves the use of generalized Maxwell's equations and some physical assumptions. We then present the conditions which allow us to make a decision as to whether the approximation is valid for DBS.

3.1. Low frequency range, time-varying fields

The large variety of electromagnetic phenomena can all be described by a unique system of field equations known as Maxwell's equation [34]. Some particular forms of these equations have been used by other authors to model the electric propagation produced by DBS [17, 22 − 24]. These equations can be simplified when slow electromagnetic fields are analyzed, i.e. fields in the so called *low frequency range* (up to 30kHz), when wave propagation does not play a fundamental role [31,34]. Before defining the situations in which wave propagation effects can be neglected, it is important to clarify some electromagnetic waves properties.

Generally, electromagnetic fields propagate with a finite velocity c [34], defined as $c = \frac{1}{\sqrt{\varepsilon\mu}}$ [m/s], where ε denotes the permittivity and μ represents the permeability of the brain tissue [31]. In addition to this, τ_{em} represents the time required for the electromagnetic field to propagate at a distance l from one region to another in a volume brain tissue, $\tau_{em} = \frac{l}{c}$ [s]. The wave propagation equation for the electrodynamic scalar potential is defined as:

$$\nabla^2 V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\varepsilon} \tag{1}$$

Where V is the electric potential function, and ρ denotes the charge density [35]. If the field problem is considered with a characteristic spatial dimension l and a characteristic time constant τ, spatial and temporal differentiations can be approximated by $(1/l)$ and $(1/\tau)$, respectively. In this case, l is related to the brain tissue volume considered, i.e. the STN and its surroundings, whereas τ is considered as the time interval for which significant changes in the field quantities arise. For time-varying electric stimulation, τ would be the reciprocal of the excitation's angular frequency, $\tau = \omega^{-1}$ [31, 34]. If these previous considerations are applied, equation (1) can be approximated by:

$$\nabla^2 V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} \approx \frac{V}{l^2}\left(1 - \left(\frac{\tau_{em}}{\tau}\right)^2\right)$$

For slow time-varying fields, the characteristic time constant τ is supposed to be much greater than the transit time τ_{em}, i.e. $\left(\frac{\tau_{em}}{\tau}\right)^2 \ll 1$. If this expression holds, then $\partial^2 V/\partial t^2 \approx 0$, and the propagation effects can be neglected.

3.2. Static and quasi-static models

When wave propagation does not play a fundamental role, the electromagnetic field simulations of slow processes are carried out by using [36].

- a static model, i.e. *electrostatics or magnetostatics,* if all variations in time can be neglected.
- a quasi-static model, i.e. *electro quasistatics or magneto quasistationary.*

The static models are just special cases of the full Maxwell's equations, whereas the quasi-static models are approximations that are not always valid [31, 32]. The quasi-static models are obtained from Maxwell's equations by neglecting either the magnetic induction, or the electric displacement current, as well as the electromagnetic waves that result from their coupling [32].

3.2.1. Electro-quasistatic model

The electro-quasistatic assumption establishes that the electric field **E** is essentially irrotational. In general, the field of gradient V (for any scalar V) is purely irrotational since $\nabla \times (\nabla V) = 0$, thus the irrotational field **E** can always be expressed in terms of a scalar field V, that is

$$\boldsymbol{E} = -\nabla V \qquad (2)$$

The negative sign shows that the direction of **E** is opposite to the direction in which V increases. The electric field **E** looks like an electrostatic field at any tissue point. Changes in the electric stimulation will immediately take effect in the whole brain tissue volume under consideration.

3.2.. *Magneto-quasistationary model*

Analogously, the magneto-quasistationary models are characterized by setting the magnetic field **H** as solenoidal. This implies that the divergence of current density **J** is zero, i.e.

$$\nabla \cdot \mathbf{J} = 0$$

3.2.3. Laplace equation

If electro-quasistatic and magneto-quasistationary approximations are simultaneously applied, then all temporal variations in Maxwell's equations are neglected. This does not mean, however, that the sources, and hence the fields, are not functions of time. But, given the sources at a certain instant, the fields at that same instant are determined without regard for what the sources of fields were an instant earlier. Using Maxwell's equations and Ohm's law, the Laplace equation used to model the electric potential in DBS can be derived. The current density **J** is related to the electric field **E** by Ohm's law as follows [31, 32]:

$$\mathbf{J} = \sigma \mathbf{E} \qquad (3)$$

Where σ is the tissue conductivity. It is measured in Siemens per meter (S/m). If the divergence is applied on both sides of (3), we have $\nabla \cdot \sigma \mathbf{E} = 0$, and using (2) we get the Laplace equation:

$$\nabla \cdot \sigma(\nabla V) = 0 \qquad (4)$$

Equation (4) corresponds to an inhomogeneous tissue. For a homogeneous tissue, equation (4) becomes:

$$\nabla^2 V = 0 \qquad (5)$$

In order to obtain Equation (5), the conductivity σ is assumed constant throughout the tissue region in which V is defined. The Laplacian operator $\nabla^2 V$ can be defined in Cartesian coordinates in the following way:

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}$$
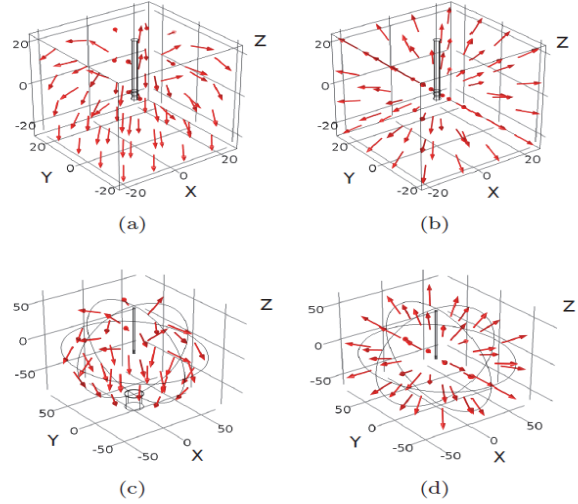
Figure 4: Electric field distribution within a specific geometry and different boundary conditions when DBS is performed. (a) Cubic geometry, ground on base side. (b) Cubic geometry, ground on whole boundary. (c) Spherical geometry, ground on base side. (d) Spherical geometry, ground on whole boundary.
Source [authors]
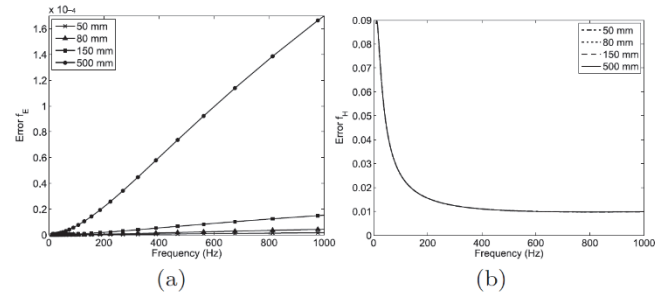
Figure 5(a) Electro-quasistatic approximation errors (f_E) for different frequencies and radius sizes. (b) Magneto-quasistationary approximation errors (f_H) for different frequencies and radius sizes (all curves are almost the same).
Source [authors].

The electric potential calculation is based on a model with a homogeneous tissue medium to reduce model complexity. Several authors have developed their experiments using this assumption [6, 17, 21, 24]. Furthermore, the STN is cytologically homogeneous, i.e., neurons are identical in every part of the nucleus [37]. We will now present four examples of the electric field (**E**) propagation obtained solving the Laplace equation (5) for a finite, homogeneous, and isotropic volume tissue, using different geometries and boundary conditions. The red arrows in Fig. 4 correspond to the electric field. Fig.4 (a) and 4(b) show a cubic geometry, in Fig. 4(a) just one side of the cube is grounded, in Fig. 4(b) all sides of the cube are grounded. Likewise, Fig. 4(c) and 4(d) show the electric field distribution (see Equation (2)) obtained for a spherical geometry. In Fig. 4(c) a small base is grounded, whereas in Fig. 4(d) all the external surface of the sphere is grounded.

3.3. Conditions for the quasistatic approximation

The validity of an approximation for a given slow time-varying field problem is determined by an analysis based on significant time constants [31]. In this sense, two constants are defined, the time constant of dielectric relaxation $\tau_e = \frac{\varepsilon}{\sigma}$, and the constant of magnetic diffusion $\tau_m = \mu\sigma l^2$. In addition, the transit time τ_{em} is the geometric average of τ_e and τ_m.

$$\tau_{em}^2 = \tau_e \tau_m = \left(\frac{l}{c}\right)^2$$

The electro-quasistatic and magneto-quasistationary approximations can be used if the relative error of the electric field and magnetic field calculated under these approximations are much smaller than one. In order to estimate this error, time derivatives in Maxwell's equations are substituted by $1/\tau$. Furthermore, only the scalar magnitudes of the fields are considered. All properties of the brain tissue are assumed to be homogeneous, linear and isotropic. The relative error f_E of the electric field within the electro-quasistatic approximation is defined as:

$$f_E = \left(\frac{\tau_{em}}{\tau}\right)^2 \left(1 + \frac{\tau}{\tau_e}\right) \ll 1 \qquad (6)$$

If this condition holds, electric fields can be calculated accurately by using the electro-quasistatic approximation [17]. Likewise, the relative error f_H of the magnetic field within the magneto-quasistationary approximation is

$$f_H = \left(\frac{\tau_{em}}{\tau}\right)^2 \left(1 + \frac{\tau}{\tau_m}\right) \ll 1 \qquad (7)$$

Magnetic fields can be calculated by using the magneto-quasistationary approximation if this condition holds.

4. Experimental background

To be allowed to use the electro-quasistatic and magneto-quasistationary approximations to model the electric potential produced by DBS, the approximation errors f_E (6) and f_H (7) have to be much less than one. To verify this, the approximation errors were calculated for different l radius and stimulation frequencies. The dielectric properties of the tissue are frequency dependent [38], and the electric field propagation time τ_{em} is a function of the spatial quantity l [34]. Therefore, the errors f_E (6) and f_H (7) depend on the stimulation frequency and the size of the brain tissue region considered. The errors obtained for different frequencies (100Hz up to 1 kHz), assuming a radius of l = 50mm, l = 80mm, l = 150mm and l = 500mm, are shown in Fig. 5. According to the Andreuccetti online dataset [39], white matter dielectric property values where considered.. Based on Fig. 5, and assuming that all properties of the brain tissue arehomogeneous, linear and isotropic, we can conclude that the electro-quasistatic and magneto-quasistationary approximations are valid for a radius of between l = 50mm

and l = 500mm, and a frequency band from 100Hz to 1kHz.

Works such as [19] and [21] use several sizes of geometrical models in 2D and 3D. These include specifications of the DBS lead shape that go into a monopolar configuration and the specification for the tissue conductivity properties of the region analyzed. Usually, two different ground configurations of the electrical models are used, one to define all the boundaries of the geometrical model, such as the ground, and the other to configure a specific area of the model, such as the ground [21]. In [40], one model is developed assuming an infinite homogeneous and isotropic medium to compute the electric propagation in different large frequencies. In [41], a detailed model of the tissue surrounding the DBS lead is built using information from magnetic resonance imaging (MRI). The model is used to assess the influence of the tissue information when the electric field surrounding the electrode is computed. It should be noted that, for future work, the patient real head shape could be included and studied in order to increase the model´s realism. Research such as [42] where a reconstruction of the head from MRI is performed could be useful.

5. Results

The propagation of the electric potential in the simulated models is obtained by solving the Laplace equation from the finite element method (FEM) using Comsol Multiphysics (COMSOL Inc., Burlington, MA). As the theoretical analysis in section 2 demonstrated how the electric potential propagation is conductivity independent when a homogeneous medium is considered, the results obtained from these models allows for the geometry to be analyzed and for building effects to be modeled in the Laplace equation solution.. The main objective of this work is to present a detailed analysis of the electrostatic process that governs the electric propagation during DBS. Several DBS simulations based on the development of geometrical models of the brain that confirm the theoretical analysis of the electric propagation were built. The presented models include more realistic geometries that allow better analysis of the stimulation results. Different ground configurations and boundary constraints are proposed to determine the influence of the ground in terms of the electric propagation results. The electrical conductivity of a homogeneous medium is not taken into account because it has no influence over the solution obtained through the Laplace equation.

(a)   (b)   (c)
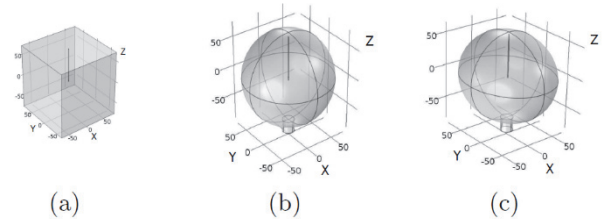
Figure 6. Geometrical forms considered to represent the volume of an adult human head. (a) Cube (50mm, 150mm, and 500mm edge length). (b) Sphere (80mm radius). (c) Ellipsoid (semi-axes x: 70mm, y: 82.5mm, and z: 65mm). Source [authors]

Three geometrical forms are considered to represent the volume of an adult human head. The first form is a cubic model (Fig. 6(a)), where the edge length is fixed to 50mm, 150mm, and 500mm, in order to study the changes in the electric propagation when the head volume is small, normal, and large. The second geometry corresponds to a spherical model with a radius 80mm (see Fig. 6(b)). Finally, as in [17], we created an ellipsoidal model with semi-axes measuring 70mm, 82.5mm and 65mm in the x, y, and z directions respectively (Fig. 6(c)). The last two geometrical forms and sizes are more realistic representations of the head, facilitating the interpretation of simulated electric potential propagation during DBS. Moreover, a Medtronic 3389 DBS lead in monopolar configuration with a stimulus voltage of -1V was used; other material properties were discarded in the idealized FEM representation by using the Laplace equation in a homogeneous medium.

All the cubic models were analyzed with two different ground configurations following the Dirichlet boundary conditions; one uses the base of the cube as ground and the second uses all the sides of the cube as ground. For the spherical and ellipsoidal models, two ground configurations were used. The first configuration has all the surface settled at $0V$. For the second configuration, a cylinder (28 mm in diameter and 20mm in height) on the base of the model was included. The cylinder represents the path that the return current should follow to the reference electrode placed in the chest cavity, then the base of the cylinder is considered as ground. The models use an adaptive mesh refinement for the FEM in order to improve the precision of particular small regions of the model: the region closer to the electrode.

Results obtained from the solution of the Laplace equation using FEM are presented as curves around the active contact of the electrode. These represent ten different levels of potential as the distance from the electrode increases in the y-z plane (coronal view). These potential curves are obtained for all the models following the above mentioned ground configurations. Fig.7 (a) and 7(b) show the results for the 50mm edge length cube. A large difference in the potential levels between ground configuration models as function of the distance is observed. When the base side of the cube is set to $0V$, higher electric potential levels can be found at larger distances from the electrode in comparison with the case in which all the sides of the cube are set to $0V$. Also, the shape of the potential curves is influenced by the position of the ground. It becomes a uniform circle when all the boundaries are used. The same calculations are undertaken for the 150mm and 500mm edge length cubes. Similar behavior to the electric potential levels is shown in Fig.7(c) and 7(d), which compares to the results for the 50mm edge length cube.

Moreover, when the size of the cube increases, the influence of the ground configuration becomes less determinant in the shape and level of the potential. Fig. 7(f) and 7(e) show the results of ten potential curves for the two different ground configurations of the spherical models. The same results are presented in Fig. 7 (h) and 7(g) for the ellipsoidal model. The influence of the ground when the

cylinder configuration is used can be noticed, and it has higher potential levels in farther regions from the electrode.

In order to better understand the results, a quantitative assessment was developed to measure the electric potential in the regions that surround the electrode in order to determine the change in the electric propagation pattern according to different geometries. According to the solution of the models, the distances from the center of the electrode to each point of a single potential curve were computed. In order to measure the distance of different potential levels in the analyzed region, the Euclidean distance from the electrode to every point within a potential curve is calculated using

$$d(p,q) = \sqrt{\left(p_y - q_y\right)^2 + (p_z - q_z)^2} \qquad (8)$$

where q is the origin and p is one point placed on a potential level curve from the coronal view; q_y, q_z, p_y, and p_z are the components on the yz plane. This Euclidean distance is calculated for every model, and 100 different potential levels of propagation are analyzed. After the distance from the center of the electrode to each point of the equipotential curve has been computed, the minimum distance for each potential curve is selected (Fig. 8 describes the methodology), using:

Fig. 9(d) shows the results from the spherical and ellipsoidal forms. In the cylinder-base grounded models, the electric potential reaches higher values at distances far from the electrode until an inflexion point is reached. After the inflection point, the potential starts to decrease linearly alongside the cylinder region. The analysis of the electric potential before the inflection point shows that it is represented by a monotonically increasing function that behaves similarly to the potential for the models without the cylinder ground configuration.
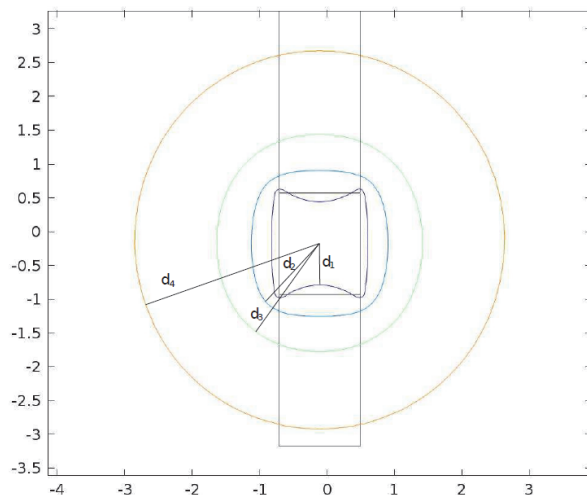
Figure 8. Black solid lines representing the minimum distance from the center of the electrode to the first 4 electric potential levels in the 50mm cubic model.
Source [authors]
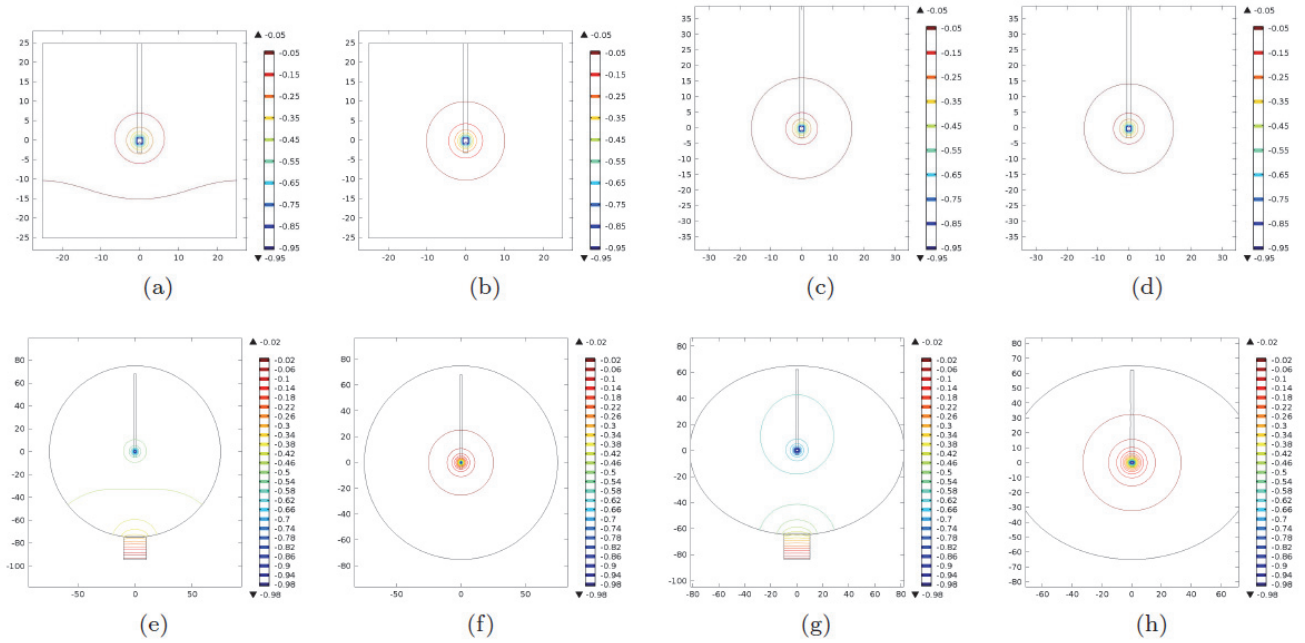
Figure 7. Electric potential propagation: Potential level curves computed on three sizes of cubical forms, one spherical form, and one ellipsoidal form, varying the ground configuration of the models. (a) Cube (50mm). Ground on base side. (b) Cube (50mm). Ground on whole boundary. (c) Cube (500mm). Ground on base side. (d) Cube (500mm). Ground on whole boundary. (e) Sphere. Ground on base of the cylinder. (f) Sphere. Ground on whole boundary. (g) Ellipse. Ground on base of the cylinder. (h) Ellipse. Ground on whole boundary.
Source [authors]

Furthermore, Table 1 presents the information regarding the percentile difference of electric potential between each model's two boundary conditions at specific distances (1, 2, 3, 4, 5, 10, 15, 20, and 30mm) from the center of the electrode. This is computed as in Equation (10):

$$d_r = \frac{v_1 - v_2}{v_1} \times 100\%, \qquad (10)$$

where v_1 and v_2 represent the value of the potential at a specific distance of the two different ground configurations of the same model, v_1 for the model with all the boundaries and v_2 for the model with the ground placed on the base side.

The value of the electric potential at the fixed distances from the electrode is obtained from linear interpolation of the curves from the minimum distances. The size of the model influences the propagation of the electric potential; lower levels of potential are reached for the smaller models in comparison with the larger models as the distance from the electrode increases. This result confirms that building a realistic model of DBS should consider size and boundary conditions due to the direct influence of these parameters on the final solution of the electric potential propagation.

6. Discussion

The results obtained in this work could be compared to studies such as [17, 18] and [19] in which simulation models were built for the same DBS electrode; however, there were a lack of real metrics that allowed a better understanding of the simulation results such as the ones presented in this work. Additionally, simulations for different ground configurations were not presented in the previously mentioned state-of-the-art studies, but they were in this present work.

Based on results, the size of the model and the ground configuration are important parameters when modeling a specific DBS simulation. The boundary conditions specified for the ground configuration and the size of the different models directly affect the shape and the magnitude of the electric potential in the region surrounding the electrode. This can be seen in all the results for the different models in Fig. 7. For the smaller models, the pattern of propagation of the potential is more influenced by the ground, more negative potential levels are reached far from the electrode, in comparison to bigger sized models. The shape of the potential levels around the electrode also changes for the two different ground configurations. When all the model's surfaces are grounded (Figs. 7(b), 7(d), 7(e) and 7(h)), a uniform potential distribution can be observed around the electrode, and a non-uniform shape of the potential levels can be found when the base side of the models is grounded (Figs. 7(a), 7(c), 7(f) and 7(g)).

For the quantification analysis presented in Fig. 9, it can be noticed that for the models with the ground configured in the whole surface, the higher potential levels reach shorter distances from the electrode than they do for the models in which only the base side is settled to $0V$.
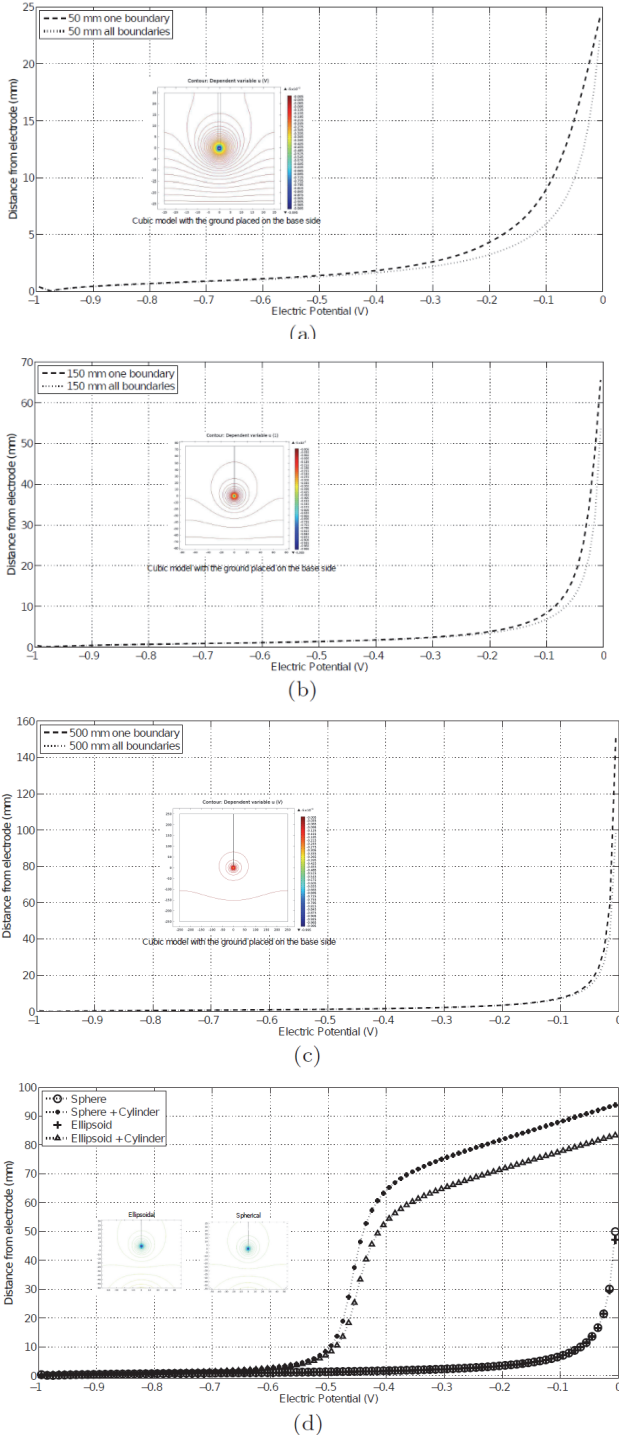
Figure 9 Curves representing the Electric Potential vs. Minimum Distance for 100 different potential levels using the cubic, spherical and ellipsoidal models and the two ground configurations. (a) 50mm edge length cube. (b)150mm edge length cube. (c) 500mm edge length cube. (d) Sphere and Ellipsoid.
Source [authors]

$$d_{min} = \overset{min}{_j} \| \boldsymbol{d}_j \|. \tag{9}$$

Table 1.
Results for the percentile difference between ground configurations in all the models.

Distance From electrode	d_r for different models [%]				
	Model				
	Cubic models			Sphere	Ellipse
	50mm	150mm	500mm		
1 mm	3.73	1.09	0.14	24.59	32.84
2 mm	12.9	3.92	0.64	87.36	92.42
3 mm	23.43	6.97	1.66	153.67	155.51
4 mm	34.24	10.13	2.55	221.28	220.38
5 mm	43.63	13.31	3.31	290.25	287.27
10 mm	79.29	28.98	7.79	668.37	653.64
15 mm	113.22	41.66	12.84	1092.10	1068.41
20 mm	137.12	51.12	19.26	1590.39	1558.30
30 mm	N/A	79.36	25.43	2965.71	2943.96

Source [authors]

From Table 1 it is possible to determine that for the cubic models the larger the size of the cube the less the influence of the ground configuration. In the case of the spherical and ellipsoidal models, since the results of the potential level propagation changes considerably when the base of the cylinder corresponds to the ground, the percentile difference between the two configurations for these models is larger than for the cubic models. Differences are reached of up to 2900% between the two different ground configurations for some distances from the electrode. Even the comparative result shows a clear difference between the ground configurations applied to the models. The development of a DBS realistic model should include tissue, electrical properties and other boundary conditions. From all of these assumptions, a DBS model could give more realistic results. From the DBS modeling presented, several applications could be derived; for example, a work presented by Michmizos et al. in [43] details the process of predicting the Parkinsonian STN spikes using the local field potentials that could be obtained using this approach.

7. Conclusion

We have described the electromagnetic phenomena that take place during DBS using classical electromagnetic theory. Moreover, we have shown that under the correct assumptions, the Laplace equation is a suitable alternative to represent the electrostatic field propagation generated after the stimulation. We have also shown through different computer simulations how factors such as the geometrical structure, size and the grounding of the conducting head volume have dramatic effects over the magnitude of the electric field, particularly for monopolar stimulation.

Acknowledgments

References

[1] Benabid, A.L., Chabardes, S., Mitrofanis, J. and Pollak, P., Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. The Lancet Neurology, 8(1), pp.67-81, 2009.

[2] de Lau, L.M. and Breteler, M.M., Epidemiology of parkinson's disease. The Lancet Neurology, 5(6), pp. 525-535, 2006. DOI: 10.1016/S1474-4422 (06)70471-9.

[3] Limousin, P., Pollak, P., Benazzouz, A., Hoffmann, D., Bas, J.F.L., Perret, J., Benabid, A.L. and Broussolle, E., Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. The Lancet Neurology, 345(8942), pp. 91-95, 1995. DOI: 10.1016/S0140-6736 (95)90062-4.

[4] Benabid, A.L., Deep brain stimulation for parkinson's disease. Current Opinion in Neurobiology, 13(6), pp. 696-706, 2003. DOI: 10.1016/j.conb.2003.11.001, URL http://www.sciencedirect.com/science/article/pii/S09594388030017 39.

[5] Visser-Vandewalle, Y.T., The functional role of the subthalamic nucleus in cognitive and limbic circuits. Progress in Neurobiology, 76, 2005. DOI: 10.1016/j.pneurobio.2005.09.005.

[6] Chaturvedi, A., Butson, C.R., Lempka, S.F., Cooper, S.E. and McIntyre, C.C., Patient-specific models of deep brain stimulation: Influence of field model complexity on neural activation predictions. Brain Stimulation, 3, pp. 65-77, 2010.

[7] Kuncel, A.M. and Grill, W.M., Selection of stimulus parameters for deep brain stimulation. Clinical Neurophysiology, 115, pp. 2431-2441, 2004.

[8] Montgomery, E.B., Deep brain stimulation programming. Principles and practice. Oxford University Press, USA, 2010.

[9] Obeso, J.A., Marin, C., Rodriguez-Oroz, C., Blesa, J., Benitez, B., Mena-Segovia, J., Rodríguez, M. and Olanow, C.W., The basal ganglia in Parkinson's disease: Current concepts and unexplained observations, Annals of Neurology, 64(S2), pp. S30-S46, 2008. DOI 10.1002/ana.2148.

[10] Obeso, J.A., Rodríguez-Oroz, M.C., Benitez-Temino, B., Blesa, F.J., Guridi, J., Marin, C. and Rodriguez, M., Functional organization of the basal ganglia: Therapeutic implications for parkinson's disease, Movement Disorders, 23(S3), pp. S548- S559, 2008. DOI 10.1002/mds.22062,

[11] Vitek, J.L., Mechanisms of deep brain stimulation: Excitation or inhibition, Movement Disorders, 17(S3), pp. S69-S72, 2002. DOI 10.1002/mds.10144.

[12] Volkmann, J., Herzog, J., Kopper, F. and Deuschl, G., Introduction to the programming of deep brain stimulators. Movement Disorders, 17(S3), pp. S181-S187, 2002. DOI: 10. 1002/mds.10162.

[13] Krack, P., Fraix, V., Mendes, A., Benabid, A.L. and Pollak, P., Postoperative management of subthalamic nucleus stimulation for parkinson's disease. Movement Disorders, 17(S3), pp. S188-S197, 2002. DOI: 10.1002/mds.10163,

[14] McIntyre, C.C., Frankenmolle, A.M., Wu, J., Noecker, A.M. and Alberts, J.L., Customizing deep brain stimulation to the patient using computational models. In: 31st Annual International Conference of the IEEE EMBS, Mineapolis, MN, USA, pp 4228-4229, 2009.

[15] Pollak, P., Krack, P., Fraix, V., Mendes, A., Moro, E., Chabardes, S. and Benabid, A.L., Intraoperative micro- and macroestimulation of the subthalamic nucleus in Parkinson's disease. Movement Disorders, 17(3), pp. S155-S16, 2002.

[16] McIntyre, C.C., Butson, C.R., Maks, C.B. and Noecker, A., Optimizing deep brain stimulation parameter selection with detailed models of the electrode-tissue interface. In: 28th Annual International

Conference of the IEEE EMBS, New York, NY, USA, pp 893-895, 2006.

[17] Grant, P.F. and Lowery, M.M., Electric field distribution in a finite-volume head model of deep brain stimulation. Medical Engineering & Physics, 31, pp. 1095-1103, 2009.

[18] Schmidt, C. and van Rienen, U., Modeling the field distribution in deep brain stimulation: The influence of anisotropy of brain tissue. IEEE Transactions on Biomedical Engineering, 59(6), pp. 1583-1592, 2012.

[19] McIntyre, C.C., Morib, S., Shermanc, D.L., Thakorc, N.V. and Vitek, J.L., Electric field and stimulating influence generated by deep brain stimulation of the subthalamic nucleus. Clinical Neurophysiology, 115, pp. 589-595, 2004.

[20] Walckiers, G., Fuchs, B., Thiran, J.P., Mosig, J.R. and Pollo, C., Influence of the implanted pulse generator as reference electrode in finite element model of monopolar deep brain stimulation. Journal of Neuroscience Methods, 186, pp. 90-96, 2010.

[21] Liberti, M., Apollonio, F., Paffi, A., Parazzini, M., Maggio, F., Novellino, T., Ravazzani, P. and D'Inzeo, G., Fundamental electrical quantities in deep brain stimulation: Influence of domain dimensions and boundary conditions. In: 29th Annual International Conference of the IEEE EMBS, Lyon, France, pp 6668-6671, 2007.

[22] McIntyre, C.C. and Grill, W.M., Extracellular stimulation of central neurons: Influence of stimulus waveform and frequency on neuronal output. Journal of Neurophysiology, 88, pp. 1592-1604, 2002.

[23] Martens, H., Toader, E., Decr, M., Anderson, D., Vetter, R., Kipke, D., Baker, K.B., Johnson, M.D. and Vitek, J.L., Quasi-stationary fields for microelectronic applications. Clinical Neurophysiology, 122, pp. 558-566, 2011.

[24] Butson, C.R. and McIntyre, C.C., Tissue and electrode capacitance reduce neural activation volumes during deep brain stimulation. Clinical Neurophysiology, 116, pp. 2490-2500, 2005.

[25] Alvarez-Alvarez, D., Rosero-Garcia, J. and Mombello, E., Core influence on the frequency response analysis (FRA) of power transformers through the finite element method. Ingeniería e Investigación, 35(1Sup), pp. 110-117. 2015. DOI: 10.15446/ing.investig.v35n1Sup.53771.

[26] Linero-Segrera, D., Oliver, J. and Huespe, A., Numerical modelling of the fracture process in reinforced concrete by means of a continuum strong discontinuity approach. Part II: application to shear panels. Ingeniería e Investigación, 30(3), pp. 16-26, 2010.

[27] Alvarado, P.A., Alvarez, M.A., Daza-Santacoloma, G., Orozco, A. and Castellanos-Dominguez, G., A latent force model for describing electric propagation in deep brain stimulation: A simulation study. In: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, pp. 2617-2620, 2014.

[28] Torres-Valencia, C.A., Daza-Santacoloma, G., Álvarez-López, M.A. and Orozco-Gutiérrez, Á.Á., Deep brain stimulation modeling for several anatomical and electrical considerations. Iteckne, 11(2), pp. 140-148. 2014.

[29] Volkmann, J., Herzog, J., Kopper, F. and Deuschl, G., Introduction to the programming of deep brain stimulators. Mov Disord, 17, pp. S181-S187, 2002.

[30] O'Suilleabhain, P., Frawley, W., Giller, C. and Dewey, R., Tremor response to polarity, voltage, pulsewidth and frequency of thalamic stimulation. Neurology, 60, pp. 786-790, 2003.

[31] Steinmetz, T., Kurz, S. and Clemens, M., Domains of validity of quasistatic and quasistationary field approximations. In: 15th International Symposium on Theoretical Electrical Engineering ISTET, Lubeck, Germnay, pp 271-275, 2009.

[32] Schmidt, K., Sterz, O. and Hiptmair, R., Estimating the eddy-current modelling error. IEEE Trans Magn, 44(6), pp. 686-689, 2008.

[33] Saville, D.A., Electrohydrodynamics: The taylor melcher leaky dielectric model. Annual Rev Fluid Mech, 29, pp. 27-64, 1997.

[34] Dirks, H.K., Quasi-stationary fields for microelectronic applications. Electrical Engineering, 79, pp. 145-155, 1997.

[35] Sadiku, M.N.O., Elements of electromagnetics. Oxford University Press, USA, 2002.

[36] Ostrowski, J., Hiptmair, R., Kramer, F., Smajic, J. and Steinmetz, T., Transient full Maxwell computation of slow processes. In: 8th Scientific Computation in Electrical Engineering conference SCEE, Toulouse, France, pp 87-95, 2010.

[37] Carpenter, M., Anatomy of the corpus striatum and brain stem integrating systems. John Wiley & Sons, Inc., 1981.

[38] Bossetti, C.A., Birdno, M.J. and Grill, W.M., Analysis of the quasi-static approximation for calculating potentials generated by neural stimulation. Journal of Neural Engineering, 59(5), pp. 44-53, 2008.

[39] Andreuccetti, D., Fossi, R. and Petrucci, C., An internet resource for the calculation of the dielectric properties of body tissues in the frequency range 10 hz - 100 ghz. IFAC-CNR, Florence, Italy, 1997, pp 87-95.

[40] Hofmanis, J., Louis-Dorr, V., Cecchin, T., Caspary, O. and Koessler, L., Propagation of electrical field in the brain using electrical intra-cerebral stimulations. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, pp 3888-3891, 2011. DOI 10.1109/IEMBS.2011.6090966.

[41] Iacono, M.I., Makris, N., Mainardi, L.T., Angelone, L.M. and Bonmassar, G., Mri-based multiscale model for electromagnetic analysis in the human head with implanted dbs. Comp Math Methods in Medicine [Online]. 2013. Available at: http://dblp.uni-trier.de/db/journals/cmmm/cmmm2013.html#IaconoMMAB13

[42] Forero, M. and Zulanga, D., Medical station for image processing and visualization of the brain electrical activity on a three-dimensional reconstruction of the patient's head. Ingeniería e Investigación, 23(3), pp. 31-38, 2010.

[43] Michmizos, K., Sakas, D. and Nikita, K., Prediction of the timing and the rhythm of the parkinsonian subthalamic nucleus neural spikes using the local field potentials. Information Technology in Biomedicine, IEEE Transactions on, 16(2), pp. 190-197, 2012. DOI 10.1109/TITB.2011.2158549.

P.A. Alvarado, is interested in probabilistic approaches for modeling music signals, with a focus on Gaussian processes and kernel methods. Alvarado holds an Electronic Engineering degree from Universidad Tecnológica de Pereira, Colombia, and a MSc in Electric Engineering from the UTP. Alvarado is currently a member of Centre for Digital Music pursuing a PhD at Queen Mary University of London.
ORCID: 0000-0002-9347-5093

C.A. Torres-Valencia, received his BSc. in Electronic Engineering in 2010 from the Universidad del Quindío, his MSc. in Electric Engineering in 2013 from the Universidad Tecnológica de Pereira, Colombia. From 2011 to date, he has been working in the Automatics research group at the Universidad Tecnológica de Pereira. Currently he is a Doctoral student at the Universidad Tecnológica de Pereira and funded by Colciencias´ "Doctorado Nacional - 647" program. His research interests include image processing, biosignal processing, neuroengineering and machine learning.
ORCID: 0000-0001-7568-6148

A.A. Orozco-Gutierrez, received a degree Electric Engineering in 1985, a MSc. degree in Electric Engineering in 2004, both from the Universidad Tecnológica de Pereira, and a PhD in Bioengineering from the Universidad Politécnica de Valencia in 2009. He is currently an Associate Professor at the Universidad Tecnológica de Pereira. His research interests include instrumentation and control, bioengineering and biosignal processing.
ORCID: 0000-0002-1167-1446

M.A. Álvarez López, received a degree BSc. in Electronic Engineering from the Universidad Nacional de Colombia in 2004, a MSc. degree in Electrical Engineering from the Universidad Tecnológica de Pereira, Colombia, and a PhD in Computer Science from the University of Manchester, UK, in 2011. He is currently an associate professor at the Universidad Tecnológica de Pereira, Colombia. His research interests include probabilistic models, kernel methods and stochastic processes.
ORCID: 0000-0002-8980-4472

G. Daza-Santacoloma received a BSc. in Electronic Engineering in 2005, a MSc. in Engineering Industrial Automation with honors in 2007, and a PhD. in Engineering - Automatics with honors in 2010, from the Universidad Nacional de Colombia. Currently, he is the R&D Manager at Neurocentro (Pereira – Colombia) where he is researching Neuroengineering. His research interests include neuroscience, feature extraction/selection for training pattern recognition systems, artificial vision, and machine learning.
ORCID: 0000-0002-1429-5925

H. Carmona-Villada, received a Medical degree in 1995 from the Universidad Tecnologica de Pereria, an MSc in Neurosurgery in 1999 from the Universidad Catolica de Chile, and a Subspecialist degree in functional neurosurgery in 2002, from the Albert Ludwig University from Freiburg, Germany. Currently, he is the Scientific Manager at Neurocentro (Pereira – Colombia) and the head of the functional neurosurgery program in Neurocentro and Colombia´s Neurological Institute where he undertakes movement disorder surgery, epilepsy surgery and pain surgery. His research interests include neuromodulation, neuroengineering, brain mapping, intraoperative monitoring.
ORCID: 0000-0002-8099-9461

A hybrid partitioning method for multimedia databases

Lisbeth Rodríguez-Mazahua [a], Giner Alor-Hernández [a], Jair Cervantes [b], Asdrúbal López-Chau [c]
& José Luis Sánchez-Cervantes [a]

[a] *Division of Research and Postgraduate Studies of the Instituto Tecnologico de Orizaba, México. lrodriguez@itorizaba.edu.mx, galor@itorizaba.edu.mx, isc.jolu@gmail.com*
[b] *Universidad Autónoma del Estado de México, Centro Universitario UAEM-Texcoco, México. chazarra17@gmail.com*
[c] *Universidad Autónoma del Estado de México, Centro Universitario UAEM-Zumpango, México. asdrubalchau@gmail.com*

Abstract
Hybrid partitioning has been recognized as a technique to achieve query optimization in relational and object-oriented databases. Due to the increasing availability of multimedia applications, there is an interest in using partitioning techniques in multimedia databases in order to take advantage of the reduction in the number of pages required to answer a query and to minimize data exchange among sites. Nevertheless, until now only vertical and horizontal partitioning have been used in multimedia databases. This paper presents a hybrid partitioning method for multimedia databases. This method takes into account the size of the attributes and the selectivity of the predicates in order to generate hybrid partitioning schemes that reduce the execution cost of the queries. A cost model for evaluating hybrid partitioning schemes in distributed multimedia databases was developed. Experiments in a multimedia database benchmark were performed in order to demonstrate the efficiency of our approach.

Keywords: hybrid Partitioning; multimedia databases, query optimization.

Un método de fragmentación híbrida para bases de datos multimedia

Resumen
La fragmentación híbrida es una técnica reconocida para lograr la optimización de consultas tanto en bases de datos relacionales como en bases de datos orientadas a objetos. Debido a la creciente disponibilidad de aplicaciones multimedia, surgió el interés de utilizar técnicas de fragmentación en bases de datos multimedia para tomar ventaja de la reducción en el número de páginas requeridas para responder una consulta, así como de la minimización del intercambio de datos entre sitios. Sin embargo, hasta ahora sólo se ha utilizado fragmentación vertical y horizontal en estas bases de datos. Este artículo presenta un método de fragmentación híbrida para bases de datos multimedia. Este método toma en cuenta el tamaño de los atributos y la selectividad de los predicados para generar esquemas de fragmentación híbridos que reducen el costo de ejecución de las consultas. También, se desarrolla un modelo de costo para evaluar esquemas de fragmentación híbridos en bases de datos multimedia. Finalmente, se presentan algunos experimentos en una base de datos de prueba con el fin de demostrar la eficiencia del método de fragmentación propuesto.

Palabras clave: fragmentación híbrida; bases de datos multimedia, optimización de consultas.

1. Introduction

Query optimization to reduce response time or to avoid the excessive use of system resources has been an active research field over the past decades [1].

Hybrid partitioning is a database design technique to improve query performance. It divides a relation or table into subsets of attributes and tuples in order to minimize the irrelevant data accessed by the queries. Hybrid partitioning has been typically applied to traditional databases (relational or object-oriented databases) to achieve query optimization.

Vertical partitioning divides a table T into a set of fragments $fr_1, fr_2, ..., fr_n$, such that each fragment fr_i contains a subset of the attributes and the primary key of table T. In contrast, horizontal partitioning splits table T into a set of fragments $fr_1, fr_2, ..., fr_n$, where each fragment fr_i has a subset of tuples of T.

There are two versions of horizontal partitioning: *primary* and *derived*. *Primary horizontal partitioning* of a table is performed by using predicates that are defined on that table. On the other hand, *derived horizontal partitioning* divides a table according to the predicates that are defined on another table. In this work, only primary horizontal partitioning is considered.

Hybrid partitioning can be accomplished in one of three ways: first, by performing vertical partitioning and then horizontally partitioning the vertical partitions (called VH partitioning), or by first performing the horizontal partitioning and then vertically partitioning the horizontal partitions (called HV partitioning), or by directly taking into consideration the semantics of the transactions [2].

Currently, multimedia applications are highly available [3-5], such as audio/video on demand, digital libraries, electronic catalogues, among others. The rapid development of multimedia applications has created a huge volume of multimedia data, which has exponentially incremented from time to time [6]. A multimedia database is crucial in these applications in order to provide efficient data retrieval.

Distributed and parallel processing on database management systems (DBMS) may improve the performance of applications that manipulate large volumes of data. This may be accomplished by removing irrelevant data accessed during the execution of the queries and by reducing the data exchange among sites, which are the two main goals of the design of distributed databases [7]. Therefore, partitioning techniques have been used in multimedia databases to improve the performance of applications.

Nevertheless, only vertical or horizontal partitioning techniques have been considered by the literature until now. Vertical partitioning reduces the irrelevant attributes accessed by the queries, but all the multimedia objects are stored in a fragment. Many of the queries issued to the multimedia databases only require some objects from the database. In order to improve the performance of the queries in multimedia databases, it is necessary to reduce access to irrelevant attributes and irrelevant objects; this is achieved with hybrid partitioning. For this reason, in this paper we propose a method for hybrid partitioning in multimedia databases. First, our method develops horizontal partitioning and then vertical partitioning, so it is therefore an HV partitioning algorithm.

This paper is structured as follows: in Section 2, the state of the art of hybrid partitioning in traditional and multimedia databases is presented. In Section 3, the Multimedia Hybrid Partitioning (MHYP) algorithm is described. In Section 4, the proposed cost model for the evaluation of different hybrid partitioning schemes is explained. Section 5 shows the performance evaluation of the queries. Finally, Section 6 presents the conclusion and future lines of research.

2. State-of-the-art

In order to clarify the difference between the related work and our approach, we classify them into two classes that are described in the following subsections.

2.1. Hybrid partitioning methods for traditional databases

Most mixed or hybrid partitioning algorithms only consider traditional databases. In [2], algorithms to generate candidate vertical and horizontal fragmentation schemes and a methodology for distributed database design using these fragmentation schemes were proposed for relational databases. They applied vertical and horizontal fragmentation schemes together to form a grid. This grid that consisted of cells was then merged to form mixed fragments.

An analysis algorithm for assisting distribution designers in the fragmentation phase of object oriented databases was proposed in [8]. The analysis algorithm indicated the most adequate fragmentation technique (vertical, horizontal or mixed) for each class in the database schema. In [9] a strategy to carry out the fragmentation phase of the distribution design of object oriented databases was proposed. Their fragmentation strategy has three steps: 1) the analysis, 2) the vertical fragmentation phase, and 3) the horizontal fragmentation phase.

In [10] a UML-based model for mixed fragmentation was presented. They validated their model using a case study with the concepts of attribute usage matrix and predicate usage matrix. A genetic algorithm for mixed fragmentation in relational databases which provides an improvement over previous works which considered vertical and horizontal partitioning separately was discussed in [11,12]. Compared to attribute partitioning only method, the mixed fragmentation design method produced database cost savings up to 69%. A mixed partitioning approach for multi-tenant data schema was provided in [13]. Their approach made a good scalability in multi-tenant shared database, while it can meet the optimal partitioning and multi-division.

Problems of the aforementioned hybrid partitioning methods in applying them to multimedia databases are the following: 1) Some techniques [2,8-10,13] do not consider the size of the attributes in the vertical partitioning stage since multimedia databases tend to be highly varied sizes (e.g., it is not the same to access an id of 8 bytes as a video of 8 MB): it is necessary to take into account the size of the attributes; 2) Some methods [2, 10] are based on affinity, which is the sum of the frequency of the attributes or predicates that are accessed together by the queries. A cost-based method is better for multimedia databases since it can incorporate more information in the creation of a fragment, such as selectivity of the predicates and size of the attributes, as well as the frequency of the queries; 3) Some techniques [2,11,12] only consider the minimization of the number of disk accesses. It is important to also reduce the transportation cost (i.e., the data exchange among sites) in order to optimize the queries in multimedia databases.

The hybrid partitioning method for multimedia databases proposed in this paper solves these problems because it takes into account the size of the attributes, the selectivity of the predicates, and the frequency of the queries to get hybrid fragments, which reduce the number of disk accesses and the transportation cost of the queries.

2.2. Partitioning methods for multimedia databases

The partitioning algorithms that take into account multimedia data only perform vertical or horizontal partitioning. In [14], primary horizontal fragmentation in distributed multimedia databases is addressed. The authors´ partitioning strategy is based on low-level multimedia features.

Table 1.
Comparison between Some Partitioning Algorithms

Approach	Multimedia Data	Hybrid Partitioning
Navathe et al. [2]	No	Yes
Baiäo and Mattoso [8]	No	Yes
Jagannatha et al. [10]	No	Yes
Ng et al. [11]	No	Yes
Gorla et al. [12]	No	Yes
Li et al. [13]	No	Yes
Saad et al. [14]	Yes	No
Getahun et al. [15]	Yes	No
Chbeir and Laurent [16]	Yes	No
Fung et al. [18]	Yes	No
Rodríguez and Li [19]	Yes	No
Rodríguez et al. [20]	Yes	No
MHYP	Yes	Yes

Source: the authors

In [15], semantic-based predicates implication required in current fragmentation algorithms is addressed in order to partition multimedia data efficiently. In [16], a formal approach dedicated to multimedia query and predicate implication is discussed. In [17], a horizontal partitioning algorithm for multimedia databases, called MHPA, is presented. MHPA is based on hierarchical agglomerative clustering.

A vertical partitioning technique was applied in an e-Learning video database system in [18] to achieve efficient query execution. The disadvantage was that this vertical partitioning technique did not consider the transportation cost of multimedia objects over the nodes of the network or the size of the multimedia objects. A vertical partitioning algorithm for distributed multimedia databases, called MAVP (Multimedia Adaptable Vertical Partitioning), is provided in [19], which takes into account the size of the attributes in the partitioning process. In [20], a system for dynamic vertical partitioning of multimedia databases, called DYMOND (DYnamic Multimedia ON line Distribution), is presented. It uses active rules for the dynamic vertical partitioning process. In Table 1, we present a comparative analysis that summarizes the relevant contributions of all these related works.

As we can see in Table 1, the implementation of hybrid partitioning in multimedia databases has two problems: (a) current hybrid partitioning algorithms do not take into account multimedia data; (b) only vertical and horizontal partitioning algorithms for multimedia databases have been developed. These deficiencies can be improved by: (a) developing a hybrid partitioning algorithm for multimedia databases, and (b) proposing a cost model to evaluate hybrid multimedia databases´ partitioning schemes. This proposal tries to solve the aforementioned deficiencies.

3. Multimedia hybrid partitioning algorithm (MHYP)

In this section, the Multimedia Hybrid Partitioning Algorithm (MHYP) is described in detail. MHYP consists of two phases:

1. **Obtaining the horizontal fragments:** The predicates of the queries are analyzed in order to obtain the initial horizontal fragments. MHPA [17] is used to obtain the horizontal fragments.

2. **Generating the hybrid fragments:** MAVP [19] is used to vertically fragment the horizontal fragments obtained in the first phase. As a result this gives the hybrid partitioning scheme.

In order to clarify our approach, we present the following scenario of a simple multimedia database used to manage equipment in a machinery sell company. The database consists of a table named $EQUIPMENT$ (id, $name$, $image$, $graphic$, $audio$, $video$) in which each tuple describes information about a specific piece of equipment, including its image, graphic, audio, and video objects. Information regarding 10,000 pieces of equipment (four different types) is stored: 2500 push mowers, 2500 string trimmers, 2500 chain saws, and 2500 water pumps. Let us also consider the following queries:

q_1:Find all chain saws images and graphics
q_2:Find name, audio and video with id "WP01"
q_3:Find all graphic, audio and video
q_4:Find all water pump images

Similarly to [16], we have considered that data that are stored in a table T can be defined by having two kinds of attributes: atomic and multimedia attributes. Also, we have assumed a fixed attribute set $U=A \cup M$, where:

- $A=\{A_1,A_2, \ldots, A_p\}$ and each A_i ($i=1, 2, \ldots, p$) is an atomic attribute associate with a set of atomic values (such as strings and numbers, among others) called the domain of A_i and denoted by $dom(A_i)$.
- $M=\{M_1,M_2, \ldots, M_q\}$ and each M_j ($j=1, 2, \ldots, q$) is a multimedia attribute, associated with a set of complex values (represented as sets of values or vectors) called multimedia features (such as, color, texture, shape, to mention a few). The domain of M_j is denoted by $dom(M_j)$.

Thus, given a table T that is defined over U, tuples t in T are denoted as $\langle a_1,a_2, \ldots, a_p,m_1,m_2, \ldots, m_q \rangle$ where a_i is in $dom(A_i)$ ($1 \leq i \leq p$) and m_j is in $dom(M_j)$ ($1 \leq j \leq q$). Every a_i (respectively m_j) is denoted by $t.A_i$ (respectively $t.M_j$).

3.1. Horizontal partitioning process

In this section, we first explain the information requirements of the horizontal partitioning process and then we present the steps of the multimedia horizontal partitioning algorithm MHPA.

3.1.1. Information requirements of horizontal partitioning

Qualitative and quantitative information about queries is required in order to develop the horizontal partitioning process [7]. Fundamental qualitative information consists of predicates used in user queries. Similarly to [16], the multimedia queries used in our approach are conjunctive projection-selection queries over T of the form $\pi_X \sigma_C (T)$, where X is a non empty subset of U and C is a conjunction of atomic select predicates, i.e., $C: P_1 \wedge \ldots \wedge P_m$ is defined as follows:

$DEFINITION\ 1$: An atomic selection predicate P_j is an expression of the form $P_j=A_i\ \theta\ a$, where $A_i \in A$, $a \in dom(A)$ and $\theta =\{=,\leq, \geq, <, >, like\}$.

Two sets are required in terms of quantitative information regarding user queries:

1. **Predicate selectivity**: number of tuples of the relation that would be accessed by a user query specified according to a given predicate. If $Pr=\{P_1, P_2, ..., P_m\}$ is a set of predicates, sel_i is the selectivity of the predicate P_i.

2. **Access frequency:** frequency with which user query access data. If $Q=\{q_1,q_2, ..., q_s\}$ is a set of user queries, f_k indicates the access frequency of query q_k in a given period.

3.1.2. The steps of the horizontal partitioning algorithm (MHPA)

Inputs: The table T is to be horizontally partitioned, and sest of queries with their frequencies are the input data of the MHPA.

Step 1: Determine the set of predicates Pr used by queries defined in the table T. These predicates are defined on a subset of attributes $A'(A' \subseteq A)$. As in [21], we call each element of A' a relevant predicate attribute.

The third query (q_3) does not have any predicate because the graphic, audio and video objects of all pieces of equipment are retrieved. Therefore, this query is not relevant for horizontal partitioning; this will be analyzed by the vertical partitioning algorithm. The predicates used by the queries q_1,q_2, q_4 in our running example are presented in Table 2.

Step 2: Build the predicate usage matrix (PUM) of table T. This matrix presents queries in rows and predicates in columns. In this matrix $PUM(q_k, P_i)=1$ if a query q_k uses a predicate P_i, otherwise it is 0. PUM also contains the fequency f_k of each query q_k and the selectivity sel_i of each predicate P_i. The PUM of our running example is shown in Table 3.

Step 3: Construct a partition tree. MHPA is based on a bottom-up approach. It first begins with single predicate fragments. It then, forms a new fragment by selecting and merging two of their fragments. This process is repeated until a fragment composed of all predicates is made. This kind of bottom-up approach generates a binary tree, which is called a partition tree (PT) [22]. Fig. 1 shows the PT of the table $EQUIPMENT$ obtained by MHPA.

When two fragments are merged, the amount of remote tuples (i.e., tuples located in another fragment) accessed is decreased while the amount of irrelevant tuples accessed by the queries is increased. For example, in Fig. 1 we can observe that in the Step 0 each predicate is located in a different fragment. Therefore, in the first fragment there are 2500 tuples (the selectivity of P_1, i.e., sel_1), in the second one there are only 1 tuple (sel_2), and in the third fragment there are 2499 tuples (sel_3-sel_2). Query q_1 has to access the first fragment, which has the 2500 relevant tuples needed to answer the query. Therefore, it does not have to access any irrelevant tuple and any remote tuple. The same happens to query q_2. Nevertheless, query q_4 has to access the second and the third fragments in order to retrieve the tuples with the name "WATER PUMP". It has to access one remote tuple (assuming that the tuple with the id="WP01" is going to be transported from the second fragment to the third fragment).

If the predicates P_1 and P_2 are merged into a fragment (as in the Step 1 of Fig. 1), query q_1 now would have to access 1 irrelevant tuple and query q_2 would have to access 2500 irrelevant tuples.

Table 2.
Predicates used by queries

Q	Pr
q_1	P_1: name="CHAIN SAW"
q_2	P_2: id="WP01"
q_4	P_3: name="WATER PUMP"

Source: the authors

Table 3.
Predicate usage matrix.

Q/Pr	P_1	P_2	P_3	f_k
q_1	1	0	0	15
q_2	0	1	0	10
q_4	0	0	1	20
sel_i	2500	1	2500	

Source: the authors
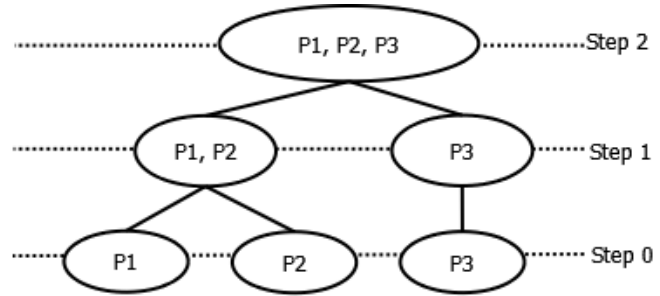
Figure 1. Partition tree of the $EQUIPMENT$ table, obtained by MHPA.
Source: the authors

If P_1, P_2, and P_3 are merged (as in the Step 2) q_4 does not have to access 1 remote tuple (i.e., the tuple with the id "WP01" located in the second fragment). On the other hand, queries q_1 and q_4 now would have to access 2500 irrelevant tuples, and query q_2 now would access 4999 irrelevant tuples. Therefore, the merged fragment will increase the amount of accesses to irrelevant tuples and it will reduce the amount of access to remote tuples.

In MHPA, in each step during the construction of a PT, two nodes (fragments) are selected that maximize the merging profit that is defined below, when they are merged into a node (fragment).

$$Merging\_Profit(HPA) = DRT - IIT \qquad (1)$$

Where,
DRT: the decreased amount of remote tuples accessed.
IIT: the increased amount of irrelevant tuples accessed.

In each step during the construction of a PT, MHPA produces an horizontal partitioning scheme ps_i, which merges two fragments that maximize the merging profit function defined in equation 1. Therefore, when the PT is finished, we have a set of horizontal partitioning schemes $PS=\{ps_1, ps_2,..., ps_m\}$, and every ps_i has a set of fragments $ps_i=\{fr_1, fr_2,..., fr_i\}$.

To select two fragments of i fragments that can maximize the merging profit $\binom{i}{2} = \left(\frac{i(i-1)}{2}\right)$, pairs should be examined. For example in Step 0 ($ps_3 =\{fr_1, fr_2, fr_3\}$) of Fig. 1 $i=m$ (where m is the number of predicates) because each predicate

Table 4.
Merging profits of the `EQUIPMENT` table in Step 0.

Pr	P_1	P_2	P_3
P_1		-25015	-87500
P_2			-25020
P_3			

Source: the authors

Table 6.
Attribute usage matrix

Q/U	id	name	image	graphic	audio	video	f_k
q_1	0	1	1	1	0	0	15
q_2	1	1	0	0	1	1	10
q_3	0	0	0	1	1	1	25
q_4	0	1	1	0	0	0	20
s_i	8	20	900	500	4100	39518	

Source: the authors

is located in a different fragment. Therefore, there are three fragments in Step 0 of Fig. 1 and it is necessary to examine the merging profits of $\left(\frac{3(3-1)}{2}\right)$ =3 pairs and merge one pair with the maximum merging profit among them. This generates the ps_2 of Step 1 in Fig. 1.

Table 4 shows MHPA Merging Profit Matrix (MPM) of the `EQUIPMENT` table in Step 0. In Algorithm 1, we show the process taken to get the MPM.

Algorithm 2 presents MHPA, it uses the PUM of the `T` and generates a set of initial horizontal fragments. Table 5 shows the horizontal partitioning schemes of the `EQUIPMENT` table that were obtained using MHPA.

Data: PUM of the table `T` (a set of predicates $Pr=\{ P_1, P_2, ..., P_m\}$, the selectivity sel_i of each predicate P_i, a set of queries $Q=\{q_1, q_2, ..., q_s\}$, the frequency f_k of each query q_k)
Result: MPM: Merging Profit Matrix
for each $P_i \in Pr$ /$1 \le i \le m$-1 **do**
    for each $P_j \in Pr$ / i+1$\le j \le m$ **do**
        DRT=0;
        IIT=0;
        $merging\_profit$=0;
        for each $q_k \in Q$ /$1 \le k \le s$ **do**
            if $PUM(q_k, P_i)$=1 & $PUM(q_k, P_j)$=1 **then**
                DRT=DRT+f_k*(sel_i+sel_j);
            else
                if $PUM(q_k, P_i)$=1 **then**
                    IIT=IIT+f_k*sel_j;
                else
                  if $PUM(q_k, P_j)$=1 **then**
                    IIT=IIT+f_k*sel_i;
                  end
                end
            end
        end
        $merging\_profit$=DRT-IIT;
        $MPM(P_i, P_j)$=$merging\_profit$;
    end
end
Algorithm 1. getMPM

Data:PUM
Result: initial horizontal partitioning schemes $PS=\{ps_1, ps_2, ..., ps_m\}$
for each step $\in$ PT **do**
    $getMPM(PUM, MPM)$
    select two nodes with maximum merging profit;
    merge the nodes;
end
Algorithm 2. MHPA

Table 5.
Resulting horizontal fragments of the table `EQUIPMENT`

PS	fr_1	fr_2	fr_3
ps_1	(P_1, P_2, P_3)		
ps_2	(P_1, P_2)	(P_3)	
ps_3	(P_1)	(P_2)	(P_3)

Source: the authors

3.2. Vertical Partitioning Process

MHYP uses the MAVP algorithm to achieve a vertical partitioning scheme (VPS). MAVP requires an Attribute Usage Matrix (AUM) as input, which has a set of atomic and multimedia attributes U=A $\cup$ M= { $A_1, A_2, ..., A_p, M_1, M_2, ..., M_q$ }. The maximum size s_i of each attribute a_i is $\in$ U, it has a set of queries $Q=\{q_1, q_2, ..., q_s\}$, the fequency f_k of each query is q_k, and it has a set of elements $AUM(q_k, a_i)$, where $AUM(q_k, a_i)$=1 if query q_k uses the attribute a_i, or, if not, $AUM(q_k, a_i)$=0. The AUM of the `EQUIPMENT` table is presented in Table 6. MAVP takes into account the size of the attributes due to its importance in the vertical partitioning process because it is not the same to access a remote or irrelevant atomic attribute as it is to access a remote or irrelevant multimedia attribute. Multimedia attributes tend to be of a lot larger size. For further details, consider [19].

MAVP finds an optimal VPS when the number of fragments is equal to two $vps_2=\{fr_1=(id, audio, video), fr_2=(name, image, graphic)\}$.

3.3. Hybrid partitioning scheme generation

Algorithm 3 shows the MHYP algorithm. MHYP takes the PUM as an input as well as AUM of table `T` of the multimedia database and generates the optimal hybrid partitioning scheme (*optimal_hps*). Algorithm 3 presents the MHYP algorithm. MHYP obtains an optimial hybrid partitioning scheme (*optimal_hps*) based on the PUM and the AUM of the table `T`. MHPA uses the PUM to obtain a set of horizontal partitioning schemes $PS=\{ps_1, ps_2, ..., ps_m\}$. In contrast, MAVP only generates one optimal vertical partitioning scheme (VPS).

HPS_Generator combines the initial horizontal partitioning schemes generated by MHPA and the VPS obtained by MAVP. The number of hybrid partitioning schemes produced by the HPS_Generator is m, i.e., the number of horizontal partitioning schemes obtained by MHPA. Therefore, the HPS_Generator obtains a set of hybrid partitioning schemes HPS=$\{hps_1, hps_2, ..., hps_m\}$. Every hps_i has a set of fragments $hps_i=\{fr_1, fr_2, ..., fr_t\}$. Each fragment fr_k has n_k attributes. We suppose that the network has nodes $N_1, N_2, ..., N_t$, the allocation of the fragments to the nodes gives rise to a mapping $\lambda:\{1, ..., t\} \rightarrow\{1, ..., t\}$, which is called location assignment [23]. Table 7 depicts the definition of the hps_1 fragments.

HPS_Generator obtains two matrices: a Fragment-Attribute Usage Matrix (FAUM) and a Fragment-Predicate Usage Matrix (FPUM). FAUM contains a set of atomic and multimedia attributes U=A $\cup$ M= { $A_1, A_2, ..., A_p, M_1, M_2,$

…., M_q }, the set of fragments of a hybrid partitioning scheme $hps_i=\{fr_1, fr_2, ..., fr_t\}$, the sum of the size of the attributes sfr_k of each fragment $fr_{k,}$ and a set of elements FAUM(fr_k, a_i)=1 if fragment fragment fr_k has the attribute a_i, or, if not, FAUM(fr_k, a_i)=0. For instance, Table 8 shows the FAUM of the $hps_1=\{fr_1, fr_2, fr_3, fr_4\}$, $sfr_1=s_{id}+s_{audio}+s_{video}=8+4100+39518=43626$. FPUM presents the fragments in rows and the predicates in columns. In this matrix, FAUM(fr_k, P_i)=1 if the fragment fr_k contains the tuples of the predicate P_i, or if not, it is 0. In addition, FPUM presents information about the cardinality cfr_k of a fragment fr_k. Table 9 presents the FPUM of the hps_1, every fragment stores 5000 tuples.

Data: *PUM, AUM*
Result: *optimal_hps*
best_cost=0;
optimal_hps=0;
MHPA(*PUM, PS*);
MAVP(*AUM, VPS*);
HPS_Generator(*PS, VPS, HPS, FAUM, FPUM)*;
for each $hps_i \in$ HPS **do**
    cost(hps_i)=IDAC(hps_i)+TC(hps_i);
    if cost(hps_i)<*best_cost* **then**
        best_cost=cost(hps_i);
        optimal_hps=i;
    end
end
Algorithm 3. MHYP

4. Cost model

The cost of a hps_i is composed of two parts: irrelevant data access cost and transportation cost.

Table 7.
Fragments of the first hybrid partitioning scheme

hps_1
$fr_1=\pi_{a1, a5, a6}(\sigma_{P1\vee P2\vee P3}(T))$
$fr_2=\pi_{a1, a5, a6}(\sigma_{\neg P1\wedge\neg P2\wedge\neg P3}(T))$
$fr_3=\pi_{a1, a2, a3, a4}(\sigma_{P1\vee P2\vee P3}(T))$
$fr_4=\pi_{a1, a2, a3, a4}(\sigma_{\neg P1\wedge\neg P2\wedge\neg P3}(T))$
a_1=id, a_2=name, a_3=image, a_4=graphic, a_5=audio, a_6=video, $T=EQUIPMENT$

Source: the authors

Table 8.
FAUM

hps_1	id	name	image	graphic	audio	video	sfr_k
fr_1	1	0	0	0	1	1	43626
fr_2	1	0	0	0	1	1	43626
fr_3	1	1	1	1	0	0	1428
fr_4	1	1	1	1	0	0	1428

Source: the authors

Table 9.
FPUM

hps_1	P_1	P_2	P_3	cfr_k
fr_1	1	1	1	5000
fr_2	0	0	0	5000
fr_3	1	1	1	5000
fr_4	0	0	0	5000

Source: the authors

$$cost(hps_i) = IDAC(hps_i) + TC(hps_i) \tag{2}$$

IDAC measures the amount of data from both irrelevant attributes and irrelevant tuples accessed during the queries. The transportation cost provides a measure for transporting between the nodes of the network.

The irrelevant data access cost is given by:

$$IDAC(hps_i) = \sum_{k=1}^{t} IDAC(fr_k). \tag{3}$$

In order to obtain the cost of an hps_i, it is necessary to use the PUM and the AUM of a table T. The irrelevant data access cost of each hybrid fragment fr_k is given by:

$$IDAC(fr_k) = IAAC(fr_k) + ITAC(fr_k). \tag{4}$$

IAAC is the irrelevant attribute access cost. ITAC is the irrelevant tuple access cost. IAAC is defined as:

$$IAAC(fr_k) = \sum_{q_j\in IQ_k} IAAC(q_j); \tag{5}$$

$$IAAC(q_j) = \begin{cases} f_j \sum_{a_i\in IA_j} s_i \sum_{P_l\in PQ_j} sel_l & if\ q_j \in PUM; \\ f_j \sum_{a_i\in IA_j} s_i\ cfr_k & otherwise \end{cases} \tag{6}$$

where IQ_k is a set of queries that uses at least one attribute and accesses at least one irrelevant attribute of the fragment fr_k. This is:

$$IQ_k = \{q_j|AUM(q_j,a_i) = 0 \wedge AUM(q_j,a_l) = 1 \wedge \{a_i,a_l\} \in fr_k\} \tag{7}$$

$$\{a_i,a_l\} \in fr_k \Rightarrow FAUM(fr_k,a_i) = 1 \wedge FAUM(fr_k,a_l) = 1 \tag{8}$$

IA_j is the set of attributes that is not used by query q_j in IQ_k. This is defined as:

$$IA_j = \{a_i|AUM(q_j,a_i) = 0 \wedge FAUM(fr_k,a_i) = 1\} \tag{9}$$

In the example $IQ_1=\{q_3\}$ because q_3 does not use the attribute *id* but it needs the attributes *audio* and *video* from the fragment fr_1 of the hps_1. Therefore, $IA_3=\{id\}$.

PQ_j has the predicates used by a query q_j and is located in the fragment fr_k.

$$PQ_j = \{P_l|PUM(q_j,P_l) = 1 \wedge FPUM(fr_k,P_l) = 1\} \tag{10}$$

For fr_1, $PQ_1=\{P_1\}$, $PQ_2=\{P_2\}$, $PQ_3=\{\varnothing\}$, $PQ_4=\{P_3\}$ because the predicates P_1, P_2, P_3 are accessed by the queries q_1, q_2, q_4 and the tuples required by the predicates are located in the fragment fr_1.

ITAC can be written as:

$$ITAC(fr_k) = \sum_{q_j \in AQ_k \wedge ITAC(q_j) \geq 0} ITAC(q_j); \qquad (11)$$

$$ITAC(q_j) = \begin{cases} f_j\left(sfr_k\left(cfr_k - \sum_{P_l \in PQ_j} sel_l\right)\right) & if\ n_p \geq 1; \\ 0 & otherwise \end{cases} \qquad (12)$$

where n_p is the number of predicates in PQ_j, and AQ_k contains the queries that access at leat one attribute of the fragment fr_k.

$$AQ_k = \{q_j | AUM(q_j, a_i) = 1 \wedge FAUM(fr_k, a_i) = 1\} \qquad (13)$$

In the example for the fragment fr_1, $AQ_1=\{q_2, q_3\}$, since the query q_2 accesses all the attributes of the fragment fr_1 (*id*, *audio*, *video*) and q_3 accesses the attributes *audio* and *video* of fr_1. Due to the fact that $PQ_2=\{P_2\}$ and $PQ_3=\{\varnothing\}$, only the query q_2 contributes to the access to irrelevant tuples.

The transportation cost of an hps_i is computed according to a given location assignment. Since transportation costs dominate the execution cost of a query [7], the TC of hps_i is the sum of the costs of each query multiplied by its frequency squared, i.e.

$$TC(hps_i) = \sum_{j=1}^{t} TC(q_j)\ f_j^2 \qquad (14)$$

The transportation cost of query q_j depends on the size of the relevant remote attributes and on the assigned locations, which decide the transportation cost factor between every pair of sites. It can be expressed by:

$$TC(q_j) = \sum_h \sum_{h'} c_{\lambda(h)\lambda(h')} s(h') sel(h') \qquad (15)$$

where h ranges over the nodes of the network for q_j, $s(h')$, which are the sizes of the relevant remote attributes, $sel(h')$ is the number of relevant remote tuples accessed by the query, q_j, $\lambda(h)$ indicates the node in the network at which the query is stored, and c_{ij} is a transportation cost factor for data transportation from node Ni to node Nj $\{i,j \in \{1,...,t\}\}$[23]. For instance, Table 10 presents the IAAC, ITAC, and IDAC of the hps_1 in megabytes. IAAC(fr_1)=IAAC(q_3), IAAC(q_3)=$f_3*s_1*cfr_1$=25*8*5000=1 million of bytes=1 MB. ITAC(fr_1)=$f_2*sfr_1*(cfr_1-sel_2)$=10*43626*(5000-1)=2180.86 MB.

The TC(hps_1) of MHYP is calculated as follows: there are four fragments, so we suppose that there are four nodes N_1, N_2, N_3, N_4, and each fragment fr_i is located in each node N_i. We also assume that each query is located in the node in which the larger attribute that it uses is located, and c_{ij}=1.

Table 10.
Costs of the first hybrid partitioning scheme

hps_1	IAAC	ITAC	IDAC
fr_1	1	2180.86	2181.86
fr_2	1	0	1
fr_3	141.71	196.34	338.05
fr_4	116.00	0.00	116.00

Source: the authors
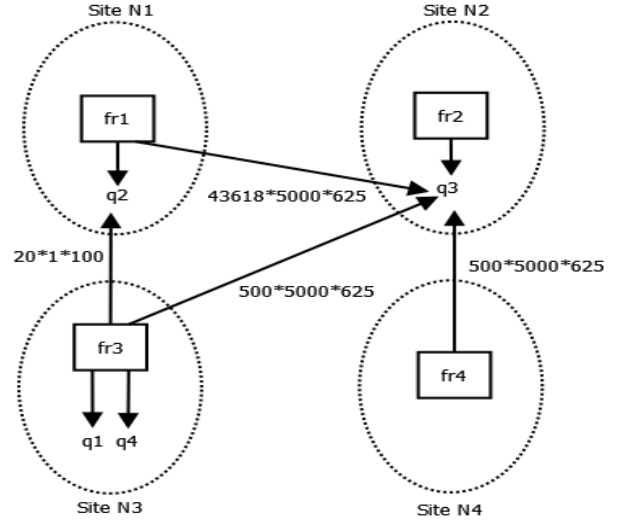
Figure 2. Location assignment of hps_1.
Source: the authors

Table 11.
Costs of the hybrid partitioning schemes

HPS	IDAC	TC	Cost
hps_1	2636.91	139431.25	142068.16
hps_2	1386.50	139431.62	140818.12
hps_3	259.70	139431.61	139691.31

Source: the authors

Fig. 2 illustrates the local assignment of hps_1. The query q_2 requires one tuple with the attribute *name*, which is located in the fragment fr_3. Therefore TC(q_2)=$s_{name}*sel_2*f_2^2$=20*1*10²=2000 bytes.

Table 11 contains the costs of the hybrid partitioning schemes generated by MHYP. The optimal scheme is hps_3 and has a cost of 139691.31 MB.

5. Evaluation

This section presents and compares the hybrid partitioning schemes obtained using MHYP, the vertical partitioning scheme generated by MAVP, and the horizontal partitioning generated by MHPA. The benchmark used for the comparison was the database of a machinery sales company used in [19, 20] and described in Section 3. Some hybrid partitioning methods, such as [2, 11, 12] consider that the response time of a query is strongly affected by the amount of data accessed from secondary storage (disk). Hence, the objective functions of these methods are to minimize the number of disk accesses. The cost model proposed in this paper is used to compare the schemes obtained by MHPA, MAVP and MHYP since the cost to perform queries in distributed systems is dominated by the remote network communication as well as by local disk accesses.

Tables 12, 13 and 14 compare the costs of the queries of MHPA, MAVP and MHYP. As it can be observed, the scheme obtained with MHYP has a lower cost in most queries. This is because MHPA only takes into account

information about the irrelevant tuples accessed by queries, MAVP focuses on the reduction of irrelevant attributes, and MHYP considers the size of the irrelevant attributes and the selectivity of the predicates in order to reduce both irrelevant attributes and tuples accessed by the queries Using this information, MHYP considerably reduces the cost of the queries.

The cost of the query q_3 is increased in MHYP because it needs all graphic, audio and video of the multimedia database. In the scheme of MHPA, this query only accesses 5000 remote graphic, audio and video objects. The transportation cost of this query is considerably reduced in the MAVP scheme because it only has to access 10000 remote graphic objects. The scheme obtained by MHYP accesses 5000 remote audio and video objects and 10000 remote graphic objects, so its transportation cost is increased. Most of the queries executed in multimedia databases tend to access only a subset of attributes and tuples of the database; therefore, hybrid partitioning is suitable for these databases in order to reduce query execution cost.

6. Conclusion and future work

Hybrid partitioning optimizes query execution cost because it reduces the irrelevant data accessed by the queries. The novel aspects of our work include the following research

Table 12.
Comparison of the execution cost of the queries

MHPA			
Query	IDAC	TC	Cost
q_1	1635.97	0	1635.97
q_2	0.014	0	0.014
q_3	116	137868.75	137984.75
q_4	2205.42	0.368	2205.79

Source: the authors

Table 13.
Comparison of the execution cost of the queries

MAVP			
Query	IDAC	TC	Cost
q_1	0.31	0	0.31
q_2	4504.96	0	4504.96
q_3	234	3125	3359
q_4	239.6	0	239.6

Source: the authors

Table 14.
Comparison of the execution cost of the queries

MHYP			
Query	IDAC	TC	Cost
q_1	0.3	0	0.3
q_2	0.014	0	0.014
q_3	234	139431.24	139665.24
q_4	25.39	0.368	25.76

Source: the authors

contributions: first, a hybrid partitioning algorithm for distributed multimedia databases has been developed, which takes into account the size of the attributes and the selectivity of the predicates to generate an optimal hybrid partitioning scheme. Second, a cost model for distributed multimedia databases has been proposed. This cost model considers that the overall query processing cost in a distributed multimedia environment consists of irrelevant data access cost and transportation cost. An experimental evaluation shows that the algorithm proposed in this paper outperforms both a horizontal and a vertical partitioning only algorithm in most cases.

In this research we assumed that the queries that run against the multimedia database are static. Distributed multimedia databases are accessed by many users simultaneously, therefore queries tend to change over time and a good hybrid partitioning scheme can be degraded, resulting in very long query response time. Present research could be extended to derive the hybrid partitioning dynamically in multimedia databases (MMDBs) based on the changes in the queries. Thus, the hybrid partitioning scheme of the multimedia database can be adaptively modified to always achieve efficient retrieval of multimedia objects.

In the future, we also wish to consider low-level features of multimedia data and similarity-based (range and k-nearest neighbor) queries in the hybrid partitioning process. These kinds of queries are needed for content-based retrieval, which consists of obtaining information from the MMDB according to the characteristics of the multimedia objects, such as color, texture, and shape (in the case of images).
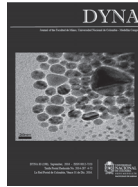
Acknowledgments

References

[1] Moreno, F.J., Ospina-Romero G., yLarios-Restrepo R. Desempeño de consultas relacionales y objeto-relacionales en Oracle, Revista Ingeniería e Investigación, 25(3), pp. 4-12, 2005.

[2] Navathe, S., Karlapalem, K. and Ra, M., A mixed fragmentation methodology for initial distributed database design, Journal of Computer and Software Engineering, 3, pp. 1-34, 1995.

[3] Motato-Toro O.F. y Loaiza-Correa H., Identificación biométrica utilizando imágenes infrarrojas de la red vascular de la cara dorsal de la mano, Revista Ingenierìa e Investigación, 29(1), pp. 90-100, 2009.

[4] Atencio, P., Sánchez G.T. and Branch J.W., Automatic visual model for classification and measurement of quality of fruit: Case mangifera INDICA L, DYNA 76(160), pp. 317-326, 2009.

[5] Álvarez M.J, González E., Bianconi F., Armesto J. and Fernández A., Colour and texture features for image retrieval in granite industry, DYNA 77(161), pp. 121-130, 2010.

[6] Rahman, M.N.A., Lazim, Y.M., Mohamed, F., Saany, S.I.A. and Yusof M.K.M., Rules generation for multimedia data classifying usin rough sets theory, International Journal of Hybrid Information Technology, 6(5), pp. 209-218, 2013. DOI: 10.14257/ijhit.2013.6.5.19

[7] Özsu, M.T. and Valduriez, P., Principles of distributed database systems. New York: Springer, third edition, 2011.

[8] Baiäo, F. and Mattoso, M., Towards an inductive design of distributed object oriented databases, Proceedings of the Third IFCIS Conference of Cooperative Information Systems (CoopIS'98), New York, USA, IEEE CS Press, pp. 88-197, 1998.

[9] Baiäo, F., Mattoso, M. and Zaverucha, Z., A Distribution design methodology for object DBMS, Distributed and Parallel Databases, 16(1), pp. 45-90, 2004. DOI: 10.1023/B:DAPD.0000026268.04288.b9

[10] Jagannatha, S., Mrunalini, M., Kumar, T.V.S. and Kanth, K.R., Modeling of mixed fragmentation in distributed database using UML 2.0, Proceedings of the Int. Conf. on Computer Engineering and Applications, pp. 190-194, 2009.

[11] Ng, V., Gorla, N. and Law, D.M., Applying genetic algorithms in database partitioning, Proceedings of the 2003 ACM Symposium on Applied Computing (SAC), pp. 544-549, 2003. DOI: 10.1145/952532.952639

[12] Gorla, N., Ng, V. and Law, D.M., Improving database performance with a mixed fragmentation design, Journal of Intelligent Information Systems, 34, pp. 559-576, 2012. DOI: 10.1007/s10844-012-0203-x

[13] Li, H., Yang, D. and Zhang, X., A mixed partitioning approach for multi-tenant database approach, Journal of Information & Computational Science, 10(15), pp. 4869-4878, 2013. DOI: 10.12733/jics20102341

[14] Saad, S., Tekli J., Atnafu, S., Chbeir, R. and Yetongnon, K., Towards multimedia fragmentation, Advances in Databases and Information Systems, Lecture Notes in Computer Science, 4152, pp. 415-429, 2006. DOI: 10.1007/11827252_31

[15] Getahun, F., Tekli, J., Atnafu, S. and Chbeir, R., The use of semantic-based predicates implication to improve horizontal multimedia database, Proceedings of the MS'07 Workshop on Multimedia Information Retrieval on The Many Faces of Multimedia Semantics, New York, USA: ACM, pp. 29-39, 2007. DOI: 10.1145/1290067.1290073

[16] Chbeir, R. and Laurent, D., Towards a novel approach to multimedia data mixed fragmentation, Proceedings of the Int. Conf. on Management of Emergent Digital EcoSystems, New York, USA: ACM, pp. 200-204, 2009. DOI: 10.1145/1643823.1643860

[17] Rodríguez, M., Alor-Hernández, G., Abud-Figueroa M.A. and Peláez-Camarena S.G., Horizontal partitioning of multimedia databases using hierarchical agglomerative clustering, in: Gelbuk A. et al. (Eds.), MICAI 2014, Part II, LNAI 8857, Springer, pp. 296-309, 2014. DOI: 0.1007/978-3-319-13650-9_27

[18] Fung, C.-W., Leung, W.-C. and Li, Q., Efficient query execution techniques in a 4Dis video database system for eLearning. Multimedia Tools and Applications, 20(1), pp. 25-49, 2003. DOI: 10.1023/A:1023418316038

[19] Rodríguez, L. and Li, X., A vertical partitioning algorithm for distributed multimedia databases, in: Proceedings of DEXA 2011, 6861, Springer Verlag, pp. 544-558, 2011. DOI: 10.1007/978-3-642-23091-2_48

[20] Rodríguez, L., Li, X., Cervantes, J. and García-Lamont, F., DYMOND: An active system for dynamic vertical partitioning of multimedia databases, Proceedings of the 16th International Database Engineering & Applications Sysmposium, New York, USA: ACM, pp. 71-80, 2012. DOI: 10.1145/2351476.2351485

[21] Bellatreche, L., Karlapalem, K. and Simonet, A., Algorithms and support for horizontal class partitioning in object oriented databases, Distributed and Parallel Databases, 8, pp. 155-179, 2000. DOI: 10.1023/A:1008745624048

[22] Son, J.H. and Kim, M.H., An adaptable vertical partitioning method in distributed systems, Journal of Systems and Software, 73(3), pp. 551-561, 2004. DOI: 10.1016/j.jss.2003.04.002

[23] Ma, H., Distribution design for complex value databases. PhD Thesis, Massey University, Palmerston North, New Zeland, 2007.

L. Rodríguez-Mazahua, received her BSc in Information Technology and her MSc in Computer Science from the Instituto Tecnológico de Orizaba, Veracruz, Mexico, in 2004 and 2007, respectively. In 2012 she obtained a PhD in Computer Science from the Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Mexico. From 2012 to 2014, she was a professor of computer science at the Universidad Autónoma del Estado de México, Centro Universitario UAEM Texcoco, Mexico. Since February 2014 she has been undertaking postdoctoral research at the Instituto Tecnológico de Orizaba. Her current research interests include distribution design of databases, database theory, autonomic database systems, multimedia databases, and Big Data.
ORCID: 0000-0002-9861-3993

G. Alor-Hernández, is a full-time researcher at the Division of Research and Postgraduate Studies in Orizaba´s technological institute, the Tecnológico de Orizaba, Mexico. He received an MSc and a PhD in Computer Science from the Center for Research and Advanced Studies at the National Polytechnic Institute (CINVESTAV), Mexico. He has led 10 Mexican research projects granted by CONACYT, DGEST and PROMEP. He is author/coauthor of around 130 journal and conference papers on computer scienceHis research interests include Web services, e-commerce, Semantic Web, Web 2.0, service-oriented and event-driven architectures, and enterprise application integration. He is an IEEE and ACM Member. He is a National Researcher recognized by the National Council of Science & Technology of Mexico (CONACYT).
ORCID: 0000-0003-3296-0981, Scopus Author ID: 17433252100.

J. Cervantes, received his BSc. in Mechanical Engineering from Orizaba Technologic Institute, Veracruz, Mexico in 2001, his MSc. and PhD. from CINVESTAV-IPN, Mexico, in 2005 and 2009 respectively. His research interests include support vector machine, pattern classification, neural networks, fuzzy logic and clustering.
ORCID: 0000-0003-2012-8151, Scopus Author ID:23033927200

A. López-Chau, received his BSc. degree in Electronic Engineering from the Instituto Politécnico Nacional, México; his MSc. in computer science from the Centro de Investigación en Computación at the Instituto Politécnico Nacional, México, in 1997 and 2000, respectively. In 2013 he obtained a PhD. in Computer Science from CINVESTAV-IPN. Since 2011, he has been a professor of computer science at Universidad Autónoma del Estado de México, Centro Universitario UAEM Zumpango. His current research interests include data mining, machine learning and embedded systems.
ORCID: 0000-0001-5254-0939

J.L. Sánchez-Cervantes, obtained a PhD in Computer Science and Technology from the Universidad Carlos III de Madrid. He received an MSc. in Computer Systems and is an engineer in Computer Systems at the Instituto Tecnológico de Orizaba. His research interests include Semantic Web, Linked Data (Linked Open Data), Social Media, Big Data and Internet of things.
ORCID: 0000-0001-5194-1263

Characterization of supply chain problems

Rafael Guillermo García-Cáceres [a] & John Wilmer Escobar [b]

[a] Universitaria Agustiniana - Uniagustiniana, Bogota, Colombia. rafael.garcia@uniagustiniana.edu.co
[b] Departamento de Contabilidad y Finanzas, Universidad del Valle, Cali, Colombia. john.wilmer.escobar@correounivalle.edu.co

Abstract
The current contribution intends to identify, characterize and provide context to the most usual supply chain problems. One hundred and twenty three SC problems were identified and addressed in a specific area that locates them within the context of the fundamental SC components. The conceptual framework developed here identifies different degrees of similarity among SC problems. Finally, as a practical example, the production scheduling problem is characterized. By making the interactions among SC problems and the implications of the available decisions that are likely to solve them clear, the present contribution is not only a useful scorecard for decision makers, but it is also an upgraded conceptual development for theoreticians on the topic. As they are the central object of business competitiveness, SC improvements such as the present one are greatly valuable in terms of the social and economic profit of more efficient chains, sustainable organizations and wealthy stakeholders.

Keywords: supply chain management; logistics.

Caracterización de las problemáticas de la cadena de abastecimiento

Resumen
El trabajo identifica 123 problemáticas cada una de las cuales es vinculada en la estructura mediante una notación. Para ilustrar el desarrollo de trabajos futuros se usa como ejemplo de caracterización el problema de programación de la producción. La estructura aquí definida establece una base que facilita el despliegue teórico de nuevos desarrollos en gerencia de la cadena de abastecimiento y a su vez facilita la labor de los tomadores de decisión. El trabajo representa un esfuerzo por establecer una visión holística para el estudio de la cadena y su mejoramiento en términos de competitividad y sostenibilidad en pro del beneficio de sus *stakeholders*.

Palabras clave: gerencia de la cadena de abastecimiento; logística.

1. Introduction

The Supply Chain (SC) can be seen as an integrated process in which raw materials are transformed into a final product that is delivered to the consumer through distribution centers, retailers or by both [1]. The SC has traditionally been understood as having three stages, namely suppliers (related to procurement), manufacturers (production) and retailers - consumers (delivery). Each of these stages may take place at several facilities or companies distributed in echelons, which are homogeneous groups of specific facilities [2]. Companies and materials within the SC, as well as the information and financial resources flowing through them, are integrated in such a way that changes in one of them affect the performance of the whole chain [1,3].

The importance of the SC lies in its comprehensive and holistic character, which intends to generate and add value. The possibilities of improving the SC are related to effectively managing competitiveness, an aspect the remarkable impact of which has made it object that is sought to constantly be improved. SC management seeks to impact value generation by promoting both efficiency and customer service, which implies component operation and coordination improvement.

Supply chain management (SCM) promotes the efficient planning, execution and control of these systems' operations. The SC is currently becoming more and more complex because of continuous changes in customer needs, the emergence of a global economy and, overall, because the demands set on its performance have been increasing due to rapid escalation of

competitiveness; all of these constitute a major management challenge. SCM encompasses raw material storage and movement, stock processing, and finished product handling until final delivery [4,5]. It always wants to benefit the stakeholders and efficiently fulfill the chain's purpose.

One of the main objectives of SCM is promoting SC performance improvement. Deficiencies in problem identification and in measuring the performance of solutions have been recognized as being some of the main current limitations of the management processes [6,7]. Conversely, improvements in troublesome areas have been found to improve the whole system [8].

SCM requires a deep understanding of the problems and implications of the SC. Thus, establishing an effective SCM system requires the fundamental characterization of an adequate decision framework to face SC problems, which is actually a basic scorecard that allows for broad contextual vision and controlled and effective management. From this current work's standpoint, the control of SC processes implies that the relation between the different problems they deal with be determined [9]. The SC problems are present at three decision levels, namely strategic, tactical and operational. Strategy influences higher decision levels, frequently requiring thorough research into areas such as business policy, financial planning, competitiveness and organizational goal achievement. Tactical problems deals with two functions: the assignation of resources and the development of strategic objectives. Problem solving at this level requires valuable information about middle level decision management. In turn, the solution to operational problems requires precise data to evaluate the impact of those decisions taken by low-level administrative personnel. This paper intends to develop a comprehensive framework for the decision-making of specific SC problems. This framework allows not only for SC problems to be determined together with their implications and interrelations, but also it identifies gaps in the literature on SCM and provides an new approach to solving them. After a thorough literature review, the first part of the paper defines the fundamental units of the SC, while the second part introduces the characterization of potential SC problems within the framework mentioned.

2. State of the art

We conducted a thorough literature search on the SC problems in question, and we found 82 relevant books from the 1993-2011 period. These are shown in Table 1.

Table 1.
Reviewed books and book chapters.

Year	Author(s)	Title
1993	American Telephone and Telegraph Company	Design's impact on logistics
1999	Anupindi, R., Chopra, S., Deshmukh, S.D., Van Mieghem, J.A. and Zemel, E.	Managing business process flows
1998	Arnold, J.R.T.	Introduction to materials management
2004	Ballou, R.H,	Business logistics/Supply chain management
2001	Bauer, M.J., Poirier, C.C.,	e-Business: The strategic impact on

Year	Author(s)	Title
	Lapide, L. and Bermudez, J.	supply chain and logistics
1976	Bender, P.	Design and operation of customer service systems
2004	Blanchard, B.S.	Logistics engineering and management
2007	Blanchard, D.	Supply Chain Management – Best Practices.
2002	Bloomberg, D.J. Lemay S. and Hanna J.B.	Logistics
2000	Bovet, D. and Martha, J.	Value nets: Breaking the supply chain to unlock hidden profits
1978	Bowersox, D.J.	Logistical management
1996	Bowersox, D.J. and Closs, D.J.	Logistical management: The integrated supply chain process
2002	Bowersox, D.J., Closs, D.J. and Cooper, M.B.	Supply chain logistics management
1992	Bowersox, D.J., Daugherty, P.J., Dröge, C.L., Germain, R.N. and Rogers, D.S.	Logistical excellence: It's not business as usual
1999	Boyson, S., Corsi, T.M., Dresner, M.E. and Harrington, L.H.	Logistics and the extended enterprise: Benchmarks and best practices for the manufacturing professional
2006	Bozarth, C.C., Handfield, R.B.	Introduction to Operations and Supply Chain Management
1997	Bramel, J. and Simchi-Levi, D.	The logic of logistics: Theory, algorithms, and applications for logistics management
1990	Brunet, H. an d Le Denn, Y.	La démarche logistique
2010	Burt, D. N., Petcavage, S. D., & Pinkerton, R. L.	Supply management
2004	Cohen, S. and Roussel, J.	Strategic Supply Chain Management: The five Disciplines for Top Performance.
2004	Chopra, S. and Meindl, P.	Supply chain management Strategy, planning, and operation
1983	Colin, J., Mathe, H. and Tixier, D.	La logistique au service de l'entreprise
1997	Copacino, W.C.	Supply chain management: The basics and beyond
2003	Coyle, J.J., Bardi, E.J., Langley Jr., C.J.	The management of business logistics: A supply chain perspective
2004	Chang, Y.S., Makatsoris, H.C., Richards, H. D.	Evolution of Supply Chain Management – Symbiosis of adaptive value networks and ICT.
1998	Dornier, P.-P., Ernst, R., Fender, M. and Kouvelis, P.	Global operations and logistics: Text and cases
1997	Eymery, P.	La logistique de l'entreprise
1992	Fawcett, P., Mcleish, R. an d Ogden, I.	Logistics management
2000	Fernandez-Ranada, M., Gurrola-Gal, F.X. and Lopez-Tello, E.	3C: A proven alternative to MRPII for optimizing supply chain performance
2001	Fleischmann, M.	Quantitative models for reverse logistics
1992	Francis, R.L., McGinnis, L.F. Jr. an d White, J.A.	Installation layout and location: An analytical approach
2001	Fredendall, L.D. and Hill, E.	Basics of supply chain management
1998	Gattorna, J., editor	Strategic supply chain alignment: Best practice in supply chain management
1990	Gattorna, J., Trost, G. and Kerr, A., editors	The Gower handbook of logistics and distribution management
2003	Giard, V.	Gestion de la production et des flux

Year	Author	Title
2012	Giese, A.	Differenziertes performance measurement in supply chains
2001	Gourdin, K.N.	Global logistics management: A competitive advantage for the new millennium
1993	Graves, S.C., RinnooyKan, A.H.G. and Zipkin, P.H.	Handbook in Operations Research and Management Science. Volume 4 — Logistics of production and inventory
1999	Handfield, R.B. and Nichols, E.L. Jr.	Introduction to supply chain management
1973	Heskett, J.L., Glaskowsky, N.A. and Ivie, R.M.	Business logistics
2007	Hubner, R.	Strategic supply chain Management in Process Industries – An application to Specialty Chemicals Production Netrwork Design.
2008	Huo, Y., Jia, F.	Supply Chain - The Way to Flat Organization
1987	Hutchinson, N.E.	An integrated approach to logistics management
2012	Jacobs, F. R., & Chase, R. B.	Operations and Supply Chain Management: The Core
1999	Johnson, J.C., Wood, D.F., Wardlow, D.L. and Murphy, P.R. Jr.	Contemporary logistics
1998	Kasilingam, R.G.	Logistics and transportation: Design and planning
1978	Kearney, A.T.	Measuring productivity in physical distribution
2007	Kogan K. and Tapiero C.S.	Supply chain games: Operations management and risk valuation
2008	Kordic, V.	Supply Chain - Theory and Applications
2005	Riopel, D., Langevin, A., and Campbell, J.F.	The Network of Logistics Decisions, in Logistics Systems: Design and Optimization
1998	Lambert, D.M., Stock, J.R. and Ellram, L.M.	Fundamentals of logistics management
1976	Lambillotte, D.	La fonction logistique dans l'entreprise
1995	Langford, J.W.	Logistics: Principles and applications
2003	Lawrence, F.B., Jennings, D.F. and Reynolds, B.E.	E-Distribution
2005	Lawrence, F.B., Jennings, D.F. and Reynolds, B.E.	ERP in distribution
2007	Lee, H., Lee, C.	Building Supply Chain Excellence in Emerging Economies
1993	Leenders, M.R. and Fearon, H.E.	Purchasing and materials management
2011	Li, J., Chen, J., and Wang S.	Risk management of supply and cash flows in supply chains
1999	Lowson, B., King, R. and Hunter, A.	Quick response: Managing the supply chain to meet consumer demand
2000	Lynch, C.F.	Logistics Outsourcing: A management guide
2002	Miller, T.C.	Hierarchical Operations and Supply Chain Planning.
2002	Monczka, R., Trent, R. and Handfield, R.	Purchasing and supply chain management
2004	Murphy Jr., P.R. and Wood, D.F.	Contemporary logistics
1973	Muther, R.	Systematic layout planning
2001	Pimor, Y.	Logistique: Techniques et mise en œuvre
1993	Pons, J. and Chevalier, P.	La logistique intégrée
2000	Ptak, CA and Schragenheim, E.	ERP: Tools, techniques and applications for integrating the supply chain
2002	ReVelle, J.B., editor	Manufacturing handbook of best practices: An innovation, productivity, and quality focus
1994	Robeson, J.F., Copacino, W.C. and Howe, R.E., editors	The logistics handbook
2009	Röderstein, R.	Erfolgsfaktoren im supply chain management der DIY-Branche
2003	Seifert, D.	Collaborative planning, forecasting, and replenishment: How to create a supply chain advantage
2001	Shapiro, J.F.	Modeling the supply chain
2003	Simchi-Levi, D. and Kaminsky, P., Simchi-Levi, E.	Designing and managing the supply chain: Concepts, strategies and case studies
1973	Smykay, E.W.	Physical distribution management
1997	Southern, R.N.	Transportation and logistics basics: A handbook for transportation and logistics, professionals and students
2008	Stadtler, H. and Kilger, C.	Supply Chain Management and Advanced Planning, concepts, models, software and case studies.
2004	Stevenson, W.J. and Hojati, M.	Operations management
1998	Stock, J.R.	Development and implementation of reverse logistics programs
2001	Stock, J.R and Lambert, D.M.	Strategic logistics management
2011	Swink, M., Melnyk, S. A., Cooper, M. B., & Hartley, J. L.	Managing operations across the supply chain
1999	Tayur, S., Ganeshan, R. and Magazine, M., editors	Quantitative models for supply chain management
1997	Tilanus, B.	Information systems in logistics and transportation
2010	Waters, D	Global logistics: New directions in supply chain management
2011	Waters, D	Supply chain risk management: Vulnerability and resilience in logistics
2005	Wisner J.D. Leong, G.K. and Tan, K.-C.	Principles of supply chain management: A balanced approach
1992	Womack, J.P., Jones, D.T. and Roos, D.	Le système qui va changer le monde
2007	Zhang. Q.	E-Supply Chain Technologies and Management

Source: The authors

In summary, these works present concepts, definitions, functionalities, planning processes and case developments, among other considerations. Nevertheless, none of them includes a framework that allows the problems confronted by the SC to be identified and organized.

According to the literature review presented in Table 2, the most frequent problems that are associated with the SC are: (i) customer satisfaction, (ii) delivery management, and (iii) costs/finances. The least covered areas are those dealing with markets and production.

Table 2.
Studied Supply Chain areas.

Item	Supply chain area or measurement construct	Related references
1	Markets	(Frohlich and Westbrook, 2001), (Li et al., 2005).
2	Production	(Brun et al., 2006), (Berrah and Cliville, 2007).
3	Customer (satisfaction)	(Frohlich and Westbrook, 2001), (Sabri and Beamon, 2000), (Bhagwat and Sharma, 2007), (Beamon, 1999), (Brewer and Speh, 2000), (Li et al., 2005), (Lai et al., 2002), (Angerhofer and Angelides, 2006), (Yeh et al 2007), (Berrah and Cliville, 2007), (Chen and Paulraj, 2004), (Otto and Kotzab, 2003).
4	Delivery management	(Frohlich and Westbrook, 2001), (Beamon, 1999), (Li et al., 2005), (Van der Vorst et al., 2000), (Lai et al., 2002), (Brun et al., 2006), (Yeh et al 2007), (Berrah and Cliville, 2007), (Chen and Paulraj, 2004), (Otto and Kotzab, 2003).
5	Costs/Finances	(Sabri and Beamon, 2000), (Bhagwat and Sharma, 2007), (Beamon, 1999), (Brewer and Speh, 2000), (Van der Vorst et al., 2000), (Krause et al., 2001), (Angerhofer and Angelides, 2006), (Chen and Paulraj, 2004), (Otto and Kotzab, 2003).
6	Flexibility in volume or delivery	(Sabri and Beamon, 2000), (Beamon, 1999), (Angerhofer and Angelides, 2006).
7	**Internal business processes: a set of measures related to production, human resources, quality, procurement and distribution.**	(Bhagwat and Sharma, 2007), (Beamon, 1999), (Brewer and Speh, 2000).
8	Process and/or product innovation.	**(Beamon, 1999), (Brewer and Speh, 2000), (Yeh et al 2007), (Otto and Kotzab, 2003).**
9	Communication with customers and suppliers	(Barut et al., 2002), (Li et al., 2005), (Schmitz and Platts, 2004).
10	Supplier flexibility	(Krause et al., 2001), (Chen and Paulraj, 2004).
11	**Supplier efficiency (quality and fulfillment)**	(Krause et al., 2001), (Lai et al., 2002), (Berrah and Cliville, 2007), (Chen and Paulraj, 2004).
12	Human resource management	(Bhagwat and Sharma, 2007), (Yeh et al 2007), (Berrah and Cliville, 2007).

Source: The authors

A few works cover the three stages of the SC (delivery, manufacturing and procurement). When they do, they usually focus on a specific logistic or manufacturing area [8,10-12]. Table 3 classifies SCM contributions according to the stage of the chain they deal with, namely customers, suppliers or internal functioning of the company.

Table 3.
Supply Chain Areas are classified according to the supply chain stage(s) they deal with.

Reference	Problems Studied	Supply Chain Stages		
		Customers	Manufacturing or internal operation of the company as part of the supply chain	Suppliers
[1]	Markets, customers, delivery management.	x		
[2]	Customers, costs, volume or delivery flexibility.	x	x	
[10]	**Customers, costs, human resource management, internal business processes.**	x	x	x
[13]	**Customers, delivery management, costs, flexibility, internal business processes.**	x	x	
[14]	Customers, costs, internal business processes, product innovation.	x	x	
[15]	**Customers, markets, delivery management.**	x	x	1
[16]	Delivery management, costs.		x	
[17]	Costs, supplier flexibility, supplier efficiency.		x	x
[10]	Customers, delivery management, supplier efficiency.	x	x	x
[18]	Customers, costs, volume or delivery flexibility.	x	x	
[19]	**Customers, delivery management,**	x	x	

	product innovation, human resource management.			
[20]	Production, delivery		x	
[11]	**Production, customers, delivery**	x	x	x
[12]	Customers, delivery management	x	x	x
[21]	Customers, delivery management	x	x	
[22, 23]	Communication with the customer and	x		x

Source: The authors

The work of [24] identifies a hierarchy of decisions within the SC and analyzes the relations between them, without specifying the associated problem. The said hierarchy, which is further developed throughout this paper through additional dimensions is comprised of: (i) necessary previous conditions for a given choice (upstream decisions); (ii) considerations on the impact of some decisions on others (downstream decisions); and (iii) an approach to the information required by the process. However, the fact it focuses on decisions and not on the challenges faced by the system certainly limits its scope. However, these authors' work constitutes part of the current development, which includes the focus of this paper.

The literature review revealed the lack of studies that seek to develop a broad and comprehensive perspective of the SC and cover a significant part of its components. Likewise, no works have been found that propose a formal structure capable of characterizing the SC's problems. Such a structure would be useful in those cases whose specificities call for a more detailed study; this is when more formal and technical decision-making proves to be useful.

3. Characterization of supply chain problems

The current SC problem is divided in two parts: the first one identifies a series of logistic units within the chain and proposes a corresponding notation system. The second one outlines troublesome instances and, from a SC standpoint, introduces the notation that situates a given problem within the chain.

3.1. Fundamental units

The fundamental units are the entities through which the SC actually performs, that is, the objects of their activity and decisions. In turn, these fundamental units address type, quantity and size specifications of the goods (or services) produced by the chain, as well as the location of the productive activities in question. The fundamental units identified by the current research are: the SC itself, companies, echelons, stages, links and facilities. Table 4 presents the notation system in question.

Table 4.
Notation of the supply chain's fundamental units.

Z|	Chain	Company	Echelon
Detail	Description: industrial sector to which it belongs, global or domestic character, countries or regions in which it performs, etc. a: constituent companies (1, 2,..) b: constituent echelons (1, 2,..) c: stages it comprises (procurement, production, distribution) d: constituent links (1, 2,..).	Description: corporate name, industrial sectors to which it belongs, global or domestic character, countries or regions in which it performs, etc. a1: chains to which it belongs a2: links to which it belongs a3: stage types to which it belongs. a4: echelon to which it belongs.	Description: associated companies; level or position within the chain, etc. b1: chain to which it belongs b2: companies belonging to this echelon (1, 2, …) b3: links of this echelon (1, 2, …) b4: stages associated with this echelon (procurement, production, distribution)
Notation	SC: (a, b, c, d)	O: (a1, a2, a3, a4)	E: (b1, b2, b3, b4)
Level	**Stage**	**Link**	**Facilities**
Detail	Description: type (procurement, production, distribution), associated companies; associated echelons, industrial sector to which it belongs; global or domestic character; countries or regions where it performs; etc. c1: chains to which it belongs c2: companies operating at this stage (1, 2, …) c3: links making up this stage (1, 2, …) c4: echelons making up this stage (1,2,)	Description: organization mode (vertical integration, strategic alliance, market), global or domestic character; countries or regions in which it exerts activity, etc. d1: chain to which it belongs d2: constituent companies (1,2,) d3: link organization (vertical integration, strategic alliance, market) d4: stage associated to with link (procurement, production, distribution)	Description: type of facility (factory, stock, distribution center, etc.), location. e1: chain to which it belongs e2: company to which it belongs e3: echelon to which it belongs e4: stage to which it belongs
Notation	S: (c1, c2, c3, c4)	L: (d1, d2, d3)	F: (e1, e2, e3, e4)

Source: The authors

3.2. Problem set featuring notation

Pursuing the social objective of the SC implies adequate management of its constitutive units and diverse functions, as well as the timely making of decisions in terms of problematic aspects. Many of these problems are associated with each SC unit – they are organizational and functional in nature. Two concepts have been coined to characterize them, namely level of decision [25,26] and characteristic matter: product flows (both, finished products and raw materials) and information flows [27]. The independent or coupled application of these concepts modulates the approach to the specific problems faced by the SC.

Decision-making within the SC is based on a clear notion of the specific nature of the problems that are to be solved. According to Decision Theory, this implies conceiving a structure made up of alternatives and associated criteria in which the latter intends to describe and even weigh up the former. A qualified decision making process must also include a mechanism to measure or estimate the criteria and look after the internal consistency of the information and the correct selection of the Decision Support System (DSS).

Although the logistic unit, the organizational function and the characterizing concepts are the main components used to define a problem, they are not the only ones. The full set of problem featuring components is the following: the elemental unit (U), the logistic functions (F), the characterizing concept defining a given set of problems (C), the levels of decision (N) and/or the characteristic flows (L), the decision (D), the previous conditions (B), the impact (I) and the descriptive alternatives, criteria and/or assumptions (J). Among these components, U and F are independent; C depends on the latter two and D depends on all the three of these, while B, I, and J depend on all of the above mentioned components.

This set of characterizing features intends to provide a comprehensive framework to fully characterize the SC problems in their own context, through the following detailed notation: $P^i_{(U, F, C, D, B, I, J)}$,

Where super index i is the label of a particular problem P. Although not all the different fields of a particular P label (U, F, C, D, B, I, J) are always active, the first four ones must be. Table 5 shows the Ps associated with the SC's logistic functions, as reported in the literature review.

As an example, we shall analyze P71: the scheduling problem, which has been an object of study for decades and remains a very active research field. Several reviews have contributed to this field [28-32], which include more than 200 papers that featuring this P:

U: Plant
F: Production
C: Operational (associated to N: the decision level)
D: Decision (scheduling)
B: Production planning, layout, routing of materials
I: Carrier type, handling of materials, determination of Stock Keep Units - SKUs)
J: Determined from notation by [33]:

The alternatives are: – work stations combinations (and associated machines) used to undertake the works (items) that could satisfy the problem's requirements.

Objective (makespan criteria to be optimized): total finishing time, total delay, total weighted delay (taking into account the relative importance of the client).

Parameters:
- Number of stations
- Number and homogeneity of parallel machines per station.
- Processing times

Assumptions and restrictions:
- Any work that can only be executed by one machine at a time.
- Any machine that can only execute one work at a time.
- Machines are constantly available.
- Any work that can and should only be processed once at any station.
- No work can be dismounted from the machine before it has been finished.
- Storage capacity among stations is unlimited.
- All works must follow the same route: from station 1 to station 2, and so on.
- The sequence in which works are processed must be the same at all stations.
- From time zero, all works are available to enter the sequence
- DSS has traditionally corresponded to heuristics and metaheuristics that have adequate CPU time solutions but only moderate acceptable optimality gaps.

The following section presents the decision framework in which each one of the problems is characterized. For practical purposes, Table 5 does not include elements B, I and J, which can be found in [24]. The table shows that the most frequent characterizing concept is "level", while "flow" is rarely taken into account. Through the literature review we detected 123 Ps that are associated with five fundamental units (that are in turn related to facilities), three decision levels, 12 logistic functions and the 48 decisions determined by [24]. Table 5 shows the problem featuring framework that was determined:

Table 5.
Characterization of SC problems.

Problem $P^i_{(u,n,f,d)}$	Logistic unit (U)	Decision level (N)	Logistic function (F)	Decision (D) and description of the P
$P^1_{(1,1,1,2)}$			1. Strategic planning	2. Degree of vertical integration and outsourcing.
$P^2_{(1,1,3,7)}$		1. Strategic	3. Communication and information (C&I) network	7. C&I network strategy
$P^3_{(1,2,3,6)}$	1. Supply channel		3. Communication and information network)	6. C&I network design
$P^4_{(1,2,8,27)}$		2. Tactical	8. Procurement and supply management	27. Procurement type
$P^5_{(1,2,8,24)}$				24. Specifications of goods

Code	Zone	Level	Function	Item
				procured
$P^6_{(1,2,8,26)}$				26. Selection of suppliers
$P^7_{(1,2,8,25)}$				25. Order intervals and quantities
$P^8_{(1,2,8,23)}$				23. Quality control
$P^9_{(1,2,11,39)}$				39. Degree of consolidation
$P^{10}_{(1,2,11,41)}$			*11. Transportation*	41. Transportation modes
$P^{11}_{(1,2,11,44)}$				44. Types of carriers
$P^{12}_{(1,2,11,43)}$				43. Carrier selection
$P^{13}_{(1,2,11,40)}$				40. Transportation fleet mix
$P^{14}_{(1,2,12,46)}$				46. Order entry procedures
$P^{15}_{(1,2,12,45)}$			*12. Order processing*	45. Order transmission means
$P^{16}_{(1,2,12,48)}$				48. Order follow-up procedures
$P^{17}_{(1,3,8,27)}$				27. Procurement type
$P^{18}_{(1,3,8,24)}$			*8. Procurement and supply management*	24. Specifications of goods procured
$P^{19}_{(1,3,8,26)}$				26. Selection of suppliers
$P^{20}_{(1,3,8,25)}$		3. Operational		25. Order intervals and quantities
$P^{21}_{(1,3,11,41)}$				41. Transportation modes
$P^{22}_{(1,3,11,44)}$			*11. Transport*	44. Types of carriers
$P^{23}_{(1,3,11,43)}$				43. Carrier selection
$P^{24}_{(1,3,11,40)}$				40. Transportation fleet mix
$P^{25}_{(2,1,1,2)}$	2. Warehouse	1. Strategic	*1. Strategic planning*	2. Degree of vertical integration and outsourcing
$P^{26}_{(2,1,2,5)}$			*2. Physical Facility network (PF)*	5. PF network strategy
$P^{27}_{(2,1,3,7)}$			*3. Communication and information network*	7. C&I network strategy
$P^{28}_{(2,1,5,9)}$			*5. Inventory management*	9. Inventory management strategy

Code	Zone	Level	Function	Item
$P^{29}_{(2,1,5,11)}$				11. Safety stock
$P^{30}_{(2,1,10,34)}$			*10. Warehousing*	34. Warehousing mission and functions
$P^{31}_{(2,2,2,4)}$			*2. Physical Facility network*	4. PF network design
$P^{32}_{(2,2,3,6)}$			*3. Communication and information network*	6. C&I network design
$P^{33}_{(2,2,5,10)}$			*5. Inventory management*	10. Relative importance of inventory
$P^{34}_{(2,2,5,12)}$				12. Control methods
$P^{35}_{(2,2,5,13)}$				13. Desired inventory level
$P^{36}_{(2,2,6,15)}$		2. Tactical	*6. Material handling*	15. Material handling fleet mix
$P^{37}_{(2,2,6,14)}$				14. Material handling fleet control
$P^{38}_{(2,2,10,33)}$				33. Warehouse layout
$P^{39}_{(2,2,10,36)}$			*10. Warehousing*	36. Stock location
$P^{40}_{(2,2,10,32)}$				32. Receiving/shipping dock design
$P^{41}_{(2,2,10,35)}$				35. Safety systems
$P^{42}_{(2,2,12,46)}$				46. Order entry procedures
$P^{43}_{(2,2,12,45)}$			*12. Order processing*	45. Order transmission means
$P^{44}_{(2,2,12,47)}$				47. Order picking procedures
$P^{45}_{(2,2,12,48)}$				48. Order follow-up procedures
$P^{46}_{(2,3,6,17)}$		3. Operational	*6. Material handling*	17. Unit loads
$P^{47}_{(2,3,6,16)}$				16. Types of material handling equipment
$P^{48}_{(3,1,1,1)}$				1. Definition of customer service
$P^{49}_{(3,1,1,3)}$			*1. Strategic planning*	3. Customer service objectives
$P^{50}_{(3,1,1,2)}$	3.Plant	1. Strategic		2. Degree of vertical integration and outsourcing
$P^{51}_{(3,1,2,5)}$			*2. Physical Facility network*	5. IF network strategy
$P^{52}_{(3,1,3,7)}$			*3. Communication and information network*	7. C&I network strategy
$P^{53}_{(3,2,2,4)}$		2. Tactical	*2. Physical Facility network*	4. IF network design
$P^{54}_{(3,2,3,6)}$			*3.*	6. C&I network

Code			Category	Description
			Communication and information network	design
P$^{55}_{(3,2,6,15)}$			*6. Material handling*	15. Material handling fleet mix
P$^{56}_{(3,2,6,14)}$				14. Material handling fleet control
P$^{57}_{(3,2,7,20)}$			*7. Packaging*	20. Information media
P$^{58}_{(3,2,7,19)}$				19. Information to be provided with the product
P$^{59}_{(3,2,7,22)}$				22. Type of packaging
P$^{60}_{(3,2,7,18)}$				18. Packaging design
P$^{61}_{(3,2,7,21)}$				21. Level of protection needed
P$^{62}_{(3,2,9,29)}$			*9. Production*	29. Product routing
P$^{63}_{(3,2,9,28)}$				28. Facilities layout
P$^{64}_{(3,2,9,31)}$				31. Master production schedule
P$^{65}_{(3,2,12,46)}$			*12. Order processing*	46. Order entry procedures
P$^{66}_{(3,2,12,45)}$				45. Order transmission means
P$^{67}_{(3,2,12,47)}$				47. Order picking procedures
P$^{68}_{(3,2,12,48)}$				48. Order follow-up procedures
P$^{69}_{(3,3,6,17)}$		3.Operational	*6. Material handling*	17. Unit loads
P$^{70}_{(3,3,6,16)}$				16. Types of material handling equipment
P$^{71}_{(3,3,9,30)}$			*9. Production*	30. Production schedule
P$^{72}_{(4,1,1,1)}$	4.Distribution Center	1. Strategic	*1. Strategic planning*	1. Definition of customer service
P$^{73}_{(4,1,1,3)}$				3. Customer service objectives
P$^{74}_{(4,1,1,2)}$				2. Degree of vertical integration and outsourcing
P$^{75}_{(4,1,2,5)}$			*2. Physical Facility network*	5. IF network strategy
P$^{76}_{(4,1,3,7)}$			*3. Communication and information network*	7. C&I network strategy
P$^{77}_{(4,1,5,9)}$			*5. Inventory management*	9. Inventory management strategy
P$^{78}_{(4,1,5,11)}$				11. Safety stock
P$^{79}_{(4,1,10,34)}$			*10.*	34. Warehousing

Code			Category	Description
			Warehousing	mission and functions
P$^{80}_{(4,2,2,4)}$			*2. Physical Facility network*	4. IF network design
P$^{81}_{(4,2,3,6)}$			*3. Communication and information network*	6. C&I network design
P$^{82}_{(4,2,4,8)}$			*4. Demand forecasting*	8. Forecasts of demand magnitude, timing and locations
P$^{83}_{(4,2,5,10)}$			*5. Inventory management*	10. Relative importance of inventories
P$^{84}_{(4,2,5,12)}$				12. Control methods
P$^{85}_{(4,2,5,13)}$				13. Desired inventory level
P$^{86}_{(4,2,6,15)}$		2. Tactical	*6. Material handling*	15. Material handling, fleet mix
P$^{87}_{(4,2,6,14)}$				14. Material handling, fleet control
P$^{88}_{(4,2,10,33)}$			*10. Warehousing*	33. Warehouse layout
P$^{89}_{(4,2,10,36)}$				36. Stock location
P$^{90}_{(4,2,10,32)}$				32. Receiving/shipping dock design
P$^{91}_{(4,2,10,35)}$				35. Safety systems
P$^{92}_{(4,2,12,46)}$			*12. Order processing*	46. Order entry procedures
P$^{93}_{(4,2,12,45)}$				45. Order transmission means
P$^{94}_{(4,2,12,47)}$				47. Order picking procedures
P$^{95}_{(4,2,12,48)}$				48. Order follow-up procedures
P$^{96}_{(4,3,6,17)}$		3.Operational	*6. Material handling*	17. Unit loads
P$^{97}_{(4,3,6,16)}$				16. Types of material handling equipment
P$^{98}_{(5,1,1,2)}$	5. Distribution channel	1. Strategic	*1. Strategic planning*	2. Degree of vertical integration and outsourcing
P$^{99}_{(5,1,3,7)}$			*3. Communication and information network*	7. C&I network strategy
P$^{100}_{(5,2,3,6)}$		2. Tactical	*3. Communication and information network*	6. C&I network design
P$^{101}_{(5,2,6,15)}$			*6. Material handling*	15. Material handling fleet

the relation between the Ps and their components. Columns 1 and 3 show the SC components, while columns 2 and 4 present the typologies of each component, their labels, and the number of related problems.

		mix
$P^{102}_{(5,2,6,14)}$		14. Material handling fleet control
$P^{103}_{(5,2,7,22)}$		22. Type of packaging
$P^{104}_{(5,2,7,18)}$	7. Packaging	18. Packaging design
$P^{105}_{(5,2,7,21)}$		21. Level of protection needed
$P^{106}_{(5,2,11,39)}$		39. Degree of consolidation
$P^{107}_{(5,2,11,41)}$		41. Transportation modes
$P^{108}_{(5,2,11,44)}$	11. Transportation	44. Types of carriers
$P^{109}_{(5,2,11,43)}$		43. Carrier selection
$P^{110}_{(5,2,11,40)}$		40. Transportation fleet mix
$P^{111}_{(5,2,12,46)}$		46. Order entry procedures
$P^{112}_{(5,2,12,45)}$	12. Order processing	45. Order transmission means
$P^{113}_{(5,2,12,47)}$		47. Order picking procedures
$P^{114}_{(5,2,12,48)}$		48. Order follow-up procedures
$P^{115}_{(5,3,6,17)}$		17. Unit loads
$P^{116}_{(5,3,6,16)}$	6. Material handling	16. Types of material handling equipment
$P^{117}_{(5,3,11,41)}$		41. Transportation modes
$P^{118}_{(5,3,11,44)}$		44. Types of carriers
$P^{119}_{(5,3,11,43)}$	3. Operational	43. Carrier selection
$P^{120}_{(5,3,11,40)}$	11. Transportation	40. Transportation fleet mix
$P^{121}_{(5,3,11,37)}$		37. Assignment of customers to vehicles
$P^{122}_{(5,3,11,38)}$		38. Vehicle routing and scheduling
$P^{123}_{(5,3,11,42)}$		42. Load plans

Source: The authors

Table 5 also shows how any given decision is often associated with the same logistic function (column 4), but not always with the same decision level (column 3). Only 11 decisions (8, 19, 20, 23, 28, 29, 30, 31, 37, 38 and 42) hold univocal relations with Ps, whereas the rest are associated with more than one P, but never exceed five. Table 6 shows

Table 6.
The relations between supply chain problems and their components.

Supply Chain component (id. Type of SC Component : # Ps)		
U	1. Supply channel level: 24	14. Fleet control: 4
	2. Stock level: 23	15. Fleet mix: 4
	3. Plant level: 24	16. Types of material handling equipment:4
	4. Distribution center level: 26	17. Unit loads: 4
	5. Distribution channel level 26	18. Packaging design: 2
N	1. Strategic level: 23	19. Information to be provided with the product: 1
	2. Tactical level: 76	20. Information media: 1
	3. Operational level: 24	21. Level of protection needed: 2
F	1. Related to warehousing	22. Type of packaging: 2
	2. Related to packaging	23. Quality control: 1
	3. Related to inventory management	24. Specifications of goods procured: 2
	4. Related to material handling	25. Order intervals and quantities: 2
	5. Related to strategic planning level	26. Selection of suppliers: 2
	6. Related to order processing	27. Procurement type: 2
	7. Related to production	28. Facilities layout: 1
	8. Related to demand projections	29. Product routing: 1
	9. Related to the Communication and Information (C&I) network	
	10. Related to Physical Facility network (PF)	30. Production program:1
	11. Related to transport	31. Master production schedule: 1
	12. Related to warehousing	32. Receiving / shipping dock design: 2
	13. Related to packaging	33. Warehouse layout: 2
	14. Related to inventory management	34. Warehousing mission and functions: 2
		35. Safety systems: 2
D	1. Definition of customer service: 2	36. Stock location: 2
	2. Degree of vertical integration and outsourcing: 5	37. Assignment of customers to vehicles: 1
	3. Customer service objectives: 2	38. Vehicle routing and scheduling: 1
	4. PF network design: 3	39. Degree of

Where **D** (on the right column near item 28) labels the right-hand list.

5. PF network strategy: 3	40. Transportation fleet mix: 4
6. C&I network design: 5	41. Transportation modes: 4
7. C&I network strategy: 5	42. Load plans: 1
8. Periodicity, magnitude and location projections: 1	43. Carrier selection: 4
9. Inventory management strategy: 2	44. Types of carriers : 4
10. Relative importance of inventory: 2	45. Order transmission means: 5
11. Safety stock: 2	46. Order entry procedures: 5
12. Control methods: 2	47. Order picking procedures: 4
13. Desired inventory level: 2	48. Order follow-up procedures: 5

Source: The authors

Finally, the concept of supply chain orientation (SCO) has been recently introduced as a philosophical approach to the implications of flow management in the supply chain [34]. However, flow management has been addressed independently of this philosophical approach, as can be observed in Arrupindi et al.'s work (1999). This can also be seen in Li et al.'s (2011) work on financial flow risk. SCO has been conceived as "the recognition, on the part of companies, of the systemic and strategic implications of the activities and processes involved in the management of the diverse flows of a supply chain". This concept has been slowly making its way in practical contexts, and has ended up becoming a significant SCM support. In this regard, one of the most relevant achievements has been the study of the implications of SCO on the procurement function through Key Supplier Relationship Management (KSRM). This approach has led to better Organizational Buying Effectiveness (OBE) as a way of measuring effective procurement behavior. SCO has also been incorporated to the supply chain structure through emphasizing the willingness of the companies to conceive the structure as an integrated entity [36]. Additionally, SCO has been used to manage uncertainty in business environments, in which it has been applied to the development of more efficient and flexible supply chains [37]; and to promoting a better willingness to "satisfy its needs by traveling along a common path" [38] among the supply chain agents. The Supply Chain Orientation concept and its framework have been developed since twenty-first century [39]. To summarize, the study of supply chain flow issues identified in the current work corresponds to the SCO paradigm. Finally, the characterization of supply chains as a research tendency has only emerged in the last decade [40].

4. Conclusions

The current paper proposes a SC problem featuring a holistic framework that is intended as a SC management and organization support tool, in which each P conveys an issue to be dealt with in the SC. This particular notation system not only allows the P in question to be specified within the context it shares with other SC aspects, but it also synthesizes its most outstanding features and sets the foundations for future developments in technical decision-making. This implies that the development of information parameterization systems allow adequate links between the inputs and outputs that modulate the SC P featuring framework. Future research perspectives are the following: the characterization of the Ps that are studied here or that may be identified in the future; the introduction of new Ps and elements into previously defined Ps; the development of specialized DSSs for SC issues; and the development of methodologies to identify P featuring parameters.

References

[1] Frohlich, M.T. and Westbrook, R., Arcs of integration: an international study of supply chain strategies. Journal of Operations Management, 19(2), pp 185-200, 2001. DOI: 10.1016/S0272-6963(00)00055-3.

[2] Sabri, E.H. and Beamon, B.M., A multi-objective approach to simultaneous strategic and operational planning in supply chain design. Omega, 28(5), pp 581-598, 2000. DOI: 10.1016/S0305-0483(99)00080-8

[3] Özbayrak, M., Papadopoulou, T.C. and Akgun, M., Systems dynamics modelling of a manufacturing supply chain system. Simulation Modelling Practice and Theory, 15(10), pp 1338-1355, 2007. DOI: 10.1016/J.SIMPAT.2007.09.007.

[4] CSCMP., Council of Supply Chain Management Professionals, [On line], pp. 3-17, 2007. [Accessed on June 20th of 2013]. Available at: http://cscmp.org.

[5] Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E., Managing the supply chain: The definitive guide for the business professional, McGraw-Hill, New York, 2004, pp 4-18.

[6] Davenport, T.H., Jarvenpaa, S.L. and Beers, M.C., Improving knowledge work processes. Sloan Management Review, 37, pp 53-66, 1996. DOI: 10.1016/0024-6301(93)90316-8.

[7] Dreyer, D.E., Performance measurement: A practitioner's perspective. Supply Chain Management Review, 4(4). pp. 63-68, 2000. DOI: 10.1108/13598541211258573

[8] Bhagwat, R. and Sharma, M.K., Performance measurement of supply chain management: A balanced scorecard approach. Computers & Industrial Engineering, 53(1), pp 43-62, 2007. DOI: 10.1016/j.cie.2007.04.001

[9] Gunasekaran, A., Patel, C. and McGaughey, R.E., A framework for supply chain performance measurement. International Journal of Production Economics, 87(3), pp 333-347, 2004. DOI: 10.1016/J.IJPE.2003.08.003

[10] Lai, K., Ngai, E.W.T. and Cheng, T.C.E., Measures for evaluating supply chain performance in transport logistics. Transportation Research Part E: Logistics and Transportation Review, 38(6), pp 439-456, 2002. DOI: 10.1016/S1366-5545(02)00019-4

[11] Berrah, L. and Cliville, V., Towards an aggregation performance measurement system model in a supply chain context. Computers in Industry, 58(7), pp 709-719, 2007. DOI: 10.1016/j.compind.2007.05.012

[12] Chen, I.J. and Paulraj, A., Towards a theory of supply chain management: the constructs and measurements. Journal of Operations Management, 22(2), pp 119-150, 2004. DOI: 10.1016/j.jom.2003.12.007

[13] Beamon, B.M., Measuring supply chain performance. International Journal of Operations and Production Management, 19(3), pp 275-292, 1999. DOI: 10.1108/01443579910249714

[14] Brewer, P.C. and Speh, T.W., Using the balanced scorecard to measure supply chain performance. Journal of Business Logistics, 21(1), pp 75-94, 2000. DOI: 10.1111.jbl.

[15] Li, S., Subba Rao, S., Ragu-Nathan, T.S. and Ragu-Nathan, B., Development and validation of a measurement instrument for studying supply chain management practices. Journal of Operations Management, 23(6), pp 618-641, 2005. DOI: 10.1016/j.jom.2005.01.002.

[16] Van der Vorst, J.G., Beulens, A.J. and van Beek, P., Modelling and simulating multi-echelon food systems. European Journal of Operational Research, 122(2), pp 354-366, 2000. DOI: 10.1016/S0377-2217(99)00238-6

[17] Krause, D.R., Pagell, M. and Curkovic, S., Toward a measure of competitive priorities for purchasing. Journal of Operations Management, 19(4), 497-512, 2001. DOI: 10.1016/S0272-6963(01)00047-X

[18] Angerhofer, BJ. and Angelides, M.C., A model and a performance measurement system for collaborative supply chains. Decision Support Systems, 42(1), pp 283-301, 2006. DOI: 10.1016/j.dss.2004.12.005

[19] Yeh, D.Y., Cheng, C.H. and Chi, M.L., A modified two-tuple FLC model for evaluating the performance of SCM: By the Six Sigma DMAIC process. Applied Soft Computing, 7(3), pp 1027-1034, 2007. DOI: 10.1016/j.asoc.2006.06.008

[20] Brun, A., Caridi, M., Salama, K.F. and Ravelli, I., Value and risk assessment of supply chain management improvement projects. International Journal of Production Economics, 99(1), pp 186-201, 2006. DOI: 10.1016/j.ijpe.2004.12.016

[21] Otto, A. and Kotzab, H., Does supply chain management really pay? Six perspectives to measure the performance of managing a supply chain. European Journal of Operational Research, 144(2), pp 306-320, 2003. DOI: 10.1016/S0377-2217(02)00396-X

[22] Barut, M., Faisst, W. and Kanet, J.J., Measuring supply chain coupling: An information system perspective. European Journal of Purchasing & Supply Management, 8(3), pp 161-171, 2002. DOI: 10.1016/S0969-7012(02)00006-0

[23] Schmitz, J. and Platts, K.W., Supplier logistics performance measurement: Indications from a study in the automotive industry. International Journal of Production Economics, 89(2), pp 231-243, 2004. DOI: 10.1016/S0925-5273(02)00469-3

[24] Riopel, D., Langevin, A. and Campbell, J.F., The network of logistics decisions. In Langevinand, A. and Riopel, D. (eds.). Logistics Systems: Design and optimization. New York: Springer, 2005, pp. 1-38. DOI: 10.1007/0-387-24977-X_1

[25] Miller, T.C., Hierarchical operations and supply chain planning. London: Springer Verlag, 2002. DOI: 10.1007/978-1-4471-0305-9

[26] Stadtler, H. and Kilger, C., Supply chain management and advanced planning: Concepts, models, software and case studies" (4thed.). Berlin: Springer, 2008. DOI: 10.1007/b106298

[27] Stewart, G., Supply-chain operations reference model (SCOR): The first cross-industry framework for integrated supply-chain management. Logistics Information Management, 10(2), pp 62-67, 1997. DOI: 10.1108/09576059710815716

[28] Linn, R. and Zhang, W., Hybrid flow shop scheduling: A survey. Computers and Industrial Engineering, 37(1), pp 57-61, 1999. DOI: 10.1016/S0360-8352(99)00023-6

[29] Kis, T. and Pesch, E., A review of exact solution methods for the non-preemptive multiprocessor flowshop problem. European Journal of Operational Research, 164(3), pp 592-608, 2005. DOI: 10.1016/j.ejor.2003.12.026

[30] Quadt, D. and Kuhn, H., A taxonomy of flexible flow line scheduling procedures. European Journal of Operational Research, 178(3), pp 686-698, 2007. DOI: 10.1016/j.ejor.2006.01.042

[31] Ribas, I., Leisten, R. and Framiñan, J., Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective. Computers & Operations Research, 37(8), pp 1439-1454, 2010. DOI: 10.1016/j.cor.2009.11.001

[32] Ruiz, R. and Vázquez-Rodríguez, J. ,The hybrid flow shop scheduling problem. European Journal of Operational Research, 205(1). pp. 1-18, 2010. DOI: 10.1016/j.ejor.2009.09.024

[33] Graham, R.L., Lawler, E.L., Lenstra, J.K. and Rinnoo Kan, A.H.G., Optimization and approximation in deterministic sequencing and scheduling: A survey. Annals of Discrete Mathematics, 5, pp 287-326, 1979. DOI: 10.1016/S0167-5060(08)70356-X

[34] Mentzer, J.T., Supply chain management. Thousand Oaks, CA, USA: Sage Publications Inc. 2001.

[35] Miocevic, D. and Crnjak-Karanovic, B., The mediating role of key supplier relationship management practices on supply chain orientation - The organizational buying effectiveness link. Industrial Marketing Management, 41(1), pp. 115-124. 2011. DOI: 10.1016/j.indmarman.2011.11.015

[36] Mello, J. and Stank, T., Linking firm culture and orientation to supply chain succes. International. Journal of Physical Distribution & Logistics Management, 35(8), pp. 542-554. 2005. DOI: 10.1108/09600030510623320

[37] Shanmugan, J. and Kabiraj, S.A., Case study approach for understanding supply chain orientation in Indian pharmaceutical firms. Kuwait Chapter of Arabian Journal of Business and Management Review, 1(9), pp. 45-78. 2012. DOI: 10.4172/2224-8358.10000106

[38] Hult, G.T., Ketchen, D.J., Adams, G.L. and Mena, J.A., Supply chain orientation and balance scorecard performance. Journal of Managerial Issues, [Online]. 20(4), pp. 526-544. 2008. Available at: http://labs.jstor.org/demo/stable/40604627

[39] Esper, T.L., Defee, C.C. and Mentzer, J.T., A framework of supply chain orientation, The International Journal of Logistics Management, 21(2), pp. 161-179, . 2010. DOI: 10.1108/09574091011071906

[40] García-Cáceres, R.G., Perdomo, A., Ortiz, O.O., Beltrán, P. and López, K., Characterization of the supply and value chains of Colombian cocoa. DYNA, 81(187), pp. 30-40, 2014. DOI: 10.15446/dyna.v81n187.39555

R.G. García-Cáceres, is Ph.D, MSc, II., Vice-president for Research – Universitaria Agustiniana (UNIAGUSTINIANA), Bogota Colombia.
ORCID: 0000-0003-0902-1038

J.W. Escobar, is Ph.D., MSc, II., Associated professor, Universidad del Valle, Cali, Colombia.
ORCID: 0000-0001-6175-9553

Algorithm for wideband spectrum sensing based on sparse Fourier transform

Alexander López-Parrado [ab] & Jaime Velasco-Medina [b]

[a] GDSPROC Research Group, Universidad del Quindío, Armenia, Colombia. parrado@uniquindio.edu.co
[b] Bionanoelectronics Research Group, Universidad del Valle, Santiago de Cali, Colombia. jaime.velasco@correounivalle.edu.co

Abstract

In this paper we present a novel sub-Nyquist algorithm to perform Wideband Spectrum Sensing (WSS) for Cognitive Radios (CRs) by using the recently developed Sparse Fast Fourier Transform (sFFT) algorithms. In this case, we developed a noise-robust sub-Nyquist WSS algorithm with reduced sampling cost, by modifying the Nearly Optimal sFFT algorithm; this was accomplished by using Gaussian windows with small support. Simulation results show that the proposed algorithm is suitable for hardware implementation of WSS systems for sparse spectrums composed of highly-noisy multiband-signals.

Keywords: Cognitive Radio; Compressed Sensing; Sparse Fourier Transform; Spectrum Sensing.

Algoritmo para sensado de espectro de banda ancha basado en transformada dispersa de Fourier

Resumen

En este trabajo se presenta un nuevo algoritmo sub-Nyquist para realizar Sensado de Espectro de Banda Ancha (WSS) para Radios Cognitivos (CR) mediante el uso de los algoritmos de Transformada Dispersa de Fourier (sFFT) recientemente desarrollados. En este caso, hemos desarrollado un algoritmo sub-Nyquist robusto ante el ruido para WSS con reducción en el costo de muestreo, mediante la modificación del algoritmo sFFT casi óptimo; esto se logró mediante el uso de ventanas Gaussianas con soporte pequeño. Los resultados de simulación muestran que el algoritmo propuesto es adecuado para la implementación hardware de sistemas WSS sobre espectros dispersos compuestos por señales multibanda altamente ruidosas.

Palabras clave: Radio Cognitiva; Sensado Compresivo; Transformada Dispersa de Fourier; Sensado de Espectro.

1. Introduction

Cognitive Radio (CR) is becoming the new paradigm for developing the next generation of radio communication systems. CR addresses the issue of spectrum misuse of current radio communication systems by adding cognitive features to the radios such as: spectrum sensing (SS), power control and spectrum management [1,2]. These features are initially presented in the IEEE 802.22 [3] standard, which was developed in 2011 by the Institute of Electrical and Electronics Engineers (IEEE) and defines a Wireless Regional Area Network (WRAN) that uses the Very High Frequency (VHF) and Ultra High Frequency (UHF) television (TV) bands by considering cognitive radio capabilities.

One technical challenge in CR is the efficient implementation of the SS function for a higher bandwidth by minimizing the required sampling rate. The SS function detects Primary Users (PUs) or Secondary Users (SUs) in some regions of the spectrum, allowing an opportunistic usage of the available bands. The IEEE 802.22 standard has an informative annex that defines two categories of SS techniques: blind sensing and signal specific sensing. Blind sensing techniques use energy measures, and signal specific
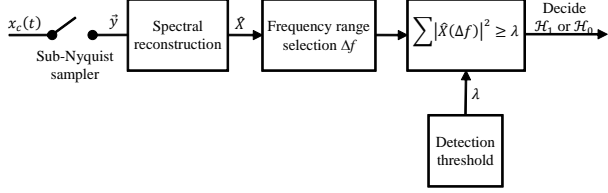
Figure 1. Block diagram of a Sub-Nyquist WSS system.
Source: [9].

sensing techniques use preambles and pilot signals. However, these sensing techniques are narrowband because they can only sense a single carrier frequency.

On the one hand, measurements carried out by the Microsoft Spectrum Observatory at Washington DC [4] show that around 3% of the band between 30 MHz and 3 GHz is used sparsely to host most of the worldwide radio services [4]. This sparse occupancy of the radio electric spectrum has motivated the theoretical research about Wideband Spectrum Sensing (WSS) techniques during the last five years [5-7]. In this case, the research results have shown that sub-Nyquist sampling techniques such as Analog to Information Conversion (AIC) [6],[8], Modulated Wideband Conversion (MWC) [6][10] and Multi Coset (MC) sampling [6],11,12] are promissory candidates for developing WSS systems as the shown in Fig. 1

The sub-Nyquist WSS systems are usually composed of a sub-Nyquist sampler, a spectral reconstruction block and a decision stage [5,6[9]. The spectral reconstruction block is usually constructed by using Compressive Sensing (CS) techniques [13,14].

On the other hand, spectral reconstruction considering the Sparse Fast Fourier Transform (sFFT) algorithms [15-19] has not been well developed as these algorithms either require a high sampling cost for typical spectrum occupancy [15,16,[19], they are very noise sensitive [17], or they are too complex [18].

In the context of WSS algorithms, the challenge is to achieve sub-Nyquist sampling rates using low-power and low-speed ADCs for highly-noisy signals. Thus, considering the above, the main contribution of this paper is the design of a new sub-Nyquist WSS algorithm that reduces the sampling cost by using a modified version of the sFFT algorithm with Gaussian small support windows. This proposed algorithm is very suitable for hardware implementation of WSS systems using ASICs or FPGAs, and to the best of our knowledge, it is the first that uses the recently developed Nearly Optimal Sparse Fourier Transform.

The rest of the paper is organized as follows: Section 2 presents some mathematical basics about the sub-Nyquist WSS algorithm we developed, Section 3 describes the proposed sub-Nyquist WSS algorithm, Section 4 presents simulation results performed on scenarios composed of highly-noisy multiband-signals, and, Section 5, presents our conclusions and suggestions for future work.

2. Mathematical background

In this section, we present some fundamental concepts about the sub-Nyquist sFFT algorithm we developed. First,

we explain some basics about DFT and sparse signals, and second, we describe the mathematical tools pseudo-random spectral permutation, filtering window, and hashing function.

2.1. Discrete Fourier transform and sparsity

Given a discrete time signal $\boldsymbol{x} \in \mathbb{C}^N$ of length N, its N-point Discrete Fourier Transform (DFT) $\hat{\boldsymbol{x}} \in \mathbb{C}^N$ is defined in Eq. (1).

$$\hat{\boldsymbol{x}}_k = \frac{1}{N} \sum_{n \in [N]} \boldsymbol{x}_n \, \omega^{kn}, k \in [N] \qquad (1)$$

Where N is a power of two, $[N]$ denotes the set of indexes $\{0, 1, \ldots, N-1\}$, and $\omega = e^{-i\,2\pi/N}$ is the N-th root of unity. In this case, the number of non-zero elements of the vector $\hat{\boldsymbol{x}}$ is called the sparsity order K and is defined in Eq. (2).

$$K = |\text{supp}(\hat{\boldsymbol{x}})|_0 \qquad (2)$$

Where $\text{supp}(\hat{\boldsymbol{x}})$ is the set of indexes of the non-zero elements of the vector $\hat{\boldsymbol{x}}$, and $||_0$ represents the l_0-norm of the vector. Then, a time domain signal $\boldsymbol{x}$ is sparse in the DFT domain if $K \ll N$.

In this context, a set of algorithms called sFFT takes advantage of the signal sparsity in the DFT domain to speed up the runtime of the Fast Fourier Transform (FFT) algorithms used to calculate the DFT [15-18]. These sFFT algorithms, like the Nearly Optimal sFFT algorithm presented in [16], use the following mathematical tools: pseudo-random spectral permutation [15-19], filtering window [16] and hashing function [15,16].

2.2. Pseudo-random spectral permutation

This permutation isolates spectral components from each other [19] and is performed as described in Eq. (3).

$$\boldsymbol{xp}_n = \boldsymbol{x}_{\sigma(n-a) \bmod N},$$
$$\widehat{\boldsymbol{xp}}_{\pi_p(k,\sigma,N)} = \hat{\boldsymbol{x}}_k \omega^{\sigma ka} \qquad (3)$$

Where $\boldsymbol{xp}$ and $\widehat{\boldsymbol{xp}}$ are the permuted spectrum signals in the time domain and the DFT domain, respectively; $\pi_p(k,\sigma,N) = \sigma k \bmod N$ is the spectral permutation function; and $\sigma \in \{2c+1 | c \in [N/2]\}$ and $a \in [N]$ are the spectral permutation parameters. The spectral permutation function translates the frequency bin from the k-th location to the $\pi_p(k,\sigma,N)$-th location, in this case $\sigma^{-1} \bmod N$ exists for all odd σ if N is a power of two. The sFFT algorithm randomly chooses the spectral permutation parameters σ and a from a uniform distribution. Thus, the spectral permutation with these pseudo-random parameters is related to a pseudo-random sampling scheme [15-19].

2.3. Filtering window

The filtering window is a new mathematical tool that reduces the size of the FFT from points N to B. This is accomplished in

the Nearly Optimal sFFT algorithm by extending a flat passband region of width N/B around each sparse component; this approach replaces the filter bank of previous sFFT algorithms [19,20] and avoids the use of non-equispaced data FFTs [21]. Nonetheless, this flat window has a support that is not small enough to achieve sub-Nyquist sampling rates.

Thus, in order to reduce the sampling cost of the sFFT algorithm, we designed a small support window $G, \widehat{G}'$ such that $|\text{supp}(G)|_0 = B$; nevertheless, this small support can reduce the accuracy, which is not a big issue in the case of WSS systems. In this case, the window in the time domain is designed with an ideal low-pass filter using a Gaussian window with finite duration to truncate the impulse response. The cutoff frequency of the low-pass filter is $2C$, where $C = 1/(2B)$, and the standard deviation σ_g of the Gaussian window is obtained from the 68-95-99.7 rule [22], as described in Eq. (4).

$$\sigma_g = \frac{B}{6} \tag{4}$$

Eq. (5)-(6) describe the Gaussian window in the time domain and the DFT domain, respectively.

$$G_{n+\frac{N}{2}\text{mod } N} = 2\,C\,e^{-\frac{(n-N/2)^2}{2\sigma_g^2}}\,\text{sinc}(2C(n - N/2))/\widehat{G}_{N/2}\ ,n \in [N] \tag{5}$$

$$\widehat{G}'_{k+\frac{N}{2}\text{mod } N} =$$
$$\begin{cases} 0, \text{if } |k - N/2| \geq N(C + 1/\sigma_g) \\ \dfrac{\text{ncdf}\left(2\pi\sigma_g\left(\dfrac{k-\frac{N}{2}}{N}+C\right)\right) - \text{ncdf}\left(2\pi\sigma_g\left(\dfrac{k-\frac{N}{2}}{N}-C\right)\right)}{\widehat{G}_{\frac{N}{2}}}, \text{otherwise} \end{cases}, k \in [N] \tag{6}$$

Where vector G' is the approximated window and $\text{ncdf}(x) = \text{erfc}(-x / \sqrt{2})/2$ is the Normal Cumulative Distribution Function [23].

The Gaussian window is normalized both in the time domain and the DFT domain in order to achieve unit DC gain, and its total bandwidth in DFT domain is given by Eq. (7).

$$BW_{G'} = 13N/B \tag{7}$$

Finally, the windowing process is described in Eq. (8), and it is performed in the time domain after the pseudo-random spectral permutation is carried out.

$$y = xp \circ G \tag{8}$$

Where, y is the windowed signal in the time domain.

2.4. Hashing function

The hashing function obtains B points from the N-point spectrum of the signal y, these points are separated by N/B bins, and they are obtained by calculating the B-point DFT of the sub-sampled signal obtained from y. The vector that has the hashes of signal y is $\widehat{u} \in \mathbb{C}^B$ and it is calculated using Eq. (9) [15,16].

$$\widehat{u}_j = DFT\left\{\sum_{i\in[N/B]} y_{j+Bi}\right\}, j \in [B] \tag{9}$$

From Eq. (7)-(9), it is possible to note that for each sparse component of the signal x there are 14 non-zero hashes located in the offsets given by Eqs. (10) and (11).

$$o_{fk}(j,\sigma,N,B) = \pi_p(j,\sigma,N) - (h_f(j,\sigma,N,B) - k)N/B, k \in \{0,1,...,6\} \tag{10}$$

$$o_{ck}(j,\sigma,N,B) = \pi_p(j,\sigma,N) - (h_c(j,\sigma,N,B) + k)N/B, k \in \{0,1,...,6\} \tag{11}$$

Where $h_f(j,\sigma,N,B) - k \bmod B, k \in \{0,1,...,6\}$ and $h_c(j,\sigma,N,B) + k \bmod B, k \in \{0,1,...,6\}$ are the 14 indexes of the hashes $\widehat{u}$ for each hashed single sparse component, and $h_f(j,\sigma,N,B) = \lfloor \pi_p(j,\sigma,N)B/N \rfloor$ is the floor-hash function and $h_c(j,\sigma,N,B) = \lceil \pi_p(j,\sigma,N)B/N \rceil$ is the ceil-hash function.

Finally, it has to be noted that the use of pseudo-random spectral permutation and small support windows leads to a sub-Nyquist random sampling scheme; where, under certain conditions, the average sampling rate is below Nyquist.

3. Sub-Nyquist wideband spectrum sensing algorithm

In this section, we describe the Sub-Nyquist WSS algorithm, called *SNSparseWSS*, which presents a reduced sampling rate compared to the sFFT algorithm described in [16]. This improvement is achieved by using the window described in the past section and by performing several modifications to the procedures presented in [16].

The *SNSparseWSS* algorithm, described in Alg. 1, calculates the spectrum occupancy X_{rfca} with constant False Alarm Probability (P_{FA}) [2],[11], and has the following input parameters: the sparse bandwidth BW in Hz, the total bandwidth W in Hz, the duration of the sensing window τ in seconds, the noise power σ_n, and the constants of the sFFT algorithm R_{est} and s [16].

The algorithm calculates the spectrum occupancy using four processing stages: the first one performs sub-Nyquist sampling on the wideband complex signal $x(t)$; the second one locates the sparse components by using the vector x_l, the set of permutation parameters P_l, and the modified *LocateSignal* procedure [16],[24]; the third one estimates the DFT values by using the vector x_e, the set of permutation parameters P_e, and the modified *EstimateValues* procedure [15]; and the fourth one detects the occupied bands with constant P_{FA} using the *ConstantPfaRecovery* procedure, which determines whether a channel is occupied or not by a PU. Therefore, it is necessary to test two spectrum sensing hypotheses, H_0 for vacant channel and H_1 for occupied channel [2], by using a detector with constant P_{FA} [2],[11] as described in Eq. (12) [11].

Algorithm 1. *SNSparseWSS* algorithm.

Input: Sparse bandwidth BW in Hz, total bandwidth W in Hz, sensing time τ in seconds, number R_{est} of estimation loops, threshold s for location, noise power σ_n

Output: X_{rcfa}

SNSparseWSS procedure $(BW, W, \tau, R_{est}, s, \sigma_n)$

    //Window generation

    $N = 2^{\lfloor \log_2(W \times \tau) \rfloor}$;

    $K = \lceil BW/W \times N \rceil$;

    $B = 2^{\lfloor \log_2 K \rfloor + 1}$;

    $R_{loc} = \lfloor \log_2 \log_2 N \rfloor$;

    Calculate Gaussian Window $(G, \widehat{G}')$ by using Eqs. (5) and (6).

    //Sub-Nyquist sampling set genration

    Pre-calculate all permutation parameters in P_l for 1 location loop [16];

    Pre-calculate all permutation parameters in P_e for R_{est} estimation loops [16];

    $S_l = \bigcup\limits_{(\sigma,a,\beta) \in P_l, n \in \text{supp}(G)} \sigma(n - a + \beta) \bmod N$;

    $S_e = \bigcup\limits_{(\sigma,a) \in P_e, n \in \text{supp}(G)} \sigma(n - a) \bmod N$;

    //Parallel random sampling of signal

    $x_{S_l} = x(S_l/W)$;

    $x_{S_e} = x(S_e/W)$;

    //sFFT calculation

    //Location of Components

    $L = \textbf{\textit{LocateSignal}}(x_l, \mathbf{0}, B, P_l, G, \widehat{G}', s, R_{loc})$;

        //Estimation of Components

    $\{\widehat{w}, J\} = \textbf{\textit{EstimateValues}}(x_e, \mathbf{0}, B, P_e, G, \widehat{G}', L, 3K_r, R_{est})$;

    $X_r = |\widehat{w}_J|_2^2$

    $X_{rfca} = \textbf{\textit{ConstantP}}_{fa}\textbf{\textit{Recovery}}(X_r, \sigma_n, N, P_{FA})$;

    end

return X_{rfca}

Source: Authors.

$$X_{rcfa_j} = \begin{cases} H_1, & X_{r_j} \geq \psi_j \\ H_0, & \text{otherwise} \end{cases} \qquad (12)$$

Where, $\psi_j = \sigma_n^2 \left(\sqrt{\dfrac{2}{N^3}} Q^{-1}(P_{FA}) + 1/N \right)$ is the decision threshold.

3.1. Modified procedures

We modified the HashToBins, LocateSignal, LocateInner, and EstimateValues procedures described in [16] in order to use the proposed filtering window and to reduce the execution time.

3.1.1. HashToBins procedure

This procedure, presented in Alg. 2, calculates the hashes-error by subtracting the hashes of the instantaneous estimation $\widehat{z}$ from the hashes $\widehat{u}$ [16], and it has the following input parameters: the time domain signal x; the instantaneous estimation $\widehat{z}$ of $\widehat{x}$; the parameter B; the spectral permutation parameters σ, and a; and the vectors $G, \widehat{G}'$ of the filtering window in the time domain and the DFT domain respectively.

Algorithm 2. Modified Hash to bins function.

Input: $x \in \mathbb{C}^N$, $\widehat{z} \in \mathbb{C}^N$, $B \in \{2^c | c \in [\log_2 N]\}$, $\sigma \in \left\{2c + 1 \middle| c \in \left[\frac{N}{2}\right]\right\}$, $a \in [N]$, $G \in \mathbb{R}^B$, $\widehat{G}' \in \mathbb{R}^{13N/B}$

Output: $\widehat{u} \in \mathbb{C}^B$

    HashToBins procedure $(x, \widehat{z}, B, \sigma, a, G, \widehat{G}')$

    $u_j = 0 \ \forall \ j \in [B]$;

    //Spectral permutation and sub-sampling

    for $j \in \{N - |\text{supp}(G)|_0/2, \ N + |\text{supp}(G)|_0/2 - 1\}$ **do**

        $u_{j \bmod B} = u_{j \bmod B} + x_{\sigma(j-a) \bmod N} G_{j-N+|\text{supp}(G)|_0/2}$;

    end

    //Sub-sampled DFT

    $\widehat{u} = \text{FFT}_B(u)$;

    //Efficient convolution calculation

    for $j \in supp(\widehat{z})$ **do**

        for $k \in [6]$ **do**

            $\widehat{u}_{h_f(j,\sigma,N,B)-k \bmod B} = \widehat{u}_{h_f(j,\sigma,N,B)-k \bmod B}$

                 $- \widehat{G}'|_{o_{fk}(j,\sigma,N,B)} |\widehat{z}_j \omega^{\sigma a j}$;

           $\widehat{u}_{h_c(j,\sigma,N,B)-k \bmod B} = \widehat{u}_{h_c(j,\sigma,N,B)-k \bmod B}$

                 $- \widehat{G}'|_{o_{ck}(j,\sigma,N,B)} |\widehat{z}_j \omega^{\sigma a j}$;

        end

    end

    return $\widehat{u}$

Source: Adapted from [16].

This procedure calculates the hashes-error using three processing stages: the first one simultaneously calculates in the time domain the pseudo-random spectral permutation, the windowing, and the hashing process; the second calculates the DFT domain hashes $\widehat{u}$ by performing the B-point FFT of the time domain hashes u; and the third one calculates the hashes-error by subtracting the DFT domain hashes of $\widehat{z}$ from the hashes $\widehat{u}$, in this case there are 14 hashes for each sparse component in $\widehat{z}$.

3.1.2. LocateSignal procedure

This procedure, presented in Alg. 3, calculates the set $L \in \mathbb{N}^{O(B)}$ of frequency bins corresponding to $O(B)$ sparse components found in x; and it has the following input parameters: the time domain signal x; the instantaneous estimation $\widehat{z}$ of $\widehat{x}$; the parameter B; the spectral permutation parameters P_l; the vectors $G, \widehat{G}'$ of the filtering window in the time domain and the DFT domain respectively; the threshold constant for location s [16]; and the number of location iterations R_{loc} [16].

This procedure locates the sparse components of x using four processing stages: the first one sets the initial conditions, the second one adjusts the frequency locations, the third one reduces the search region, and the fourth one inverts the spectral permutation. The setting of initial conditions is performed by first calculating a reference hashes vector $\widehat{u}$; second pre-calculating an initial guess of value $l_j^{(0)} = jN/B \ \forall \sim j \in [B]$ of frequency locations in the permuted spectrum; and third by pre-calculating the initial values of w, t, and D_{max}, where w is the width of the region for searching the frequency adjustment, t is the number of candidate adjustments in w, D_{max} is the number of

Algorithm 3. Modified locate signal function.

Input: $x \in \mathbb{C}^{O(K)}$, $\hat{z} \in \mathbb{C}^N$, $B \in \{2^c | c \in [\log_2 N]\}$, $P_l \in \mathbb{N}^{3 \times O(K)}$, $G \in \mathbb{R}^B$, $\hat{G}' \in \mathbb{R}^{\frac{13N}{B}}$, $s \in \mathbb{R}$, $R_{loc} \in \mathbb{N}$

Output: $L \in \mathbb{N}^{O(B)}$

LocateSignal procedure $(x, \hat{z}, B, P_l, G, \hat{G}', s, R_{loc})$

    Choose a and σ from P_l;

    $\hat{u} = HashToBins(x, \hat{z}, B, \sigma, a, G, \hat{G}')$

    $l_j^{(0)} = jN/B \; \forall \sim j \in [B]$;

    $w_0 = N/B$;

    $t = \log_2 N$;

    $t' = t/4$;

    $D_{max} = \lceil \log_{t'}(w_0 + 1) \rceil$;

    //Main loop

    for $D \in [D_{max}]$ **do**

        $l^{(D+1)} = LocateInner(x, \hat{z}, B, \sigma, a, P_l, G, \hat{G}', s, R_{loc}, w_0 / (t')^D, t, \hat{u}, l^{(D)})$;

    end

    $L = \left\{ \sigma^{-1} \left\lfloor l_j^{(D_{max})} + 0.5 \right\rfloor \middle| j \in [B] \right\}$;

    return L

Source: Adapted from [16].

Algorithm 4. Modified locate signal inner function.

Input: $x \in \mathbb{C}^N$, $\hat{z} \in \mathbb{C}^N$, $B \in \{2^c | c \in [\log_2 N]\}$, $\sigma \in \left\{ 2c + 1 \middle| c \in \left[\frac{N}{2} \right] \right\}$, $a \in [N]$, $G \in \mathbb{R}^B$, $\hat{G}' \in \mathbb{R}^{\frac{13N}{B}}$, $s \in \mathbb{R}$, $R_{loc} \in \mathbb{N}^+$, $w \in \mathbb{R}$, $t \in \mathbb{N}^+$, $\hat{u} \in \mathbb{C}^B$, $l \in \mathbb{N}^{O(B)}$

Output: $l' \in \mathbb{N}^{O(B)}$

LocateInner procedure $(x, \hat{z}, B, \sigma, a, P_l, G, \hat{G}', s, R_{loc}, w, t, \hat{u}, l)$

    $v_{j,q} = 0 \; \forall (j, q) \in [B] \times [t]$;

    //Main loop

    for $i \in [R_{loc}]$ **do**

        Choose β from P_l;

        $\hat{u}' = HashToBins(x, \hat{z}, B, \sigma, a + \beta, G, \hat{G}')$;

        for $j \in [B]$ **do**

            if $l_j \neq \perp$ **then**

                $r_j = \hat{u}_j / \hat{u}'_j$;

                $c_j = \arctan2(\Im\{r_j\}, \Re\{r_j\})$;

                if $c_j < 0$ **then**

                    $c_j = c_j + 2\pi$;

                end

                for $q \in [t]$ **do**

                    $m_{j,q} = l_j + \frac{q + 1/2}{t} w$;

                    $\theta_{j,q} = \frac{2\pi \beta m_{j,q}}{N} \mod 2\pi$;

                    if $\min\{|\theta_{j,q} - c_j|, 2\pi - |\theta_{j,q} - c_j|\} < s\pi$ **then**

                        $v_{j,q} = v_{j,q} + 1$;

                    end

                end

            end

        end

    end

    for $j \in [B]$ **do**

        $Q = \{q \in [t] \mid v_{j,q} > R_{loc}/2\}$;

        if $Q \neq \emptyset$ **then**

            $l'_j = l_j + \min_{q \in Q} qw/t$;

        else

            $l'_j = \perp$;

        end

    end

    return l'

Source: Adapted from [16].

adjustments, and R_{loc} is the number of location iterations for each adjustment. The adjustment of frequency locations is performed by using D_{max} times the procedure *LocateInner*. The reduction of the search region is performed by dividing w by a factor $1/(t')^D$ at the D-th adjustment, this reduction allows a systematic refining of the frequency location. Finally, the spectral permutation inversion is performed by calculating the function $\pi_p^{-1}(k, \sigma, N)$ on each located frequency.

3.1.3. LocateInner procedure

This procedure, presented in Alg. 4, calculates the adjustment $l' \in \mathbb{N}^{O(B)}$ of the frequency locations in the permuted spectrum, and it has the following input parameters: the time domain signal x; the instantaneous estimation $\hat{z}$ of $\hat{x}$; the parameter B; the spectral permutation parameters σ, a, and P_l; the vectors G, $\hat{G}'$ of the filtering window in the time domain and the DFT domain respectively; the threshold constant for location s; the number of location iterations R_{loc}; the width of search region w; the number of candidate adjustments t; the reference hashes vector $\hat{u}$; and the current estimation of frequency locations l.

This procedure performs the adjustment of the frequency location using five processing stages: the first one sets the initial conditions, the second calculates the hashes-error, the third calculates the angles of the hashes-error and candidate frequency bins, the fourth performs the voting stage, and the fifth locates the frequency bins. The setting of initial conditions clears the vote counters of the t candidate adjustments for the B candidate frequency bins. The hashes calculation stage obtains R_{loc} hashes vectors $\hat{u}'$, which are calculated from the signal x by using the *HashToBins* procedure with pseudo-random permutation parameters of the form $(\sigma, a + \beta)$. The angle calculation stage obtains the vector c, which represents the angle differences between the reference hashes $\hat{u}$ and the hashes $\hat{u}'$. The voting stage increments the vote counter $v_{j,q}$ corresponding

to the j-th frequency bin and q-th adjustment, if Eq. (13) is satisfied.

$$\min\{|\theta_{j,q} - c_j|, 2\pi - |\theta_{j,q} - c_j|\} < s\pi \qquad (13)$$

Where $\theta_{j,q}$ is the angle of the j-th candidate frequency bin for the q-th adjustment, that is, $\theta_{j,q}$ is related to the q-th candidate frequency adjustment qw/t. Additionally, the above calculation converges if β is chosen at random from the set $\{\lfloor sNt/4w \rfloor, \ldots, \lfloor sNt/2w \rfloor\}$ with enough small threshold s. Finally, the frequency location stage selects the minimum q from the set $Q = \{q \in [t] \mid v_{j,q} > R_{loc}/2\}$; thus the estimated j-th permuted-frequency bin is refined using Eq. (14).

$$l'_j = l_j + \min_{q \in Q} qw/t \qquad (14)$$

Algorithm 5. Modified estimate values function.

Input: $x \in \mathbb{C}^N$, $\hat{z} \in \mathbb{C}^N$, $B \in \{2^c | c \in [\log_2 N]\}$, $P_e \in \mathbb{N}^{2 \times O(K)}$, $G \in \mathbb{R}^B$, $\hat{G}' \in \mathbb{R}^{\frac{13N}{B}}$, $L \in \mathbb{N}^{O(B)}$, $K' \in \mathbb{N}^+$, $R_{est} \in \mathbb{N}^+$

Output: $L \in \mathbb{N}^{O(B)}$

***EstimateValues* procedure** $\left(x, \hat{z}, B, G, \hat{G}', L, K', R_{est}\right)$

    $\hat{w}_j = 0 \; \forall \, j \in [|L|_0]$;

    //Main loop

    for $i \in [R_{est}]$ **do**

        Choose σ_i and a_i from P_e;

        $\hat{u} = HashToBins\left(x, \hat{z}, B, \sigma_i, a_i, G, \hat{G}'\right)$;

        for $j \in [|L|_0]$ **do**

            $\hat{u}'_{i,j} = \hat{u}_{h(L_j, \sigma_i, N, B) \bmod B} \, \omega^{-\sigma_i a_i L_j} / \hat{G}'_{|o(L_j, \sigma_i, N, B)|}$;

        end

    end

    for $j \in [|L|_0]$ **do**

        // Median for real and imaginary parts
        separately across the i dimension

        $\hat{w}_j = \text{median}_i \, \hat{u}'_{i,j}$

    end

    $J = \arg \max\limits_{|J|_0 = \min\{|L|_0, K'\}} |\hat{w}_J|_2$

  return $\{\hat{w}_J, J\}$

Source: Adapted from [16].

3.1.4. EstimateValues procedure

This procedure, presented in Alg. 5, calculates the DFT estimation adjustment $\hat{w}_J$, and it has the following input parameters: the time domain signal x; the instantaneous estimation $\hat{z}$ of $\hat{x}$; the parameter B; the spectral permutation parameters P_e; the vectors $G, \hat{G}'$ of the filtering window in the time domain and the DFT domain respectively; the set of located sparse components L; the number of sparse components to estimate K'; and the number of estimation iterations R_{est} [16]. This procedure calculates the DFT estimation adjustment $\hat{w}_J$ using three processing stages: the first one calculates the R_{est} sets of hashes-error from the signal x by using the HashToBins procedure and considering different pseudo-random permutation parameters; the second one separately calculates the median of the real and imaginary parts of the calculated hashes-error by only considering the set of located sparse frequency bins in L, and by cancelling the pseudo-random spectral permutation and the effect of the windowing in the DFT domain; and the third one saves the K' most energetic components $\hat{w}_J$.

4. Simulation results of SNSparseWSS algorithm

This section presents the simulation results of the *SNSparseWSS* algorithm for sub-Nyquist sampling and verification results for Wideband Spectrum Sensing.

4.1. Sub-Nyquist capabilities

In order to verify the sampling cost, we need to know all the spectral permutation parameters σ_i and a_i that are pseudo-randomly generated by the SNSparseWSS algorithm; thus, the set of sampling points can be calculated using Eq. (15).
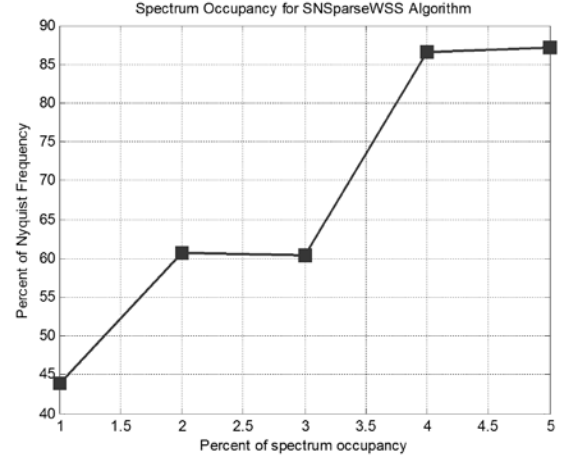
Figure 2. Percent of Nyquist Sampling Rate versus Spectrum occupancy percent for *SNSparseWSS* algorithm.
Source: Authors.

$$S = \bigcup_{n \in \text{supp}(G)} \{(\sigma_i (n - a_i)) \bmod N\} \tag{15}$$

Considering the set of sampling points S, it is possible to calculate the average sampling rate f_{sa} of the *SNSparseWSS* algorithm using Eq. (16).

$$f_{sa} = f_{Nyq} \frac{|S|_0}{\sum_{k \in \text{supp}(S)} \nabla S_k} \tag{16}$$

Where, f_{Nyq} is the Nyquist frequency, and $\nabla S_k = S_k - S_{k-1}$ is a finite-backward difference to estimate the average separation between samples. Fig. 2 shows the percentage of Nyquist frequency versus the percentage of spectrum occupancy for the *SNSparseWSS* algorithm.

From Fig. 2, we can see that the algorithm reaches a sampling rate of close to $0.6 \times f_{Nyq}$ for a spectrum occupancy between 2% and 3%; thus, for typical scenarios where the spectrum occupancy is close to 2% the *SNSparseWSS* algorithm is very suitable for implementing sub-Nyquist WSS systems and its performance is comparable to the WSS systems based on MC sampling [11].

4.2. Verification of the SNSparseWSS algorithm

In order to verify the WSS capabilities of the SNSparseWSS algorithm, we considered as test vehicle a highly-noisy multiband-signal scenario composed of 5 time domain signals located in the center frequencies 317-576-1300-984-163 MHz, and a total bandwidth W of 1.5 GHz. Each signal has a bandwidth of 5 MHz, is composed of random 4-QAM symbols that are filtered using a raised cosine filter with roll-off factor r = 0.5, and has SNR values from -5 dB to 5 dB. The test signal with SNR=-5 dB is shown in Fig. 3.
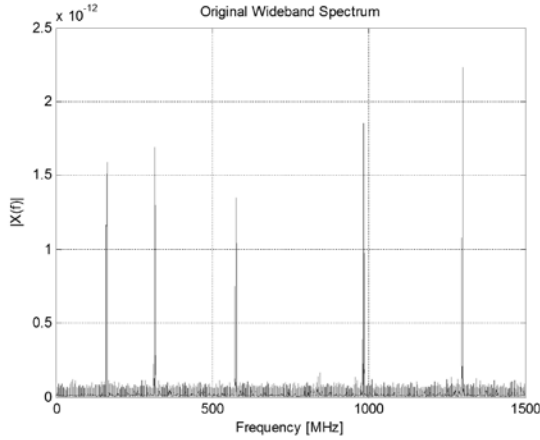
Figure 3. Wideband spectrum of 1.5 GHz containing 5 signals each with bandwidth **BW** = 5 MHz and SNR = -5 dB.
Source: Authors.

Figure 5. Wideband spectrum sensing result by using the constant false alarm probability detector.
Source: Authors.

In this case, the duration of the spectrum sensing window is $\tau = 160 \ \mu s$ which implies that an FFT with $N = 2^{17}$ must be used, and the total sparse bandwidth is 25 MHz (5 signals $\times$ 5 MHz) which leads to a sub-sampled FFT with $B = 8192$. The algorithm was parameterized with $R_{est} = 10$ estimation iterations and a location threshold of $s = 0.1$, with these settings the *SNSparseWSS* algorithm achieves an average sampling rate of $f_{sa} = 0.599 \times f_{Nyq}$.

Fig. 4 shows the recovered spectrum using the *SNSparseWSS* algorithm. In this figure, we can see that the Gaussian small support window increases the total error of estimated DFT value; this issue can be mitigated using appropriate settings for the constant P_{FA} detector, which implies that the occupied channels can be detected with constant P_{FA} [11].

Fig. 5 shows the WSS simulation results using the constant P_{FA} detector with $P_{FA} = 0.01$. In this figure, we can see that the algorithm can perform WSS by detecting the occupied channels by the Pus. If additional information about the multiband-signal is available, such as the minimum channel separation Δ_f [11], the detection can be improved by reducing the P_{FA}.

Figure 6. False alarm probability performance for different SNRs.
Source: Authors.

Fig. 6 shows simulation results for the false alarm probability with SNR values of -5 dB, -2.5 dB, 0 dB, 2.5 dB, and 5 dB. In this figure we can see that the P_{FA} is approximately constant regardless the SNR of the multiband-signals.

5. Conclusions and future work

In this paper, we present the design of a novel algorithm for sub-Nyquist Wideband Spectrum Sensing based on a modified Nearly Optimal sFFT algorithm. This WSS algorithm was verified using several tests. From the verification results we can conclude that the proposed algorithm is suitable for implementing the spectrum sensing function of wideband cognitive radios in highly-noisy environments. To the best of our knowledge, the proposed WSS algorithm is the first that uses the new Nearly Optimal Sparse Fourier Transform algorithm, and it has a reduced sampling cost by using flat Gaussian small support windows

Figure 4. Recovered spectrum for $\tau = 160 \ \mu s \ (N = 2^{17}) \ B = 8192$, , 1 location iteration, 10 estimation iterations, and Gaussian window with $|\text{supp}(\boldsymbol{G})|_0 = B$.
Source: Authors.

and modified procedures.

Future work will addressed efficient hardware implementation of the modified Nearly Optimal sFFT algorithm using an FPGA.

Acknowledgements

References

[1] Mitola, J. and Maguire, G.Q., Cognitive radio: Making software radios more personal. IEEE Personal Communications, 6(4), pp. 13-18, 1999. DOI: 10.1109/98.788210.

[2] Arslan, H., Cognitive radio, software defined radio, and adaptive wireless systems. Dordrecht: Springer, 2007. DOI: 10.1007/978-1-4020-5542-3.

[3] IEEE. IEEE 802.22-2011, wireless regional area networks (wran) – specific requirements part 22: Cognitive wireless ran medium access control (mac) and physical layer (phy) specifications: Policies and procedures for operation in the tv bands. 2011. DOI: 10.1109/IEEESTD.2011.5951707.

[4] M.T.P. Group, Microsoft spectrum observatory, [online], Seattle, [Date of reference, Nov. 2013.]. Available at: http://spectrum-observatory.cloudapp.net/.

[5] Vito, L.D., A review of wideband spectrum sensing methods for cognitive radios, Proceedings of Instrumentation and Measurement Technology Conference (I2MTC), pp. 2257-2262, 2012. DOI: 10.1109/I2MTC.2012.6229530.

[6] Hongjian, S., Nallanathan, A., Wang, C.X. and Chen, Y., Wideband spectrum sensing for cognitive radio networks: a survey. IEEE Wireless Communications, 20(2), pp. 74-81, 2013. DOI: 10.1109/MWC.2013.6507397.

[7] Axell, E., Leus, G., Larsson, E.G. and Poor, H.V., Spectrum sensing for cognitive radio: State-of-the-art and recent advances. IEEE Signal Processing Magazine, 29(3), pp. 101-116, 2012. DOI: 10.1109/MSP.2012.2183771.

[8] Laska, J., Kirolos, S., Duarte, M., Ragheb, T., Baraniuk, R. and Massoud, Y., Theory and implementation of an analog-to-information converter using random demodulation, Proceedings of IEEE International Symposium on Circuits and Systems, pp. 1959-1962, 2007. DOI: 10.1109/ISCAS.2007.378360.

[9] Sun, H., Chiu, W.Y., Jiang, J., Nallanathan, A. and Poor, H.V., Wideband spectrum sensing with Sub-Nyquist sampling in cognitive radios. IEEE Transactions on Signal Processing, 60(11), pp. 6068-6073, 2012. DOI: 10.1109/TSP.2012.2212892.

[10] Mishali, M. and Eldar, Y.C., From theory to practice: Sub-nyquist sampling of sparse wideband analog signals. IEEE Journal of Selected Topics in Signal Processing, 4(2), pp. 375-391, 2010. DOI: 10.1109/JSTSP.2010.2042414.

[11] Yen, C.-P., Tsai, Y. and Wang, X., Wideband spectrum sensing based on Sub-Nyquist sampling. IEEE Transactions on Signal Processing, 61(12), pp. 3028-3040, 2013. DOI: 10.1109/TSP.2013.2251342.

[12] Tsui, J.B., Digital techniques for wideband receivers, 2nd Edition. Raleigh: Scitech, 2004. DOI: 10.1049/SBRA005E.

[13] Donoho, D.L., Compressed sensing. IEEE Transactions on Information Theory, 52(4), pp. 1289-1306, 2006. DOI: 10.1109/TIT.2006.871582.

[14] Lobato-Polo, A.P., Ruiz-Coral, R.H., Quiroga-Sepúlveda, J.A. and Recio-Vélez, A.L., Sparse signal recovery using orthogonal matching pursuit (OMP). Ingeniería e Investigación, 29(2), pp. 112-118, 2009.

[15] Hassanieh, H., Indyk, P., Katabi, D. and Price, E., Simple and practical algorithm for sparse fourier transform, Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 1183-1194, 2012. DOI: 10.1137/1.9781611973099.93.
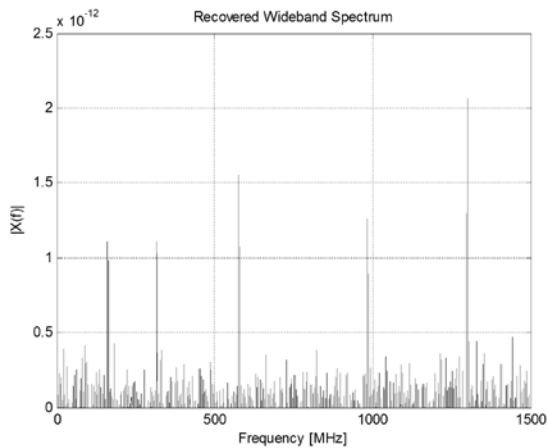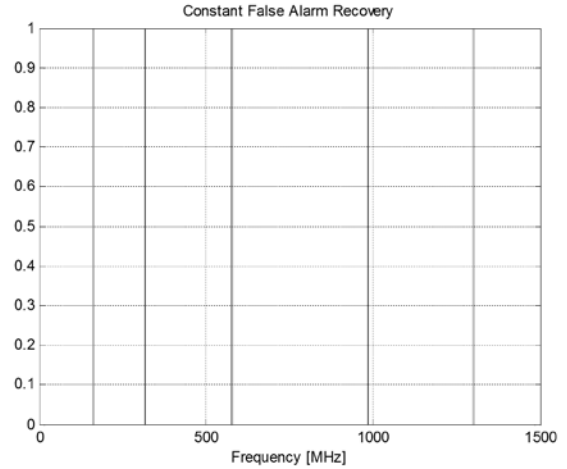
[16] Hassanieh, H., Indyk, P., Katabi, D. and Price, E., Nearly optimal sparse fourier transform, Proceedings of the 44th symposium on Theory of Computing (STOC), pp. 563-578, 2012. DOI: 10.1145/2213977.2214029.

[17] Hassanieh, H., Shi, L., Abari, O., Hamed, E. and Katabi, D., Ghz-wide sensing and decoding using the sparse fourier transform, Proceedings of IEEE INFOCOM, pp. 2256-2264, 2014. DOI: 10.1109/INFOCOM.2014.6848169.

[18] Indyk , P., Kapralov, M. and Price, E., (Nearly) Sample-optimal sparse fourier transform, Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 480-499, 2014. DOI: 10.1137/1.9781611973402.36.

[19] Gilbert, A.C., Strauss, M.J. and Tropp, J.A., A tutorial on fast fourier sampling. IEEE Signal Processing Magazine, 25(2), pp. 57-66, 2008. DOI: 10.1109/MSP.2007.915000.

[20] Gilbert, A.C., Muthukrishnan, S. and Strauss, M.J., Improved time bounds for near-optimal sparse fourier representations, Proceedings of SPIE Wavelets XI, pp. 1-15, 2005. DOI: 10.1117/12.615931.

[21] Dutt, A. and Rokhlin, V., Fast fourier transforms for nonequispaced data, ii. Applied and Computational Harmonic Analysis, 2(1), pp. 85-100, 1995. DOI: 10.1006/acha.1995.1007.

[22] Tanton, J., Encyclopedia of Mathematics. New York: Facts on File, 2005.

[23] Abramowitz, M. and Stegun, I., Handbook of mathematical functions with formulas, graphs, and mathematical tables, 10th printing. Washington, D.C.: Dover, 1972. DOI: 10.1063/1.3047921.

[24] Gilbert, A.C., Li, Y., Porat, E. and Strauss, M.J., Approximate sparse recovery: Optimizing time and measurements, Proceedings of the 42nd ACM symposium on Theory of computing, pp. 475-484, 2010. DOI: 10.1145/1806689.1806755.

A. López-Parrado, completed his BSc. Eng. in Electronics Engineering in 2002 at Universidad del Quindío, Armenia, Colombia, and his MSc. degree in Electronics in 2009 at Universidad del Valle, Cali, Colombia. He is currently a PhD. candidate in Electrical and Electronics Engineering at Universidad del Valle, Cali, Colombia. His research interests are the FPGA design of complex digital systems, design of baseband processors, DSP complex functions, embedded systems, and compressive sensing. He is an assistant professor in the Electronics Engineering Program at Universidad del Quindío, Colombia.
ORCID: 0000-0002-0274-6901

J. Velasco-Medina, completed his BSc. Eng. in Electrical Engineering in 1985 at Universidad del Valle, Cali, Colombia, his MSc. degree in Microelectronics 1995 at Universite De Grenoble I (Scientifique Et Medicale - Joseph Fourier), and his PhD. degree in Microelectronics in 1999 at Institut National Polytechnique De Grenoble, France. His research interests are the FPGA design of complex digital systems, design of baseband processors, Cryptosystems, DSP complex functions, DNA processor, bionanomachines, bionanosensors, biological systems modeling, and Citocomputation. He is a titular professor in the Electrical and Electronics Engineering School of Universidad del Valle, Cali, Colombia.
ORCID: 0000-0003-4091-1055
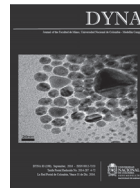
Transient analysis of mixed wind parks with different turbine types

Helleson J. B. da Silva [a] & Carolina M. Affonso [a]

[a] Faculty of Electrical Engineering, Federal University of Para, Pará, Brasil. helleson_jorthan@hotmail.com, carolina@ufpa.br

Abstract

The variable-speed wind turbine is widely used in new wind parks and became the dominant type of wind turbine due to its technological advantages. However, in the case of already installed wind parks based on fixed speed wind turbines, the immediate migration from one technology to another becomes economically impracticable; the process of gradual modernization is more suitable. In this sense, this paper presents a detailed analysis of a mixed wind park containing both doubly fed and squirrel cage induction generators. It considers different rates of capacity for each technology and implementing the gradual replacement of some fixed speed units by variable speed units. The study addresses aspects of voltage stability, fault analysis, wind speed variation and critical fault clearance time. The network used to perform simulations is a real system from Brazil. The results show that DFIG improves the wind park behavior in many aspects and can be a cost-effective solution.

Keywords: wind generation; squirrel cage induction generator; doubly fed induction generator; stability.

Análisis de transitorios de parques eólicos mixtos con diferentes tipos de turbinas

Resumen

La turbina eólica de viento variable se volvió el tipo dominante de turbina en nuevos parques eólicos debido a sus ventajas tecnológicas. En caso de parques eólicos ya en funcionamiento con turbinas basadas en viento constante, la migración de una tecnología hacia la otra se vuelve económicamente inviable, lo que requiere su modernización gradual. En ese sentido, este artículo presenta un análisis de un parque eólico mixto de generadores de inducción con jaula de ardilla y doblemente alimentados, simulando el reemplazo gradual de algunas unidades de velocidad fija por otras de velocidad variable. El estudio abarca aspectos de estabilidad de tensión, análisis de faltas, variación de la velocidad del viento y tiempo crítico de eliminación de faltas. El sistema de potencia empleado es un sistema real brasileño y los resultados muestran que generadores de inducción doblemente alimentados mejora el desempeño del parque eólico en distintos aspectos, siendo una solución económicamente efectiva.

Palabras clave: generación eólica; generador de inducción de jaula de ardilla; generador de inducción doblemente alimentado; estabilidad.

1. Introduction

Wind generation is gaining increasing interest around the world today since it can provide a clean, cheap and renewable source of energy for many people. In Brazil, wind is the fastest growing source of power generation with an installed capacity of 6 GW at the end of 2014, representing 4.5% of the country's total installed capacity [1].

There are two main popular wind power technologies in the world: fixed speed turbines and variable speed turbines. The fixed speed turbines use squirrel cage induction generators (SCIG) and have variations in output power, mechanical stress and limited power quality control. Also, they always draw reactive power from the grid and it is necessary to use capacitor banks. The advantage of this scheme is the low cost and robust structure [2].

The variable speed turbines are modern systems and can use doubly fed induction generators (DFIG), permanent magnet synchronous generators (PMSG) and others [2]. These technologies use power electronic converters and offer extensive controllability of both active and reactive power. They show less fluctuation in output power and can meet most grid code requirements. The disadvantage is their high costs [3]. The doubly fed induction generator (DFIG) is currently one of the most common wind turbine technologies installed in wind farm projects.

Recently, due to their technological advantages, the variable-speed wind turbines began to be widely used in new wind parks

and became the dominant type used among installed wind turbines. However, the SCIGs are still used in many existing wind parks because of their lower price and higher operation reliability. Therefore, in the case of already installed wind parks, the immediate migration from one technology to another becomes economically impractical. The process of gradual modernization is more suitable.

Most papers study the operation characteristics of each wind park concept separately [4,5,6], and there is little research on the coordinated operation characteristics of wind parks containing both DFIGs and SCIGs. The authors in reference [7] investigate the coordinated operation of a wind park with different wind turbine concepts (fixed and variable speed) in order to guarantee the operational set points of active and reactive power specified by the Spanish transmission system operator. In [8], a coordinated control strategy is proposed for a wind park composed by DFIG and fixed speed induction generators to improve system stability during and after faults. Reference [9] analyzes the transient performance of a wind park with DFIG and SCIG with different rates of capacity; it considers DFIG operation with and without voltage control on rotor side converter. In [10], an investigation of the flicker emission is presented, considering a wind park with different types of wind turbines, fixed and variable speeds that operate under a variety of wind conditions and network characteristics. The results show the replacement of some fixed speed wind turbines with DFIG wind turbines can reduce the flicker emission in the system. Reference [11] investigates the control and operation of doubly fed induction generator and fixed-speed induction generator based wind farms under unbalanced grid conditions. The simulations are implemented in Matlab/Simulink.

Given what has already been investigated, we deem it necessary to undertake research that pays more attention to the operation characteristics of wind parks containing fixed speed turbines and variable speed turbines and how they affect the grid. In this context, this paper presents a detailed analysis of a mixed wind park containing both DFIGs and SCIGs. It follows a trend where an existing SCIG-based wind farm is being extended with DFIG turbines by the gradual replacement of some SCIG units by DFIG units. Different rates of capacity of DFIG and SCIG are considered, and the study addresses aspects of voltage stability, fault analysis, wind speed variation and critical fault clearance time. The simulations were conducted using the ANAREDE and ANATEM software that were developed by CEPEL [12,13]. The network used to perform the simulations is a real system in Brazil.

2. Models

This section presents the wind turbine models and induction generators that are used in this study. Two technologies are considered: squirrel cage induction generator and doubly fed induction generator. This section presents the turbines' technical details.

2.1. Wind turbine model

The wind turbine converts wind flow energy into mechanical energy. Depending on the wind speed, the turbine power output varies. The power extracted from a wind turbine (P_w) is proportional to the cube of the wind speed, and
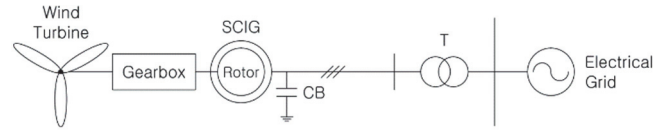
Figure 1. Squirrel cage induction generator.
Source: The authors.

can be represented by the mathematical expression as follows:

$$P_W = \frac{\rho}{2}.A.C_p(\lambda,\beta).v_w^{\ 3} \quad , \qquad (1)$$

where ρ is the air density, A is the area swept by the rotor blades, vw is the wind speed, Cp is the power coefficient, β is the blade pitch angle, and λ is the tip speed ratio [14]. Tip speed ratio can be expressed as the ratio between the speed of a blade tip (ωT) and the wind speed (vw) as show in eq. (2), which considers the turbine radius R.

$$\lambda = \frac{R.\omega_T}{v_w} \qquad (2)$$

2.2. Squirrel Cage Induction Generator

Squirrel cage induction generators (SCIG) are fixed-speed wind turbines directly connected to the grid. They have a soft-starter and a capacitor bank to reduce reactive power compensation. Fig. 1 shows the general structure of a model of a constant-speed wind turbine. For power flow studies, the squirrel cage induction generator is modeled as a PQ bus. The induction generator model, which neglects the stator transients, is the same that is used in [15].

2.3. Doubly fed induction generator

Doubly fed induction generator (DFIG) is a variable-speed wind turbine that can achieve maximum aerodynamic efficiency over a wide range of wind speeds. In this configuration the stator is directly connected to the grid and the rotor is fed from a back-to-back AC/DC/AC converter set, as is shown in Fig. 2 [2]. The converter circuit allows the production or consumption of reactive power, making it different from the SCIG, which can only consume reactive power. Since the DFIG has reactive power controllability, it can operate on power factor control mode (PQ representation in load flow studies) or voltage control mode (PV representation in load flow studies). In this paper, voltage control mode was adopted, and DFIG was represented by a simplified third-order model in which electromagnetic transients of the stator are neglected [16].

The reactive power capability of DFIG changes with wind speed variation, and can be determined based on:

$$Q_s^{\max} = \sqrt{S_s^2 - P_s^2} \qquad (3)$$
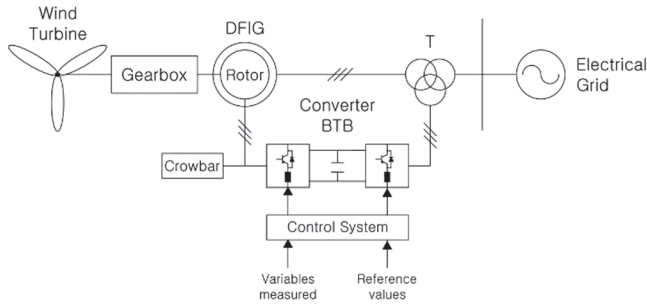
Figure 2. Doubly fed induction generator.
Source: The authors.

where Q_S^{max} is the stator reactive power limit, S_s is the stator apparent nominal power and P_s is the active power injected at the stator. A more detailed representation can be found in [5]. Then, during low wind speed conditions, which translate to lower values of active power, the reactive power injection limits are bigger. On the other hand, during high wind speed conditions, which translate to higher values of active power, the reactive power injection limits are lower.

3. Test system

The wind park used in this paper is shown in Fig. 3. This project is under construction in the city of Santana do Livramento, which is located in the Rio Grande do Sul state, Brazil [17]. This wind farm will be integrated into the Brazilian main system by two 230kV transmission lines that connect the cities of Alegrete and Santana do Livramento [18].

The wind system is compounded by DFIGs and SCIGs, which operate in parallel. The equivalent induction generators system has a 67MW capacity with 67 1 MW rated power wind turbines. The equivalent variable speed system has a 53.55MW capacity with 63 wind turbines based on a DFIG of 850kW rated power. All SCIGs were aggregated into a single equivalent system, and the incoming wind speed is considered to be identical on all the wind turbines. The DFIGs adopt the same equivalent method as that of SCIGs.

The loads are connected at substation Alegrete and Livramento, 20.0 MW/19.4 MVAr and 67.6 MW/19.3 MVAr respectively, and they are both modeled as constant power.
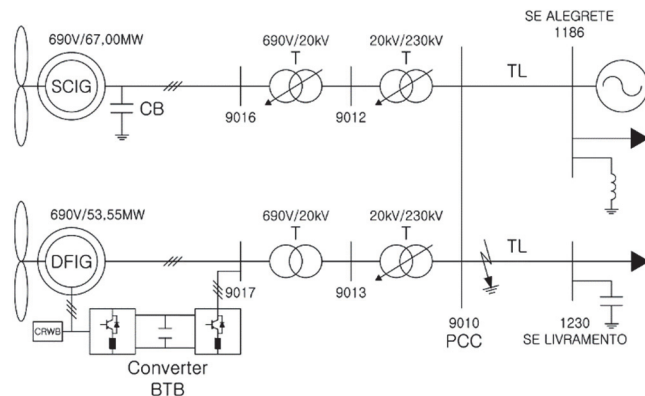
Figure 3. Test system diagram.
Source: The authors.

4. Simulation results

In order to analyze the behavior of mixed wind parks with different turbine types, four studies were conducted in this paper: voltage stability, fault analysis, wind speed variation and critical fault clearance time.

In most studies, five scenarios were analyzed to simulate the gradual modernization of the wind park, that is, the gradual replacement of SCIG wind turbines for DFIG wind turbines. The penetration level of wind generation SCIG and DFIG were modified simultaneously in order to maintain the active power –fixed at 50 MW– at the point of common coupling (PCC). The group of SCIG machines was connected to bus 9016 and the group of DFIG machines was connected to bus 9017. The initial wind speed was 10.14 m/s for both technology groups.

The results will now be presented, and the five scenarios considered were the following:
- Scenario 1: SCIG = 45.3 MW / DFIG = 4.7 MW;
- Scenario 2: SCIG = 35.2 MW / DFIG = 14.8 MW;
- Scenario 3: SCIG = 25.0 MW / DFIG = 25.0 MW;
- Scenario 4: SCIG = 14.9 MW / DFIG = 35.1 MW;
- Scenario 5: SCIG = 4.7 MW / DFIG = 45.3 MW.

4.1. Voltage stability

Voltage stability margin (VSM) is the percentage difference between the system maximum loading point and the forecasted load, as is shown in eq. (4). It can be expressed in p.u., as a percentage, or in MW. This index is the most widely used indicator of system voltage security and is determined by using PV curve methods [19]. The PV curve can be obtained by using the continuation power flow method [20]. For each load increase, a load flow problem is solved, and the set of obtained equilibrium points defines the PV curve. The maximum loading point (critical point) corresponds to the system bifurcation point from the PV curve.

$$VSM\% = \left(\frac{P_{max} - P_0}{P_0}\right) \times 100 \quad , \quad (4)$$

where P_0 is the system initial load and P_{max} is the maximum load.

Fig. 4 shows the PV curve for all scenarios analyzed. The results show that the system voltage stability margin increases as the DFIG penetration level also increases. This can be explained by the fact that voltage instability is mainly associated with reactive power support, and DFIG technology can supply reactive power to the system when operating in PV mode. It is important to mention that DFIG generators' maximum and minimum reactive power limits were modified for each scenario analyzed since the active power was also modified.

4.2. Fault analysis

A three-phase short-circuit fault was applied to 20% of the transmission line, which connects bus 9010 to bus 1230. The fault is applied at t = 2 seconds, and after 200 ms the fault is cleared. The simulation time is 10 seconds. In all simulations, the DFIG crowbar protection was activated.
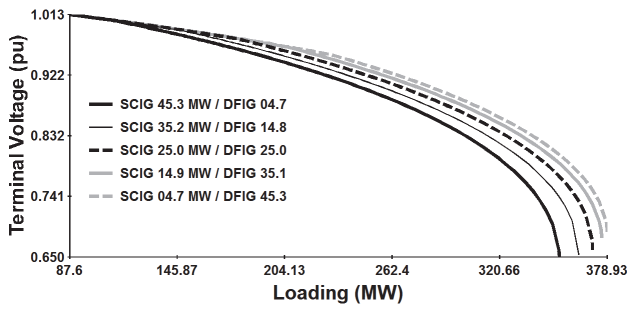
Figure 4. PV Curve.
Source: The authors.

increases as the capacity of DFIG increases due to the reversal of reactive power flow. Moreover, the system recovery time is smaller due to DFIG reactive power support, which quickly reestablishes the electromagnetic torque. Also, as the capacity rate of DFIG increases, the reactive power absorbed by SCIG is smaller due to the DFIG's capacity to supply reactive power to the grid.

Fig. 6 shows the results obtained for the DFIG group that was connected at bus 9017. The fault also causes voltage sag, which is more severe due to the actuation of the crowbar protection scheme. This is responsible for opening the circuit between the rotor windings and rotor side converter, and interrupting the terminal voltage control. When the fault is cleared, the connection between the rotor side converter and the generator rotor circuit is restored, and terminal voltage control action is undertaken with immediate effect, which translated to the significant recovery of terminal voltage.

Figure 5. Simulation results for the SCIG group (fault analysis).
Source: The authors.

Fig. 5 shows the results obtained by the SCIG group that was connected at bus 9016. During the fault, the terminal voltage drops very quickly. The voltage drop severity
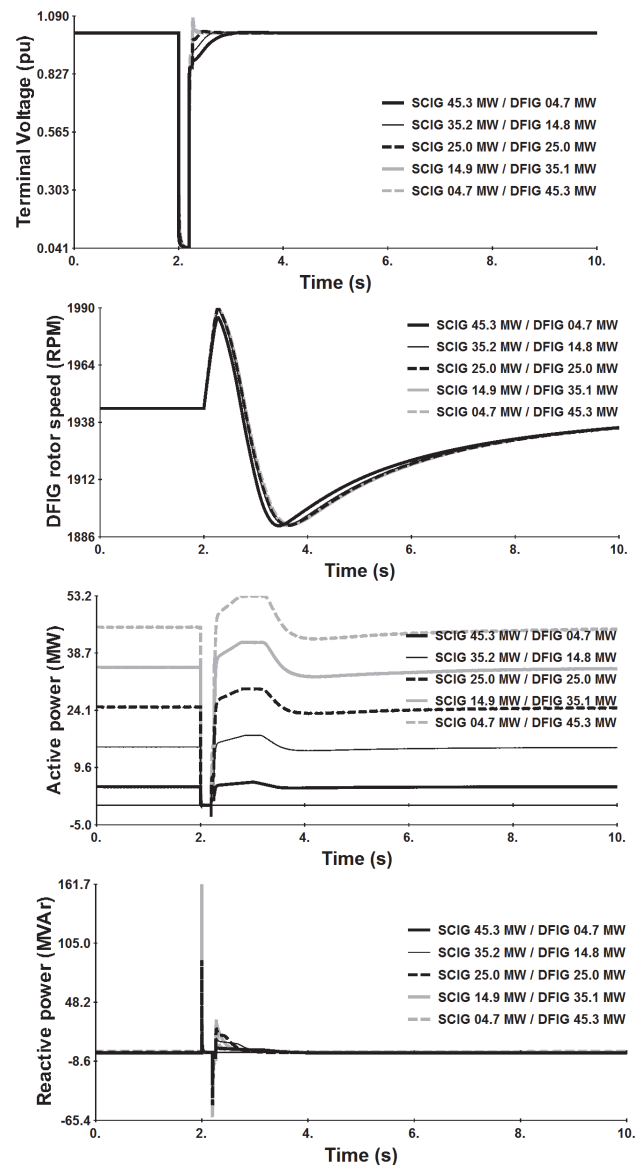
Figure 6. Simulation results for the DFIG group (fault analysis).
Source: The authors.
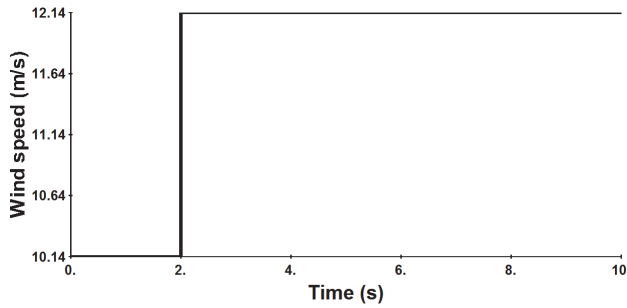
Figure 7. Wind speed variation.
Source: The authors.

Table 1.
Critical Clearance Time.

	Scenario 1 (SCIG / SCIG)	Scenario 2 (SCIG / DFIG)
CCT	700ms	1000ms

Source: The authors.

4.3. Wind speed variation

In all scenarios, the system was submitted to the same wind speed increase as is shown in Fig. 7: a positive step function of 2m/s applied at t = 2 seconds. The simulation time is 10 seconds. The crowbar protection of DFIG was not activated in all simulations. We considered the five scenarios previously presented with different capacity rates of DFIG and SCIG.

Fig. 8 shows the results obtained for the SCIG group that was connected at bus 9016. When the DFIG capacity rate increases and the SCIG decreases, the terminal voltage recovers more rapidly after the fault. Moreover, with the capacity rate of DFIG increasing, the recovery time and magnitudes of the oscillations are smaller. This is mainly because SCIG absorbs reactive power and DFIG can provide reactive power to the grid. As the simulation shows, the modernization of the wind park minimizes the undesirable effect of fixed speed wind turbines, which is typical oscillatory behavior. It is important to mention that all scenarios start with different initial conditions of reactive power consumption due to the progressive reduction on the number of SCIG generator units. Also, after the wind speed variation, the voltage does not return to its initial value as a result of SCIG capacitor banks.

Fig. 9 shows the results obtained for the DFIG group that was connected at bus 9017. The larger the DFIG rate of capacity, the more rapidly the voltage recovers after the fault. Also, in all scenarios, the voltage returned to its initial value due to DFIG capacity injecting reactive power that controlled terminal voltage and improved transient operation characteristics.

As wind speed increases, the active power produced by the wind generator also increases. Therefore, the DFIG reactive power limits decreases, as is shown in Fig. 10. It is imp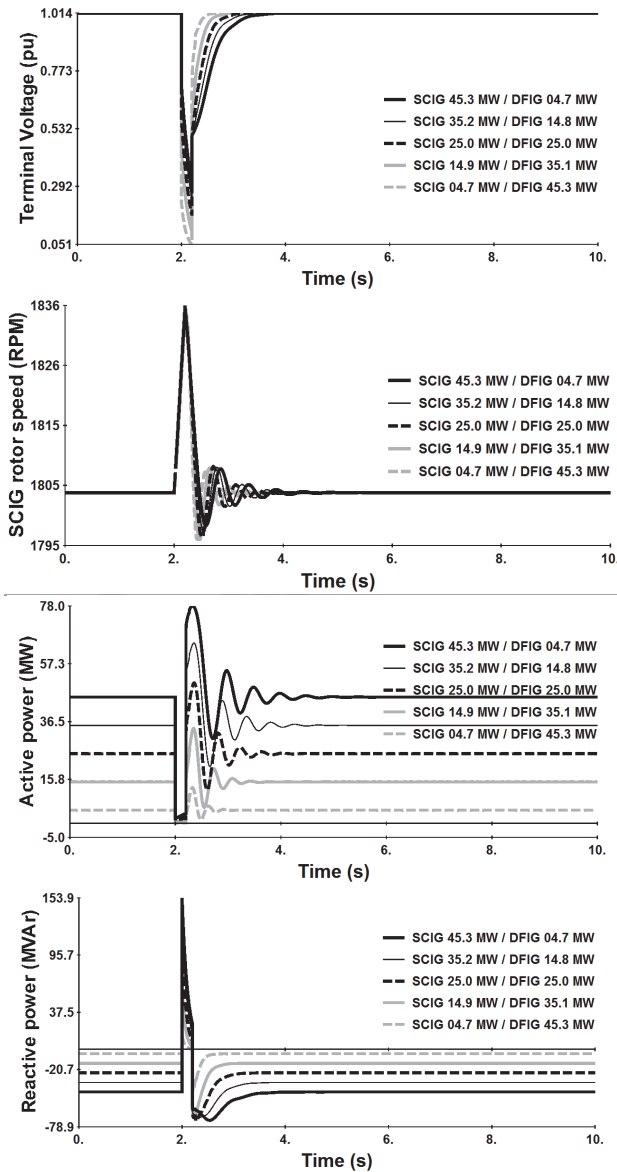ortant to mention that as DFIG penetration level increases, as expected, the reactive power available to provide reactive power support to the grid also increases. Also, wind park modernization provides an improvement in power system operation at the point of common coupling (PCC), as is shown in Fig. 11.

4.4. Critical fault clearance time

During severe faults, the electromagnetic torque of induction generators decreases when the mechanical torque

related to wind speed is considered constant. In this case, the rotor shaft accelerated. When the fault is eliminated before the rotor speed reaches critical speed (when the electromagnetic torque corresponds to the same amount as before the disturbance), the rotor speed starts to decrease and the generator response is stable. Otherwise, when the fault is eliminated after the rotor speed reaches the critical speed, the rotor speed continues to increase and the generator will become unstable [21]. The time duration starting from the fault time until the critical slip point is the critical clearance time (CCT). The CCT is defined as the maximal fault duration for which the system remains transiently stable. The concept of the critical speed of induction generators was first proposed in [22] and further analyzed in [23]. During major faults in SCIG and DFIG, critical clearance time (CTT) is an important issue to consider.

The critical clearance time is conventionally evaluated using time-domain simulations, which involve several trials and repetitive runs of transient programs [24]. In this paper, the critical fault clearance time was calculated by gradually increasing the fault clearance time in steps of 5 ms. Two scenarios were then analyzed:
• Scenario 1: SCIG = 25.0 MW / SCIG = 25.0 MW;
• Scenario 2: SCIG = 25.0 MW / DFIG = 25.0 MW.

The first scenario considers that wind park is only based on SCIG technology. The second scenario considers that wind park is based on SCIG and DFIG technologies. The penetration level of wind generation SCIG and DFIG was fixed –at 50 MW– at the point of common coupling (PCC) for both scenarios. The group of SCIG machines, which remained installed in the wind park, was connected at bus 9016.

The results are presented in Table 1, and they reveal that the presence of DFIG technology in the wind park increases the critical clearance time, which is positive in terms of system stability.

Also, two simulations were performed in order to confirm the results. First, a network fault with the same duration of CCT (T_f = CCT) was considered. Then, a network fault with duration larger than CTT (T_f >> CCT) was assumed and the time-domain simulation was run once again. The results are presented in Fig. 12: they show scenario 1 (wind park composed by SCIG only) and scenario 2 (wind park composed by SCIG and DFIG). This figure reveals that, in the first scenario, when the short circuit is eliminated in or faster than 700ms, the wind generator does not lose its stability. Otherwise, when the fault clearance time is longer than 700ms, the wind generator becomes unstable. The same behavior is observed in scenario 2; this confirms that the critical fault clearance time was met.
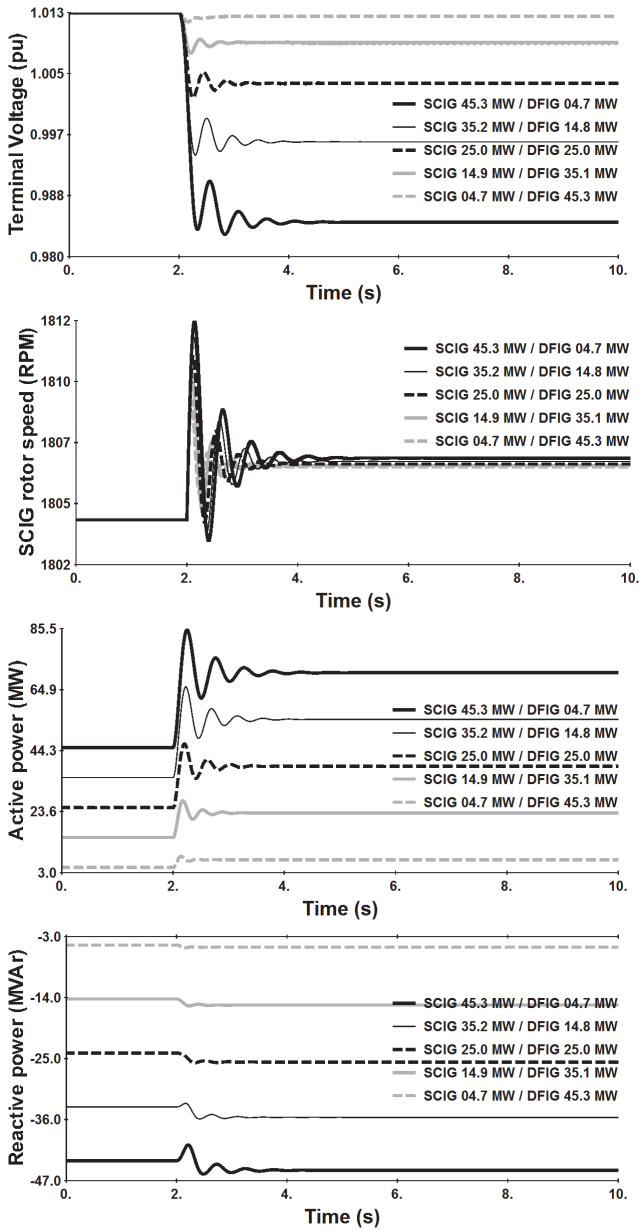
Figure 8. Simulation results for the SCIG group (wind speed variation).
Source: The authors.
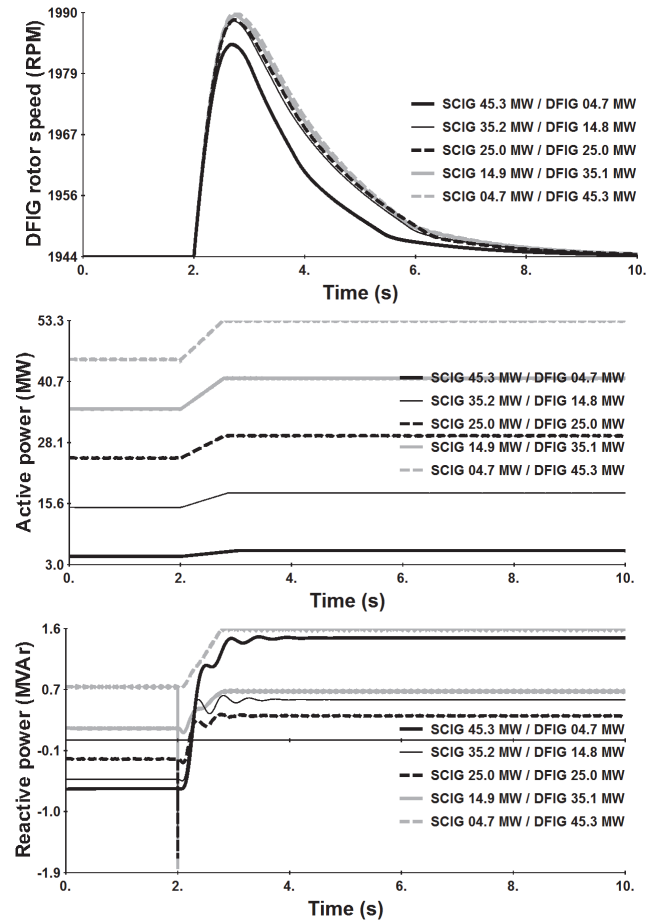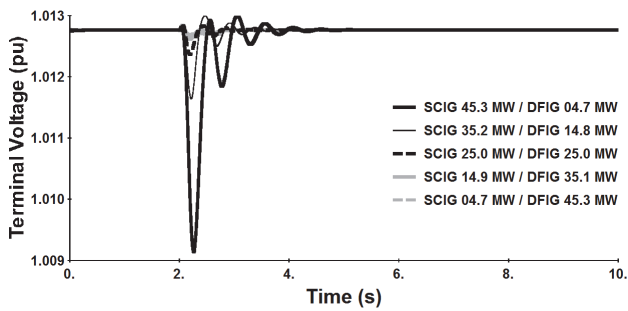
Figure 9. Simulation results for the DFIG group (wind speed variation).
Source: The authors.
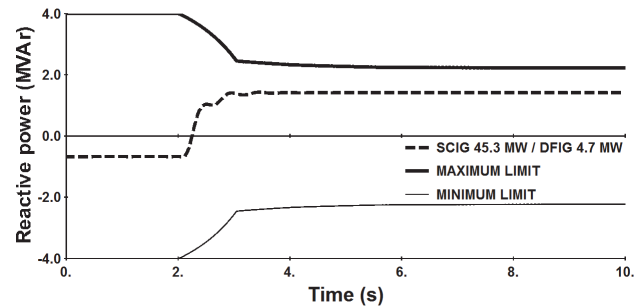
Figure 10. DFIG reactive power limits (wind speed variation).
Source: The authors.
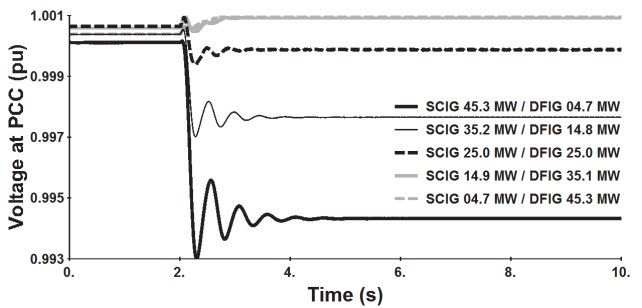
Figure 11. Voltage at the point of common coupling (PCC).
Source: The authors.
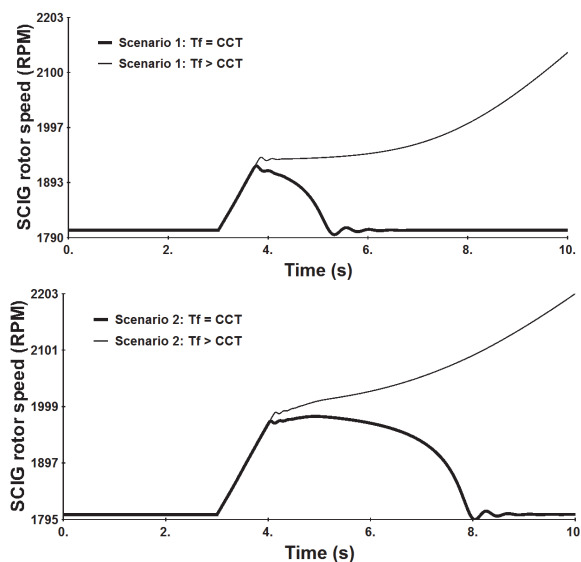
Figure 12. SCIG rotor speed for different fault clearing times.
Source: The authors.

5. Conclusions

This paper presents a behavioral analysis of a mixed wind park that is composed of doubly fed induction generators and squirrel cage induction generators. It considered different rates of capacity for each technology and simulated the gradual replacement of some SCIG units by DFIG units. The study addresses aspects of voltage stability, fault analysis, wind speed variation and critical fault clearance time. The network used to perform simulations is a real system in Brazil.

The results show that the parallel operation of DFIG and SCIG can improve voltage stability, increase fault clearance time, and improve the transient operation characteristics of the wind park and also the point of common coupling. According to the results obtained, gradual wind park modernization could be a cost-effective solution.

References

[1] Brasil Energy. A fast-growing market – Power special report. [Online]. February 2015. Available at: http://brasilenergy.editorabrasilenergia.com/.

[2] Jenkins N., Allan R., Crossley P., Kirschen D. and Strabac G., Embedded generation. The Institution of Electrical Engineers, London, United Kingdom, 2000. DOI: 10.1049/PBPO031E

[3] Alvarez J.M.G., Schwickardt G. and Targarona J.C.G., An overview of wind energy taking into consideration several important issues including an analysis of regulatory requirements for the connection of wind generation into the power system, DYNA, 79(172), pp. 108-117, 2012.

[4] Alvarez J.M.G. and Targarona J.C.G., Wind generation using different generators considering their impact on power system, DYNA, 78(169), pp. 95-104, 2011.

[5] Londero R.R., Affonso C.M. and Vieira J.P.A., Long-term voltage stability analysis of variable speed wind generators, IEEE Transactions on Power Systems, 30, pp. 439-447, 2015. DOI: 10.1109/TPWRS.2014.2322258

[6] Xu, L.Y.Z., Østergaard, J., Dong, Z.Y, Wong, K.P. and Ma, X., Oscillatory stability and eigenvalue sensitivity analysis of A DFIG wind turbine system. IEEE Transactions on Energy Conversion, 26, pp. 328-339, 2011. DOI: 10.1109/TEC.2010.2091130

[7] Rodríguez-Amenedo J.L., Arnaltes S. and Rodríguez M.A., Operation and coordinated control of fixed and variable speed wind farms. Renewable Energy, 33, pp. 406-414, 2008. DOI: 10.1016/j.renene.2007.03.003

[8] Foster, S., Xu, L. and Fox, B., Coordinated control and operation of DFIG and FSIG based wind farms, Proceedings of IEEE Power Tech, 2007, pp. 522-527. DOI: 10.1109/pct.2007.4538371

[9] Lin, L., Zhang, Y. and Yang, Y., Transient characteristics of the grid-connected wind power farm with DFIGs and SCIGs, Proceedings of Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies - DRPT, 2008, pp. 2676-2681. DOI: 10.1109/drpt.2008.4523864

[10] Gogdare, A.N., Doroudi, A. and Ghaseminejad, M., A new method to mitigate voltage fluctuation of a fixed speed wind farm using DFIG wind turbine. Proceedings of 17th Conference on Electrical Power Distribution Networks (EPDC), 2012, pp. 1-6.

[11] Wang, Y. and Xu, L., Coordinated control of DFIG and FSIG-based wind farms under unbalanced grid conditions. IEEE Transactions on Power Delivery, 25, pp. 367-377, 2010. DOI: 10.1109/TPWRD.2009.2033966

[12] CEPEL, Centro de Pesquisas de Energia Elétrica. ANAREDE: Programa de Análise de Redes. Manual Guide, 07-08 Version, August 1999.

[13] CEPEL, Centro de Pesquisas de Energia Elétrica. ANATEM: Análise de Transitórios Eletromecânicos. Manual Guide, 09 Version, December 2002.

[14] Slootweg, J.G., Polinder, H. and Kling, W.L., Dynamic modeling of a wind turbine with doubly fed induction generator. Proceedings of IEEE PES Summer Meeting, 2001, pp. 644-649.

[15] Rangel, D.R., Ferraz, J.C.R., Gomes Jr., S. e Henrique, R.M., Modelagem de geração eólica em programas de fluxo de potência e de análise de estabilidade eletromecânica. Proceedings of Seminário Nacional de Produção e Transmissão de Energia Elétrica - XVIII SNPTEE, 2005, pp. 1-8.

[16] Anaya-Lara, O., Jenkins, N., Ekanayake, J., Cartwright, P. and Hughes, M., Wind generation, modeling and control. John Wiley & Sons, Ltd., 2009.

[17] Soares, J., Jornal $$ - Parque eólico deve receber investimento de R$ 4,8 bi.. Porto Alegre, RS, [Online]. 2013. Available at: http://www.valor.com.br/brasil/3017264/parque-eolico-deve-receber-investimento-de-r-48-bi

[18] Pavan, A.P., Impacto Da geração eólica no desempenho dinâmico do sistema elétrico. MSc. Thesis, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil, 2006.

[19] Taylor, C., Power system voltage stability. New York: McGraw-Hill Inc., 1994.

[20] Kundur, P., Power system stability and control. New York: McGraw-Hill, 1994.

[21] Grilo, A.P., Mota, A.A., Mota, L.T.M. and Freitas, W., An analytical method for analysis of large-disturbance stability of induction generators. IEEE Transactions on Power Systems, 22, pp. 1861-1869, 2007. DOI: 10.1109/TPWRS.2007.907346

[22] Akhmatov, V., Knudsen, H., Bruntt, M., Nielsen, A.H., Pedersen, J.K. and Poulsen, N.K., A dynamic stability limit of grid-connected induction generators. Proceedings of International Conference on Power and Energy - IASTED, Spain, 2000, pp. 235-244.

[23] Akhmatov, V., Analysis of dynamic behavior of electric power systems with large amount of wind power. PhD. dissertation, Tech. Univ. Denmark, Lyngby, Denmark, 2003.

[24] Salman, S.K. and Teo A.L.J., Investigation into the estimation of the critical clearing time of a grid connected wind power based embedded generator, Proceedings of IEEE PES Transmission and Distribution Conference and Exhibition, 2002, pp. 975-980. DOI: 10.1109/TDC.2002.1177609

H.J.B. da Silva, received his BSc. in Technical Industrial from IFPA, Brazil, in 2011, and his MSc. degree from the Federal University of Para, Brazil, in 2014. At present, he is working towards a BSc. degree in Control and Automation Engineering at IFPA, Brazil. His research interests include wind power generation and power system stability.
ORCID: 0000-0002-8260-1381

C. de M. Affonso, received her BSc. in Electrical Engineering from the Federal University of Para, Brazil, in 1998, her MSc. degree from the Federal University of Santa Catarina, Brazil, in 2000 and her PhD. degree in Electrical Engineering from State University of Campinas, Brazil, in 2004. Currently, she is a professor of electrical engineering at the Federal University of Para, Brazil. Her research interests are in power system planning and stability and distributed generation.
ORCID: 0000-0002-2241-1711

New robust capability ratios approaches for quality control

Salvador Naya [a], Andrés Devia-Rivera [b], Javier Tarrío-Saavedra [a] & Miguel Flores [c]

[a] Escuela Politécnica Superior, Departamento de Matemáticas. Grupo MODES, Universidad de A Coruña, España. salva@udc.es, jtarrio@udc.es
[b] Departamento de Marketing Comercial, Avon Cosmetics S.A.U., Madrid, España. andres.devia@avon.com
[c] Facultad de Ciencias, Departamento de Matemáticas. Escuela Politécnica Nacional, Quito, Ecuador. miguel.flores@epn.edu.ec

Abstract
Robustness of process capability measurements is a very important matter in statistical quality control. In this paper, two new classes of capability measurements are studied as robust mechanisms to detect the influence of factors that may cause large departures from the process' engineering specifications. The behavior of the new indices was analyzed by comparing their performance to other capability measures that have been widely studied in literature. The paper aims to investigate the robustness of the new capability ratios under the presence of outliers and a lack of normality. For this purpose, bootstrap techniques were applied to detect the true potential capability of a process via statistical inference methods. The accuracies of the proposed indices are discussed by means of numerical results from a real data example.

Keywords: Process capability ratio, robust statistics, bootstrap hypothesis testing, bootstrap confidence intervals.

Nueva propuesta de índices de capacidad robustos para el control de la calidad

Resumen
En este trabajo, se proponen dos nuevos índices de capacidad robustos para detectar la influencia de los factores que pueden causar grandes desviaciones de las especificaciones técnicas del proceso. El comportamiento de estos nuevos índices se analizó mediante la comparación de su rendimiento con respecto a otras medidas de capacidad ampliamente estudiados en la literatura. El trabajo tiene como objetivo investigar la robustez de estos nuevos índices de capacidad bajo la presencia de valores extremos y de falta de normalidad. Para este propósito, se aplicaron técnicas de remuestreo Bootstrap para detectar la verdadera capacidad potencial de un proceso a través de los métodos de inferencia estadística. La precisión de los índices propuestos es discutida por medio de resultados numéricos con un ejemplo de datos reales.

Palabras clave: índices de capacidad; estadística robusta, contraste de hipótesis bootstrap, intervalos de confianza bootstrap.

1. Introduction

A process capability ratio (PCR) is a numerical score that helps the manufacturers to know whether the output of a process meets the engineering specifications. Large values of the ratio indicate that the current process is capable of producing items that meet or exceed customer requirements. Unfortunately, traditional assumptions of the data, such as normality or independence are often violated in many real situations. A common scenario, in which the assumptions of

normality or independent and identically distributed data (i.i.d.) does not hold, is, for example, when the data are autocorrelated or when they belong to non-centered and skewed distributions. Specifically, if the assumption of normality is violated, it could then be very difficult, or even impossible, to obtain closed expressions for the probability distribution of the PCR estimator. This means that, in many cases, it is not possible to derive exact confidence intervals for the estimates of process capability. As a consequence of this, capability estimates may be far away from the true

parameters of interest, and manufacturers could, therefore, be making the wrong decisions about the quality management of the process.

Many authors have studied different estimators for process capability under various distributional settings. Recent advances in inferential analysis applied to quality control techniques have motivated more theoretical research into the distribution theory of estimated PCR (see, for instance, the works by Chou and Owen [1], Clements [2], Pearn et al. [3], Ebadi and Shahriari [4], Kotz and Johnson [5], and Chien-Wei et al. [6]). The last two presented an exhaustive discussion on a number of capability indices, their sampling properties and practical applications. Moreover, there are some studies that address capacity indexes applied to autocorrelated data, such as Pan et al.'s [7] work that is applied to environmental features. Additionally, in terms of the particular case that deals with non-normal data and processes with unilateral specifications, extensive discussions can be found in the work undertaken by the following authors: Somerville and Montgomery [8], Kotz and Lovelace [9], Shore [10], Tang and Than [11], Chang et al. [12], Pearn and Chen [13], and Kotz and Johnson [14]. Most of the literature devoted to the study of process capability analysis frequently considers four indices, C_p, C_{pk}, C_{pm}, C_{pmk}, which are defined as:

$$C_p = \frac{USL - LSL}{6\sigma} \tag{1}$$

$$C_{pk} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} \tag{2}$$

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - \tau)^2}} \tag{3}$$

$$C_{pmk} = \min\left\{\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - \tau)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - \tau)^2}}\right\} \tag{4}$$

$$= \frac{d - |\mu - m|}{3\sqrt{\sigma^2 + (\mu - \tau)^2}},$$

Where *USL* and *LSL* are the upper and the lower specification limits for the variations in the process, μ is the process mean, σ is the process standard deviation, τ is some target value of interest in the process, d is the half of the spread between the upper and lower engineering specifications, and *m* is the midpoint between the specification limits.

It is well known that indices C_p and C_{pm} are improved when the process data comes from symmetric distributions, e.g. the normal one. These are used to estimate process capability when two-sided tolerance limits are of concern. In the case of one-sided specifications, capability indices C_{pk} and C_{pmk} are preferred to obtain the desired process capability estimates [15, 16]. The C_{pmk} index deals with the departure of the process mean μ from the target value τ. It does this faster than the indices, C_p, C_{pk} and C_{pm} while remaining sensitive to changes in the total deviation of the process (see the

expression in equation (4)). A handicap of most capability measures is related to their efficiency, which strongly depends on the appropriate estimation of the process variability. They are also influenced by the shape of the underlying distribution function that characterizes it (see, for instance, the papers by Heavlin [17], Chou and Owen [1], Pearn et al. [3] and Borges and Ho [18] for a more theoretical discussion).

The present paper is concerned with the estimation of process capability measurement when the data are possibly affected by contamination, hereafter the outliers, which may be an alert from an out-of-control process. A real data example is presented to compare the performance of the new capability ratios relative to the standard indices (1) to (4), that are under various schemes of distributions, sample sizes, and percentages of contamination of data.

The analysis consisted of a double inferential procedure. Bootstrap techniques for statistical inference were used to study the degree of potential capability of processes generated by distributions that have been affected by different outliers in different percentages. Bootstrap techniques for statistical inferences were used to study the degree of potential capability of processes that were generated by distributions affected by the different percentage of outliers. Inferential methods consisted of testing a standard null hypothesis in the context of process capability analysis. Moreover, approximately 95 percent bootstrap confidence intervals were obtained for the indices.

The present study is organized as follows: in section 2, definitions and mathematical formulations to compute the new capability indices are given. In section 3, theoretical basis of the bootstrap approach for statistical inference are introduced in a process capability analysis context. Section 4 is devoted to results obtained by applying the indices and methods described in section 3. Finally, numerical results and concluding remarks are discussed in section 5.

2. Robust process capability ratios

2.1. Robust capability measures for quality process

In a broad, but non theoretical sense, since outliers are data that commonly come from distributions different from the main set of data, it could be thought that the presence of outliers could be evidence that the process is out of statistical control. In that sense, there are a wide range of mathematical methods to deal with the problem of outliers. All these methods are connected to robust statistics [19]. Robust statistics are used in many applications of statistical process control analysis. Abu-Shawiesh and Abdullah [20] studied control limits for control charts by using robust estimates of process parameters (location, scale, shape, etc.). In Grznar et al.'s paper [21], the authors present a routine for outlier detection based on the smoothing methods. Kocherlakota and Kocherlakota [22] discuss different methods to obtain confidence intervals for PCR based on robust estimates under non normal data. Prasad and Bramorski [23] studied the interactions between outliers and correlation structures as unknown sources of variability under the scope of time series. Also, Yeh and Bhattacharya [24] proposed an index

based on the idea of estimating non-conforming proportions. They also discussed a methodology to obtain bootstrap confidence intervals.

In the following subsections, an ordered sample of n independent identically distributed (i.i.d.) random variables (r.v.) with the same distribution of the process X is represented by $X_{1n}, \ldots, X_{1n}$, and a corresponding realization of n items taken from X is denoted by $x_{1n}, \ldots, x_{1n}$. Some robust estimators that will be used in the rest of this paper are the standard deviation, σ, the sample range, R, the median, M, and the first and third quantiles, Q_1 and Q_3.

2.2. Definition of C_{pk} and C_{prk} indices

The introduction of these new process capability estimates is justified since there are systems characterized by the existence of sources of deviations that make large departures from the engineering specifications. We assume that such sources of deviations take place due to the presence of outliers in the data.

The first class of ratio that we introduce, C_{pr}, is based on the idea of the outlier detection criteria that is defined by the boxplot charts. It is expected that C_{pr} will be robust for detecting outliers when it is used to estimate the process capability of two-sided specifications systems. The second class of ratio, denoted by C_{prk}, was developed to estimate process capability for unilateral (one-sided) specifications in a similar way to that of the equations (2) and (4). C_{pr} measure the potential process capability while C_{prk} estimate the real capability of a process.

We define the new indices C_{pr} and C_{prk} as:

$$C_{pr} = \frac{USL - LSL}{\lambda_U - \lambda_L} \tag{5}$$

$$C_{pru} = \frac{USL - M}{\lambda_U - M} \tag{6}$$

$$C_{prl} = \frac{M - LSL}{M - \lambda_L} \tag{7}$$

$$C_{prk} = \min\left\{ \frac{USL - M}{\lambda_U - M}, \frac{M - LSL}{M - \lambda_L} \right\} \tag{8}$$

We stress that λ_U and λ_L are, respectively, two robust measures for the upper and the lower bounds of an in-control process, defined as:

$$\lambda_U = \min\left\{ X_{nn}, L_{upper} \right\} \tag{9}$$

and $\lambda_L = \max\left\{ X_{1n}, L_{lower} \right\}$ (10)

Where $L_{upper} = Q_3 + 1.5 \times IQR$, $L_{lower} = Q_1 - 1.5 \times IQR$ and $IQR = Q_3 - Q_1$ is the sample interquartile range. Thus, the quantity $\lambda_U - \lambda_L$ represents a new robust measure for the process width.

Figure 1. Box-plot chart.
Source: The authors.

Fig. 1 shows a basic representation of a univariate process distribution based on the idea of box-plot charts, i.e. based on a robust definition for outliers into the data set. Thus, if we denote the domain of the process by $D_x \subseteq \mathbb{R}$, then any point belonging to O_x, where O_x is defined as the set $O_x = \{x \in D_x : x < L_{lower} \lor x > L_{upper}\}$, will represent an outlier for the process X. Let us analyze situations represented in A, B or C, and A', B' or C'. Under the new approach, three elemental criteria can be used to construct process capability measures based on the (natural) process variability estimation. The setting can be defined as follows:

i) For bilateral specifications, the process width is estimated by:

$$\widehat{\lambda_U - \lambda_L} = \begin{cases} L_{upper} - L_{lower}, & \text{if } (A \lor B) \land (A' \lor B') \text{ occurs} \\ x_{n:n} - L_{lower}, & \text{if } (A \lor B) \land C' \quad\quad \text{occurs} \\ L_{upper} - x_{1:n}, & \text{if } C \land (A' \lor B') \quad \text{occurs} \\ \hat{R}, & \text{if } C \land C' \quad\quad\quad\quad \text{occurs} \end{cases} \tag{11}$$

ii) For one upper specification, the process width is estimated by:

$$\widehat{\lambda_U - M} = \begin{cases} L_{upper} - M, & \text{if } (A' \lor B') \quad \text{occurs} \\ x_{n:n} - M, & \text{if } C' \quad\quad\quad\quad \text{occurs} \end{cases} \tag{12}$$

iii) For one lower specification, the process width is estimated by:

$$\widehat{M - \lambda_L} = \begin{cases} M - L_{lower}, & \text{if } (A \land B) \quad \text{occurs} \\ M - x_{1:n}, & \text{if } C \quad\quad\quad\quad \text{occurs} \end{cases} \tag{13}$$

3. Bootstrap analysis

3.1. Bootstrap based inference methods

In recent years, due to the analytical advances in industry control, not only point estimation but also hypothesis testing and interval estimation, is often obligatory for the producer to demonstrate process capability as part of the contract. In that sense, an important part of the analysis included in the present paper is based on the behavior results of hypothesis testing and confidence intervals for the indices and C_{prk}, relative to the indices (1) to (4), that are obtained via bootstrap techniques. Informally, bootstrap methods are based on sampling with replacement following the next idea. Let $X_1, \ldots, X_n$, a sequence of n i.i.d. r.v. with the same distribution of the process X, and $x_1, \ldots, x_n$ a realization of size n; then a uniformly distributed random variable X^*

taking values on the set $\{x_1, \ldots, x_n\}$ is defined by the probability distribution

$$P(X^* = x^*) = \frac{1}{n}, i = 1, \ldots, n \tag{14}.$$

Thus, a bootstrap sample, denoted by $\{x_1^*, \ldots, x_n^*\}$, is a sample drawn with replacement from the original sample $\{x_1, \ldots, x_n\}$ by using the law of probability distribution defined in (14). In the next sections we describe the methods used in this paper in the context of statistical inference for PCR.

3.2. Hypothesis testing

From industry experience, it is frequently necessary to demonstrate that the capability ratio C meets or exceeds a particular target value, say C_0. This may be formulated as a hypothesis testing problem, i.e. $H_0 : C \le C_0$ (Process is not capable) and $H_1 : C > C_0$ (Process is potentially capable). We would like to test H_0 against H_1 in the case of the indices C_{pr} and C_{prk} relative to the indices (1) to (4), in different scenarios. Several authors have investigated this test (see for instance Kane [25], Montgomery [26], Lin and Pearn [27], Shu and Lu [28], Mathew et al. [29] and Albing [30]). In all mentioned works, the authors have dealt with this problem by finding parametric distributions for test statistics under the null hypothesis. Regarding the above, we have avoided complicated mathematical processes to derive the null distribution of any pivotal test statistic based on the C_{pr} and C_{prk} indices. Thus, we have proposed a reasonable alternative based on bootstrap techniques to test the null hypothesis H_0. The approach implemented in this paper is based on the ideas discussed by Hall and Wilson [31] and Becher et al. [32]. The method is defined as follows.

Let us denote $C_u(X|\theta) = C_u$ as the value of the process capability ratio of the process X, with θ representing a non-stochastic set of intrinsic parameters of the quality process X, and u representing one of the following classes of indices: C, C_{pk}, C_{pm}, C_{pmk}, C_{pr} and C_{prk}. In what follows, we shall consider that θ contains (non-identical) subsets of intrinsic parameters such as: USL, LSL, μ, σ, τ, M, Q_1, Q_3, π_L, and π_U, where $\pi_U = P(X > USL)$, and $\pi_L = P(X < LSL)$ are the upper and lower proportions of non-conformity detected in the process, respectively. The corresponding sample estimator and bootstrap estimator of the index $C_u(X|\theta) = C_u$ are $C_u(x_1, \ldots, x_n|\hat\theta) = \hat{C}_u$ and $C_{ubl} = (x_{1b}^*, \ldots, x_{nb}^*|\hat\theta^*) = \hat{C}_{ub}^*$, where sub index b represents the b-th bootstrap replicate $(b = 1, \ldots, B)$.

The study of the probability distribution of the estimator $\hat{C}_{pr}$ is equivalent to finding the distribution of the inverse of a random variable. As it can be seen from the equations (11) to (13), this random variable contains a highly nonlinear transformation of order statistics. To solve this problem, long and tedious algebra are necessary, as well as the application of asymptotic results from the distribution theory of order

statistics. This is outside the scope of this paper and has been left to future work.

A common alternative to constructing useful pivotal statistics, avoiding such a theoretical analysis, is based on the idea of using bootstrap techniques. Hall and Wilson [31] proposed two guidelines to test hypotheses for the population mean based on percentiles of the null distribution of a bootstrap test statistic. Due to every u-th class of capability, ratio $\hat{C}_u$ is an univariate real-valued function of the sample $x_1, \ldots, x_n$. We can adapt the first guideline in the above mentioned paper to test $H_0 : C_u(X|\theta) \le C_0$ against $H_1 : C_u(X|\theta) > C_0$ for a specific value C_0. The implementation of the test is synthesized as follows:

i) *Null statistic*: Computing the ratio $C_u(.|\hat\theta)$ with the original sample $x_1, \ldots, x_n$, a natural null statistic T_0 is given by

$$T_0 = \frac{C_u(x_1, \ldots, x_n|\hat\theta)}{\hat\sigma_u^*} = \frac{\hat{C}_u - C_0}{\hat\sigma_u^*} \tag{15},$$

where $\hat\sigma_u^*$ is a suitable estimator of the squared root of $var(\hat{C}_u|x_1, \ldots, x_n)$. In the present work, we propose to use a large enough number B_0 of bootstrap samples of $\{x_1, \ldots, x_n\}$ to compute

$$\hat\sigma_u^* = \left[\frac{1}{B_0 - 1} \sum_{b=1}^{B_0} \left(\hat{C}_{ub}^* - \overline{\hat{C}_u^*} \right)^2 \right]^{1/2} \tag{16},$$

where $\overline{\hat{C}_u^*} = \frac{1}{B_0} \sum_{b=1}^{B_0} \hat{C}_{ub}^*$ is the bootstrap estimator of $E(\hat{C}_u|x_1, \ldots, x_n)$. It can be proved that $\overline{\hat{C}_u^*} \to E(\hat{C}_u|x_1, \ldots, x_n)$ with probability one [31,33].

ii) *Test statistic*: The statistic $C_u(.|\hat\theta)$ is computed with the b-th resample $x_{1b}^*, \ldots, x_{nb}^*$. Thus, the b-th bootstrap test statistic is obtained by computing

$$\hat{T}_{ub}^* = \frac{C_u(x_{1b}^*, \ldots, x_{nb}^*|\hat\theta_{ub}^*) - \hat{C}_u}{\hat\sigma_{ub}^*} = \frac{\hat{C}_{ub}^* - \hat{C}_u}{\hat\sigma_{ub}^*} \tag{17}$$

for $b = 1, \ldots, B$. Here, $\hat\sigma_u^*$ is the bootstrap estimator of the squared root of $var(\hat{C}_u^*)$, which is obtained in the same way as

$var(\hat{C}_u|x_1, \ldots, x_n)$, i.e., $\hat\sigma_u^* = \left[\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{C}_{ub}^* - \overline{\hat{C}_u^*} \right)^2 \right]^{1/2}$ with $\overline{\hat{C}_u^*} = \frac{1}{B_0} \sum_{b=1}^{B} \hat{C}_{ub}^*$

Informally, the basic idea of the expression (17) is based on the assumption that the distribution of the statistic $\hat{T}_u^*$ accurately mimics the distribution of $\hat{T}_u^0$. Thus, the decision rule for the test at the b replication is to reject H_0 t if $\hat{T}_{ub}^* > \hat{T}_u^0$ or not to reject H_0 if $\hat{T}_{ub}^* \le \hat{T}_u^0$.

iii) *P value*: Se $\delta_b^* = I(\hat{T}^* > \hat{T}^0|H_0)$t, where the probability or p-value in the expression (17) is computed by:

$$\hat{p}_{value}^* = \frac{1 + \sum_{b=1}^{B} \delta_b^*}{1 + B} \tag{18}$$

3.3. Confidence intervals

Bootstrap methods to construct confidence intervals for PCR, have been widely studied by several authors, e.g.

Franklin and Wasserman [34,35], Choi et al. [36], Tong and Chen [37], Yeh and Bhattacharya [24], Balamurali and Kalyanasundaram [38], Mathew et al. [29], and Wang et al. [39] among others. The conventional parametric approach would suggest that the probability distribution of capability ratios $\hat{C}_{pr}$ and $\hat{C}_{prk}$ need to be investigated, but due to the same reasons explained in section 3.2, from a practical point of view, it makes more sense to try to approximate the distribution of every class of ratio, C_u, via bootstrap techniques. In this paper, we have used the method of bias corrected percentile bootstrap (BCPB). A complete justification of this method is in Efron [40].

The method is summarized as follows: firstly, using the ordered distribution of $\hat{C}_{u1}^*,...,\hat{C}_{uB}^*$, a significance level is fixed, α, and then the following quantities are calculated:

$$q_0 = P(\hat{C}_u^* < C_u), \qquad z_0 = \Phi^{-1}(q_0), \qquad q_L = \Phi^{-1}(2z_0 - z_{1-\alpha})$$

$q_U = \Phi^{-1}(2z_0 + z_{1-\alpha})$, where Φ^{-1} is the inverse of the standard normal probability distribution and $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$ is the $(1-\alpha)$th-quantile of the standard normal distribution. Then, using a large number B of bootstrap resamples, a $[(1-\alpha) \times 100]$ percent *BCBP* confidence interval for C_u, is given by $[\hat{C}_{u[q_L \times B]}^*, \hat{C}_{u[q_U \times B]}^*]$.

Efron and Tibshirani [41] indicated that a rough minimum of 1000 bootstrap samples is usually sufficient to compute reasonably accurate confidence interval estimates. Nevertheless, numerical results included in the present paper have been obtained by using *B*=10000 bootstrap samples.

4. Numerical results

The following subsections presents the results corresponding to the application of the proposed capability indices (and their quality assessment methodology) to the real data.

4.1. Real data example

For the purpose of investigating whether outliers can affect decisions in a quality management process, in this section we present numerical results obtained from the application of the new approach to a real data application.

Original experimental data were collected from a study of 150 test steel pipes, which were analyzed in the Science and Engineering of Materials labs at the University of A Coruña in Spain. Originally, the study consisted of developing statistical quality control measurements on traction-resistance by using the European norm UNE-EN 10002-140. Previous analyses of goodness of fit confirmed that the data were normally distributed with a 516 mean and standard deviation of 20 for 95% of confidence level. Fig. 2, shows a histogram describing the real data set.

Fig. 2 shows that the data are centered to its mean value. In this case, the sample mean is very close to the sample median, 516.3 and 517.6 respectively, and the sample standard deviation is 20.81. Moreover, the histogram also shows that the data set is a little skewed to the left: the skewness coefficient is -0.1641, whereas the kurtosis excess is 3.0112.
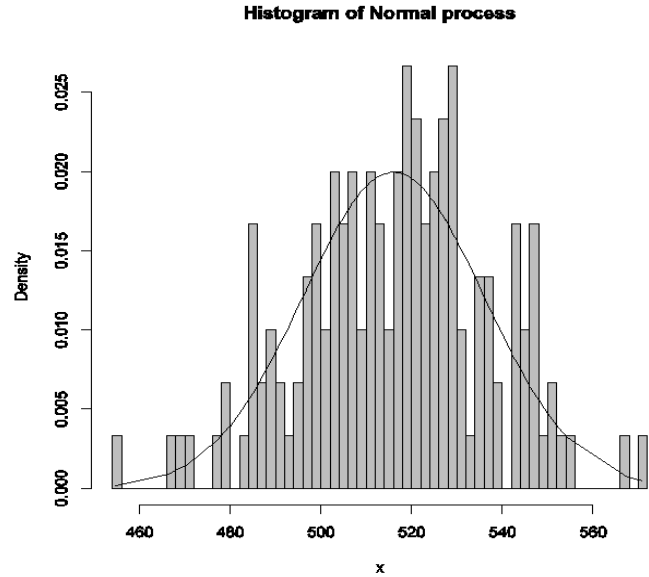
Figure 2. Histogram for real data set.
Source: The authors.

In terms of capability, several indices were computed to characterize the process *X* and for which the estimates of *LSL* and *USL* are 453.87 and 578.73 respectively. These values remained fixed during the whole bootstrap analysis and their results will be discussed in subsection 4.2.

For the first classical approach, the estimate $C_p = 1.00007$. This value tells us that the process seems to be potentially possible under the 3σ criteria. The other classical ratios showed little differences with respect to the C_p measure. In fact, the estimates were $\hat{C}_{pk} = 0.99936$, $C_{pm} = 0.99779$, and $\hat{C}_{pmk} = 0.99715$, which suggest that process is not possible. In this case, $\hat{C}_{pm}$ and $\hat{C}_{pmk}$ were computed by choosing the value $\tau = \hat{M}$ (target = median). Finally, the proposed measures, $\hat{C}_{pr}$ and $\hat{C}_{prk}$ suggest the same conclusion: the $\hat{C}_p$ index, i.e., the process is potentially possible, and $\hat{C}_{pr} = 1.20012$ and $\hat{C}_{pmk} = 1.19545$.

4.2. Bootstrap-based inference results

In this section, we present the results obtained by the bootstrap based statistical inference approach. The results collected in the following tables were obtained after varying the sample size and also the percentage of outliers in the original experimental data. After having applied the goodness of fit tests for contaminated data, we conclude that both 5 percent and 10 percent of outliers resulted in lack of normality for the original data. This fact added a third source of lack of robustness for all capability measures.

Table 1 shows results of the bootstrap hypothesis testing that was described in section 3.1. Firstly, allow us to draw your attention to the columns in Table 1. It can be seen that in case of normally distributed data (0% outliers), p-values seem to decrease as the sample size increases for all indices even though, as it was expected, p-values are less than the

Table 1.
Bootstrap p-value obtained with approximately a 95% confidence level.

Sample size	Outliers	C_p	C_{pk}	C_{pm}	C_{pmk}	C_{pr}	C_{prk}
n =100	0%	0.0079	0.0059	0.0479	0.0069	0.0100	0.0026
	5%	0.0095	0.0071	0.8036	0.7524	0.1130	0.0651
	10%	0.4819	0.3857	0.9999	0.9993	0.2442	0.0689
n =150	0%	0.0001	0.0001	0.0064	0.0051	0.0020	0.0013
	5%	0.0085	0.0059	0.7859	0.7414	0.0060	0.2300
	10%	0.4918	0.4009	0.9989	0.9995	0.1541	0.3560
n =300	0%	0.0001	0.0001	0.0055	0.0028	0.0002	0.0001
	5%	0.0134	0.0103	0.9948	0.9826	0.0159	0.0612
	10%	0.4957	0.4148	0.9983	0.9981	0.0529	0.0597

Source: The authors.

0.05 nominal significance level. Thus, in this experiment, the tests do not fail by rejecting the null hypothesis.

When data significantly departs from normality (5% and 10% outliers), some differences can be observed in the performance behavior of indices $\hat{C}_p$, $\hat{C}_{pk}$ and $\hat{C}_{pr}$, relative to the indices $\hat{C}_{pm}$, $\hat{C}_{pmk}$ and $\hat{C}_{prk}$.

For a moderate level of contamination of data (5% outliers), indices $\hat{C}_p$, $\hat{C}_{pk}$ do not reject the null hypothesis. However, it seems that the test tends are not rejected when sample size increase. For a higher level of outliers (10% outliers), results show that all indices behave as expected and they gain in robustness. In fact, this happens because they do not reject H_0, no matter the sample size.

Results for indices $\hat{C}_{pm}$, $\hat{C}_{pmk}$ show that these capability measures were more robust than the other indices. These measurements do not fail by rejecting H_0.

Finally, results concerning the proposed indices $\hat{C}_{pr}$ and $\hat{C}_{prk}$ show some evidence that the indices are robust capability measures. This is because these indices do not reject the null hypothesis when data are completely out-of-control.

Table 2 shows results on approximately 95 percent confidence intervals obtained via the bias corrected percentile bootstrap method. It can be seen that some results are somewhat surprising, relative to the results collected in Table 1.

Firstly, for normal distributed data (0% outliers), natural intuition tells us that confidence intervals should contain the hypothetical value $C_0 = 1$ when sample sizes increase. As it can be seen in Table 2, the results of our experiment show that the proposed indices $\hat{C}_{pr}$ and $\hat{C}_{prk}$ showed the expected behavior when the sample size was greater than or equal to 300. This result suggests that these two new classes of process capability measurements have robust properties in normal conditions.

Nevertheless, when there are non-normal conditions (5% and 10% outliers), the 95% confidence intervals for the proposed indices are surprisingly unified (C_{pr} and C_{prk} lies within the confidence interval).

For both, moderate and highly contaminated data (5% and 10% outliers), upper limits of obtained confidence intervals tend to be less than 1 (C_p, C_{pk}, C_{pm} and C_{pmk} are out of the interval). This fact allows to consider that classical measures $\hat{C}_r$ and $\hat{C}_{pk}$ show a better behavior when sample size increase ($n >100$).

Table 2.
Approximate 95% bias of corrected percentile bootstrap confidence intervals.

Sample size	Capability ratio	0% outliers Lower lim.	0% outliers Upper lim.	5% outliers Lower lim.	5% outliers Upper lim.	10% outliers Lower lim.	10% outliers Upper lim.
n =100	C_p	0.8844	1.1674	0.8002	1.2531	0.8539	1.1090
	C_{pk}	0.8698	1.1645	0.8139	1.2290	0.8439	1.1084
	C_{pm}	0.7929	1.0335	0.8484	0.9544	0.3954	0.7033
	C_{pmk}	0.7927	0.9741	0.8502	0.8965	0.3338	0.6845
	C_{pr}	1.0836	1.2034	0.7618	1.2612	0.7511	1.2781
	C_{prk}	1.0261	1.2127	0.7436	1.2788	0.7107	1.2942
n =150	C_p	0.9047	1.1338	0.8927	1.1390	0.6556	0.9425
	C_{pk}	0.8935	1.1280	0.8931	1.1241	0.6235	0.9186
	C_{pm}	0.7930	0.9708	0.7862	0.9211	0.5983	0.7659
	C_{pmk}	0.7937	0.9246	0.7927	0.9110	0.5701	0.7458
	C_{pr}	1.0836	1.2062	0.8217	1.2289	0.9452	1.3784
	C_{prk}	1.0264	1.2146	0.8052	1.2383	0.9047	1.3477
n =300	C_p	0.6556	0.9425	0.4605	0.6853	0.3850	0.5813
	C_{pk}	0.6235	0.9186	0.4226	0.6708	0.3576	0.5689
	C_{pm}	0.5983	0.7659	0.4181	0.6048	0.3615	0.5302
	C_{pmk}	0.5701	0.7458	0.3796	0.5958	0.3311	0.5182
	C_{pr}	0.9452	1.3784	0.9738	1.3352	0.9109	1.2466
	C_{prk}	0.9047	1.3477	0.9738	1.3352	0.8747	1.2369

Source: The authors.

Finally, $\hat{C}_{pm}$ and $\hat{C}_{pmk}$ indices can be seen to be more conservative than the other indices, no matter the sample size. These two measurements suggest that when data are not normal the process is not possible.

Taking the results into account, the proposed robust indices could be applied in a wide range of study cases, e.g. energy efficiency evaluation, academic evaluation, or the assessment of technology and innovation in companies [42-46].

5. Conclusions

In this paper, two classes of capability measurements inspired by the idea of robustness derived from the theory of construction box-plot charts have been studied. These new measurements were formulated by defining two new classes of robust process capability ratios that were then compared with traditional ratios in the literature, under several experimental schemes.

Due to the difficulty of obtaining the probability distribution of the new capability ratios, bootstrap methods were applied to study the robustness of the indices via the statistical inference approach. Thus, bootstrap hypothesis testing and bootstrap confidence intervals were used to test and estimate the true level of capability of a quality process.

The bootstrap experimental analysis was implemented by using a set of laboratory data that came from the analysis of the traction-resistance of steel pipes by using the European norm UNE-EN 10002-1 at the University of A Coruña, Spain.

The inference based results showed that the proposed capability measurements were comparable with the traditional process capability ratios. Comparative analyses suggested that the new ratios are robust measures to estimate the true level of process capability under normality.

Moreover, the new capability ratios were shown to be less conservative than some traditional ratios under the presence of outliers; this produced a moderate lack of normality that seems to improve when increasing the sample size.

Acknowledgements

References

[1] Chou, Y.M. and Owen, D.B., On the distribution of the estimated process capability indices. Communications in Statistics – Theory and Methods, 18, pp. 4549-4560, 1989. DOI: 10.1515/EQC.2003.5

[2] Clements, J.A., Process capability indices for non normal calculations. Quality Progress, 22, pp. 49-55, 1989. DOI: 10.1080/08982119408918772

[3] Pearn, W., Kotz, S. and Johnson, N.L., Distributional and inferential properties of process capability indices. Journal of Quality Technology, 24, pp. 216-231, 1992.

[4] Ebadi, M. and Shahriari, H., A process capability index for simple linear profile. The International Journal of Advanced Manufacturing Technology, 64, pp. 857-865, 2013. DOI: 10.1007/s00170-012-4066-7

[5] Kotz, S. and Johnson, N.L., Process Capability Indices. Chapman & Hall, London, 1993.

[6] Chien-Wei, W., Pearn, W.L. and Kotz, S., An overview of theory and practice on process capability indices for quality assurance. International Journal of Production Economics, 117, pp. 338-359, 2009. DOI: 10.1016/j.ijpe.2008.11.008

[7] Pan, J.N., Li, C.I. and Chen, F.Y., Evaluating environmental performance using new process capability indices for autocorrelated data. Environmental Monitoring and Assessment, 186, pp. 6369-6384, 2014. DOI: 10.1007/s10661-014-3861-z

[8] Somerville, S.E. and Montgomery, D.C., Process capability indices and non-normal distributions. Quality Engineering, 9, pp. 305-316, 1996. DOI: 10.1080/08982119608919047

[9] Kotz, S. and Lovelace, C.R., Introduction to Process Capability Indices: Theory and Practice. Arnold, London, 1998.

[10] Shore, H., A new approach to analysing non-normal quality data with applications to process capability analysis. International Journal of Production Research, 36, pp. 1917-1933, 1998. DOI: 10.1080/002075498193039

[11] Tang, L.C. and Than, S.E., Computing process capability indices for non normal data: A review and comparative study. Quality and Reliability Engineering International, 15, pp. 339-353, 1999. DOI: 10.1002/(SICI)1099-1638(199909/10)15:5<339::AID-QRE259>3.0.CO;2-A

[12] Chang, Y.S., Choi, I.S. and Bai, D.S., Quality and Reliability Engineering International, 18, pp. 383-393, 2002.

[13] Pearn, W.L. and Chen, K.S., One-sided capability indices Cpu and Cpl: decision making with sample information. International Journal of Quality and Reliability Management, 19, pp. 221-245, 2002. DOI: 10.1108/02656710210421544

[14] Kotz, S. and Johnson, N.L., Process capability indices -A review, 1992-2000 with discussion. Journal of Quality Technology, 34, pp. 2-53, 2002.

[15] Grau, D., New process capability indices for one-sided tolerances. Quality Technology & Quantitative Management, 6, pp. 107-124, 2009.

[16] Grau, D., Process yield, process centering and capability indices for one-sided tolerance processes. Quality Technology & Quantitative Management, 9, pp. 153-170, 2012.

[17] Heavlin, W.D., Statistical properties of capability indices. Technical Report 320. Technical Library, Advanced Micro Devices, Inc., Sunnyvale, CA, 1988.

[18] Borges, W. and Ho, L.L., A fraction defective based capability index.

[19] Quality and Reliability Engineering International, 17, pp. 447-458, 2001. DOI: 10.1002/qre.438

[19] Huber, P., Robust Statistics. John Wiley & Sons, New York, 1981. DOI: 10.1007/978-3-642-04898-2_594

[20] Abu-Shawiesh, M.O. and Abdullah, M.B., New robust statistical process control chart for location. Quality Engineering, 12, pp. 149-159, 1999. DOI: 10.1080/08982119908962572

[21] Grznar, J. and Booth, D.E., A robust smoothing approach to statistical process control. Journal of Chemical Information and Computer Sciences, 37, pp. 241-248, 1997. DOI: 10.1021/ci9600317

[22] Kocherlakota, S. and Kocherlakota, K., Confidence intervals for the process capability ratio based on robust estimators. Communications in Statistics – Theory and Methods, 23, pp. 257-276, 1994. DOI: 10.1080/03610929408831252

[23] Prasad, S. and Bramorski, T., Robust process capability indices. International Journal of Management Science, 26, pp. 425-435, 1998. DOI: 10.1016/S0305-0483(97)00075-3

[24] Yeh, A.B. and Bhattacharya, S., A robust process capability index. Communications in Statistics – Simulation and Computation, 27, pp. 565-589, 1998. DOI: 10.1080/03610919808813495

[25] Kane, V.E., Process capability indices. Journal of Quality Technology, 18, pp. 41-52, 1986.

[26] Montgomery, D.C., Introduction to Statistical Quality Control, 6th edition. John Wiley & Sons, New York, 2009. DOI: 10.1002/qre.4680070316

[27] Lin, P.C. and Pearn, W.L., Testing process capability for one-sided specification limit with application to the voltage level translator. Microelectronics Reliability, 42, pp. 1975-1983, 2002. DOI: 10.1016/S0026-2714(02)00103-8

[28] Shu, M.H., Lu, K.H., Hsu, B.M. and Lou, K.R., Testing quality assurance using process capability indices CPU and CPL based on several groups of samples with unequal sizes. Information and Management Sciences, 17, pp. 47-65, 2006.

[29] Mathew, T., Sebastian, G. and Kurian, K.M., Generalized confidence intervals for process capability indices. Quality and Reliability Engineering International, 23, pp. 471-481, 2007. DOI: 10.1002/qre.828

[30] Albing, M., Process capability indices for Weibull distributions and upper specification limits. Quality and Reliability Engineering International, 25, pp. 317-334, 2009. DOI: 10.1002/qre.972

[31] Hall, P. and Wilson, S.R., Two guidelines for bootstrap hypothesis testing. Biometrics, 47, pp. 757-762, 1991. DOI: 10.2307/2532163

[32] Becher, H., Hall. P. and Wilson, S.R., Bootstrap hypothesis testing procedures. Biometrics, 49, pp. 1268-1272, 1993. DOI: 10.2307/2532271

[33] Hall, P., The Bootstrap and Edgeworth Expansion. Springer, NewYork, 1992. DOI: 10.1007/978-1-4612-4384-7

[34] Franklin, L.A. and Wasserman, G., Bootstrap confidence intervals of Cpk: An introduction. Communications in Statistics – Simulation and Computation, 20, pp. 231-242, 1991. DOI: 10.1080/03610919108812950

[35] Franklin, LA. And Wasserman, G., Bootstrap lower confidence limits for process capability indices. Journal of Quality Technology, 24, pp. 196-210, 1992. DOI: 10.1108/02656710210442875

[36] Choi, K.C., Nam, K.H. and Park, D.H., Estimation of capability index based on bootstrap method. Microelectronic Reliability, 36, pp. 1141-1153, 1996. DOI: 10.1016/0026-2714(95)00211-1

[37] Tong, L.I. and Chen, J.P., Lower confidence limits of process capability indices for nonnormal process distributions. International Journal of Quality Reliability Management, 15, pp. 907-919, 1998. DOI: 10.1108/02656719810199006

[38] Balamurali, S. and Kalyanasundaram, M., Bootstrap lower confidence limits for the process capability indices Cp, Cpk, and Cpm. International Journal of Quality & Reliability Management, 19, pp. 1088-1097, 2002. DOI: 10.1108/02656710210442875

[39] Wang, D.S., Koo, T.Y. and Chou, C.Y., On the bootstrap confidence intervals of the capability index Cpk for multiple process streams. International Journal for Computer-Aided Engineering and Software, 24, pp. 473-485, 2007. DOI: 10.1108/02644400710755870

[40] Efron, B., The Jackknife, the Bootstrap and other Resampling Plans. Society for Industrial and Applied Mathematics, Pennsylvania, 1982.

[41] Efron, B. and Tibshirani, R.J., Bootstrap methods for standard errors,

confidence intervals, and other measures of statistical accuracy. Statistical Science, 1, pp. 54-77, 1986. DOI: 10.1214/ss/1177013815

[42] Torres, L.M., Castellanos, Ó.F. and Fúquene, A.M., Evaluating Colombian SMEs technological innovation: Part 1: Conceptual basis, evaluation methodology and characterisation of innovative companies. Ingeniería e Investigación, 27(1), pp. 158-167, 2007.

[43] Jiménez, C., Cristancho, A. and Castellanos, O., The role of capability in technology valuation. Ingeniería e Investigación, 31(2), pp. 112-123, 2011.

[44] Cortés, H., Gallego, L. and Rodríguez, G., The engineering faculty today: An approach towards consolidating academic indicators. Ingeniería e Investigación, 31, pp. 74-90, 2011.

[45] Zaragoza-Fernández, S., Tarrío-Saavedra, J., Naya, S., López-Beceiro, J. and Álvarez-García, A., Impact estimates of the actions for the rehabilitation of energy efficiency in residential building. DYNA, 81(186), pp. 200-207, 2014. DOI: 10.15446/dyna.v81n186.39930

[46] Castrillón, R.D.P., González, A.J. and Ciro-Quispe, E., Energy efficiency improvement in the cement industry by wet process through integral energy management system implementation. DYNA, 80(177), pp. 115-123, 2013.

S. Naya, is professor in statistics at the University of A Coruña, Spain. He received a BSc. and MSc. in Mathematics from the University of Santiago de Compostela, Spain and a PhD in Industrial Engineering from the University of A Coruña, Spain with a with a Cum Laude mention of excellence. His main scientific-technical fields of interest are Statistical Quality Control, Reliability and Statistical Applications in Materials Engineering. He is an Elected Member of the ISI (International Statistical Institute) and was given the International Educator Award in 2013 from the Juarez Lincoln Martí Project. Currently he is Vice rector for Science Policy, Research and Transfer at the University of A Coruña.
ORCID: 0000-0003-4931-9859.

A. Devia-Rivera, is an Excess and Obsolescence Planner at the Commercial Marketing Department of Avon Cosmetics in Madrid, Spain. He is a Statistical Engineer graduate from the University of Santiago de Chile, Chile. In 2008 he received a Diploma of Advanced Studies in Statistics and Operations Research from the University of A Coruña, Spain. He obtained his PhD in Statistics and Operations Research in January 2016 with a Cum Laude mention of excellence. His main scientific fields of interest are Statistical Modeling in Finance and Engineering. He has taught several courses in statistics, probability and time series as assistant professor in Chile, and he has published papers and short communications in the field of Statistical Modeling of Credit Risk.
ORCID: 0000-0002-6499-5233

J. Tarrío-Saavedra, is an Industrial Engineer, has a MSc. and PhD in Statistics and Operations Research from the University of A Coruña (UDC). He was given a special PhD award in the area of Mathematics and Computer Science. Currently he is a Professor in Statistics, Statistical Quality Control, and Thermomechanical Fatigue on undergraduate and Master's degrees in the UDC. He has participated in 14 projects and research contracts with public and private entities. He is author of 34 scientific articles, 4 book chapters and 60 conference papers related to statistics, materials science, engineering and bibliometrics. He has conducted research at the University Paris Diderot - Paris 7.
ORCID ID: 0000-0002-9584-127X

M.A. Flores, is a professor at the National Polytechnic School and a researcher at the Center for Modeling Mathematics at the National Polytechnic School in Quito, Ecuador. He is a BSc. in Statistical Computing Engineer from the Polytechnic School of the Coast. In 2006 he received a in MSc. in Operations Research from the National Polytechnic School, and in 2013 received a MSc. in Technical Statistics from the University of A Coruña. He is currently a doctoral student at the University of A Coruña in the area of Statistics and Operations Research. He has over 14 years professional experience in various areas of Statistics, Computing and Optimization, multivariate data analysis, econometric, Market Research, Quality Control, definition and construction of systems indicators, development of applications and optimization modeling.
ORCID ID: 0000-0002-7742-1247

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN

FACULTAD DE MINAS

Área Curricular de Ingeniería Administrativa e Ingeniería Industrial

Oferta de Posgrados

Especialización en Gestión Empresarial
Especialización en Ingeniería Financiera
Maestría en Ingeniería Administrativa
Maestría en Ingeniería Industrial
Doctorado en Ingeniería - Industria y Organizaciones

Mayor información:

E-mail: acia_med@unal.edu.co
Teléfono: (57-4) 425 52 02
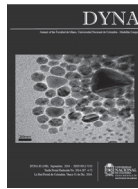
Effect of elastoplastic behavior on the impact response of expanded metal tubes

Carlos Graciano [a], Gabriela Martínez [b] & Edwar Saavedra [b]

[a] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia, cagracianog@unal.edu.co
[b] Departamento de Mecánica, Universidad Simón Bolívar, Caracas, Venezuela, gabrielamb@usb.ve, saavedra.edwar@gmail.com

Abstract
This paper presents a numerical study on the effect of the elastoplastic behavior of the material on the structural response of expanded metal tubes subjected to impact loads. The study is carried out using the finite element method, considering two material models. Firstly, a numerical model is built assuming a bilinear behavior with strain hardening. Secondly, a model is elaborated considering the speed of deformation using the correlation of Cowper-Symonds. These models are compared and validated with experimental results taken from the literature. Subsequently, a parametric study is conducted, varying the dimensions and orientation of the mesh as well as the impact speed. The results show that the response of expanded metal tubes is sensitive to changes in the orientation of the geometrical axes of the mesh. Increasing the dimensions of the strands increase the material available for plastic deformation, this, in turn, produces an increase in the energy absorbed per mass unit. In addition, the loading speed causes a delay in the response of the specimens.

Keywords: Expanded metal, Plasticity, Axial impact, Energy absorption, Dynamic response.

Efecto del comportamiento elastoplástico sobre la respuesta al impacto de tubos de metal expandido

Resumen
Este trabajo presenta un estudio numérico sobre el efecto del comportamiento elastoplastico del material sobre la respuesta estructural de tubos fabricados con mallas de metal expandido sometidos a cargas de impacto. El estudio se lleva a cabo utilizando el método del elemento finito, considerando dos modelos de material. En primer lugar, se elabora un modelo numérico asumiendo un comportamiento bilineal con endurecimiento por deformación. En segundo lugar, se emplea un modelo que considera la velocidad de deformación utilizando la correlación de Cowper-Symonds. Estos modelos son comparados y validados con resultados experimentales tomados de la literatura. Posteriormente, se lleva a cabo un estudio paramétrico variando las dimensiones y orientación de las mallas, y la velocidad de impacto. Los resultados muestran que la respuesta de los tubos de metal expandido es sensible a los cambios en la orientación de los ejes geométricos. Incrementando las dimensiones de las venas aumenta el material disponible para deformación plástica, lo cual a su vez produce un incremento en la energía absorbida por unidad de masa. Adicionalmente, la velocidad causa un retraso en la respuesta de los especímenes.

Palabras clave: Metal expandido, Plasticidad, Impacto axial, Absorción de energía, Respuesta dinámica.

1. Introduction

Many investigations have been conducted on devices designed to absorb energy during an impact scenario and therefore protect the structure in consideration [1]. An energy absorber is a system that converts, completely or partly, the kinetic energy into another form of energy, which can be either reversible or irreversible [2].

Since the second half of the twentieth century, there has been a remarkable growth in the number of vehicles on the roads. This has resulted in a large number of accidents and fatalities due to collisions and impacts that occur in various

ways. The need to improve the safety of vehicles and roads has led the scientific community to research and develop impact attenuation devices in order to absorb the kinetic energy of impacts and mitigate impacts on people and structures [3].

Significant state-of-the art reviews [2-4] in the field of crashworthiness have been conducted, summarizing over 40 years of research and developments of various types of collapsible structures that are designed to absorb kinetic energy during impact situations.

Nagel and Thambiratnam [1] used a finite element model to compare the response of conical and straight tubes subjected to impact loading tubes. They concluded that the lateral inertia of the tube affects the conical tubes less than the straight ones. Reid and Reddy [5] investigated the quasi-static and dynamic collapse of thin-walled tapered rectangular tubes. They developed analytical solutions of average for both collapse mode strengths. The dynamic solution took into account the effects of the strain-rate constitutive equation that was proposed by Cowper Symonds [6].

Wang and Lu [7] conducted experimental tests in which cylindrical tubes were impacted at speeds of about 300 m/s; their results showed a mushroom deformation mode. They mentioned that impacts at such high speeds are high complexity problems that can be analyzed only by numerical simulations and experimental tests.

Langseth et al. [8] performed finite element analyses on the collapse of square aluminum tubes. The numerical models were validated by comparing the results with experimental data from dynamic and quasi-static tests. It was concluded that, an increase in impact speed causes a significant increase in the breakdown strength. Hsu and Jones [9] performed quasi-static and dynamic tests on thin-walled tubes made with low carbon steel, stainless steel and aluminum alloys. The effects of strain-rate on material properties of the steel tubes were analyzed.

Rossi et al. [10] numerically analyzed the dynamic collapse of hexagonal extruded aluminum tubes with symmetric and asymmetric modes of collapse. They obtained good correlation with experimental results. Jones [11,12] classified the dynamic collapse of the tubes into two types: a) dynamic progressive buckling b) dynamic plastic buckling.

Recently, Graciano et al. [13-16] conducted a series of experimental and numerical investigations on the axial collapse of expanded metal tubes under quasi-static loading. It was found that expanded metal tubes under certain configurations may undergo a progressive and stable collapse and failure. Nouri et al. [17-18] performed a series of experimental and numerical studies on flattened expanded metal tubes that were subjected to impact.

In spite of the amount of research on the impact behavior of thin-walled structural member, expanded metal tubes under impact loading has received little attention. Hence, this research aims to investigate the effect of the elastoplastic behavior of the material on the impact response of expanded metal tubes. This investigation was carried out by means of nonlinear finite element analysis, and the numerical models are validated using results taken from the relevant literature. Thereafter, a parametric analysis is performed in order to
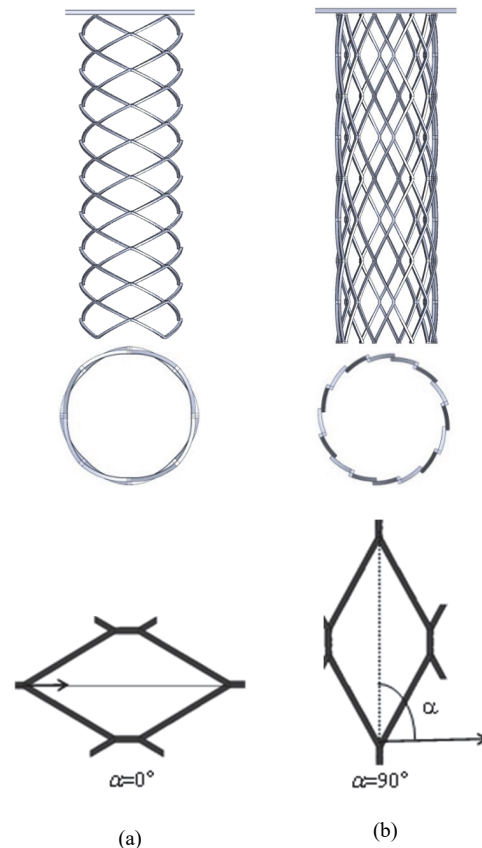
Figure 1. Numerical models of expanded metal tubes.
Source: The authors.

investigate the influence of: a) two material models, a bilinear and strain-rate dependent model; b) the orientation of the expanded metal cells; and c) size of the cells on the impact response of the expanded metal tubes.

2. Numerical model

2.1. Description

This section presents the development of a numerical model, using a finite element method for expanded metal tubes subjected to axial impact. The numerical analysis is conducted using the explicit analysis software ANSYS LS-DYNA [19]. The expanded metal cells in the tube were modeled with hexahedral high order SOLID 164 elements [19] that are suitable for large strain analyses. Two mesh orientations are considered in the analysis, one with $\alpha = 0°$ (Fig. 1a), and another with $\alpha = 90°$ (Fig. 1b).

Fig. 2 shows a schematic view of an EM cell; the pattern is characterized by a rhomb like shape. Table 1 presents the dimensions of the strands and cells used in the analysis. All tubes have an average length $L = 400$ mm, with a diameter $\phi = 120$ mm. In general, the models were built depending on the cell orientation, for $\alpha = 0°$ the models have 4 cells in the circumferential direction and 9 cells in the longitudinal (Fig. 1a). For $\alpha = 90°$ the models have 8 cells in the circumferential direction and 4 cells in the longitudinal (Fig. 1b).
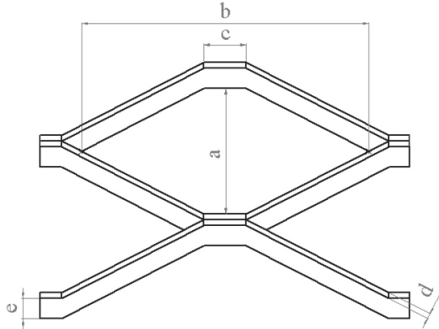
Figure 2. Nomeclature for EM cells.
Source: The authors.

Table 1.
Dimensions used in the parametric analysis.

Cell	a [mm]	b [mm]	d [mm]	e [mm]
H-24	44.70	90.00	1.90	2.80
H-26	44.20	89.60	3.00	3.20
H-27	36.80	81.60	6.00	6.00

Source: The authors.

The impact condition is generated using plates separated 1 mm from the upper end of the tube (Fig. 3a), which were modeled as a rigid body with a fixed mass M=20 kg. Initially, the rigid plate has an initial speed in the axial direction of the tube that causes axial impact. At the lower end the test specimen is clamped (Fig. 3b), $i.e.$ it has all the degrees of freedom are fixed.
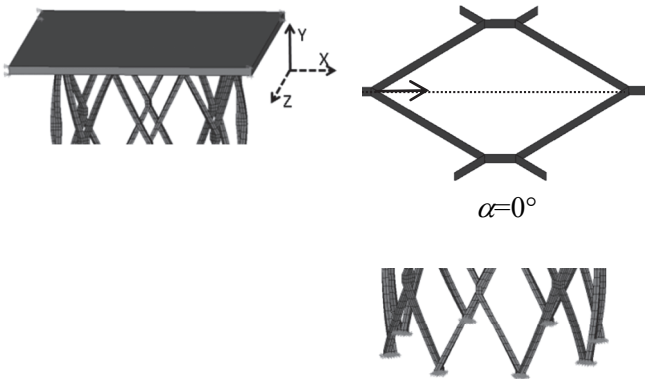
(a) Top end   (b) Bottom end
Figure 3. Boundary conditions.
Source: The authors.

The material used was an ASTM A569 steel; however, two material models are used to investigate their effect on the elastoplastic response of the expanded metal tubes under impact:

Model 1: classical bilinear isotropic hardening model (strain-rate independent) that uses two slopes (elastic and plastic) to represent the stress–strain behavior of the material [20]. It has the following mechanical properties: Young's modulus E = 205 GPa; Poisson's ratio v = 0.3; density ρ =

Table 2.
Material properties used for validation.

E [GPa]	v	ρ [kg/m³]	σ_y [MPa]	E_t [MPa]
207	0.3	7385	330	892

Source: Adapted from [1].
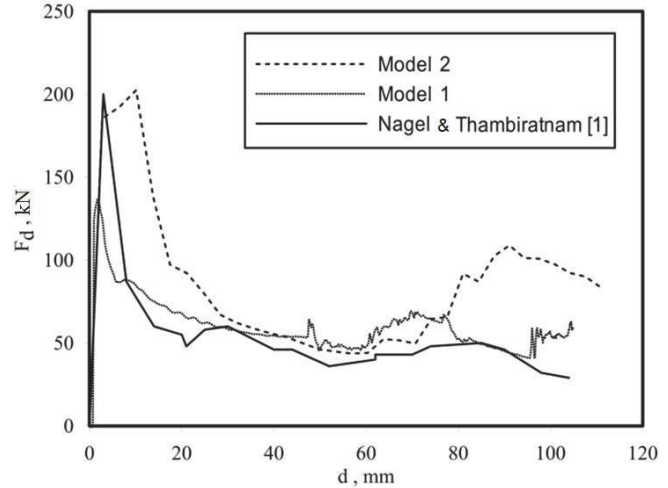
Figure 4. Load-displacement responses (Validation).
Source: The authors.

7835 kg/m³; yield strength σ_y = 250 MPa; tangent modulus E_t = 660 MPa.

Model 2: considering dynamic effects, the dynamic yield strength needs to be calibrated depending on the deformation speed with Eq. (1). For carbon steels, the experimentally determined values [21] for the coefficients c and P are: c = 6844, and P = 3.91.

$$\sigma_{yd} = \sigma_y \left[1 + \left(\frac{\dot{\varepsilon}}{c} \right)^{1/P} \right] \qquad (1)$$

2.2. Validation

The numerical procedure is validated by comparing the experimental and numerical results obtained by Nagel and Thambiratnam [1]. The geometry used was a rectangular solid tube with the following measurements: 48.31mm x 79.88mm, length L= 199mm and a thickness t= 1.5 mm. Table 2 shows the material properties employed. The tubes were subjected to an impact mass M= 90 kg, at a speed v= 15m/s. Due to symmetry in loading, geometry and boundary conditions, only one-fourth of the geometry was modeled.

Fig. 4 shows a comparison of the load–displacement responses obtained in [1] and those computed using the two material models depicted in the previous section, namely Model 1 and Model 2. It was observed that the best agreement with the experimental response from [1] was attained with Model 2. It is important to note that, in this case, Model 1 was unable to fully capture dynamic effects such as those presented in impact scenarios
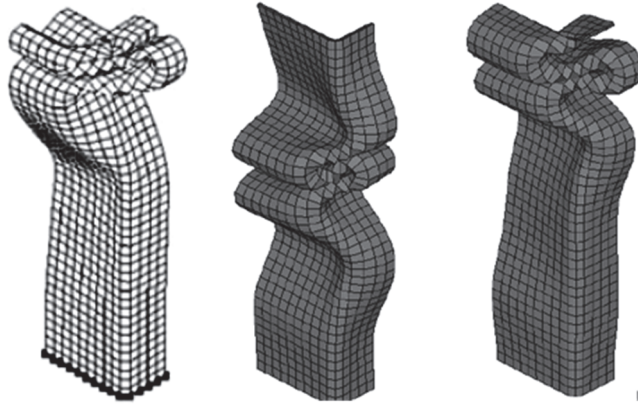
a) Taken from [1]   b) Model 1   c) Model 2

Figure 5. Comparison of the deformed shapes for the straight tubes.
Source: The authors.

Comparing the peak loads for the three models, it seems that the best agreement is also achieved with Model 2. The peak load computed in [1] was P_{peak} = 200 kN; the corresponding for Model 1 is P_{peak} = 136.84 kN, for Model 2 it is P_{peak}= 202.46 kN. By using Model 1, the peak load is considerably underestimated.

Fig. 5 shows the deformed shapes obtained using the two material models. Once again, the results obtained with Model 2 (Fig. 5c) best agree with those attained in [1] (Fig. 5a). The deformation patterns appear rather close to the applied load. Additionally, for Model 1, a larger deformation occurs in the middle of the tube.

It is important to point out that the authors in [1] use a shell element formulation. The numerical model developed herein makes use of a solid element formulation. Recent studies have demonstrated an enhanced capability of the latter to represent impact effects in comparison to the former [22].

3. Parametric analysis

The numerical procedure was validated in the previous section with solid tubes [1]. A convergence analysis was conducted to calibrate the model for expanded metal tubes, hence a mesh with 8605 elements was chosen for the tubes with α = 0°, and 7169 elements for α = 90° [21].

Next, a parametric study was conducted in order to investigate the influence of geometric variables on the impact response of expanded metal tubes:

(a) cell orientation (0 ° and 90 °);
(c) the loading speed (v =10, 15, 25, 35, 45, 50 m/s), and
(b) cell dimensions (see Table 1)

Load-displacement responses, deformed patterns at various displacement levels, and the energy absorption characteristics were the output parameters of the experimental investigation. To measure the energy absorption capacity, the following parameters were calculated: initial peak load (F_{peak}), mean load (F_m), absorbed energy (E_{200}), specific energy absorbed (E_e), and the structural efficiency (η).

The absorbed energy is primarily calculated by integrating the load-displacement curves, such as:

$$E_{200} = \int_0^{l_c} F dl \qquad (2)$$

where F is the measured force and l_c is the crushing length. The mean force F_m based on the area under the curve over the crushing distance l_c is calculated by:

$$F_m = \frac{\int_0^{l_c} F dl}{l_c} \qquad (3)$$

Afterward, the specific energy absorbed (E_e) was calculated by dividing the absorbed energy E_{200} by its weight W, according to:

$$E_e = \frac{E_{200}}{W} \qquad (4)$$

Finally, the structural efficiency, defined as the ratio between the mean and the peak force is calculated by:

$$\eta = \frac{F_m}{F_{peak}} \qquad (5)$$

All these parameters were calculated up to a crushing length l_c=200 mm.

3.1. Influence of the mesh orientation (α)

The load-displacement responses for the tubes with cells oriented at α = 0°, and α = 90° are shown in Figs. 6a and 6b, respectively. In this analysis, a mesh type H-26 (Table 1) was used, the speed was v=15 m/s, and the mass was M= 20 kg.

From the load-displacement responses plotted in Fig. 6a, it can be observed that tubes with α = 0° exhibit a stable and controlled response, which is desirable for energy absorption purposes. In contrast, Fig. 6b shows rather instable behavior for tubes with α = 90°. Regarding material behavior, Model 2 offers a greater resistance to the applied load than Material 1, and this effect is translated into a higher proportion of energy absorbed (see Table 3).

(a) α = 0°

(b) $\alpha = 90°$
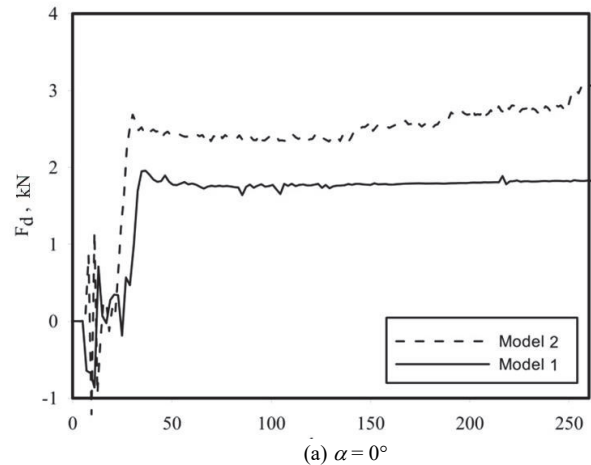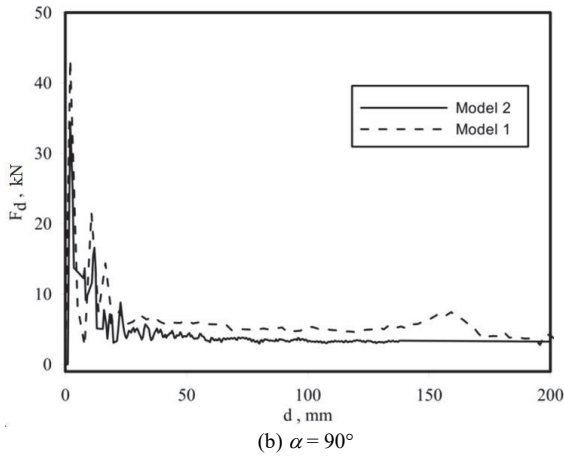
Figure 6. Load-displacement responses (mesh H-26 v=15 m/s).
Source: The authors.

Table 3.
Effect of the strand cross section.

	Model	F_{peak} [kN]	F_m [kN]	E_{200} [J]	η
0°	1	1.96	1.59	318.50	0.81
	2	2.69	2.29	458.21	0.85
90°	1	34.41	4.30	859.71	0.12
	2	43.29	5.98	1195.58	0.14

Source: The authors.

Table 3 summarizes the results for peak loads F_{peak}, mean loads F_m, energy absorbed E_{200}, and crush efficiency η for both the material models and the geometries analyzed herein. The tubes with $\alpha = 90°$ achieved a peak load up to 17 times greater than that for tubes with $\alpha = 0°$. At the same time, their mean load was just about 2.7 times higher than for $\alpha = 0°$. Nevertheless, the latter exhibits a better plastic performance and stability, which was shown in its crush efficiency that was close to 1. Moreover, the tubes with $\alpha = 90°$ absorb much more energy despite their instability: until almost 3 times more than the tubes with $\alpha = 0°$.

3.2. Influence of the loading speed

3.2.1. Tubes with $\alpha = 0°$

The impact speed varied from 10 m/s to 50 m/s. The nonlinear impact responses of the tubes with $\alpha = 0°$ for the two material models are plotted in Fig. 7. The mesh geometry corresponds to type H26, according to Table 1.

From the responses plotted in Fig. 7, it can be observed that an increase in impact speed leads to a decrease in the energy absorption capacity within the studied crushing length. This may be because during the process of cutting and stretching of the meshes, the loading direction is aligned with the corresponding direction of deformation, and, therefore, the cells collapsed more easily in this direction.

It is also observed that an increasing speed leads to a delay in the load-displacement responses, thus the amount of energy absorbed is reduced.
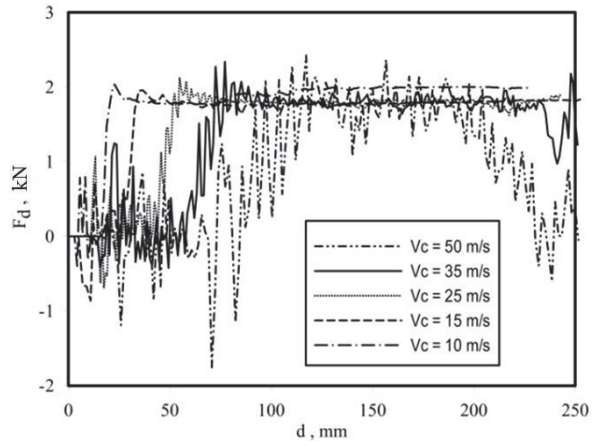
(a) α=0°/ Material Model 1
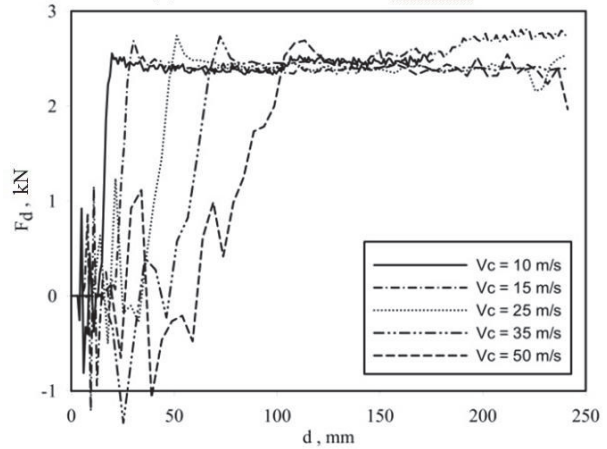
(b) α=0°/ Material Model 2

Figure 7. Effect of the loading speed α=0°.
Source: The authors.

As seen in Fig 8a, for speeds ranging from 10 m/s to 25 m/s, the cell collapses mainly in the middle part of the tube; this plastic behavior is more similar to the quasi-static case, as obtained in [13-16,22,23]. Furthermore, for higher speeds, from 35 m/s up to 50 m/s (Fig. 8b) the collapse propagates throughout the whole tube.

To summarize, the collapse modes of this configuration, within the investigated speed ranges, are dominated by local crushing of the cells.

3.2.1. Tubes with $\alpha = 90°$

The tubes' nonlinear impact responses, with $\alpha = 90°$ for the two material models, are plotted in Fig. 9. The mesh geometry corresponds to type H26 (Table 1). In this case, the strain-rate effects are also evidenced in the load-displacement responses, which are enhanced by the energy absorption characteristics that are observed after comparing the responses for Material 1 (Fig. 9a) and for Material 2 (Fig. 9b).

Furthermore, it can be observed that increasing speed produces an increase in the energy absorbed, in contrast to the previous case with $\alpha = 90°$. This leads to the conclusion that the orientation of the cell produces different effects in the structural responses, with respect to changes in the loading speed.

(a) Speed range: 10 m/s y 25 m/s

(b) Speed range: 35 m/s y 50 m/s
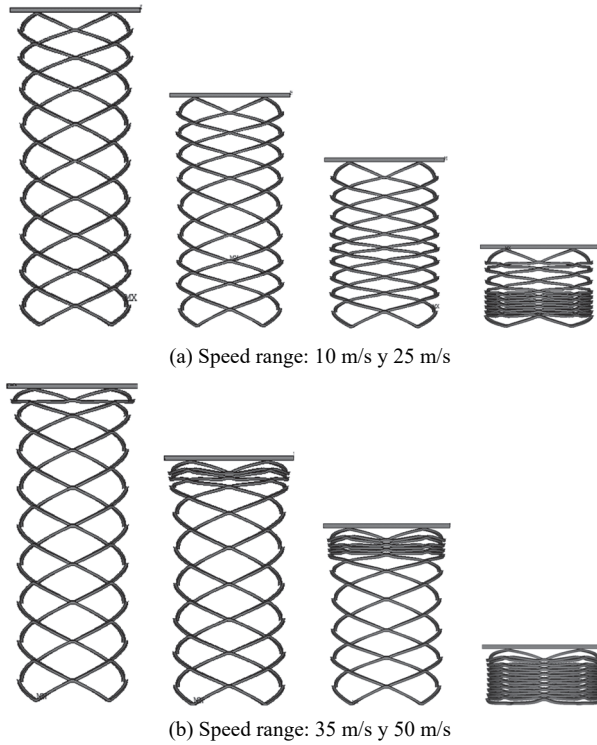
Figure 8. Progressive collapse for various speed ranges (α =0° - Model 2).
Source: The authors

(a) 90°/ Material Model 1

(b) 90°/ Material Model 2

Figure 9. Effect of the loading speed α =90°.
Source: The authors

Fig. 10a shows that at collapse, for speeds ranging from 10 m/s to 25 m/s, tubes with α =90° exhibit a widening in the middle, which originates from a buckling of the strands in the cells. This widening reaches its limit when the cells already involved are completely crushed, and results in the formation of a rigid circular ring near the middle section. Thereafter, the adjacent cells undergo progressive crushing. This ring reached a diameter of approximately two (02) times the initial tube diameter, as shown in Fig. 10a. Moreover, after the ring is shaped, the deformation of the specimen stops in the radial direction. It then begins to collapse axially through a combining crushing and buckling process of the cells, starting with those closest to the ring.

The deformation pattern changes for speeds ranging from 35 m/s up to 50 m/s (Fig. 10b). It can clearly be observed that, as the speed increases, the area of maximum deflection, due to local buckling, is the one closer to the impact zone. At low speeds, it approaches the area in the middle of the specimen (Fig. 10a). This seems reasonable, since in crash situations local deformation is first manifested in the impact zone: similar to the mushroom pattern observed for circular solid tubes in [7]. In the final crushing state, the load is transmitted throughout the whole structure.

3.3. Influence of the cell dimensions

In order to study the influence of the cell dimensions, three geometries are used in the analysis: namely H-24, H-26, and H-27A (Fig. 11). Both, the impact speed and mass remain constant: v=15 m/s, and M=20 kg. In this analysis, only material Model 2 is considered.

(a) Speed range: 10 m/s y 25 m/s

(b) Speed range: 35 m/s y 50 m/s

Figure 10. Progressive collapse for various speed ranges (α =90°-Model 2).
Source: The authors

H-24   H-26   H-27A
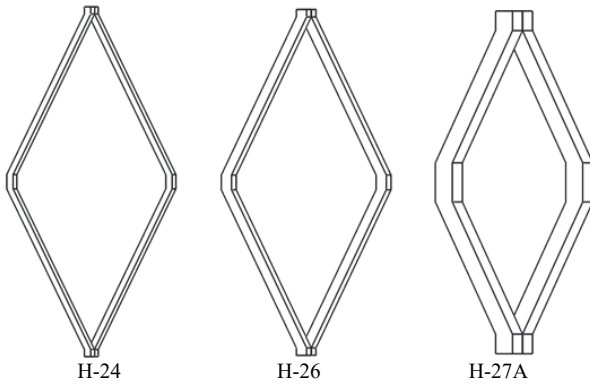
Figure 11. Cell sizes used in the analysis.
Source: The authors

Table 4.
Effect of the strand cross section.

α	Model	W [kg]	F_p [kN]	F_m [kN]	E_{200} [J]	E_e [J/kg]
0°	H-24	0.294	1.8	1.22	244.48	831.56
	H-26	0.508	2.69	2.29	458.21	901.99
	H-27A	1.068	19.14	15.25	3050.13	2855.9
90°	H-24	0.261	20.73	2.81	562.79	2156.3
	H-26	0.451	43.29	5.98	1195.58	2650.9
	H-27A	0.949	145.29	35.86	7172.22	7557.7

Source: The authors

Increasing the dimensions of the cross-section of the strands increases their bending strength; it can be expected that the impact strength also increases. Moreover, Table 4 shows that this increase in strength is very sensitive to changes in dimensions, for instance, the cross section for mesh H-27A is 3.75 times larger than for mesh H-26, but the increase in the mean loads increases from 6.5 times for α = 90°, to 7 times for α =0°.

Consequently, the absorbed energy per the tubes' unit mass increases, which is highly convenient when designing compact, lightweight and economic energy absorbing systems that are of great the interests for the automotive industry (guardrails and crash attenuators). Table 4 shows these values and manifests the energy per unit mass (E_e) or specific energy.

4. Conclusions

This paper investigated the effect of the elastoplastic behavior of the material on the impact response of expanded metal tubes. Two material models, one bilinear with strain hardening and other depending on strain-rate were analyzed. We observed that the best results were achieved for the strain-rate dependent model (Model 2). Regarding the failure mechanism, we observed that at low-speed ranges impact scenarios the mechanism is similar to the one observed in quasi-static testing. Conversely however, for high-speed ranges the failure mechanism changes substantially; the failures occur rather close to the impact zone and then propagate to the bottom of the tubes. This conclusion is valid for the two cell orientations that were investigated in this paper. In addition, an increase in the cell cross section leads to a higher degree of the expanded metal tubes' energy absorption capacity.

References

[1] Nagel, G.M. and Thambiratnam, D.P., A numerical study on the impact response and energy absorption of tapered thin-walled tubes. International Journal of Mechanical Sciences, 46(2), pp. 201-216, 2004. DOI: 10.1016/j.ijmecsci.2004.03.006.

[2] Alghamdi, A.A.A., Collapsible impact energy absorbers: An overview, Thin-Walled Structures, 39(2), pp. 189-213, 2001. DOI: 10.1016/S0263-8231(00)00048-3.

[3] Olabi A.G., Morris, E. and Hashmi, M.S.J., Metallic tube type energy absorbers: A synopsis, Thin-Walled Structures, 45(7-8), pp. 706-726, 2007. DOI:10.1016/j.tws.2007.05.003.

[4] Qiu, X.M. and Yu, T.X., Some topics in recent advances and applications of structural impact dynamics. Applied Mechanic Review, 64(3), pp. 1-12, 2011. DOI: 10.1115/1.4005571.

[5] Reid, S.R. and Reddy, T.Y., Static and dynamic crushing of tapered tubes of rectangular cross-section, International Journal of Mechanical Sciences, 28(9), pp. 623-637, 1986. DOI: 10.1016/0020-7403(86)90077-9.

[6] Cowper, G.R. and Symonds, P.S. Strain-hardening and strain-rate effects in the impact loading of cantilever beams, Brown University Providence, R. I., nº TR-C11-28, 1957.

[7] Wang, B. and Lu, G., Mushrooming of circular tubes under dynamic axial loading, Thin-Walled Structures, 40(2), pp. 167-182, 2002. DOI: 10.1016/S0263-8231(01)00057-X.

[8] Langseth, M., Hopperstad, O. and Berstad, T., Crashworthiness of aluminium extrusions: Validation of numerical simulation, effect of mass ratio and impact velocity, International Journal of Impact Engineering, 22(9-10), pp. 829-854, 1999. DOI: 10.1016/S0734-743X(98)00070-0.

[9] Hsu, S. and Jones, N., Quasi-static and dynamic axial crushing of thin-walled circular stainless steel, mild steel and aluminum alloy tubes, International Journal of Crashworthiness, 9(2), pp. 195-217, 2004. DOI: 10.1533/ijcr.2004.0282

[10] Rossi, A., Fawaz, Z. and Behdinan, K,. Numerical simulation of the axial collapse of thin-walled polygonal section tubes, Thin-Walled Structures, 43(10), pp. 1646-1641, 2005. DOI: 10.1016/j.tws.2005.03.001

[11] Jones, N., Energy-absorbing effectiveness factor. International Journal of Impact Engineering, 37, pp. 754-65, 2010. DOI: 10.1016/j.ijimpeng.2009.01.008

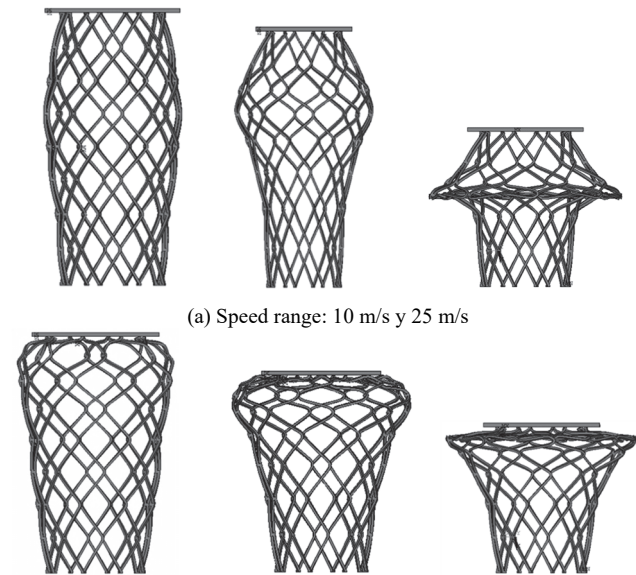[12] Jones, N., Structural Impact, New York: Cambridge University Press, 2011.

[13] Graciano, C., Martínez, G. and Smith, D., Experimental investigation on the axial collapse of expanded metal tubes, Thin-Walled Structures, 47, pp. 953-961, 2009. DOI: 10.1016/j.tws.2009.02.002

[14] Graciano, G., Martínez, G. and Gutiérrez, A., Failure mechanism of expanded metal tubes under axial crushing. Thin-Walled Structures, 51, pp. 20-24, 2012. DOI: 10.1016/j.tws.2011.11.001

[15] Martínez, G., Graciano, C. and Teixeira, P., Energy absorption of axially crushed expanded metal tubes. Thin-Walled Structures, 71, pp. 134-146, 2013. DOI: 10.1016/j.tws.2013.05.003

[16] Smith, D., Graciano, C. and Martínez, G., Quasi-static axial compression of concentric expanded metal tubes. Thin-Walled Structures, 84, pp. 170-176, 2014. DOI: 10.1016/j.tws.2014.06.012

[17] Nouri, M.D., Hatami, H. and Jahromi, A.G., Experimental and numerical investigation of expanded metal tube absorber under axial impact loading. Structural Engineering & Mechanics, 54(6), pp. 1245-1266, 2015. DOI: 10.12989/sem.2015.54.6.1245

[18] Hatami, H. and Nouri, M., Damghani. Experimental and numerical investigation of lattice-walled cylindrical shell under low axial impact velocities. Latin American Journal of Solids and Structures, 12(10), pp. 1950-1971, 2015. DOI: 10.1590/1679-78251919

[19] ANSYS LS-DYNA user's guide release 12.0 (2009).

[20] Sánchez, R., Determinación de las propiedades mecánicas de láminas de metal expandido. MSc. Tesis, Coordinación de Postgrado en Ingeniería Mecánica, Universidad Simón Bolívar, Caracas, Venezuela.

[En línea]. 2005. Disponible en: http://159.90.80.55/tesis/000129658.pdf

[21] Abramowics, W. and Jones, N., Dynamic progressive buckling of circular and square tubes. International Journal of Impact Engineering, 4(4), pp. 243-270. 1986. DOI: 10.1016/0734-743X(86)90017-5

[22] Kohar, C., Mohammadi, M., Raja, K. Mishra, R.K. and Inal, K., Effects of elastic-plastic behaviour on the axial crush response of square tubes. Thin-Walled Structures, 93, pp. 64-87. 2015. DOI: 10.1016/j.tws.2015.02.023 2015.

[23] Saavedra, E., Estudio numérico de estructuras fabricadas con mallas de metal expandido sometidas a cargas de impacto. Proyecto Final, Coordinación de Ingeniería Mecánica, Universidad Simón Bolívar, Caracas, Venezuela, [En línea]. 2011. Disponible en: http://159.90.80.55/tesis/000152468.pdf

[24] Smith, D., Estudio experimental de estructuras fabricadas con láminas de metal expandido sometidas a compresión axial. MSc. Tesis, Coordinación de Postgrado en Ingeniería Mecánica, Universidad Simón Bolívar, Caracas, Venezuela, [En línea]. 2008. Disponible en: http://159.90.80.55/tesis/000144470.pdf

[25] Gutiérrez, A., Análisis de los mecanismos de falla en estructuras de metal expandido sometidas a compresión. MSc. Tesis, Coordinación de Postgrado en Ingeniería Mecánica, Universidad Simón Bolívar, Caracas, Venezuela, 2010.

C. Graciano, received both his BSc. in 1992 and MSc. in 1995, in Mechanical Engineering from the Simon Bolivar University, Cafracas, Venezuela. He later moved to Sweden where he obtained a Licentiate of Engineering in 2001 and a PhD. in 2002, in Structural Engineering from Chalmers University of Technology and Luleå University of Technology, respectively. From 1997 to 2013, he served as assistant, associate and full professor in the Mechanical Engineering Department at the Simon Bolivar University, Caracas, Venezuela. Currently, he is an associate professor in the Civil Engineering Department at the Universidad Nacional de Colombia - Medellin campus. His research interests include: finite element modeling, structural stability, piping stress analysis and crashworthiness among others.
ORCID: 0000-0003-0659-7963.

G. Martínez, received a BSc. in 1996 in Civil Engineering from the Universidad Francisco de Miranda in Coro, Venezuela, an MSc. in 2000 in Civil Engineering and a PhD. in 2005 in Mechanical Engineering from the Central University of Venezuela. She is currently a full professor in the Mechanical Engineering Department at the Simon Bolivar University, Caracas, Venezuela. Her research interests include: finite element modeling, biomechanics, structural stability and crashworthiness among others.
ORCID: 0000-0003-0515-9400.

E. Saavedra, received a BSc. in 2011 in Mechanical Engineering from the Universidad Simon Bolivar, Caracas, Venezuela. He is currently a research assistant at the Universidad Simon Bolivar, in Caracas, Venezuela. His research interests include: finite element modeling, mechanical design, structural stability and crashworthiness among others.
ORCID: 0000-0002-1151-0677

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería Mecánica

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Mecánica

Mayor información:

E-mail: acmecanica_med@unal.edu.co
Teléfono: (57-4) 4259262

Phenomenological modeling and parametric identification applied to the monitoring of a belt conveyor system in a copper mine

Renan Landau Paiva-de Medeiros [a], Walter Barra Júnior [a], José Ruben Sicchar [b],
Florindo Antonio de Carvalho-Ayres Júnior [a] & José Augusto Lima-Barreiros [a]

[a] Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Pará, Belén, Brasil.renanlandau@yahoo.com.br
[b] Faculdade de Controle e Automação, Universidade do Estado do Amazonas, Amazonas, Brasil. jvilchez@uea.edu.br

Abstract
This paper investigates the modeling of a belt conveyor system for use in an automated monitoring system aiming at obtain improved energetic performance and fault diagnosis capability of the system. Firstly, a phenomenological model of the process is discussed, which is based on the laws of mechanics considering various movement opposing forces. The main parameters of the belt conveyor system model were estimated by using a non-recursive least mean square identification techniques. The results of computational studies show a good efficiency of the proposed strategies when applied to a set of data collected in a real world belt conveyor system, located at copper mine in south of Pará state, Brazil.

Keywords: Parametric identification, phenomenological model, Belt Conveyor, non-recursive least mean square, Monitoring and Diagnostics.

Modelado fenomenológico e identificación paramétrica aplicada al monitoreo de una correa transportadora en una mina de cobre

Resumen
En este se investiga el modelado de un sistema del tipo correa transportadora con la finalidad de uso en el monitoreo y en el diagnóstico de faltas. Primeramente es discutido un modelo fenomenológico del proceso, el cual es basado en la aplicación de las leyes de la mecánica y considerando los diversos tipos de fuerza de oposición al movimiento que actúan sobre la correa. Los principales parámetros del transportador fueron estimados utilizando técnicas de identificación basadas en Mínimos Cuadrados No Recursivos. Los resultados obtenidos en estudios computacionales muestran el buen desempeño de la estrategia propuesta, cuando aplicada a datos reales colectados en el sistema de transporte de minerales de una mina de cobre al sur de Pará, Brasil.

Palabras clave: Identificación Paramétrica, Modelado Fenomenológico, Correa Transportadora, Mínimos Cuadrados No Recursivo, Monitoreo y Diagnóstico.

1. Introducción

Las máquinas y equipos de transporte son utilizados para mover los más diversos tipos de cargas en el interior de fábricas. Los procesos de transporte no se limitan apenas en mover materiales de un lugar para otro, pero incluyen también, operaciones de carga y descarga, de modo que su mal funcionamiento puede ocasionar pérdidas de producción, de calidad y de rentabilidad [6,8].

Los sistemas del tipo correa transportadora, poseen una posición dominante cuanto al transporte de cargas al grano. Esto se debe a su versatilidad de uso y algunas ventajas, como: Integrar innúmeros procesos garantizando el flujo continuo del material entre dos operaciones sucesivas, transportarlos más diversos tipos de materiales, adaptación a los más variados tipos de terreno, tienen también algunas ventajas ambientales, tales como: poder operar dentro de pequeños túneles, evitando la formación de polvo, y operar silenciosamente [2,4,8].

La necesidad de mejorarla eficiencia de procesos de

transformación de minerales ha abierto un importante campo de investigaciones. En Cores et. al, (2015) [13], los autores investigan estrategias de automación y control aplicados a un proceso de sinterización de minerales de hierro. La calidad del proceso de sinterización de minerales de hierro fue investigada en Mochón et. al., (2014) [12].

Los sistemas del tipo correa transportadora de minerales están entre los más eficientes medios de transporte en cortas y medias distancias. Sin embargo, este sistema tiene un costo elevado, en lo que dice respecto al consumo energético [5]. Según [1], se estima que cerca de 40% del costo operacional es debido al gasto energético. Debido a tal hecho, trabajos visando la mejoría de la eficiencia energética de los equipos constituyentes de este tipo de sistema son de extrema importancia para la operación económica de los procesos productivos, contribuyendo de modo a reducción de los costos operacionales y preservación de los equipos [5,8]. En este trabajo se investiga el modelado de un sistema del tipo correa transportadora, así como una metodología de estimación de los parámetros constituyentes del sistema. El objetivo es detectar un punto óptimo de operación, bien como desarrollar un sistema de monitoreo y de detección de faltas. Inicialmente es discutido un modelo fenomenológico del proceso, basándose en las leyes de la mecánica, focalizando en los diversos tipos de fuerzas de oposición al movimiento. En seguida, es propuesta una metodología de estimación paramétrica, basada en el método de Mínimos Cuadrados No Recursivo (MCNR). Para realizar la validación del algoritmo desarrollado, técnicas basadas en análisis de correlación fueron utilizadas, de modo que estudios iníciales fueron realizados utilizándose un modelo propuesto en [1,2]. Por fin, teniendo el modelo fenomenológico y el algoritmo de estimación calibrado y validado, fue investigado un sistema real del tipo correa transportadora, utilizado en una mina de cobre al sur del estado del Pará- Brasil.

En este trabajo, son presentados los resultados obtenidos de los modelos aproximados, atravesá de pruebas experimentales, los modelos obtenidos serán usados en un algoritmo de detección de faltas para ser investigado en estudios posteriores a ser desarrollado por los autores.

Este artículo está organizado de la siguiente forma: en la sección 2 se presenta el modelado fenomenológico del sistema del tipo correa transportadora, en la sección 3 se presenta el desarrollo del algoritmo de estimación paramétrica del modelo fenomenológico utilizando MCNR, en la sección 4 son presentados los resultados obtenidos para la realización de la calibración del algoritmo de estimación paramétrica, en la sección 5 son presentados los resultados obtenidos de la estimación paramétrica de un sistema real del tipo correa transportadora, en la sección 6 se tiene la conclusión del trabajo.

2. Modelado fenomenológico de la correa transportadora

El modelado fenomenológico del Sistema del tipo correa transportadora es de importancia fundamental para diversas situaciones, tales como: obtención del punto óptimo de operación del sistema buscándose mayor eficiencia energética, constitución de herramientas que operen como
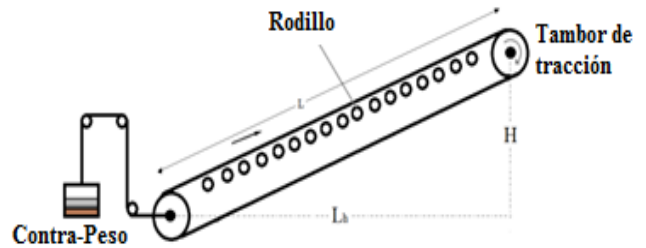
Figura 1. Esquemático del modelo genérico de una correa transportadora.
Fuente: Adaptado [1].

supervisores del sistema así como elaboración de algoritmos para detectar y/o diagnosticar faltas ocurridas en el sistema. En la literatura [1, 2, 3] son difundidos dos métodos distintos para la obtención de un modelo fenomenológico. El primer método es basado en el cálculo de las fuerzas de resistencia que se oponen al movimiento de la correa el otro método se basa en la conservación de la energía atravesá de una longitud de compensación. Este trabajo fue elaborado utilizando el modelo propuesto por [1], donde se propone una inter-ligación entre los dos enfoques. La Fig. 1 muestra el diagrama esquemático simplificado de una correa transportadora típica de minerales, con sus partes que la constituyen.

En la Fig. 1, L(m) es la longitud de transporte de la correa, la cual es medida en relación a los centros de los rodillos de las extremidades, las cuales promueven la tracción y el movimiento del sistema; L_h(m) es la distancia horizontal y H(m) es la altura de elevación.

Según [1-4], en el régimen permanente, el consumo de potencia eléctrica activa es predominante debido a diversas fuerzas de resistencia al movimiento de la sección cargada y de la sección de retorno de la correa transportadora. Equipos tales como "*belt cleaners*" y "*skirt boards*" también contribuyen considerablemente para el consumo de la potencia total de la energía. Las normas ISO 5048 y DIN 22101 clasifican a las fuerzas de resistencia al movimiento de acuerdo con las siguientes clases:

2.1. *Fuerza de resistencia primaria – (FH)*

La fuerza de resistencia primaria es una composición de las fuerzas de fricción a lo largo de todo el trayecto de la correa. Esta fuerza puede ser calculada atravesá de la siguiente formula:

$$F_H = fgL[(Q_G + 2Q_B)\cos(\delta) + Q_{RO} + Q_{RU}] \qquad (1)$$

Donde:
f – Factor de fricción artificial.
L – Longitud de la correa transportadora.
g – Aceleración de la gravedad
Q_G - Masa por unidad de longitud de la carga debida al material siendo transportado en la estera (kg/m).
$2Q_B$ - Masa por unidad de longitud de la estera vacía (kg/m).
δ-Ángulo de inclinación de la correa (º).

Q_{RO} - Masa por unidad de longitud de las partes rotativas de los rodillos en los trayectos de ida (kg/m).

Q_{RU}- Masa por unidad de longitud de las partes rotativas de los rodillos en los trayectos de retorno (kg/m).

2.2. Fuerza de resistencia secundaria – (FN)

La fuerza de resistencia secundaria es compuesta por fuerzas de naturaleza inercial y también de fricción. Son fuerzas de fricción del material con la correa, en las estaciones de alimentación, así como en las *skirt boards*. La norma ISO 5048 propone el cálculo de la fuerza secundaria, através de una relación con la fuerza primaria, como presentada en (2)-(3). Conforme puede ser observado F_N varia con la longitud de la correa transportadora.

$$F_N = (C(L) - 1)F_H \qquad (2)$$

$$C(L) = \begin{cases} 0.85 + 13.31L^{-0.576} & , \text{para } (10 \le L < 1840) \\ 1.025 & , \text{para } (L \ge 1840) \end{cases} \qquad (3)$$

2.3. Fuerza de resistencia de slop – (F_{st})

La fuerza de resistencia de slop se debe a la variación relativa del punto de alimentación para el punto de flujo, de tal manera que esta fuerza puede ser agregada o reducida dependiendo de la variación de la altura, esta fuerza puede ser expresada como se indica en (4). Esta depende de la carga y de la altura de elevación de la correa.

$$F_{st} = Q_G g H \qquad (4)$$

2.4. Fuerza de resistencia especial – (Fs)

La fuerza de resistencia especial es una composición de las demás fuerzas de resistencia al movimiento, pudiendo ser expresada conforme la Ec. (5).

$$F_S = k_1 \frac{T^2}{V} + k_2 \frac{T}{V} + k_3 \qquad (5)$$

Donde los parámetros k_1, k_2 e k_3 dependen de la estructura y de la geometría de la correa transportadora [1,2].Las variables T y V son respectivamente la tasa de alimentación y la velocidad tangencial de la correa transportadora.

2.5. Fuerza de resistencia total – (F_U)

La fuerza de resistencia total es compuesta por la sumatoria de todas las fuerzas de resistencia que actúan sobre la correa transportadora. Esta es definida según la Ec. (6).

$$F_U = F_H + F_N + F_{st} + F_S \qquad (6)$$

De acuerdo con las leyes de la mecánica, la potencia mecánica total necesaria para equilibrar las fuerzas de resistencia al movimiento, puede ser expresada conforme la Ec. (7).

$$P_M = F_U V \qquad (7)$$

Donde V es la velocidad tangencial de la correa transportadora. La potencia eléctrica activa total que los motores eléctricos y sus respectivos *drives* de accionamiento necesitan consumir de la red eléctrica, para suministrarla potencia mecánica total da correa. Puede ser calculada através de la Ec. (8).

$$P_T = \frac{1}{\eta} P_M \qquad (8)$$

Donde η es el rendimiento del conjunto motor-*driver*.

Substituyendo las Ec. (1)-(5) en la Ec. (6) y en seguida en la Ec. (7)-(8),la potencia eléctrica consumida por el sistema correa transportadora y carga, puede ser expresada conforme se indica en (9)-(12).

$$P_{ec} = \frac{1}{\eta} \left[\begin{array}{l} fg(L\cos(\delta) + L_{01})QV + \\ fg(L\cos(\delta) + L_{02})\frac{T}{3.6} + gH\frac{T}{3.6} + Pacs \end{array} \right] \qquad (9)$$

$$Q = 2Q_B + Q_{RO} + Q_{RU} \qquad (10)$$

$$L_{01} = L(1 - \cos(\delta))(1 - \frac{2Q_B}{Q}) \qquad (11)$$

$$L_{02} = \frac{V}{gf} \left(V + \frac{T}{1.8b_1^2 \rho} + \frac{3.6C_{ft}}{T} \right) \qquad (12)$$

3. Algoritmo de estimación de mínimos cuadrados no recursivos (MCNR)

La identificación paramétrica de sistemas, puede ser vista como un problema de optimización que envuelve algunas restricciones para la adecuación de modelos candidatos para representar un determinado proceso real. La selección de modelos y el ajuste de sus parámetros son influenciados por diversos factores, tales como: Conocimiento a *priori* del sistema, propiedades del modelo a ser identificado, selección de la medida del error a ser minimizado y la presencia de ruidos [4,11]. Como el problema de optimización en muchos casos se torna un problema interesante, dado que se enfoca la minimización de un determinado criterio de desempeño. Si todas las restricciones impuestas al sistema fueron satisfechas, el modelo encontrado pode ser considerado satisfactorio. Caso contrario, si por lo menos una de las restricciones fue violada, todos los procedimientos de identificación y estimación paramétrica y diagnóstico del modelo deberán ser reevaluados hasta que un modelo apropiado sea encontrado [4,9-11].

Considerando un determinado proceso físico, operando en un punto de operación, un modelo de tiempo discreto puede ser expresado bajo la forma:

$$y(t) = -a_1 y(t-1) - \ldots - a_{na} y(t-na)$$
$$+ b_0 u(t-d) + b_1 u(t-d-1) + \ldots \qquad (13)$$
$$+ b_{nb} u(t-d-nb)$$

Donde:
y(t) – Salida del sistema
u(t) – Entrada del sistema
a_n e b_n – Parámetros del sistema.

Puede reescribirse la Ec. (13), en un formato matricial, conforme se indica en (14)-(16).

$$\overline{\Phi} = [-y(t-1),\ldots,-y(t-na),$$
$$u(t-d),\ldots,u(t-d-nb)] \qquad (14)$$

$$\overline{\theta}^T = [a_1, a_2, \ldots, a_{na}, b_0, b_1 \ldots, b_{nb}] \qquad (15)$$

$$\overline{Y} = \overline{\Phi}^T . \overline{\theta} + \overline{E} \qquad (16)$$

Donde θ y Φ, son respectivamente, el vector de parámetros y el vector de regresores.

La estimativa del vector de parámetros puede ser obtenida através del procedimiento de mínimos cuadrados. Utilizando la estimativa, la mejor previsión de la salida del sistema, es calculada como se muestra en las Ec. (17)-(18) [4,9-11].

$$\hat{\theta} = \left[\hat{\Phi}^T \hat{\Phi}\right]^{-1} \hat{\Phi}^T \overline{Y} \qquad (17)$$

$$\hat{Y}(t) = \Phi(t) . \hat{\theta} \qquad (18)$$

De modo que el error de previsión puede ser determinado conforme se indica en (19).

$$\varepsilon = \overline{Y} - \hat{Y}(t) = \overline{Y} - \Phi(t) . \hat{\theta} \qquad (19)$$

3.1. Estimación y calibración de los parámetros del modelo fenomenológico de la correa transportadora

El modelo matemático que describe el comportamiento estático del sistema, basado en el cálculo del equilibrio energético de la correa transportadora, es obtenido através de la relación entre las Ec. (9)-(12).

Actualmente, muchas empresas aún utilizan métodos convencionales de operación, en los cuales los puntos de operación de las correas transportadoras no son optimizados desde el punto de vista del gasto energético. Ese hecho, se debe principalmente la complejidad del sistema y la ausencia de modelos confiables de la planta, visto que diversos parámetros de éste modelo son de difícil obtención vía medición. De esta forma, se plantea un modelo de cálculo energético modificado, el cual viabiliza la parametrización y calibración del modelo planteado, de forma fácil y objetiva [1,4].

De acuerdo con la norma ISO 5048, la fuerza de resistencia total, puede ser obtenida através de la relación entre la tasa de alimentación de la correa transportadora (T) y la velocidad de la correa transportadora (V), para representar esta fuerza, se puede despejar los coeficientes de los parámetros a ser estimados y reescribiendo la Ec.(9) obteniendo la Ec.(20).

$$P_T = \frac{V^2 T}{3.6} + T^2 V \theta_1 + V \theta_2 + \frac{T^2}{V} \theta_3 + T \theta_4 \qquad (20)$$

Donde $\theta_1, \theta_2, \theta_3$ e θ_4 son parámetros empíricos a determinar vía pruebas. Tales parámetros, están relacionados con los parámetros físicos del sistema de la correa transportadora, definidos por las expresiones:

$$\theta_1 = \frac{1}{6.48 b_1^2 \rho} \qquad (21)$$

$$\theta_2 = gf(L_h + L_{01})Q + k_s + C_{Ft} \qquad (22)$$

$$\theta_3 = k_1 \qquad (23)$$

$$\theta_4 = \frac{gH + gfL_h}{3.6} + k_2 \qquad (24)$$

Con el intuito de aplicar la técnica de identificación MCNR, se reescribió la Ec.(20) conforme se indica en (25)-(30).

$$\frac{P_T}{M} - \frac{V^2 T}{3.6M} = \left(\frac{m_1}{M}\theta_1\right)\frac{T^2 V}{V m_1} + \left(\frac{m_2}{M}\theta_2\right)\frac{V}{m_2} +$$
$$\left(\frac{m_3}{M}\theta_3\right)\frac{T^2}{m_3 V} + \left(\frac{m_4}{M}\theta_4\right)\frac{T}{m_4} \qquad (25)$$

$$m_1 = \max(T^2 V) \qquad (26)$$

$$m_2 = \max(V) \qquad (27)$$

$$m_3 = \max(\frac{T^2}{V}) \qquad (28)$$

$$m_4 = \max(T) \qquad (29)$$

$$M = \max(P_T - \frac{V^2 T}{3.6}) \qquad (30)$$

Reescribiendo, entonces, la Ec. (25) en el formato matricial se obtiene:

$$P_T = \Phi^T . \theta \qquad (31)$$

$$\Phi^T = \left[\frac{T^2 V}{m_1} \quad \frac{V}{m_2} \quad \frac{T^2}{m_3 V} \quad \frac{T}{m_4}\right]^T \qquad (32)$$

$$\theta = \left[\frac{m_1}{M} \theta_1 \quad \frac{m_2}{M} \theta_2 \quad \frac{m_3}{M} \theta_3 \quad \frac{m_4}{M} \theta_4 \right] \qquad (33)$$

La Fig. 2 presenta el flujograma detallado del procedimiento utilizado para la realización de la estimación paramétrica de los parámetros físicos y mecánicos de la correa transportadora. Primeramente los datos de potencia eléctrica consumida, tasa de alimentación y velocidad tangencial de la correa son colectados en el campo (A). En seguida, es efectuada la estimación de los parámetros del modelo (B). La Ec. (20) es entonces utilizada para estimar la potencia consumida por el sistema (C). El valor de potencia estimada es comparada al valor real medido, P_T (D). Caso el valor del error calculado, esté abajo de una tolerancia previamente determinada, el modelo es considerado ajustado (E), de modo que los parámetros encontrados pueden ser utilizados en simulaciones de laboratorio.
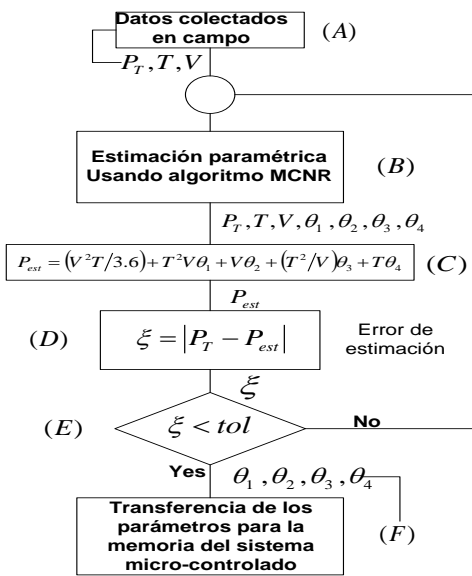
Figura 2. Flujograma de la metodología de estimación paramétrica. Fuente: Los Autores.

Tabla 1.
Valores de los parámetros de una correa transportadora.

Param.	Valor	Uni.	Descripción
L	1000	m	Longitud de la correa
ρ	900	kg/m³	Densidad del material
G	9,8	m/s²	Aceleración de la gravedad
Q_{RU}	7,76	kg/m	Unidad de masa de las partes rotativas de los rodillos de retorno
V	3,15	m/s	Velocidad tangencial de la correa transportadora
H	9,98	m	Altura de elevación
b_1	1,4	m	Intervalo de las *skirt-boards*
Q_B	18,73	kg/m	Unidad de masa de la correa
Q_G	176,3	kg/m	Unidad de masa de carga
P_{acs}	0,0	W	Potencia consumida debido a los equipos accesorios
δ	1,825	°	Angulo de inclinación
f	0,024	-	Factor de fricción artificial
T	2000	t/h	Tasa de alimentación de la correa
Q_{RO}	15,75	kg/m	Unidad de masa de las partes rotativas de los rodillos del camino directo

Fuente: Zhang, 2009.
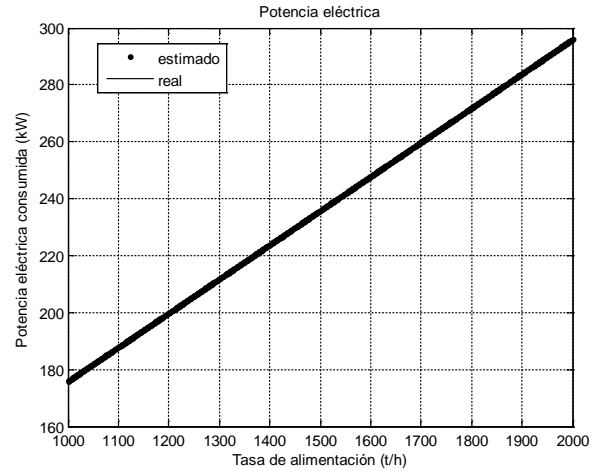
Figura 3. Potencia eléctrica consumida cuando el sistema es sometido a una variación de la tasa de alimentación de 1000-2000 t/h.
Fuente: Los Autores.

3.1. Validación del algoritmo de estimación paramétrica

Con el objetivo de realizar la validación y calibración del algoritmo de estimación paramétrica desarrollada, se utilizó un conjunto de datos obtenido en [1], donde se presentan datos constructivos a cerca de los parámetros físicos y mecánicos del sistema, de modo que este posibilite realizar el cálculo de los parámetros. La Tabla 1 presenta los valores de los parámetros obtenidos en [1], referente a una correa transportadora.

Con el intuito de realizar la validación del algoritmo desarrollado, se utilizó los datos de la Tabla 1, y se varió la tasa de alimentación de la correa transportadora de 1000 t/h a 2000 t/h, y se verificó el consumo da potencia eléctrica y el valor estimado de potencia por el algoritmo MCNR

La Fig. 3 presenta el gráfico de la potencia consumida y estimada cuando el sistema es sometido a una variación de la tasa de alimentación de la correa transportadora.

Se puede observar, en la Fig. 3, que los valores estimados de potencia eléctrica, se igualaron al valor real generado por el simulador, tal cual como deseado.

La Fig. 4 presenta el error absoluto de estimación de la potencia eléctrica de la correa transportadora. Se puede observar que este error es bastante pequeño, mostrando la adecuación del modelo estimado.

Para averiguar la robustez de la estimación *off-line* desarrollada, fue fijada la velocidad y la tasa de alimentación como muestra la Tabla 1 y fue adicionada a la tasa de alimentación un componente aleatorio de 1% de su valor nominal y en seguida fueron estimados los valores de potencia eléctrica consumida y de los parámetros θ's relacionados. La Fig. 5 presenta la potencia eléctrica consumida cuando adicionado un componente aleatorio de 1% a la tasa de alimentación.

Se puede verificar, en la Fig. 5, que mismo con una variación aleatoria de 1% de la tasa de alimentación, el algoritmo consigue estimar con gran precisión los valores de potencia consumidos del sistema. La Fig. 6 muestra el error absoluto de estimación cuando el sistema es sometido a una variación de 1% de su tasa de alimentación.
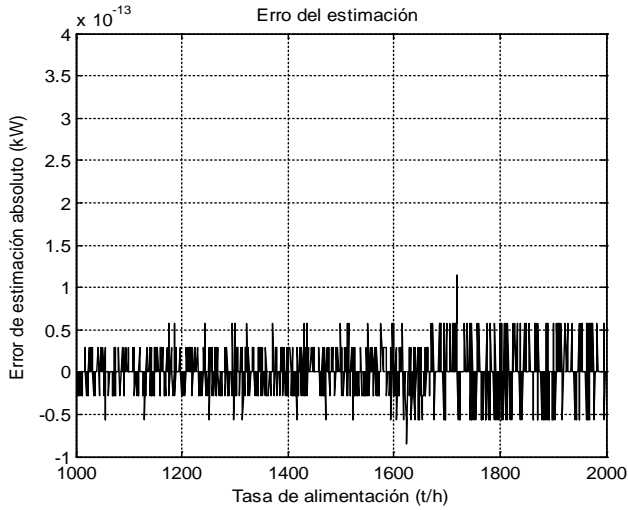
Figura 4. Error absoluto de estimación de la potencia eléctrica consumida.
Fuente: Los Autores.
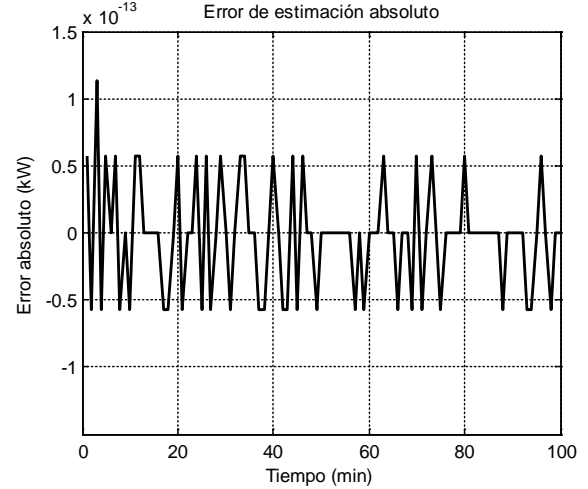
Figura 6. Error absoluto de la potencia eléctrica consumida.
Fuente: Los Autores.
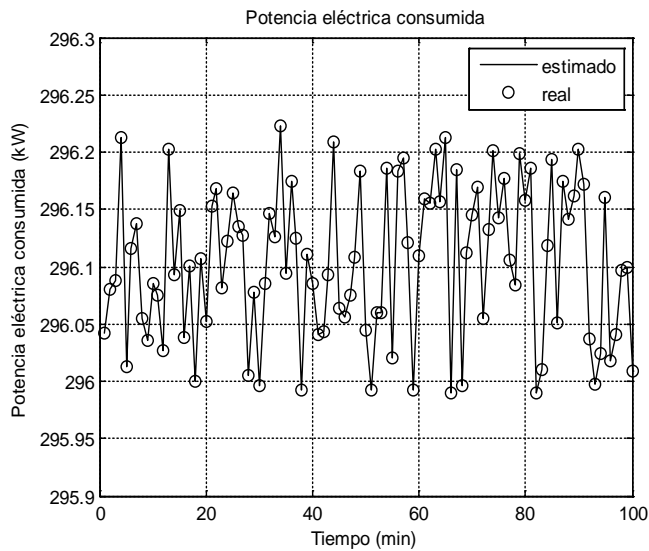
Figura5. Potencia eléctrica consumida cuando adicionado un componente aleatorio de 1% a la tasa de alimentación.
Fuente: Los Autores.
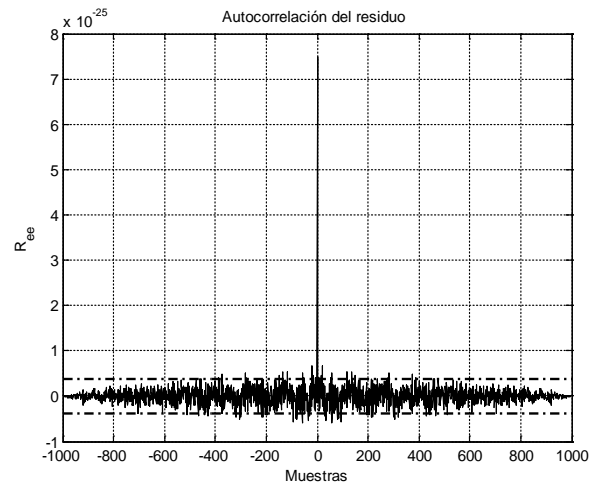
Figura 7. Autocorrelação de los residuos del sistema estimado.
Fuente: Los Autores.

En la Fig. 6, se verifica que el error de estimación, está muy pequeño, indicando cierta robustez del algoritmo desarrollado. En la Fig. 7 se presenta el resultado del análisis de correlación del vector de residuos del modelo.

Puede ser observado en la Fig. 7, que la auto-correlación presenta un impulso en (k = 0) y aproxima el valor cero para los valores donde k es diferente de cero, dentro de una faja de tolerancia estadística de orden de 5%.En la Fig. 8se presentan los valores estimados de los parámetros teta, con el intuito de validar el algoritmo desarrollado, fueron trazados los valores calculados de los parámetros utilizando los valores obtenidos en la Tabla 1 y fueron también trazados, los valores de los parámetros estimados, en ella referida prueba.
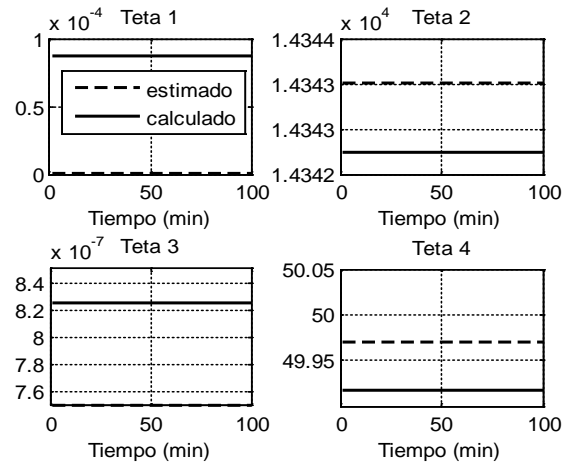
Figura 8. Relación de los parámetros calculados y de los parámetros estimados por el algoritmo desarrollado.
Fuente: Los Autores.
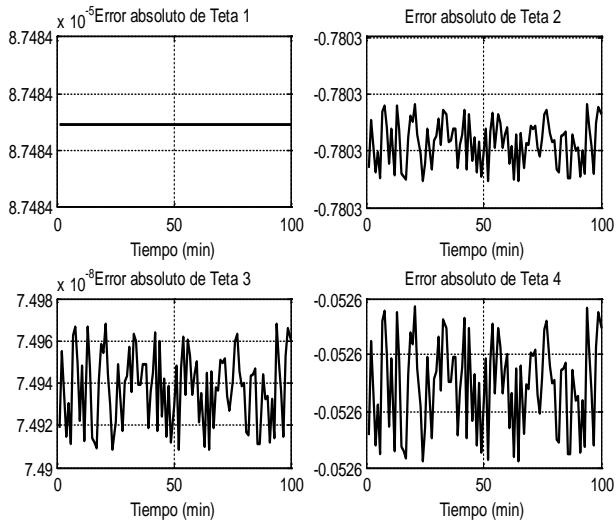
Figura 9. Error relacionado a las estimativas de cada parámetro de la correa transportadora.
Fuente: Los Autores.

En la Fig. 9 se muestra el error absoluto referente a los valores de los parámetros θ estimados. Se verifica que el mayor error de estimación obtenido fue en la variable θ_2, con el fin de rectificar la validación del algoritmo de estimación y denotar su eficacia.

4. Aplicación del algoritmo de estimación paramétrica en una correa transportadora de minerales de cobre al sur del estado de Pará- Brasil

Inicialmente fue desarrollada una metodología de estimación paramétrica *off-line*, con el intuito de estimar un modelo que represente de forma satisfactoria y fidedigna a un sistema del tipo correa transportadora de una mina de cobre localizada al sur del estado de Pará, en Brasil. Este sistema tiene las siguientes características: El transportador tiene una longitud de 4,1 km, operando de forma continua a una velocidad fija de 4,17 m/s, no existe un control rígido en la tasa de alimentación de la correa transportadora, debido a esto el valor estimado de tasa varía en una faja de error estimada, el sistema de accionamiento de la correa es compuesta por 4 motores trifásicos de 200 kW, de modo que todos operen sincronizados, resultando en una potencia consumida para el accionamiento de aproximadamente 800 kW, todos los motores son accionados vía inversor de frecuencia visando mantener fija a la velocidad de la correa. Este sistema también dispone de diversos equipos que se toman como accesorios, tales como: raspadores, viradores de correa, *skirtboards* de modo que consumen una parcela de potencia adicional que es prevista en el modelo fenomenológico através del término P_{acs}, como indica la Ec. (9). Con el intuito de obtener un modelo que represente adecuadamente al sistema de la correa transportadora investigada, fue constituido un banco de datos con medidas de las principales variables del sistema, como: Potencia eléctrica consumida por los motores, velocidad tangencial de la correa transportadora y la variación de la tasa de alimentación de minerales de cobre del sistema. La Tabla 2 muestra las características operacionales del sistema investigado.

Tabla2.
Características operacionales de la correa transportadora de una mina de cobre al sur del estado de Pará – Brasil.

Param.	Valor	Uni.	Descripción
L	4100	m	Longitud de la correa transportadora
ρ	8960	kg/m³	Densidad del Cobre
g	9,8	m/s²	Aceleración gravitacional
V	4,17	m/s	Velocidad tangencial de correa
T	2000±500	t/h	Tasa de alimentación de minerales en la correa

Fuente: Los Autores.

La Fig. 10 muestra la correa transportadora investigada en este estudio.

Figura 10. Correa transportadora de una mina de cobre al sur del estado de Pará - Brasil.
Fuente: Los Autores.

La Fig. 11 muestra a los sistemas de accionamiento de la correa transportadora.

Figura 11. Sistema de accionamiento de una correa transportadora de una mina de cobre al sur del estado de Pará - Brasil.
Fuente: Los Autores.

La Fig. 12 presenta los datos colectados en campo referentes a la potencia eléctrica consumida por el sistema, mostrando también la estimativa de potencia eléctrica consumida
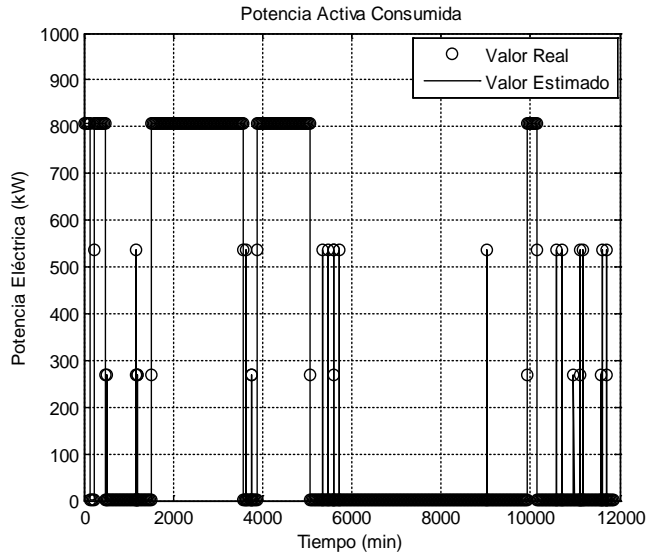
Figura 12. Potencia eléctrica consumida y potencia eléctrica estimada por el modelo propuesto.
Fuente: Los Autores.
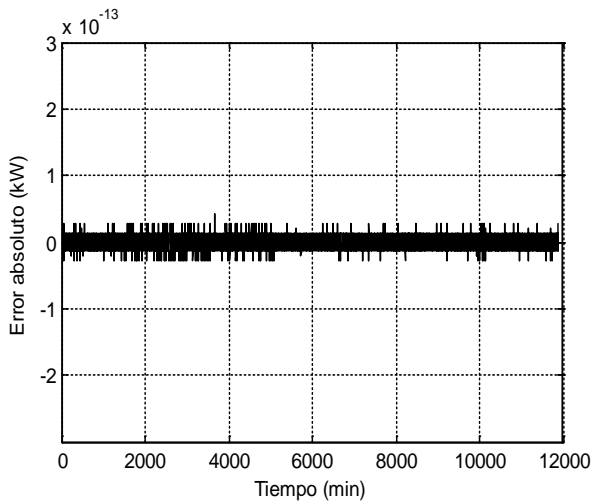
Figura 13. Error absoluto relacionado ala estimativa de potencia eléctrica del modelo propuesto.
Fuente: Los Autores.

utilizando el algoritmo desarrollado. Se verifica que la potencia eléctrica consumida es representada através de la sumatoria de todas las potencias de los motores responsables por el movimiento de la correa transportadora, proporcionando una potencia consumida máxima de 800 kW.

La Fig. 13muestra el error absoluto que relaciona la estimativa de la potencia eléctrica consumida por el sistema, se verifica que el error obtenido de la estimativa de potencia es muyo pequeño denotando un modelo estimado preciso y que representa adecuadamente al sistema investigado.

En la Fig. 14 se presenta el resultado del análisis de correlación del vector de residuos del modelo basado en datos colectados en campo.

En la Fig. 14 se verifica, que la auto-correlación presenta un impulso en (k = 0) y aproxima el valor cero para los valores
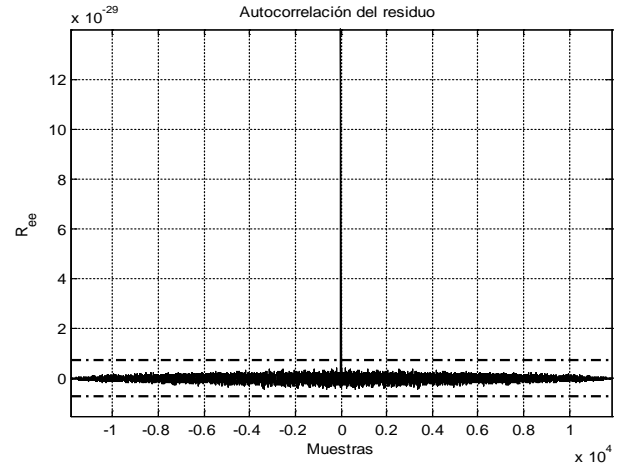
Figura 14. Auto-correlación de residuo del sistema estimado.
Fuente: Los Autores.
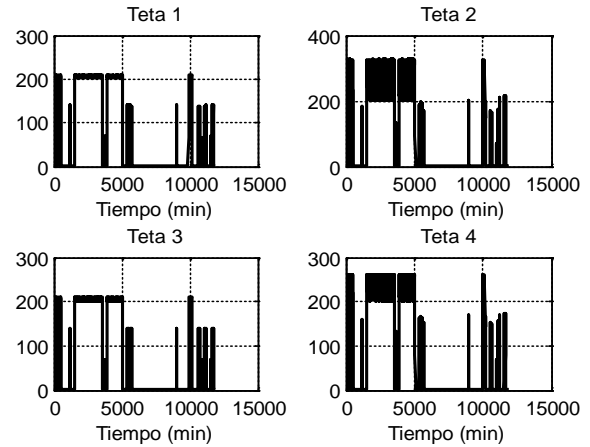
Figura 15. Valores estimados de los parámetros tetas para los datos reales.
Fuente: Los Autores.

donde k es diferente de cero, dentro de una faja de tolerancia estadística de orden de 5%.La Fig. 15 muestra las estimativas de los parámetros tetas referentes al modelo estimado de los valores reales adquiridos, se verifica que los valores referentes a θ_1 e θ_3 son valores bajos como ya esperados debido a sus formulaciones matemáticas.

5. Conclusión

En este trabajo fueron presentados los resultados referentes a pruebas computacional es referentes al modelado fenomenológico de un sistema del tipo correa transportadora, seguido del desarrollo de una estrategia, basada en mínimos cuadrados con el objetivo de realizar la estimación paramétrica del modelo propuesto, permitiendo la realización de estimación de los parámetros físicos y mecánicos del sistema que son de difícil obtención vía medición, de datos de campo en el sistema real, así como realizar la calibración del simulador para representar adecuadamente al

comportamiento del sistema. Los datos utilizados en esta investigación fueron de dos tipos y en dos etapas. Inicialmente fueron utilizados datos obtenidos de la literatura, teniendo como objetivo validar al algoritmo de estimación y validar el simulador desarrollado para representar al sistema en análisis. En seguida con el simulador y el algoritmo de estimación validados, fueron utilizados datos reales colectados en campo, de una correa transportadora de minerales de cobre de una mina al sur del estado de Pará – Brasil. Los resultados obtenidos mostraron un buen desempeño de los modelos estimados, tales modelos están siendo incorporados a un sistema industrial de monitoreo y supervisión, así como auxiliar en la detección y en el diagnóstico de faltas en el sistema.

Agradecimientos

Bibliografía

[1] Zhang, S. and Xia, X., A new energy calculation model of belt conveyor. IEEE AFRICON 2009, Nairobi-Kenya, 2009.

[2] Zhang, S. and Xia, X., Modeling and energy efficiency optimization of belt conveyors. Applied Energy, 88(9), pp. 3061-3071. 2011. DOI: 101016/japenergy201103015

[3] Zhang, S. and Tang, Y., Optimal control of operation efficiency of belt conveyor. IEEE Power and Energy Engineering Conference (APPEEC), Asia-Pacific, 2011.

[4] Medeiros, R.L.P., Estratégia de identificação paramétrica aplicada à modelagem fenomenológica de um sistema do tipo correia transportadora industrial para fins de detecção de faltas. MSc. Thesis, Programa de Pós-Graduação em Engenharia Elétrica – PPGEE, Universidade Federal do Pará – UFPa. Belém, Pará, Brasil. 2014

[5] Rudenko, N., Máquinas de elevaçãos e transporte. Livros Técnicos e Científicos editora, S.A.

[6] Nuttal, A.J.G. and Lodewijks, G., Dynamics of multiple drive belt conveyor system. WILLEY-VHC Verlag GmbH & Co. KGaA, Weinheim, ppsc-journal, pp. 365-369, 2007. DOI: 10.1002/ppsc.200601118.

[7] Wiid, A.P., Sithole, F., Bagus, M. e Khosa, T.H., Constant speed versus variable speed operation for belt conveyor system. IMHC. 2009.

[8] Nogueira, F.J.H., Programa de eficientização industrial: Módulo de correias transportadoras. Eletrobrás. s.a.

[9] Ljung, L., System idetification: Theory for the user. Prentice-Hall, Upper Saddler River, NJ, 1999.

[10] Aguirre, L.A., Introdução à identificação de sistemas: Técnicas lineares e não_lineares aplicadas a sistemas reais, Editora UFMG, Belo Horizonte – MG, 2007.

[11] Coelho, A.A.R. and Coelho, L.S., Identificação de sistemas dinâmicos lineares. Editora da UFSC, Florianópolis – SC, Brasil, 2004.

[12] Mochón, J., Cores, A., Ruiz-Bustinza, I., Verdeja, L.F., Robla, J.I. and Garcia-Carcedo, F., Iron ore sintering Part 2. Quality indices and productivity. DYNA, 81(183), pp. 168-177, 2014.

[13] Cores, A., Verdeja, L.F., Ferreira, S., Ruiz-Bustinza, I., Mochón, J., Robla, J.I., Gasca, C.G., Iron ore sintering Part 3. Automatic and control system. DYNA, 82(190), pp. 227-236, 2015. DOI: 10.15446/DYNA.v82n190.44054

R.L.P. de Medeiros, received his BSc. Eng. and MSc. degrees from the Federal University of Pará - UFPA, Brazil, in 2013, with an emphasis in control and monitoring industrial systems. Currently pursuing a PhD. in Electrical Engineering at UFPA. He has experience in electrical engineering with emphasis on automation and control of electrical power systems. His main research topics are: robust control, fuzzy control, control and stability of electric power systems, control of power converters and industrial systems.
ORCID: 0000-0002-1645-2736

W. Barra Junior, received his BSc., MSc. and Dr. degrees in Electrical Engineering, in 1991, 1997, and 2001, from Federal University of Pará, Brazil. He is currently an associate professor at the Faculty of Electrical Engineering of UFPA. He has experience in electrical engineering with emphasis on automation and control of electrical power systems. His main research topics are: adaptive control, fuzzy control, control and stability of electric power systems, and industrial automation.
ORCID: 0000-0003-0246-7264

J.R. Sicchar, received his BSc., in Electrical Engineering from Federal University of Amazonas –UFAM, Brazil, in 1999; the MSc. in 2007 from Federal Universityof Pará- UFPA also in Electrical Engineering. Nowadays, he is doing the Phd. degree also from UFPA and electrical engineering with an emphasis on control and automation of electrical power systems. He has experience in Electrical Engineering with emphasis on control and intelligent supervisors applied in low voltage smart grid.
ORCID: 0000-0002-6497-2496

F.A. de C. Ayres Júnior, received his BSc. Eng. and MSc. degrees from the Federal University of Pará - UFPA, Brazil, in 2013, with an emphasis in fractional control of a power electrical systems. Currently pursuing a PhD. in Electrical Engineering at UFPA. He has experience in electrical engineering with emphasis on automation and control of electrical power systems. His main research topics are: Robust control, fractional control and fractional control and stability of electric power systems.
ORCID: 0000-0002-0266-1019

J.A.L. Barreiros, received the BSc. degree in Electrical Engineering from Federal University of Pará (UFPA), Brazil, in 1974; the MSc. in 1989 from University of Manchester Institute of Science and Technology (UMIST), in the United Kingdom; and the PhD. in 1995, from Federal University of Santa Catarina (UFSC), in Brazil. He has been working in the field of control systems, and applications of control in power systems, for over 20 years. He is currently a titular professor of electrical engineering and Director of the Technological Center at UFPA-Brazil.
ORCID: 0000-0002-9458-0098

State observer design for biomass and ethanol estimation in bioreactors using cybernetic models

Diseño de un observador de estado para estimación de biomasa y etanol en biorreactores a partir de modelos cibernéticos

Pablo De Villeros [a], Héctor Botero [b] & Hernán Alvarez [c]

[a] Escuela de Química, Universidad Nacional de Colombia, Medellín, Colombia. padea@unal.edu.co
[b] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. habotero@unal.edu.co
[c] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. hdalvare@unal.edu.co

Abstract
Over recent years, the problem of multi-substrate microbial-growth modeling has received more and more attention. However, there is no evidence in the literature referring to state observer design that is applied to microbial growth described by cybernetic models. This paper focuses on the design of an asymptotic observer applicable to the estimation of biomass and ethanol in a two-substrate anoxic-fermentation process described by cybernetic modeling. Simulation results for the proposed observer show good performance for continuous, fed-batch and batch regimes.

Keywords: asymptotic observer; cybernetic modeling; microbial growth; bioreactor.

Resumen
En los últimos años, el modelado del crecimiento microbiano en múltiples sustratos ha recibido especial atención. Sin embargo, no hay evidencia en la literatura acerca de la aplicación de observadores de estado en sistemas de crecimiento microbiano descritos por modelos cibernéticos. Este trabajo se enfoca en el diseño de un observador asintótico aplicable a la estimación de biomasa y etanol en una fermentación con dos sustratos y en condiciones anóxicas, bajo una perspectiva cibernética. Los resultados de simulación del observador propuesto muestran un buen desempeño en régimen continuo, semicontinuo y por lotes.

Palabras clave: observador asintótico; modelado cibernético; crecimiento microbiano; biorreactor.

1. Introduction

Measuring of process variables is mandatory in order to implement control loops, perform troubleshooting detection and monitor critical tasks. In many cases, this measurement is difficult to perform owing to the unavailability of sensors, time delays, errors in the measurement system, the high cost of devices or the environmental conditions where the measurement is to be performed [5]. For bioreactors, a precise knowledge of the state variables is vital due to the high sensitivity of microorganisms to slight changes in the culture medium.

Many state variables are difficult to measure (such as biomass), so indirect measurement techniques are needed. One of those techniques is known as state estimation, also known as: state observers, filters, software sensors and more generally, model-based virtual sensors [4]. The purpose is to estimate one or more variables of interest through directly measuring the inputs and outputs of the real system, based on a mathematical model. The extended Kalman Filter [12], Asymptotic Observers [2,7,8] and High Gain Observers [3], are some of the most commonly used in chemical and biotechnological processes.

State estimators/observers have been applied in both

continuous and batch processes. The latter is much more complicated since the absence of inflows and outflows could not excite dynamics that may be decisive for the observer performance.

The use of state observers in chemical and biotechnological processes has become a common practice in the last decade because it can be relatively easily implemented and it has an undeniable utility. Examples of successful applications can be found in [4]. However, there is no evidence in the literature referring to state observer design applied to microbial growth described by cybernetic models. These models originated in the 1980s and they have proven to be very accurate in predicting the growth of various microorganisms on two substrates (diauxic growth) and three substrates (triauxic growth) [13-15,19]. In addition, cybernetic models exhibit the following great advantages compared to other multi-substrate models:

- There is no need for a-priori specification of the order in which the substrates are consumed.
- They describe the simultaneous or sequential use of substrates depending on the culture conditions.
- They are applicable to both batch and continuous cultures.
- The inputs are the values of the parameters obtained from experiments on individual substrates.
- They are relatively easy from a mathematical point of view (ordinary differential equations).

This paper has three main objectives. One is to address the problem of the inherent discontinuity present in the cybernetic model. The second is to evaluate local observability using a linear approximation. The last is to design an observer, applicable to biomass and ethanol estimation in a two-substrate anoxic-fermentation process described by cybernetic models. Continuous, fed-batch, and batch regimes will be considered.

The paper is organized as follows: in Section 2, cybernetic modeling for microbial growth is described. Section 3 deals with the nonlinear observability of the cybernetic model and presents a methodology to estimate biomass and ethanol via an asymptotic observer. In Section 4, simulations are performed in order to test the behavior of the proposed observer. Finally, in Section 5 conclusions are presented as well as some final remark.

2. Cybernetic modeling of anoxic

The assumption that the internal mechanics of microorganisms is determined by a single nutrient is valid for very specific cases in which the culture medium is prepared in a laboratory (i.e. defined culture medium). Such a class of culture is undertaken with the purpose of decreasing the degree of uncertainty in the experiment and ensuring growth and / or production of specific metabolites. This nutritional restriction contrasts with natural ecosystems, in which there is a huge range of potential substrates that are available to suit a particular physiological requirement [9,10]. In this case, simplification suggested by one limiting substrate is inadequate and it becomes necessary to take another approach.
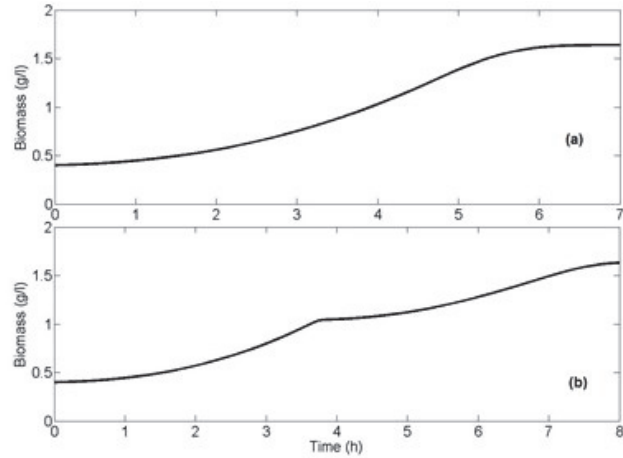
Figure 1. Microbial growth according to substrate consumption.
(a) Simultaneous. (b) Sequential
Source: The authors

The first scientist who explored the problem was Jacques Monod [18]. During an experiment with *Bacillus subtilis*, he observed two clearly distinguishable growth phases in a batch culture with a mixture of two sugars. This phenomenon was called diauxic growth and led to the systematic study of multi-substrate growth.

Fig. 1 shows the behavior of a certain microorganism in a two-substrate culture. In (a), a second exponential phase is not observed, indicating that the microorganism has a similar affinity for both substrates and consumes them simultaneously. In (b), a second exponential phase can be observed, which shows the sequential use of substrates, and it has an intermediate phase in which the microorganism changes its enzymatic machinery to degrade the second substrate. The order in which the substrates are consumed depends on the affinity substrate-microorganism. In other words, the first substrate to be degraded will be the one that provides the highest growth rate [15].

As can be seen, microbial growth on multiple substrates exhibits different behaviors from sequential to simultaneous use of all substrates, giving rise to multi-exponential growth phases. This indicates that cells have a high capacity to regulate their metabolism and optimize the consumption of substrates. Kompala et al. [15] addressed this regulatory behavior by introducing two variables: one considers the induction or repression of a key enzyme involved in the degradation of a specific substrate, while the other considers the activation or inhibition of preexisting enzymes. These two variables are described by the following expressions:

$$u_i = \frac{\mu_i}{\sum_j \mu_j} \ \ and \ \ v_i = \frac{\mu_i}{\max_j \mu_j} \tag{1}$$

When variable u_i is close to zero, the i-th enzyme is repressed, thus its corresponding substrate is not degraded, whereas if its value is close to one, the *i*-th enzyme is induced. This leads to the consumption of the associated substrate.

Variable v_i incorporates the regulatory actions of inhibition and activation of enzymes that are already present

in the cell. If v_i is close to zero, then the degradation reaction of the specific substrate does not take place, even if the said substrate and the corresponding enzyme are present. If its value is close to one, the reaction is activated with the subsequent catabolism of the substrate. Note the discontinuity in the definition of this variable due to the presence of a *max* function. This issue will be considered in Section 3.1.

μ_i is the specific growth rate in the *i*-th substrate that is obtained from a modified version of the well-known Monod equation:

$$\mu_i = \frac{\mu_{max,i} E_{R,i} S_i}{K_i + S_i} \qquad (2)$$

$\mu_{max,i}$ is the maximum growth rate for substrate *i*; S_i is the concentration of substrate *i*; K_i is the affinity constant; $E_{R,i}$ is the relative level of enzyme for the *i*-th substrate and can be calculated from the following relations:

$$E_{R,i} = \frac{e_i}{e_{max,i}} \cdot$$

$$e_{máx,i} = \frac{\alpha_i}{\mu_{max,i} + \beta_i} \qquad (3)$$

where α_i and β_i represent synthesis and degradation constants, respectively. The state variable e_i is the actual enzyme concentration in grams per gram of biomass, while $e_{max,i}$ is the maximum concentration of enzyme for the *i*-th substrate.

Based on typical dynamic balances, without considering cellular death and introducing the previous expressions, a set of six ODEs can be derived as a process model [11, 20]:

$$\frac{dX}{dt} = \left(\sum_i \mu_i v_i\right)X - \frac{F}{V}X - \frac{X}{V}\frac{dV}{dt} \qquad (4)$$

$$\frac{de_i}{dt} = \frac{\alpha_i S_i}{K_i + S_i}u_i - \beta_i e_i - \left(\sum_j \mu_j v_j\right)e_i + \alpha^* \qquad (5)$$

$$\frac{dS_i}{dt} = \frac{F_{in}}{V}S_{i,in} - \frac{F}{V}S_i - \frac{S_i}{V}\frac{dV}{dt} - \frac{1}{Y_{X/S,i}}\mu_i v_i X \qquad (6)$$

$$\frac{dP}{dt} = \sum_i \emptyset_i \frac{\mu_i v_i}{Y_{X/S,i}}X - \frac{F}{V}P - \frac{P}{V}\frac{dV}{dt} \qquad (7)$$

$$\frac{dC}{dt} = \sum_i \theta_i \frac{\mu_i v_i}{Y_{X/S,i}}X - kla_c(C - C^*) - \frac{F}{V}C - \frac{C}{V}\frac{dV}{dt} \qquad (8)$$

$$\frac{dV}{dt} = F_{in} - F \qquad (9)$$

X, *S*, *P* and *C* represent the concentration of biomass, substrate, product (ethanol) and dissolved CO_2 in time *t*, respectively, while *V* stands for volume. Eq. (5) illustrates the dynamics of the *i*-th enzyme, limiting the rate of consumption of the corresponding substrate. The first term on the right corresponds to enzyme synthesis. The following term represents the degradation of the enzyme. The third term represents the dilution of the enzyme concentration due to cell growth and, finally, the term α^* is included to ensure a low level of enzyme even in the absence of substrate. Parameters $\emptyset_i$ and θ_i are stoichiometric coefficients for ethanol and CO_2 production, respectively. $Y_{X/S,i}$ is the biomass yield for substrate *i*. kla_c is the mass transfer coefficient and C^* is the saturation concentration, both for CO_2. *F* represents the volumetric flow.

Jones and Kompała [13] introduced an additional dynamic of intracellular storage of carbohydrates in their aerobic model. However, for the present work it is not considered since in anoxic cultures the storage rate is significantly low compared to the growth rate. The model uses pressure, temperature and pH as controlled variables. Table 1 shows the complete parameters for the model represented by eq. (4)-(9) and applied to the growth of *Saccharomyces cerevisiae* on a mixture of glucose (substrate 1) and galactose (substrate 2). Continuous, fed-batch and batch regimes are considered individually.

Within the development of their mathematical model, Kompała et al. presented four basic principles to be followed by any model that addresses the multisubstrate growth [15]:

- Given multiple substrates, on which growth rates are distinctly different, the microbes prefer to grow on the fastest substrate when the growth curve showing multiauxial behavior.
- More generally, the growth behavior ranges from simultaneous utilization of multiple substrates (which occurs when growth rates are very nearly the same) to sequential utilization with intermediate lag periods.
- The growth rate on a mixture of substrates is never greater than the maximum of the growth rates on individual substrates.
- While growing on a slower substrate, if a faster substrate is added, the microbes inhibit the activity of the already available enzymes for the slower substrates.

3. Observer design

Having defined the cybernetic model for an anoxic microbial culture, the next step is to review the concepts of nonlinear observability and state observers. These two concepts are needed to establish a consistent methodology that relates cybernetic modeling and observer designing.

Section three begins with a mathematical approach to the discontinuity of the cybernetic variable v_i, in order to have a continuous system in which observability can be analyzed.

3.1. Max function approximation

This section deals with the discontinuity of the max function found in eq. (1) in order to analyze the nonlinear observability of the cybernetic system. For observability

Table 1.
Variables and parameters for *Saccharomyces cerevisiae* on Glucose and Galactose.

Parameter	Unit	Value
$\mu_{max,1}$	h^{-1}	0.44
$\mu_{max,2}$	h^{-1}	0.40
K_1	g/l	0.05
K_2	g/l	0.10
$Y_{x/s,1}$	g/g	0.16
$Y_{x/s,2}$	g/g	0.15
α_1 / α_2	h^{-1}	0.0010
$\alpha*_1 / \alpha*_2$	h^{-1}	1.0E-5
β_1 / β_2	h^{-1}	0.050
ϕ_1 / ϕ_2	g/g	0.49
θ_1 / θ_2	g/g	0.49
$S_{1,0} / S_{2,0}$	g/l	4.0
X_0	g/l	0.40
P_0 / C_0	g/l	0.00
V_0	1	6.0
$e_{1,0} / e_{2,0}$	g/g	1.0E-4
kla_c	h^{-1}	1.0E3
$C*$	g/l	0.0053
F_{in} / F	l/h	1.1

Source: Adapted from [6, 14, 16]

analysis and the designing of *Jacobian*-based observers, the system has to be differentiable. As shown in the cybernetic model presented in Section 2, variable v_i contains a discontinuous *max* function that needs to be converted into a continuously differentiable one, without substantially modifying the intrinsic dynamics of the process. The mathematical characteristics of the *max* function show that it can be considered as a vector norm. This issue suggests that it can be substituted with another vector norm. For instance, *p*-norm (which is continuous) would be a good candidate:

$$\|\vec{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p} \qquad (10)$$

Applying this approximation, variable v_i takes the following form:

$$v_i = \frac{\mu_i}{\sqrt[p]{|\mu_1|^p + |\mu_2|^p}} \qquad (11)$$

To verify this approximation, v_1 and v_2 were simulated for several values of *p* based on the parameters previously given in Table 1. The results are shown in Figs. 2, 3:

As can be seen that when *p* is increased the values of v_i converge to those established by the original cybernetic model, whereby the proposed approach is valid for the case being studied. Therefore, in this work, a value of *p* = 50 was taken for the simulations.

3.2. Nonlinear observability analysis

Observability is an important property to take into account when designing an observer. This property allows it to be discovered if there are sufficient information channels between the outputs and the states.

A dynamic system can be expressed in the form:
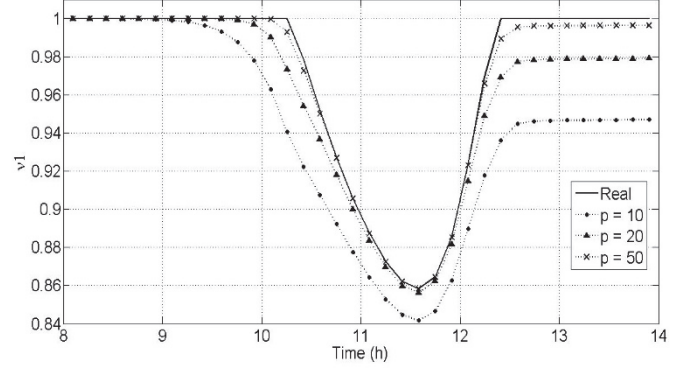
Figure 2. Behavior of v_1 for three different values of *p*.
Source: The authors
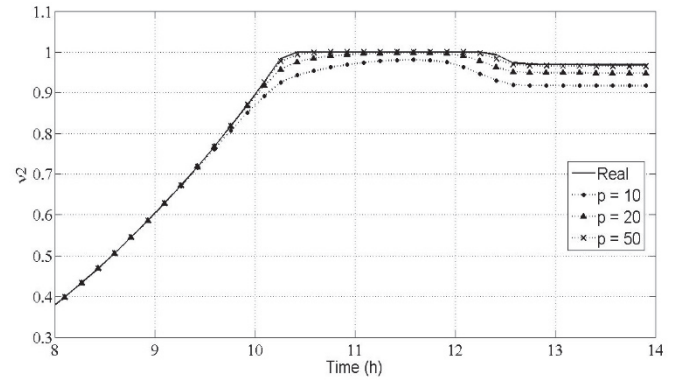
Figure 3. Behavior of v_2 for three different values of *p*.
Source: The authors

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ y(t) = h(x(t)) \end{cases} \qquad (12)$$

where $x(t) \in R^n$ denotes the state variables, $u(t) \in R^m$ denotes control inputs, and $y(t) \in R^p$ denotes the measured outputs.

In very general terms, a nonlinear system is locally observable if the initial states x_0 and x'_0 are distinguishable for every input $u(t)$ at $t \geq 0$:

$$h(x(t, x_0)) \neq h(x(t, x'_0)) \qquad (13)$$

This criterion can be verified with three different perspectives or conditions: a geometric condition, an algebraic condition, and a linear approximation condition [1]. The first two conditions are not suitable for large systems due to, firstly, the enormous computer-time required to obtain Lie derivatives and secondly, due to the degree of variable cross-linking. So, this work will be based on the linear approximation condition.

Consider the following linearized system based on the system in eq. (12). It is evaluated in the stationary state x_{ss} and u_{ss}:

$$\begin{cases} \dot{\boldsymbol{x}}_L = \boldsymbol{A}\boldsymbol{x}_L \\ \boldsymbol{y}_L = \boldsymbol{H}\boldsymbol{x}_L \end{cases}$$

$$\text{with } \boldsymbol{A} = \left(\frac{\partial f}{\partial x}\right)_{x_{ss}, u_{ss}} \text{ and } \boldsymbol{H} = \qquad (14)$$

$$\left(\frac{\partial h}{\partial x}\right)_{x_{ss}, u_{ss}}$$

From eq. (14), it is possible to generate an observability matrix $\boldsymbol{\Omega}$ of the form:

$$\boldsymbol{\Omega} = [\boldsymbol{H}, \boldsymbol{HA}, \boldsymbol{HA}^2, \dots, \boldsymbol{HA}^{n-1}]^T \qquad (15)$$

Once $\boldsymbol{\Omega}$ is calculated, the system is considered observable if $\boldsymbol{\Omega}$ is full rank [2], i.e., matrix $\boldsymbol{\Omega}$ has the same rank of the process (which is seven for the model presented in section 2). For small systems (two or three state variables), matrix $\boldsymbol{\Omega}$ rank analysis can be undertaken by visual inspection, but for the cybernetic model it is not possible due to the extension of the *Jacobian*. Therefore, we propose rank evaluation over time, by calculating the rank of matrix $\boldsymbol{\Omega}$ in each simulation-time step as an attempt to find out the local observability. It is assumed that in each simulation-time an operational equilibrium state is achieved. This approach is a necessary but not a sufficient condition for observability, and it just provides a means of identifying processes that are normally not observable.

As mentioned in Section 1, some variables in bioreactors are very difficult to measure online. These variables include biomass and sugars (substrates). Common biomass measurement techniques are dry/wet weight measurement, optical density via spectrophotometry and direct cell counting; sugar analysis usually also involves spectrophotometric methods. All of those methods are time consuming and are carried out offline. Conversely, dissolved CO_2 concentration is easily measured online through submerged sensors found in bioreactors that are standard equipment. For this reason, in this work, the dissolved CO_2 is taken as measured, without considering other aqueous species that may occur in equilibrium (H_2CO_3, HCO_3^-).

Figs. 4-6 show that the cybernetic system is not locally observable when CO_2 is used as the measured variable. This is because the observability matrix is not full rank at any point or at any regime. The preliminary conclusion is that the CO_2, despite being an easy-to-measure variable, does not offer enough conditions for local observability of the full cybernetic system. However, in this work, some additional considerations will permit the use of CO_2 measurement for observer design.

3.3. Asymptotic observer

The conclusion derived from the analysis in Section 3.2 leads us to think that the best option is to employ an observer that does not require the kinetics of the process so that CO_2 measure can be used directly. The observer that fulfills this condition is the asymptotic observer (AO) [2, 7].

The dynamics of biotechnological processes can be represented by the following general nonlinear state space model, written in matrix form:
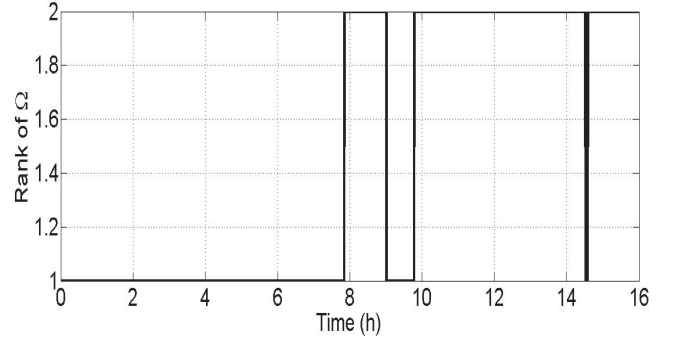
Figure 4. Matrix $\boldsymbol{\Omega}$ rank using CO_2 as measured variable (Continuous).
Source: The authors
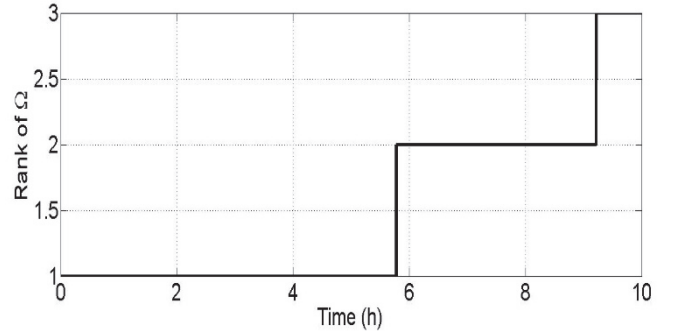
Figure 5. Matrix $\boldsymbol{\Omega}$ rank using CO_2 as measured variable (Fedbatch).
Source: The authors
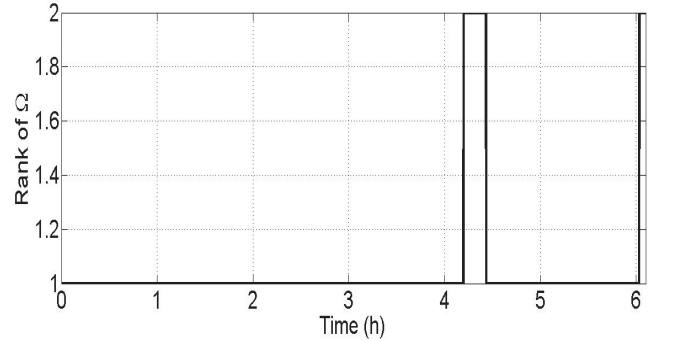
Figure 6. Matrix $\boldsymbol{\Omega}$ rank using CO_2 as measured variable (Batch).
Source: The authors

$$\frac{d\xi}{dt} = \boldsymbol{K}\varphi(\xi, t) - D\xi - \boldsymbol{Q} + \boldsymbol{F} \qquad (16)$$

State variables are presented by $\xi \in \mathrm{R}^n$. The first term $\boldsymbol{K}\varphi(\xi, t)$ describes the kinetics of the biochemical and microbiological reactions that are involved in the process. Matrix $\boldsymbol{K} \in \mathrm{R}^{n \times m}$ represents the yield coefficients while matrix $\varphi \in \mathrm{R}^m$ denotes the reaction rates.

The remaining terms $-D\xi - \boldsymbol{Q} + \boldsymbol{F}$ describe the transport dynamics of the components through the reactor. D is the dilution factor, calculated as F_{in}/V. $\boldsymbol{Q} \in \mathrm{R}^n$ is the vector of rate of mass outflow from the reactor in gaseous form. $\boldsymbol{F} \in$

R^n is the vector of mass feed rate in the reactor.

aking $\xi_a \in R^p$ to be the vector of measured variables and $\xi_b \in R^{n-p}$ as the vector of non-measured variables, the system in eq. (16) can be rewritten:

$$\frac{d\xi_a}{dt} = K_a\varphi(\xi,t) - D\xi_a - Q_a + F_a.$$

(17)

$$\frac{d\xi_b}{dt} = K_b\varphi(\xi,t) - D\xi_b - Q_b + F_b$$

Now, consider vector Z as a linear state transformation:

$$Z = A_0\xi_a + \xi_b$$

(18)

where A_0 is the unique solution of the matrix equation:

$$A_0K_a + K_b = 0$$

(19)

So, the state space model is equivalent to:

$$\frac{d\xi_a}{dt} = K_a\varphi(\xi,t) - D\xi_a - Q_a + F_a.$$

(20)

$$\frac{dZ}{dt} = -DZ - A_0(F_a - Q_a) + F_b - Q_b$$

Vector Z can be rewritten as a linear combination of the form:

$$Z = A_1\xi_1 + A_2\xi_2$$

(21)

with appropriate dimension for matrices A_1 and A_2. If A_2 has a left inverse, the following asymptotic observer can be derived from eq. (20):

$$\hat{\xi}_2 = A_2^+(\hat{Z} - A_1\xi_1).$$

(22)

$$\frac{d\hat{Z}}{dt} = -D\hat{Z} - A_0(F_a - Q_a) + F_b - Q_b$$

where A_2^+ is a left inverse of A_2. $\hat{Z}$ and $\hat{\xi}_2$ denote online estimates of Z and ξ_2, respectively.

3.4. Asymptotic observer design for biomass and ethanol estimation

In order to design an AO, the system must be expressed in the form of state-space representation, based on eq. (16):

$$\xi = \begin{bmatrix} X \\ e_1 \\ e_2 \\ S_1 \\ S_2 \\ P \\ C \end{bmatrix} \quad , \quad F = \begin{bmatrix} 0 \\ \alpha_1^* \\ \alpha_2^* \\ DS_{in1} \\ DS_{in2} \\ 0 \\ 0 \end{bmatrix} \quad , \quad Q = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ CTR \end{bmatrix}$$

(23)

$$K = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ -\dfrac{1}{Y_{X/S1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\dfrac{1}{Y_{X/S2}} & 0 & 0 & 0 & 0 \\ \dfrac{\emptyset_1}{Y_{X/S1}} & \dfrac{\emptyset_2}{Y_{X/S2}} & 0 & 0 & 0 & 0 \\ \dfrac{\theta_1}{Y_{X/S1}} & \dfrac{\theta_2}{Y_{X/S2}} & 0 & 0 & 0 & 0 \end{bmatrix}$$

(24)

$$\varphi(\xi) = \begin{bmatrix} \nu_1\mu_1 X \\ \nu_2\mu_2 X \\ (\nu_1\mu_1 + \nu_2\mu_2)e_1 \\ (\nu_1\mu_1 + \nu_2\mu_2)e_2 \\ \dfrac{\alpha_1 S_1}{K_1 + S_1}u_1 - \beta_1 e_1 \\ \dfrac{\alpha_2 S_2}{K_2 + S_2}u_2 - \beta_2 e_2 \end{bmatrix}.$$

(25)

$$CTR = kla_c(C - C^*)$$

As can be seen in eq. (23)-(25) and in the cybernetic model described in eq. (4)-(9), the dynamics of state variables X, P and C are similar, specifically in the term $(\sum \mu_i \nu_i)X$. This leads to the proposal for a reduced system of the form:

$$\xi_1 = [C] \quad , \quad \xi_2 = \begin{bmatrix} X \\ P \end{bmatrix}$$

(26)

$$K_r = \begin{bmatrix} 1 & 1 \\ \dfrac{\emptyset_1}{Y_{\frac{X}{S1}}} & \dfrac{\emptyset_2}{Y_{\frac{X}{S2}}} \\ \dfrac{\theta_1}{Y_{\frac{X}{S1}}} & \dfrac{\theta_2}{Y_{\frac{X}{S2}}} \end{bmatrix}$$

(27)

$$K_{r1} = \begin{bmatrix} \dfrac{\theta_1}{Y_{\frac{X}{S1}}} & \dfrac{\theta_2}{Y_{\frac{X}{S2}}} \\ 1 & 1 \end{bmatrix}, K_{r2}$$

$$= \begin{bmatrix} \dfrac{\emptyset_1}{Y_{\frac{X}{S1}}} & \dfrac{\emptyset_2}{Y_{\frac{X}{S2}}} \\ \dfrac{1}{Y_{X/S1}} & \dfrac{1}{Y_{X/S2}} \end{bmatrix}$$

(28)

However, asymptotic observers present a restriction referring to the dimension of the coefficient matrix K. As there is just one measured variable (CO_2), the dimension of K_r must be one [7]. In order to overcome this issue, the dimension of K_r is reduced to one using the following coefficient averaging:

$$K_r = \begin{bmatrix} 1 \\ k_p \\ k_c \end{bmatrix} \quad , \quad K_{r1} = [k_c] \quad , \quad K_{r2} =$$

$$\begin{bmatrix} 1 \\ k_p \end{bmatrix} .$$

(29)

$$k_c = \frac{\theta_1 + \theta_2}{Y_{X/S,1} + Y_{X/S,2}} \quad , \quad k_p$$

$$= \frac{\emptyset_1 + \emptyset_2}{Y_{X/S,1} + Y_{X/S,2}}$$

The advantage of this averaging is that it is calculated only using the yield coefficients ($Y_{X/S}$) and the stoichiometric production coefficients for CO_2 (θ_1 and θ_2) and Ethanol ($\emptyset_1$ and $\emptyset_2$).

Finally, the equations that govern the behavior of the asymptotic observer take the following form:

$$\frac{d\hat{Z}}{dt} = -D\hat{Z} + K_{r2}K_{r1}^{-1}CTR \quad .$$

(30)

$$\begin{bmatrix} \hat{X} \\ \hat{P} \end{bmatrix} = \hat{Z} + K_{r2}K_{r1}^{-1}C \quad , \quad \hat{Z} = [\hat{Z}_1, \hat{Z}_2]^T$$

Therefore, it is possible to estimate X and P without considering the kinetics of the process and using dissolved CO_2 concentration as the only measured variable.

4. Simulation results and analysis

The estimation of Biomass and Ethanol via an asymptotic observer was performed in a continuous regime ($F_{in} = F = 1.1$ l/h), a fedbatch regime ($F_{in} = 1.1$ l/h, $F = 0$) and a batch regime ($F_{in} = F = 0$) until cryptic growth was achieved (16 hours, 10 hours and 6 hours, respectively). White noise was added to the CO_2 measurement signal, with a power of 2.0E-4 and a sample time of 0.01 h. Fig. 7 illustrates the evolution of the signal with and without noise, for the continuous case:

Figs. 8-10 show the performance of the proposed observer for biomass estimation. In all simulations, the initial condition of the observer was 10% above the real initial condition of the process. F_{in} and F in continuous and fedbatch regimes had the same fixed value of 1.1 liters per hour.
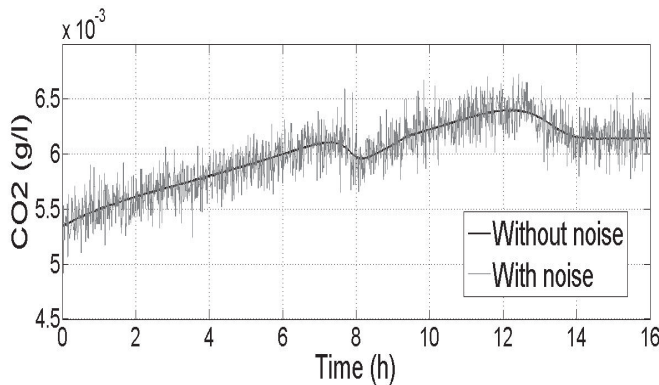
Figure 7. Evolution of CO_2 measure with and without noise (Continuous).
Source: The authors

This value was chosen to be as large as the parameters of the culture permitted (Table 1) whilst avoiding cell wash out. The initial volume of the culture was 6.0 l.

As shown in Figs. 8-10, the asymptotic observer closely follows the actual trajectory of the biomass, but it should be noted that the speed of convergence is given by the value of the dilution rate D. For this reason, in the batch regime (Fig.10), in which the dilution rate is zero, the observer presents a steady-state error and does not converge to the actual value of the biomass while in continuous and fed-batch regimes (Figs. 8 - 9). The observer converges approximately four hours after initiating the process. Taking this into account, the initial state of the observer in a batch regime should be identical to the initial state of the process. Even though the AO is not a filter, it can be seen that the noisy signal was smoothed and the noise did not induce appreciable peaks in the estimation.

For ethanol estimation, the performance is similar to that shown by the estimated biomass in all regimes, as can be seen in Figs. 11 - 13:
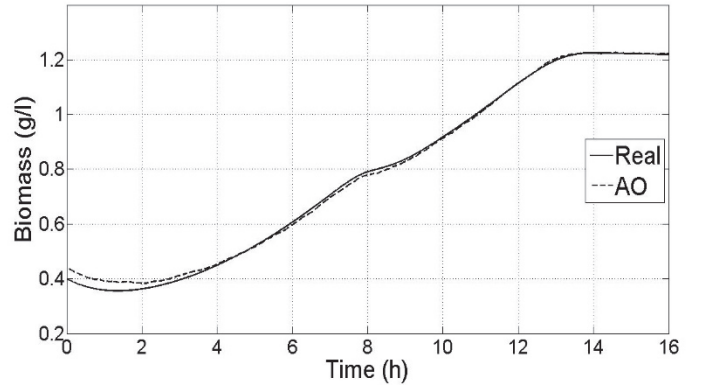
Figure 8. Biomass estimation in continuous regime.
Source: The authors
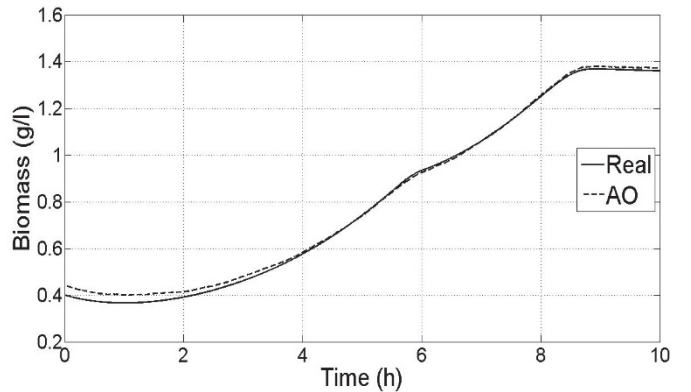
Figure 9. Biomass estimation in fedbatch regime.
Source: The authors
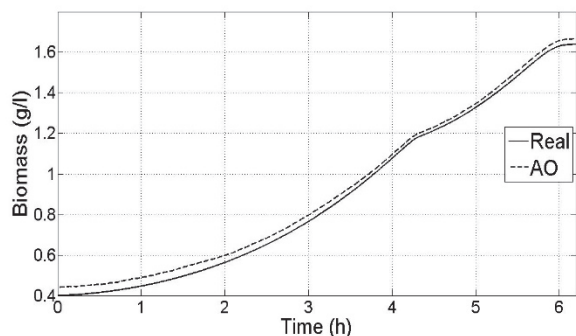
Figure 10. Biomass estimation in batch regime.
Source: The authors
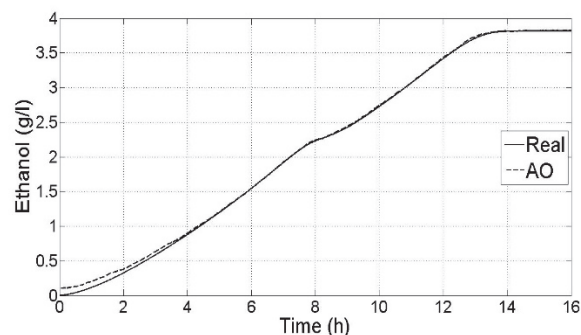
Figure 11. Ethanol estimation in continuous regime.
Source: The authors
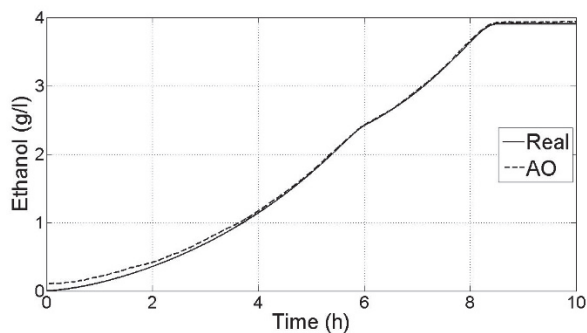
Figure 12. Ethanol estimation in fedbatch regime.
Source: The authors
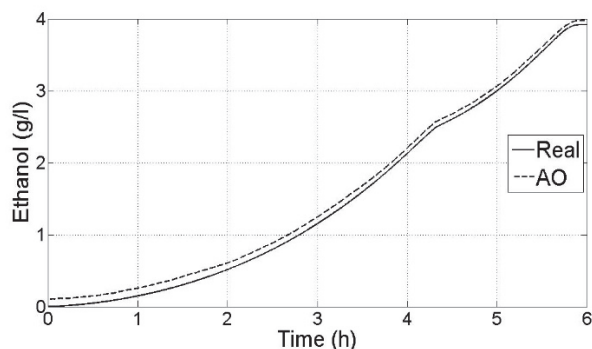
Figure 13. Ethanol estimation in batch regime.
Source: The authors

5. Conclusions

This work approached the discontinuity of the max function on the cybernetic variable v_i by using a substitution based on p-norm. This approximation did not appreciably affect the dynamics of the original process and allowed the jacobian of the system and the rank of the observability matrix of the linearized system to be calculated. Derived from observability analysis, it has been shown that dissolved CO_2 concentration is not a suitable variable for estimating all states of cybernetic system. However, it can be used to estimate biomass and product through a reduced-order asymptotic observer. In terms of the design of this observer, a mathematical approach was proposed to reduce the rank of the coefficient matrix K_r to one.

The designed observer estimated biomass and ethanol satisfactorily, but in batch regime (where the dilution rate is zero) it had to be initiated close to the initial condition of the original system in order to prevent steady-state deviations due to initial estimation error.

It is possible to improve the convergence of the asymptotic observer to the real value of biomass by increasing the dilution rate. The limit is given by the maximum dilution rate before culture washout is reached.

In the design of asymptotic observers, it is essential to have adjusted values of parameters and yield stoichiometric coefficients since the dynamics of these type of observers is totally dependent on them. This is due to the absence of feedback terms or kinetics.

References

[1] Aguilar, R., Mata, J. and Martínez, R., On the observability for a class of nonlinear (Bio)chemical Systems, International Journal of Chemical Reactor Engineering, 8(1), A3, pp. 1-27, 2010.

[2] Bastin, G. and Dochain, D., On-line estimation and adaptive control of bioreactors. Amsterdam, Elsevier Science Publishing, 1998.

[3] Biagiola, S. and Figueroa, J.L., A high gain nonlinear observer: Application to the control of an unstable nonlinear process, Computers and chemical engineering, Elsevier, ISSN 0098-1354, 2004.

[4] Botero, H. y Álvarez, H., Una revisión de los métodos más frecuentes para la estimación del estado en procesos químicos, DYNA, 76(158), pp. 135-146, 2009.

[5] De Assis, A. and Maciel, R., Soft sensors development for on-line bioreactor state estimation, Computers & Chemical Engineering, 24, pp. 1099-1103, 2000. DOI: 10.1016/S0098-1354(00)00489-0

[6] Di Serio, M., De Alteriis, E., Parascandola, P. and Santacesaria, E., A general kinetic and mass transfer model to simulate baker's yeast growth in bioreactors, Catalysis Today, 66, pp. 437-445, 2001. DOI: 10.1016/S0920-5861(00)00644-1

[7] Dochain, D., State and parameter estimation in chemical and biochemical processes: A Tutorial, Journal of Process Control, 13, pp. 801-818, 2003. DOI: 10.1016/S0959-1524(03)00026-X

[8] Dochain, D., Automatic control of bioprocesses. USA, John Wiley & Sons, 2008. DOI: 10.1002/9780470611128

[9] Egli, T., Lendermann, U. and Snozzi, M., Kinetics of microbial growth with mixtures of carbon sources, Antonie van Leeuwenhoek, 63, pp. 289-298, 1993. DOI: 10.1007/BF00871224

[10] Egli, T. and Zinn, M., The concept of multiple-nutrient-limited growth of microorganisms and its applications in biotechnological process, Biotechnology Advances, 22, pp. 35-43, 2003. DOI: 10.1016/j.biotechadv.2003.08.006

[11] Fogler, H., Elements of chemical reaction engineering 3rd edition. Upper Saddle River, New Jersey, Prentice Hall, 1999.

[12] Grewal, M. and Andrews, A., Kalman Filtering: Theory and practice using Matlab. New York, John Wiley & Sons Inc, 2001.

[13] Jones, K.D. and Kompala, D.S., Cybernetic model of the growth dynamics of Saccharomyces cerevisiae in batch and continuous cultures, Journal of Biotechnology, 71, pp. 105-131, 1999. DOI: 10.1016/S0168-1656(99)00017-6

[14] Kompala, D., Bioprocess engineering: Fundamentals and applications [Online]. University of Colorado, 2013 [date of reference June 16th of 2014], section 7.3, Microbial growth on multiple substrates. Available at: http://www.umich.edu/~elements/07chap/html/kompala.pdf

[15] Kompala, D., Ramkrishna, D. and Tsao, G., Cybernetic modeling of microbial growth on multiple substrates, Biotechnology and Bioengineering, 26, pp. 1272-1281, 1984. DOI: 10.1002/bit.260261103

[16] Lubenova, V., Rocha, I. and Ferreira, E.C., Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses, Bioprocess and Biosystem Engineering, 25, pp. 395-406, 2003.

[17] Magaña-Ramírez, L., Rubio-Nuñez, R., Jiménez-Islas, H. and Martínez-García, M., Anaerobic treatment of lactic waste and goat manure, Ingeniería e Investigación, 31(1), pp. 93-98, 2011. DOI: 10.15446/ing.investig

[18] Monod, J., The growth of bacterial cultures, Annual Review of Microbiology, 3, pp. 371-394, 1949. DOI: 10.1146/annurev.mi.03.100149.002103

[19] Narang, A., Konopka, A. and Ramkrishna, D., The dynamics of microbial growth on mixtures of substrates in batch reactors, Journal of Theoretical Biology, 184, pp. 301-317, 1997. DOI: 10.1006/jtbi.1996.0275

[20] Nielsen, J., Villadsen, J. and Lidén, G., Bioreactor engineering principles 2nd ed. New York, Kluwer Academic, 2003. DOI: 10.1007/978-1-4615-0767-3

[21] Rangel-Jara, H. et al., Biorreactores: Modelos matemáticos y su simulación sobre una hoja electrónica, Ingeniería e Investigación, 48, pp. 20-23, 2001. DOI: 10.15446/ing.investig

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN

FACULTAD DE MINAS

Área Curricular de Ingeniería
Química e Ingeniería de Petróleos

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Química
Maestría en Ingeniería - Ingeniería de Petróleos
Doctorado en Ingeniería - Sistemas Energéticos

Mayor información:

E-mail: qcaypet_med@unal.edu.co
Teléfono: (57-4) 425 5317

P. De Villeros, received his BSc. degree in Process Engineering in 2005 from Eafit University and his MSc. degree in Chemical Engineering from Universidad Nacional de Colombia, Medellin, Colombia in 2015. Currently, he is a professor in General Chemistry at the Facultad de Ciencias, Universidad Nacional de Colombia, Sede Medellin and a professor in Physical Chemistry at Eafit University, Medellin. His research interests are State Estimation and Bioprocesses.
ORCID: 0000-0002-1053-8213

H. Botero, received his BSc. degree in Electrical Engineering and his Post-Graduate Diploma in Industrial Automation from Universidad de Antioquia, Colombia, his MSc. in Engineering from the Universidad del Valle, Colombia, and his PhD. from the Universidad Nacional de Colombia, Medellín. He is currently a professor in the Departamento de Energía Eléctrica y Automática, Facultad de Minas, Universidad Nacional de Colombia, Medellín. His research interests are state estimation and engineering education.
ORCID: 0000-0003-3567-4550

H. Alvarez, received his BSc. degree in Chemical Engineering and his MSc. in Systems Engineering from the National University of Colombia. In 2000 he received his Ph.D. degree in Control Systems Engineering from the Instituto de Automática in National University of San Juan, Argentina. Currently he is a full professor at the Departamento de Procesos y Energía, Facultad de Minas, Universidad Nacional de Colombia, Medellín. His research interests are Chemical Process Modeling and Control.
ORCID: 0000-0002-2253-3583

Use of white rot fungi in the degradation of an azo dye from the textile industry

Ana Zuleta-Correa [a], Andrés Merino-Restrepo [b], Sara Jiménez-Correa [b], Angelina Hormaza-Anaguano [b] & Santiago Alonso Cardona-Gallo [c]

[a] Department of Biological and Agricultural Engineering, North Carolina State University, Raleigh, USA. azuleta@ncsu.edu
[b] Facultad de Ciencias, Universidad Nacional de Colombia, Medellín, Colombia. ramerinor@unal.edu.co, sjimenezc@unal.edu.co, ahormaza@unal.edu.co
[c] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. scardona@unal.edu.co

Abstract

Textile industry effluents—a complex mix of chemicals, among which colorants are of particular concern—impose great environmental challenges. In this study, a full 2^3 factorial design was used for determining the best conditions for the degradation of textile dye Basic Red 46 under solid state fermentation (SSF). Three white rot fungi *Trametes versicolor, Pleurotus ostreatus,* and *Pleurotus pulmonarius* were used in the fermentation process. A maximum degradation percentage of 63.0% was achieved at 17 days of incubation with *T. versicolor* under a moisture content of 90%, carbon to nitrogen ratio of 12: 1, and at 20°C. *P. ostreatus* and *P. pulmonarius* reached a maximum degradation percentage of 69.3% and 63.1%, respectively, after 25 days of fermentation. The scale-up of the fermentation process using *T. versicolor* led to a degradation percentage of 45.7% after 30 days of incubation. Additionally, the enzyme activity of laccase, manganese peroxidase and lignin peroxidase was measured. The results indicate that SSF offers a satisfactory degradation, whose efficiency depends on the optimization of process conditions.

Keywords: Solid-state fermentation; white-rot fungi; basic red 46; banana peel.

Utilización de hongos de la podredumbre blanca en la degradación de un colorante tipo azo de la industria textile

Resumen

Los efluentes de la industria textil imponen grandes retos ambientales. Estos son una mezcla compleja de productos químicos, donde colorantes son de particular interés. En este estudio, se implementó un diseño factorial completo 2^3 para establecer las condiciones más apropiadas para la degradación del rojo básico 46 bajo fermentación en estado sólido, FES. Tres hongos de podredumbre blanca *Trametes versicolor, Pleurotus ostreatus y Pleurotus pulmonarius* fueron utilizados en el proceso de fermentación. Se alcanzó una degradación máxima de 63.0% a los 17 días de incubación con *T. versicolor* con una humedad del 90%, proporción carbono:nitrógeno de 12:1 y a 20°C. *P. ostreatus y P. pulmonarius* alcanzaron porcentajes máximos de degradación del 69.3% y 63.1% respectivamente luego de 25 días de fermentación. El escalado del proceso con *T. versicolor* condujo a una degradación del 45.7% a los 30 días de incubación. Adicionalmente, se cuantificó la actividad de las enzimas lacasa, manganeso peroxidasa y lignina peroxidasa. Los resultados señalan que la FES ofrece una degradación satisfactoria, cuya eficiencia depende de la optimización de las condiciones del proceso.

Palabras clave: Fermentación en estado sólido; hongos de podredumbre blanca; rojo básico 46; cascara de banano.

1. Introduction

The textile and clothing industry is one of the main pillars of the Colombian economy. It generates more than 800,000 jobs directly and indirectly and it accounts for about 12.1% of the national industrial production, 6.0% of total exports and 13.3% of non-traditional products sales [1,2]. Colombian

textiles are recognized by high quality, color, and design. Such characteristics have allowed its success in highly competitive markets like the United States, the European Union, and the Andean Community, among others [1].

Textile effluents are usually made up of acids, bases, salts, oils, fats, surfactants and various types of dyes. Concentration of such pollutants is highly variable since it depends on the stage of the process and type of fabric. The variability of effluent composition and extreme temperatures and pH, make textile industry effluents difficult to treat [3]. The dying stage, in particular, has the greatest negative environmental impact. It uses large amounts of water; it is energy intensive, and releases highly toxic dyes into surface water [4].

It is estimated that 10 to 14% of the colorants used in the dyeing and finishing processes are discarded in effluents [5]. Unfortunately, their presence, even in minimal concentrations of 1.0 ppm, is sufficient to cause the aesthetic deterioration of the environment [6]. The most dramatic impact of their presence is the negative effect on photosynthetic processes. They reflect solar radiation reducing the self-regeneration of water resources. They also increase the biological oxygen demand (BOD) and chemical oxygen demand (COD) [1, 7] that could lead to anoxic conditions affecting the whole aquatic ecosystem [8].

Act number 3930 of 2010 in Colombian law regulates water quality standards and parameters for industrial and domestic effluents. The absence, in both the current and preceding act (Act number 1594 of 1984), of the amount of dyes permitted in industrial effluents, has led to the appearance of colored bodies of water emerging around the country. This is as a recurring and disturbing phenomenon in large cities such as Medellín.

Physicochemical treatment technologies such as activated carbon adsorption, flocculation, chemical oxidation, ozonation and filtration, among others, exhibit satisfactory decoloration efficiency. However, their use is restricted due to high costs, incomplete removal of pollutant, and sludge and the generation of toxic byproducts. Therefore, it is important to develop alternative treatment methodologies [9-11]. Between these methods, adsorption with agricultural wastes has emerged as a promising strategy; nevertheless, its study has been primarily limited to metals [12,13].

Biological strategies have been implemented as a result of the limitations associated with the traditional methods and the need to solve the persisting environmental problem. These methodologies use the microorganism's metabolic potential for transforming the pollutants into smaller molecules or for affecting the functional groups involved in their toxicity [14]. In the present study, a novel "mixed" or "combined" strategy is employed for treating a simplified simulated textile effluent. It combines a physiochemical and a biological process. The first step consists of removing dissolved dye from solution without breaking the molecule. In the following step, solid state fermentation (SSF) is used to mineralize or partially degrade the previously adsorbed pollutant.

SSF is characterized by the growth of microorganisms on solid substrates in the absence of free water (low water activity). In this case, the residue-dye complex was used as a source of support and nutrition for the microorganism [15-19]. Previous studies demonstrated the efficiency of the methodology by reaching degradation percentages higher than 90% under the best conditions during the fermentation process [20]. Contaminant degradation is attributed to the action of various enzymes produced by the microorganism.

White rot fungi (WRF) are proven the most efficient microorganisms in the treatment of xenobiotic molecules such as synthetic dyes [21]. Their metabolic capacity to mineralize complex polymers, even like lignin, is attributed to the secretion of non-specific and non-stereoselective enzymes. Laccases (EC 1.10.3.2), manganese peroxidases (EC 1.11.1.13) and lignin peroxidases (EC 1.11.1.14) stand out among the biological catalyzers [22] and they have been able to degrade recalcitrant compounds under conditions similar to the natural habitat of WRF, which are recreated under SSF conditions [23].

Banana peel (BP) was selected as an adsorbent and subsequent substrate in SSF because it allowed us to simulate the natural environment of WRF. Additionally, the lignin, cellulose and hemicellulose content of BP [24] seem to be appropriate for enhancing the production of ligninolytic enzymes [25-27]. Moreover, BP also represents an alternative, highly available, low-cost agricultural residue. On the other hand, basic red 46 (BR46) was selected because it is an azo dye widely used in the textile industry and it is characterized as a highly recalcitrant xenobiotic [28].

In the present study, the most appropriate conditions for the biodegradation of the synthetic dye BR46 under SSF were evaluated. Residues from the banana industry were implemented as substrates. *P. ostreatus*, *P. pulmonarius*, and *T. versicolor* were used to inoculate the fermentation.

2. Materials and methods

2.1. Microorganisms and culture conditions

Three WRF were evaluated, *P. ostreatus*, *P. pulmonarius* and *T. versicolor*. Microorganisms were obtained from Plant Tissue Laboratory at the University of Antioquia, Medellín, Colombia. They were preserved in Petri dishes with PDA agar at 4 °C and subcultured every two months. Before each fermentation, 1.2 cm disks containing the microorganism were subcultured and incubated at 28 °C for a period of 5 to 7 days. For degradation of RB46, small discs (diameter 1.2 cm) were taken from the edge of fungal growth zone for ensuring the exponential growth.

2.2. Adsorbent pretreatment and adsorption of BR46

Banana Peel was acquired in one of the dining halls at the Universidad Nacional de Colombia –Medellín Campus. It was washed, dried at 100 °C, and ground in an Ika mill. Then, it was sieved and particles ranging in size from 300 to 500 µm were chosen for performing the SSF experiments. To neutralize organic acids, BP was washed again with a KOH solution (83.17 mM) for 60 minutes [27,29]. The BR46 dye CI 110 825, was purchased from a local company.

The adsorption was performed in 500 mL beakers at room temperature and 150 rpm for 3 h in a Heidolph Unimax 1010

shaker; 500 mg of BP were added to 250 mL of BR46 solution at 800 ppm. A preliminary analysis of the influence of pH on dye removal showed that a pH greater than 7.0 favors adsorption. Therefore, pH=7.0 was selected for avoiding the costs of adding an alkali. The concentration of the dissolved dye was determined using an UV-Vis Lambda 35 spectrophotometer, Perkin Elmer. Absorbance before and after the adsorption were determined at the wavelength of maximum absorption (λmax = 531 nm). The amount of dye impregnated in the substrate or removal percentage, %R, was calculated using Equation 1, where c_0 and c_f correspond to the initial and final concentrations of BR46, respectively.

$$\%R = \frac{c_0 - c_f}{c_0} \times 100 \quad (1)$$

2.3. Determination of the best conditions of degradation in 50 mL erlenmeyer flasks

For each microorganism, we evaluated the effect of temperature (20 and 28 °C), ATRO moisture (90 and 142), and carbon nitrogen ratio C:N (12:1 and 20:1) in the degradation of BR46 dye adsorbed on BP. A randomization scheme was implemented in a randomized full block factorial design with four replicates resulting in a total of eight treatments, whose characteristics are detailed in Table 1.

Fifty mL Erlenmeyer flasks were used with 140 mg of colored solid substrate (BP-BR46) that were inoculated with three disks (1.2 cm in diameter) of the respective fungus grown on PDA agar.

The ATRO moisture (water percentage in the dry matter of the substrate) was reached using a nutrient solution at the beginning of the fermentation and afterwards regulated by adding sterile distilled water. The solution contained NH_4Cl 0.35 g/L, $(NH_4)_2SO_4$ 1.40 g/L, KH_2PO_4 2.00 g/L, $CaCl_2$ 0.30 g/L, $MgSO_4$ 0.30 g/L, $FeSO_4$ 5.00 mg/L, $MnSO_4$ 1.60 g/L, $ZnSO_4$ 1.40 mg/L, $CuSO_4$ 44.70 mg/L and peptone 1.00 g/L. Moisture was determined by gravimetry on a moisture balance. The C: N ratio was adjusted taking into account the results of the bromatological analysis of the BP, adding the required amount of glucose in the nutrient solution. Compositional analysis was performed at the Laboratory of Chemical and Bromatological Analysis at the Universidad Nacional de Colombia – Medellín Campus.

2.4. Optimization of the BR46 desorption

BR46 dye desorption was initially evaluated using different mixtures of polar and nonpolar solutions, finding that the combination of acetone-hydrochloric acid gave the best results. This mixture was optimized through a Box-Behnken design, in four blocks, where the concentration of acetone (84, 86 and 88% v/v), contact time (13, 16 and 19 h), and solid support dosage varied (80, 140 and 200 mg), reaching a maximum desorption of 99.5% with an acetone concentration of 86% v / v in 0.2 M HCl, a contact time of 16 hours and a dosage of 140 mg [30].

2.5. Degradation kinetics curves

BR46 dye degradation in the SSF process was quantified comparing the amount of desorbed dye in the control sample

Table 1.
Characteristics of the SSF for the three fungal species evaluated in 50 mL Erlenmeyer flask

Treatment	Temperature (° C)	C:N	ATRO moisture
1	20	12	90
2	20	12	142
3	20	20	90
4	20	20	142
5	28	12	90
6	28	12	142
7	28	20	90
8	28	20	142

Source: The authors

(substrate without the action of the microorganism) with the amount of desorbed dye in each test sample. The degradation was monitored every five days and from the tenth day, every three days for a total period of 30 days. Equation 2 was used to establish the degradation percentage of the samples at a specific time.

$$\%Degradation = \frac{(m_i - m_r)}{m_i} \times 100 \quad (2)$$

Where:
m_i: mass of the desorbed dye in the control sample [mg]
m_r: residual mass of desorbed dye in the study sample [mg]
Since the desorption achieved was close to 100%, no correction factor for the terms involved in equation 2 was implemented.

2.6. Evaluation of the degradation process in 100 mL erlenmeyer flasks

This preliminary scale-up of the SSF process was conducted with the fungal species that presented the best adaptation and growth on the solid substrate in the experiments carried out in 50 mL Erlenmeyer flasks, and under the conditions that allowed obtaining the highest degradation percentage.

Thus, 2,500 mg of BP were used. They were contained in a 100 mL Erlenmeyer flask, and were inoculated with five agar disks, 1.2 cm in diameter. The assembly included a total of 20 Erlenmeyer flasks. Ten experimental units were selected for determining the degradation percentage at 15 and 30 days of fermentation. Since the growth of the fungus is not homogeneous on the colored solid substrate, three samples of 140 mg were taken randomly from each experimental unit. They were used to perform the desorption process in triplicate and for the respective analysis.

2.7. Evaluation of the enzyme activity

In order to determine the ligninolytic activity, an enzymatic extraction was conducted by using previously reported protocols [14,25]. The experimental setup was similar to that described for the previous test of the scale-up in 100 mL Erlenmeyer flasks in terms of dosage and fermentation times.

For the extraction, 20 mL of acetate buffer (50 mM, pH = 5.5) were added to each experimental unit. These were stirred at 160 rpm, ensuring the agglomerate disintegration of the fermented substrate, in an ice bath for one hour and

centrifuged at 6,000 rpm for five minutes. The supernatant was used to determine the nonspecific enzymatic activity of the selected fungal species.

Laccase (Lac) activity was established using 2,2-azinobis (3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) 2 mM as substrate and sodium citrate buffer (0.1 M, pH = 3.0), by measuring the increase in absorbance at 420 nm (ε_{420} = 36.00 mM^{-1}cm^{-1}) [31,32]. Manganese peroxidase (MnP) activity was determined using 0.5 mM $MnSO_4$ and 0.1 mM H_2O_2 in sodium malonate buffer (50 mM, pH = 4.5), measuring the increase in absorbance at 270 nm (ε_{270} = 11.59 mM^{-1}cm^{-1}) [31,32]. Lignin peroxidase (LiP) activity was determined using 2.0 mM veratryl alcohol as substrate, and 0.5 mM H_2O_2 in sodium tartrate buffer (50 mM, pH = 2.5), by measuring the increase in absorbance at 310 nm (ε_{310} = 9.30 mM^{-1}cm^{-1}) [31,33]. The enzyme activity is reported as enzyme units (U) per gram of solid substrate, gss, defined as the amount of enzyme required to produce 1.0 µmol product/min [32,33] or, in parallel, the amount necessary to oxidize a 1.0 µmol substrate/minutes [31].

2.8. Statistical analysis

All experimental trials were assessed with 95% confidence. To compare the results an analysis of variance was conducted (ANOVA), followed by the Scheffe's test. To determine the validity of the variance analysis, Shapiro-Wilk and Levene tests were performed with a confidence level of 99%, verifying the criteria of normality and homoscedasticity, respectively. To ensure independence, the sample selection and treatments were applied randomly. All results are reported as the average of measurements plus or minus (±) the standard deviation.

3. Results and discussion

3.1. Adsorbent pretreatment and BR46 adsorption

A sufficient amount of BP was prepared to carry out the BR46 adsorption process. Thus, with the conditions previously described concerning BP dosage, initial amount of dye, pH value, temperature and contact time, it was proceeded to quantify the concentration of the remnant BR46 in the solution. A solution of 40 ppm was obtained at the end of the process, which is equivalent to 95% removal. This high percentage of removal suggests that BP represents a suitable adsorbent material for its incorporation as a solid support in

SSF processes aimed at treating colored effluents.

3.2. Determination of the best conditions of degradation in 50 mL erlenmeyer flasks

The results of BR46 degradation due to the action of the fungal species P. ostreatus, P. pulmonarius and T. versicolor are presented in Table 2. The eight treatments evaluated are shown in descending order of the mean percentage of degradation.

Treatments classified with the same letter for each microorganism exhibit no significant statistical difference. These results are derived from an ANOVA analysis with a confidence level of 95%.

The results suggest that the dye degradation process with P. ostreatus and P. pulmonarius worked best at the lowest temperature (treatments 1-4). In particular, a maximum degradation of 72.87% with P. ostreatus under treatment 1 was reached, which corresponds to a temperature of 20 °C, moisture content of 90% and C:N ratio of 12:1. That is, it includes the lowest levels for each factor, resulting in a decrease in the amount of glucose (lower C:N ratio) and in the volume of water (low moisture). For the above reasons, the conditions of this treatment were selected for constructing the kinetic curves.

It is noteworthy that under the same conditions, the highest percentage of RB46 degradation with P. pulmonarius was 63.38%, which is satisfactory but about 10 percentage units lower compared to P. ostreatus. These results are probably due to the difference in the intrinsic metabolism of the mentioned microorganisms, which belong to different species. The effect of a temperature increase (treatments 5-8), is verified with a significant decrease in the percentage of BR46 degradation; in the case of P. ostreatus, down to 1.78%. One possible explanation for this behavior is attributed to a faster drying process that the solid medium has at this temperature compared to 20 °C, which prevents the proper growth of the microorganism. At 28 °C, the inoculum dries and contracts preventing the proper colonization of the entire substrate. It is emphasized that for treatment 8, the synergy between high levels of the C:N ratio and moisture, mitigates the impact of temperature favoring the availability of nutrients, allowing appropriate development of the microorganism, and resulting in 51.5% degradation. Meanwhile, P. pulmonarius shows a greater sensitivity to the temperature rise, leading to a lack of degradation (treatment 5, corresponding to the conditions of lower moisture and less addition of carbon source).

Table 2.
BR46 degradation rates by *P. ostreatus*, *P. pulmonarius* and *T. versicolor* and grouping by Scheffe's test.

Microorganism								
P. ostreatus			P. pulmonarius			T. versicolor		
Treatment	Mean	Scheffe	Treatment	Mean	Scheffe	Treatment	Mean	Scheffe
1	72.875	A	1	63.375	A	8	64.100	A
3	69.250	A	2	63.350	A	3	61.550	A
2	66.025	A	4	62.300	A	1	60.975	A
4	65.200	A	3	61.100	A	6	58.725	A
8	51.500	B	6	7.025	B	7	57.850	A
6	6.500	C	8	4.400	BC	5	56.250	A
7	5.625	C	7	0.250	BC	2	55.125	A
5	1.775	C	5	0.000	C	4	54.825	A

Source: The authors
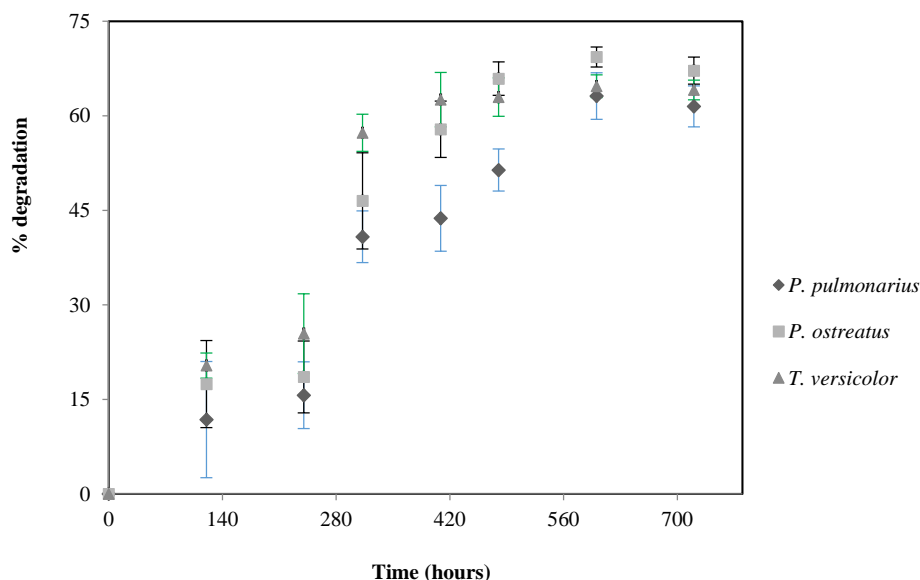
Figure 1. Progress of the BR46 dye degradation by *P. ostreatus, P. pulmonarius* y *T. versicolor*
Source: The authors

The degradation process of the BR46 dye using *T. versicolor* was not affected under the conditions of temperature, moisture or C:N ratio evaluated, since a significant statistical difference between the pairs of treatment methods was not found. These results contrast with those found with *P. pulmonarius* and *P. ostreatus*, and a possible explanation for such a trend is taxonomy. This microorganism belongs to another genus within WRF, and therefore it may adapt differently to diverse nutritional conditions, directly influencing its development [34]. In particular, for this species the additional supply of carbon source (C:N 20:1 in treatments 3, 4, 7 and 8), did not favor the degradation; it is clear that the presence of carbon can support the primary metabolism; however, ligninolytic enzymes associated with dye degradation are part of the secondary metabolism [21].

A maximum degradation of 64.1% was achieved with treatment 8 using T. versicolor, which corresponds to a temperature of 28 ° C, an ATRO moisture of 142% and a C:N ratio of 20:1. Since no significant difference occurs with treatment 1 and 8 for T. versicolor, and the highest degradation with P. ostreatus and P. pulmonarius, were reached under treatment 1 conditions, these parameters were taken to evaluate the degradation kinetics curves for the three strains.

3.3. Degradation kinetics in 50 mL erlenmeyer flasks

BR46 dye degradation by *P. ostreatus*, *P. pulmonarius* and *T. versicolor*, under the best growing conditions, determined from the previous experimental designs, is shown in Fig. 1. These conditions were as follows: temperature of 20 °C, C:N ratio of 12:1, and 90% moisture. The amount of solid substrate used in each assay was 140 mg.

It is noted that in a period of 408 hours (approximately 17 days) *T. versicolor* reached the maximum degradation percentage corresponding to 63.0±3.1%. In turn, *P. ostreatus* and *P. pulmonarius* reached their maximum values of degradation in 600 hours (approximately 25 days) with percentages of 69.3±1.6% and 63.1±3.7%, respectively. Likewise, it is emphasized that the error bars indicate quite high deviations at the beginning of the process, which decrease as time advances (Fig. 1). This occurs because in the early days of the fermentation, the microorganisms have not colonized the entire substrate and the growth on the residue is not homogeneous. As time progresses, the fungus grows throughout the substrate, acting on the BP-BR46 residue-dye complex. Furthermore, while the experimental units were treated with adequate homogeneity, the inoculum of each fermentation could have come from a different Petri dish, which may contribute to the increased in difference between the assays due to the fact that the microorganisms may have exhibited different adaptation times to the solid medium. Similar degradation results were reported by Robinson *et al.* [16], who evaluated a mixture of five reactive dyes adsorbed on barley husk. In their study, the WRF *Bjerkandera adusta* was used and a maximum degradation percentage of 53.1% after 21 fermentation days was reached. This process was performed using 5.0 g of the residue-dye mixture as substrate, with culture conditions similar to those used in the present study, an incubation temperature of 25 ° C, a moisture of 85% and a medium supplemented with glucose and micronutrients. It is noteworthy that the C:N ratio is not detailed.

The results of the kinetics allow to establish that *T. versicolor* offered the best performance in BR46 degradation.

It reached a satisfactory degradation of 63.0%, in 17 days, whereas *P. pulmonarius* and *P. ostreatus* only achieved similar percentages in 25 days.

3.4. Evaluation of the degradation process in 100 mL Erlenmeyer flasks.

T. versicolor was selected to evaluate the process on a larger scale. The selection was based on its performance in kinetic studies. The fermentation was carried out at the same temperature, moisture, and C:N ratio conditions of previous experiments. However, the amount of solid substrate was increased to 2,500 mg. The degradation percentage at 15 and 30 days of fermentation are presented in Table 3.

When results of scale-up experiments are compared with the previous tests, a considerable decrease in the maximum degradation is observed. After 15 days of fermentation, the degradation percentage was 37 percentage points below the expected value (63%). However, after 30 days of fermentation, the degradation percentage corresponds to 72% of the maximum expected value. It may be noted that standard deviation is high even though each value corresponds to the weighted average of ten experimental units with three measurements for each sample. This uncertainty can be due to the heterogeneous fungal growth on the substrate: it was inoculated in the periphery of the Erlenmeyer flask, and then the substrate located in the center will be colonized at later stages. Besides, random samples were taken from each experimental unit and they could correspond to variable degradation regions.

The preliminary results suggest that the scaling-up of this process is feasible since the mass of the BP-BR46 complex increased over 18 times, but a satisfactory degradation percentage was achieved in a period that did not double the time reported for the fermentation process at a lower scale. Results also indicate that the amount of solid support to be treated has a direct impact on the rate of degradation and the time required to carry out the fermentation process. Therefore, modifications need to be implemented to promote further fungal growth and their enzyme activity. Some alternatives include the use of different inoculation techniques that enable a more homogeneous growth and a higher rate of colonization of substrate. It has also been reported that various metal inductors and co-substrates act as enhancing agents for enzyme activity and they lead to a substantial increase in the degradation of recalcitrant compounds. In particular, Baldrian and Gabriel, indicate that copper and cadmium addition led to eight and twelve fold increase of laccase activity compared to a medium without these metallic inductors [35].

Table 3.
Degradation percentage and enzyme activity of *T. versicolor* in process evaluated at 100 mL Erlenmeyer flasks

Days	%Degradation	Enzyme activity [U/gss]		
		Lac	LiP	MnP
15	26.4±7.4	15.6±2.5	1.8±1.2	6.4±2.8
30	45.7±8.4	5.4±2.2	3.1± 1.8	1.0 ±0.4

Source: The authors

3.5. Determination of enzyme activity

The ligninolytic activity of *T. versicolor* was quantified after 15 and 30 days of the fermentative process (Table 3).

After 15 days of fermentation, enzyme activity for Lac, MnP and LiP were 15.62 ± 2.53 U/gss, 6.42 ± 2.89 U/gss and 1.83 ± 1.23 U/gss, respectively. The value obtained for laccase activity in this investigation is lower when compared to the results reported by Iandolo *et al*. [36]. They obtained an activity of 35 U/gss with *T. versicolor* at 16 days of fermentation using a mixture of tomato waste and sorghum stalks. However, it should be noted that the production of enzymes under SSF conditions without further catalytic applications was evaluated in this study. In a period of 30 days, an increase in LiP (3.1 ± 1.8 U/s) was recorded while Lac and MnP activity decreased significantly, registering values of 5.4 ± 2.2 U/gss and 1.0 ± 0.4 U/gss respectively. It is noteworthy that a greater degradation efficiency is achieved at 30 days compared to the results for 15 days, but the degradation percentage is a cumulative variable, while the enzyme activity can change from time to time.

The production of ligninolytic enzymes by different WRF depends mainly on the fungal species, the substrate composition and culture conditions [36,37]. The results obtained in this study indicate that enzyme activity varies over time, a trend also evidenced by Winquist *et al*., with enzyme production of *T. versicolor* on oat hulls [38]; Reddy *et al*. with two *Pleurotus* strains using banana waste [39]; Robinson *et al*. with grim *B. adusta* on barley husk [16]; Sun *et al*. with *Trametes sp*. AH28-2 using four agricultural byproducts [40]; and Tychanowics *et al*., with *P. pulmonarius* on corn residues [33], among others.

It is also noteworthy that Robinson *et al*. [16] evaluated the enzyme activity during the degradation of adsorbed dyes and they found that it changes over time. In the early days of fermentation, low rates of degradation were observed, which coincided with the absence or minimum enzyme activity. As the fermentation process progressed, a higher content of ligninolytic enzymes was detected, even before the degradation of the dye could be perceived. This behavior allows to suggest that the enzymes are associated initially to the lignocellulose cleavage and subsequently to the dye degradation. Enzyme content increased to a maximum value, which occurs at a different time for each enzyme, after such a maximum value is reached, the activity starts to decrease. These authors reached maximum activities of LiP, MnP and laccase of 17.0 U/gss, 15.8 U/gss and 2.0 U/gss at nine, fifteen and twelve days, respectively, obtaining a maximum degradation percentage of 53.0% after 21 days.

The results of this study point out that BR46 degradation, impregnated on an agricultural waste, with WRF, represents a highly efficient and satisfactory alternative process for its bioremediation. In particular, further studies should continue to explore appropriate conditions to achieve an efficient degradation of this azo dye with the fungal species *T. versicolor* at a larger scale.

4. Conclusions

Degradation percentages higher than 60% were achieved at 20 °C, a moisture content of 90% and a C:N ratio of 12:1

using *T. versicolor,* which proved to be the most efficient fungus between the tested species for the degradation of the azo dye BR46. Seventeen days were needed to reach the maximum level of degradation with *T versicolor*, while *P. ostreatus* and *P. pulmonarius* required 25 days.

T. versicolor is an outstanding candidate for large-scale fermentation due to its versatility against changes in the evaluated variables, which allow it to achieve efficient BR46 degradation. It is therefore a suitable species to develop an efficient and low-cost strategy with a positive environmental impact for the bioremediation of xenobiotic compounds such as synthetic dyes.

The enzymatic potential of WRF depends on their different ligninolytic enzymes, primarily laccases. Their unspecific action and great oxidative ability allow them to degrade a wide variety of organic compounds, including polymers such as lignin, chlorinated phenols, dioxins, chloroanilines and dyes. Thus, their application in several biotechnology fields and in the bioremediation of pollutants will undoubtedly be important in the future.

Acknowledgements

References

[1] Cámara de Comercio de Medellín para Antioquia. Inexmoda: Un caso exitoso de liderazgo. Documentos Comunidad Clúster. CCMA: Medellín, 2007, 210 P.

[2] Cámara de Comercio de Medellín para Antioquia. Clúster: Una estrategia para Crear Ventaja Competitiva. Documentos Comunidad Clúster. CCMA: Medellín, 2006, 110 P.

[3] Ministerio del Medio Ambiente and FUNDES, Guía de Buenas Prácticas para el Sector Textiles. 2008, 53 P.

[4] Usepa. EPA Office of Compliance Sector Notebook Project: Profile of the Textile Industry. [Online]. 1997. [date of reference February 23rd of 2015]. Available at: http://www.clu-in.org/download/toolkit/textilsn.pdf.

[5] Vaidya, A.A. and Datye, K.V., Environmental-pollution during chemical-processing of synthetic-fibers. Colourage, 29(1), pp. 3-10, 1982.

[6] Banat, I.M., Nigam, P., Singh, D. and Marchant, R., Microbial decolorization of textile-dye- containing effluents: A Review. Bioresource Technology, 58(1996), pp. 217-227, 1997.

[7] Yesilada, O., Cing, S. and Asma, D., Decolourisation of the textile dye astrazon red FBL by funalia trogii pellets. Bioresource Technology, 81(2), pp. 155-157, 2002. DOI: 10.1016/S0960-8524(01)00117-1

[8] Walsh, G.E., Bahner, L.H. and Horning, W.B., Toxicity of textile mill effluents to freshwater and estuarine algae, crustaceans and fishes. Environmental Pollution Series A, Ecological and Biological, 21(3), pp. 169-179, 1980. DOI: 10.1016/0143-1471(80)90161-0

[9] Robinson, T., McMullan, G., Marchant, R. and Nigam, P., Remediation of dyes in textile effluent: A critical review on current treatment technologies with a proposed alternative. Bioresource Technology, 77(3), pp. 247-255, 2001. DOI: 10.1016/S0960-8524(00)00080-8

[10] Rafatullah, M., Sulaiman, O., Hashim, R. and Ahmad, A., Adsorption of methylene blue on low-cost adsorbents: A review. Journal of Hazardous Materials, 177(1-3), pp. 70-80, 2010. DOI: 10.1016/j.jhazmat.2009.12.047

[11] Gupta, V.K. and Suhas., Application of low-cost adsorbents for dye removal - A review. Journal of Environmental Management, 90(8), pp. 2313-2342, 2009. DOI: 10.1016/j.jenvman.2008.11.017

[12] Vargas-Nieto, C., Garriazo, J.G. and Castillo, E., Estudio de materiales adsorbentes de bajo costo para remover Cr (VI) de efluentes acuosos. Ingeniería e Investigacón, 31(1), pp.154-162, 2011

[13] Higuera-Cobos, O.F., Florez-García, L.C. and Arroyave-Londoño, J.F., Estudio de la biosorción de cromo con hoja de café. Ingeniería e Investigacón, 29(2), pp.59-64. 2009.

[14] Rauf, M.A. and Salman-Ashraf, S., Survey of recent trends in biochemically assisted degradation of dyes. Chemical Engineering Journal, 209, pp. 520-530, 2012. DOI: 10.1016/j.cej.2012.08.015

[15] Krishna, C., Solid-state fermentation systems-an overview. Critical Reviews in Biotechnology, 25(1-2), pp. 1-30, 2005. DOI: 10.1080/07388550590925383

[16] Robinson, T. and Nigam, P.S., Remediation of textile dye waste water using a white-rot fungus Bjerkandera adusta through solid-state fermentation (SSF). Applied Biochemistry and Biotechnology, 151(2-3), pp. 618-628, 2008. DOI: 10.1007/s12010-008-8272-6

[17] Papadopoulou, K., Kalagona, I.M., Philippoussis, A. and Rigas, F., Optimization of fungal decolorization of azo and anthraquinone dyes via Box-Behnken design. International Biodeterioration and Biodegradation, 77, pp.31-38, 2013. DOI: 10.1016/j.ibiod.2012.10.008.

[18] Tisma, M., Komar, M., Rajic, M., Pavlovic, H. and Zelic, B., Decolorization of dyes by Aspergillus ochraceus cultivated under solid state fermentation on sugar beet waste. Chemical Engineering Transactions, 27, pp.145-150, 2012.

[19] Buitrago H.G., Ospina, M.P. y Rengifo, L.J.R., Construcción de un fermentador para operación en estado sólido y diseño de sistemas de control. Ingeniería e Investigacón, 18, pp.45-53, 1989.

[20] Jaramillo, A., Jimenez, S., Merino, A. y Hormaza, A., Obtención de un inóculo fúngico para la degradación de un colorante azo por fermentación en estado sólido. Rev.Udcaactual.Divulg.Cient, 17(2), pp. 577-585, 2014.

[21] Wesenberg, D., Kyriakides, I. and Agathos, S.N., White-rot fungi and their enzymes for the treatment of industrial dye effluents. Biotechnology Advances, 22(1-2), pp. 161-187, 2003. DOI: 10.1016/j.biotechadv.2003.08.011

[22] Levin, L., Papinutti, L. and Forchiassin, F., Evaluation of Argentinean white rot fungi for their ability to produce lignin-modifying enzymes and decolorize industrial dyes. Bioresource Technology, 94(2), pp. 169176, 2004. DOI: 10.1016/j.biortech.2003.12.002

[23] Kadam, A.A., Telke, A.A., Jagtap, S.S. and Govindwar, S.P., Decolorization of adsorbed textile dyes by developed consortium of Pseudomonas sp. SUK1 and Aspergillus ochraceus NCIM-1146 under solid state fermentation. Journal of Hazardous Materials, 189(1-2), pp. 486-494, 2011. DOI: 10.1016/j.jhazmat.2011.02.066

[24] Shah, M.P., Reddy, G.V., Banerjee, R., Ravindra-Babu, P. and Kothari, I.L., Microbial degradation of banana waste under solid state bioprocessing using two lignocellulolytic fungi (Phylosticta spp. MPS-001 and Aspergillus spp. MPS-002). Process Biochemistry, 40(1), pp. 445-451, 2005. DOI: 10.1016/j.procbio.2004.01.020

[25] Nigam, P., Armour, G., Banat, I.M., Singh, D. and Marchant, R., Physical removal of textile dyes from effluents and solid-state fermentation of dye-adsorbed agri-cultural residues. Bioresour. Technol., 72, pp. 219-226, 2000.

[26] Boer, C.G., Obici, L., De Souza, C.G.M. and Peralta, R.M., Decolorization of synthetic dyes by solid state cultures of Lentinula (Lentinus) edodes producing manganese peroxidase as the main ligninolytic enzyme. Bioresource Technology, 94(2), pp. 107-112, 2004. DOI: 10.1016/j.biortech.2003.12.015

[27] Osma, J.F., Toca-Herrera, J.L. and Rodríguez-Couto, S., Banana skin: A novel waste for laccase production by Trametes pubescens under solid-state conditions. Application to synthetic dye decolouration. Dyes and Pigments, 75(1), pp. 32-37, 2007. DOI: 10.1016/j.dyepig.2006.05.021

[28] Deniz, F. and Karaman, S., Removal of basic red 46 dye from aqueous solution by pine tree leaves. Chemical Engineering Journal, 170(1), pp. 67-74, 2011. DOI: 10.1016/j.cej.2011.03.029

[29] Stredansky, M. and Conti, E., Xanthan production by solid state fermentation. Process Biochemistry, 34, pp. 581-587, 1999.

[30] Jimenez, S., Merino, A., Zuleta, A. y Hormaza, A., Optimización de la desorción como metodología de cuantificación de la biodegradación del colorante rojo básico 46 por Pleurotus pulmonarius bajo fermentación en estado sólido, XIII Congreso Argentino de Microbiología, de Microbiología, 2013, 255 P.

[31] Faraco, V., Pezzella, C., Miele, A., Giardina, P. and Sannia, G., Bio-remediation of colored industrial wastewaters by the white-rot fungi Phanerochaete chrysosporium and Pleurotus ostreatus and their enzymes. Biodegradation, 20(2), pp. 209-220, 2009. DOI: 10.1007/s10532-008-9214-2

[32] Murugesan, K., Nam, I.H., Kim, Y.M. and Chang, Y.S., Decolorization of reactive dyes by a thermostable laccase produced by Ganoderma lucidum in solid state culture. Enzyme and Microbial Technology, 40(7), pp. 1662-1672, 2007. DOI: 10.1016/j.enzmictec.2006.08.028

[33] Tychanowicz, G.K., Zilly, A., De Souza, C.G.M. and Peralta, R.M., Decolourisation of industrial dyes by solid-state cultures of Pleurotus pulmonarius. Process Biochemistry, 39(7), pp. 855-859, 2004. DOI: 10.1016/S0032-9592(03)00194-8

[34] Kaushik, P. and Malik, A., Fungal dye decolourization: Recent advances and future potential. Environment International, 35(1), pp. 127-141, 2009. DOI: 10.1016/j.envint.2008.05.010

[35] Baldrian, P. and Gabriel, J., Copper and cadmium increase laccase activity in Pleurotus ostreatus. FEMS Microbiology Letters, 206(1), pp. 69-74, 2002. DOI: 10.1016/S0378-1097(01)00519-5

[36] Iandolo, D., Piscitelli, A., Sannia, G. and Faraco, V., Enzyme production by solid substrate fermentation of Pleurotus ostreatus and Trametes versicolor on tomato pomace. Applied Biochemistry and Biotechnology, 163(1), pp. 40-51, 2011. DOI: 10.1007/s12010-010-9014-0

[37] Stajić, M., Persky, L., Friesem, D., Hadar, Y., Wasser, S.P., Nevo, E., and Vukojević, J., Effect of different carbon and nitrogen sources on laccase and peroxidases production by selected Pleurotus species. Enzyme and Microbial Technology, 38(1-2), pp. 65-73, 2006. DOI: 10.1016/j.enzmictec.2005.03.026

[38] Winquist, E., Moilanen, U., Mettälä, A., Leisola, M. and Hatakka, A., Production of lignin modifying enzymes on industrial waste material by solid-state cultivation of fungi. Biochemical Engineering Journal, 42(2), pp. 128-132, 2008. DOI: 10.1016/j.bej.2008.06.006

[39] Reddy, G.V., Ravindra-Babu, P., Komaraiah, P., Roy, K.R.R.M., and Kothari, I.L., Utilization of banana waste for the production of lignolytic and cellulolytic enzymes by solid substrate fermentation using two Pleurotus species (P. ostreatus and P. sajor-caju). Process Biochemistry, 38(10), pp. 1457-1462, 2003. DOI: 10.1016/S0032-9592(03)00025-6

[40] Sun, Q.Y., Hong, Y.Z., Xiao, Y.Z., Fang, W. and Fang, J., Decolorization of textile reactive dyes by the crude laccase produced from solid-state fermentation of agro-byproducts. World Journal of Microbiology and Biotechnology, 25(7), pp. 1153-1160, 2009. DOI: 10.1007/s11274-009-9994-5

A. Zuleta-Correa, completed her BSc. Eng in Biological Engineering at Universidad Nacional de Colombia, Medellín Campus in 2009; and her MSc degree in Environment and Development at Universidad Nacional de Colombia, Medellín Campus in 2013. She was a very active member of the Research Group "Synthesis, Reactivity and Transformation of Organic Compounds" (SIRYTCOR) and her researching experience in the environmental topic allowed her to direct several projects. Currently, she is a PhD student in Biological and Agricultural Engineering at North Carolina State University.
ORCID: 0000-0001-8674-6396

A. Merino-Restrepo, completed his BSc. Eng in Biological Engineering at Universidad Nacional de Colombia, Medellín Campus in 2016. He is a MSc Biotechnology student and a member of the Research Group "Synthesis, Reactivity and Transformation of Organic Compounds" (SIRYTCOR). His research experience in the environmental topic, specifically in the biological degradation processes through Solid State Fermentation, has strengthened due to his important participation in Master's projects and his expertise in high-end equipment. He has taken part in several research projects.
ORCID: 0000-0001-6818-3251

S. Jiménez-Correa, completed her BSc. Eng in Biological Engineering at Universidad Nacional de Colombia, Medellín Campus in 2016. She is a MSc Biotechnology student and a member of the "Synthesis, Reactivity and Transformation of Organic Compounds" research group. Her researching experience in the environmental field, specifically in biological degradation processes by Solid State Fermentation, has improved thanks to her participation in Master's projects and her expertise in high-tech equipment. She has attended several worldwide congresses.
ORCID: 0000-0001-7103-7474

A. Hormaza-Anaguano, completed her BSc. in Chemistry at Universidad de Nariño in 1994; her MSc in Chemical Science at Universidad del Valle in 1997; and her PhD in Natural Sciences at the Johannes Gutenberg-University Mainz, Germany, in 2003. She began working at Universidad Nacional de Colombia Medellín Campus in 1997 as an Assistant Professor and she is currently a full-time Associate Professor in the School of Chemistry, School of Sciences, Universidad Nacional de Colombia, Medellín Campus. Since 2003 she has been the Director of the Research Group "Synthesis, Reactivity and Transformation of Organic Compounds", SIRYTCOR, whose research lines are focused on the treatment of industrial effluents, exploration and evaluation of alternative adsorbents, adsorption and desorption, and biological processes by solid state fermentation.
ORCID: 0000-0001-5825-3885

S. Cardona-Gallo, completed his BSc. Eng. in Sanitary Engineering in 1997; his MSc. degree Environmental Engineering at Universidad Nacional Autónoma of Mexico in 2000; his PhD. in Environmental Engineering at Universidad Nacional Autónoma of Mexico in 2004; and his Post. PhD in Rice University in 2012. He has been a professor at the Department of Geosciences and the Environemnt, School of Mines, Universidad Nacional de Colombia, Medellín Campus, Colombia, since 2005. His research interests include water quality, soil quality, remediation, bioremediation, hazardous waste, and process design.
ORCID: 0000-0002-1875-7330

The mechanical properties of Portland cement mortars blended with carbon nanotubes and nanosilica: A study by experimental design

Oscar A. Mendoza-Reales [a], Germán Sierra-Gallego [a] & Jorge I. Tobón [a]

[a] *Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia.* oamendoz@unal.edu.co, *geasierraga@unal.edu.co, jitobon@unal.edu.co*

Abstract
The pozzolanic effect of nanosilica (NS) particles when combined with Multi Walled Carbon Nanotubes (MWCNT) was studied in $Ca(OH)_2$ pastes and Portland cement mortars. Experimental design techniques were used to plan the experiments and identify the effect of the nanoparticles on the properties of the cementing matrices by means of Analysis of Variance (ANOVA). Samples were prepared with different combinations of NS and MWCNT. $Ca(OH)_2$ pastes were used to study the effect the nanoparticles had on the Calcium-Silicate-Hydrate (C-S-H) production. Portland cement mortars were used to study the effect of the nanoparticles on the compressive and flexural strength of the cementing matrices. We found that only NS had a significant effect on the C-S-H formation for up to 21 days of hydration, and that MWCNT did not present a positive effect on the mechanical strength of mortars due to the effects of reagglomeration.

Keywords: Carbon nanotubes; nanosilica; hybrid effect; pozzolanic activity; strength.

Propiedades mecánicas de morteros de cemento Portland adicionados con nanotubos de carbono y nanosilica: Estudio por diseño de experimentos

Resumen
El efecto puzolánico de partículas de nanosílice (NS) combinadas con Nanotubos de Carbono de Pared Múltiple (MWCNT) fue estudiado en pastas de $Ca(OH)_2$ y morteros de cemento Pórtland. Técnicas de diseño experimental fueron aplicadas para planear los experimentos e identificar el efecto de las nanopartículas sobre las propiedades de las matrices cementantes usando Análisis de Varianza (ANOVA). Se prepararon muestras con diferentes combinaciones de NS y MWCNT. Las pastas de $Ca(OH)_2$ fueron usadas para estudiar el efecto de las naopartículas en la producción de Silicato de Calcio Hidratado (C-S-H), y los morteros de cemento Pórtland se usaron para estudiar el efecto de las nanopartículas en las resistencias a compresión y flexo-tracción de las matrices. Se encontró que solo la NS tuvo un efecto significativo sobre la producción de C-S-H hasta 21 días de hidratación, y que los MWCNT no presentaron un efecto positivo sobre las propiedades mecánicas de los morteros debido a efectos de reagomeración.

Palabras clave: Nanotubos de carbono; nanosílice; efecto híbrido; actividad puzolánica; resistencia mecánica.

1. Introduction

It is known that the main contribution to the mechanical properties of a hydrated Portland cement matrix comes from calcium-silicate-hydrate (C-S-H) [1], which is nanometric by nature. Due to the nanometric character of C-S-H, nanoparticles with pozzolanic or reinforcing properties have the most adequate size scale to modify its properties [2]. Additions of nanosilica (NS) are widely recognized for their positive effects on the properties of mortars, especially on its compressive strength [3]. Also, it has been found that the main interaction of carbon nanotubes within a cement matrix is via the C-S-H, which generates a "bridge effect" that improves the tensile load distribution within the matrix and its overall flexural strength [4].

To improve the properties of a cement based matrix using nanoparticles, one of the main issues that has to be addressed is the adequate dispersion of the nanoparticles throughout the matrix [5]. Depending on the type, amount and pH of the media, among other factors, nanoparticles are more prone to form agglomerations [6-7]. When the nanoparticles are not adequately dispersed in the matrix, they act as filler in the nano pores. When properly dispersed, nanoparticles act as nucleation spots for hydration products, accelerating the hydration process, increasing the viscosity of the suspension, helping to maintain the cement grains and aggregates suspended, decreasing the vulnerability to segregation, and increasing the workability of the system [1]. In some cases, it has been reported that poorly dispersed nanoparticles have a tendency to worsen the properties of the cement based matrices [8].

Previous research has been carried out with multi walled carbon nanotubes (MWCNT) dispersed in water using a policarboxilate based superplasticizer (SP) as a dispersing agent. In this research it has been found that alkaline environments, such as those formed during the hydration of cement, which are rich in Ca(OH)2, affect the stability of the MWCNT dispersions and generate agglomeration phenomena due to an interaction between MWCNT and the Ca(OH)2. This interaction prevents electrostatic repulsion between the functional groups on the MWCNT and the SP molecules [9]. Additionally, it has been found that the MWCNT and NS combinations have an accelerating effect on the kinetics of the hydration reaction during the first 24 hours. This has been associated with MWCNT working as nucleation spots for the hydration products generated by the NS. Due to the presence of Ca(OH)2 in the media since the first hours of hydration, the MWCNT dispersed in water will suffer a reagglomeration process. This decreases the specific surface area available for the MWCNT to work as nucleation spots and eventually inhibits the pozzolanic activity of NS. This is reflected in the amount of C-S-H produced and the amount of heat released during the hydration reaction [9-10]. This work studies the effect of NS and MWCNT combinations on the pozzolanic activity of NS and the mechanical properties of Portland cement using factorial experimental designs as a statistical tool to interpret the analysis and the results.

2. Methodology

The materials used in the experimental campaign were multi walled carbon nanotubes (MWCNT) functionalized with acid groups produced by Nanoamor, nanosilica particles (NS) dispersed in water produced by BASF Chemicals, reactive grade Ca(OH)2 produced by Merck, a superplasticizer additive (SP) produced by BASF Chemicals, and type III Portland cement (Cto) produced by Cementos Argos. Aqueous dispersions of MWCNT with a solids concentration of 0.25% wt were produced using an ultrasonic tip processor and SP as a dispersing agent in a 4:1 proportion of SP to MWCNT. The aqueous dispersions were sonicated in 20 second on/off cycles to avoid any overheating until a total amount of energy of 40,000 J was applied to the system. Details of how the dispersions parameters were determined

as well as the quantification of the obtained dispersion degree have been published separately [9].

Ca(OH)2 pastes with a water-to-Ca(OH)2 ratio of 1 were produced to identify the influence of MWCNT and NS on the production of C-S-H due to the pozzolanic activity of the NS particles. Due to the fact that both MWCNT and NS were dispersed in water, there is a maximum amount of solid nanoparticles that can be added to the Ca(OH)2 paste without surpassing the fixed water-to-Ca(OH)2 ratio. Taking into account that the solid concentration of the aqueous MWCNT dispersion was 0.25%, and the solid concentration of the NS aqueous dispersion was 49.5%, it was calculated that the maximum possible solid additions of MWCNT and NS were 0.025% and 9.5%, respectively. The pastes were prepared first by mixing the liquid components in a glass beaker and later by adding Ca(OH)2, which was hand-mixed with a spatula until a homogeneous paste was obtained. Finally, the samples were stored in airtight plastic containers that were to be tested after 1, 3, 7, and 21 days of hydration. At each testing age the hydration of the pastes was stopped by grinding them in acetone, oven drying at them at 60 °C, and sieving them though a No 200 mesh. The response variable chosen was the mass loss at 230°C, which was measured in a Precisa Master XM 60 moisture analyzer. Each sample was heated at a rate of 85 °C/min up to the maximum temperature, which was held constant for 3.5 minutes or until the sample reached constant mass. We considered that the sample reached constant mass when in a period of 60 seconds the sample lost 2.0 mg or less.

The experiment was planned using a 23 factorial experimental design with two factors and three levels for each factor. The first factor was the amount of MWCNT, and the levels were 0.00%, 0.025% and 0.225%. The second factor was the amount of NS and the levels were 3.5%, 6.5% and 9.5%. The experimental matrix and the randomization of the experimental runs are presented in Table 1. This design was repeated and analyzed separately for the four different curing ages.

Mortars with a water-to-cementing material ratio of 0.5 were prepared according to the procedure established in the ASTM C109/C109M standard. Immediately after mixing, the flow of each mortar was measured using a flow table, following the procedure established in the ASTM 1437 standard. A necessary amount of SP was added to each mortar to obtain a $110 \pm 5\%$ flow. For each experimental run and each testing age (1, 3, 7 and 28 days), three 4.0x4.0x16.0 cm prisms were molded. After 24 hours, the prisms were removed from the molds and cured by submersion in lime

Table 1.
Experimental matrix of MWCNT/NS/Ca(OH)$_2$ pastes.

Randomization	MWCNT (%)	NS (%)
3	-	3.5
8	-	6.5
5	-	9.5
2	0.025	3.5
6	0.025	6.5
4	0.025	9.5
9	0.225	3.5
1	0.225	6.5
7	0.225	9.5

Source: The authors

Table 2.
Experimental matrix of MWCNT/NS/Cement mortars.

Randomization	MWCNT (%)	NS (%)
2	-	-
1	-	7.3
4	0.0115	-
3	0.0115	7.3

Source: The authors

saturated water until they reached the testing age. When the testing age was reached, each set of samples was taken out of the curing water and immediately tested for compressive and flexural strength according to the procedure established in the ASTM C348 y C349 standards. The experiments were carried out using a factorial 2^2 experimental design with two factors (MWCNT and NS), and each factor had two levels (0%NS, 7.3%NS, 0%MWCNT and 0.115%MWCNT). The experimental matrix and the randomization of the experiments are presented in Table 2. This design was repeated and analyzed separately for the four different curing ages.

3. Results and discussion

This section presents the results obtained in the experimental campaigns and is divided into the following sub-sections: materials characterization, pozzolanic activity, superplasticizer demand and mechanical properties.

3.1. Materials characterization

Images were taken of MWCNT using transmission electronic microscopy (TEM) in a model TECNAI 20kV FEI microscope that was operated at 200 kV. The sample was dispersed in ethanol and sonicated in an ultrasonic bath for 30 seconds, then a drop of the sample was placed in a No 300 Formvar cupper mesh. The images obtained are presented in Fig. 1. The structures of MWCNT were confirmed and found to be nanometric in diameter and micrometric in length. MWCNT formed agglomerations and entanglements as a consequence of the Van der Waals forces that cause them to attract and agglomerate [11]. A more detailed characterization of the MWCNT and NS that were used in this research can be found in previous publications [9-10].
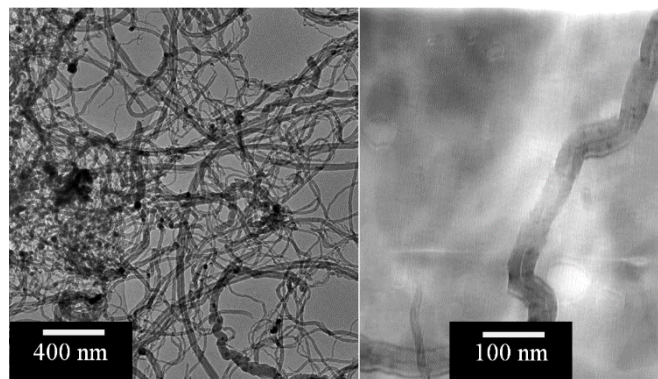
Figure 1. Transmission Electron Microscopy images of the MWCNT
Source: The authors

Table 3.
X-Ray fluorescence results of MWCNT, NS, and Cement.

Component (%)	MWCNT	Cto	NS
Al_2O_3	-	8.64	-
SiO_2	-	19.4	93.56
S	-	1.44	-
K_2O	-	0.41	0.01
CaO	-	59.20	0.22
TiO_2	-	0.25	-
MnO	-	0.05	0.01
Fe_2O_3	-	2.81	0.39
Cu	-	0.01	-
Cl	-	0.15	-
P_2O_5	-	1.27	0.13
Ni	98.00	-	-
Sr	-	0.05	-
Rh	1.41	-	-
Pd	0.63	-	-
MgO	-	-	0.13
Na_2O	-	-	0.62
SO_3	-	-	0.30

Source: The authors

The chemical composition of MWCNT, NS, and cement was studied by X-Ray fluorescence (FRX), using portable FRX equipment from Bruker model S1. Tests were performed in powder samples without further preparation. The results obtained are presented in Table 3. The main component of MWCNT is carbon, which is not detected by the FRX equipment. From the elements that were detected, the main constituent was Ni, which was used as a precursor in the production of the nanotubes [9]. As expected, cement was found to be composed mainly of CaO, SiO_2, and Al_2O_3. NS was found to be mainly SiO_2.

3.2. Pozzolanic activity

$Ca(OH)_2$ pastes blended with NS and MWCNT in the amounts presented in Table 1 were produced according to the procedure described in the methodology section. After 1, 3, 7, and 21 days of curing, the mass loss of each paste at 230 °C was measured using a moisture analyzer. The results obtained are presented in Figs. 2-5. To validate the use of the moisture analyzer as a method to identify pozzolanic activity, $Ca(OH)_2$ pastes that were blended with different types and different amounts of pozzolans were produced. After the different curing ages the mass loss of each paste was tested at the same temperature by thermogravimetry analysis and by a moisture analyzer. Statistical analysis showed a good correlation between the two testing methods. This validation has been separately published and details can be found in reference [12].

The mass loss results show that for the four ages studied, the higher the amount of NS the higher the mass loss of the pastes. This is caused by the pozzolanic activity of NS, whish is identified as a function of the amount of C-S-H formed by the fixation of $Ca(OH)_2$ by NS [13]. Additionally, for all the NS and MWCNT amounts studied, the mass loss stabilizes at between 2 and 3% from the 7th to the 21st day of hydration. This indicates that in this time interval all the NS reacted to form C-S-H. The addition, MWCNT does not seem to affect the pozzolanic activity of the NS. To confirm this statement, an analysis of variance (ANOVA) was performed, and its results are presented in Table 4.
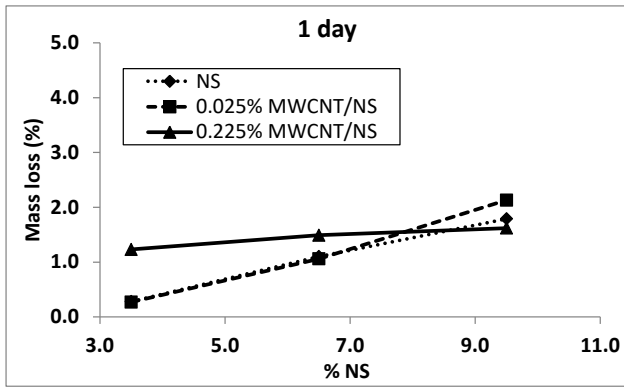
Figure 2. Mass loss results at 230 °C for Ca(OH)$_2$ pastes blended with MWCNT and NS after 1 day of hydration.
Source: The authors

Figure 3. Mass loss results at 230 °C for Ca(OH)$_2$ pastes blended with MWCNT and NS after 3 days of hydration.
Source: The authors
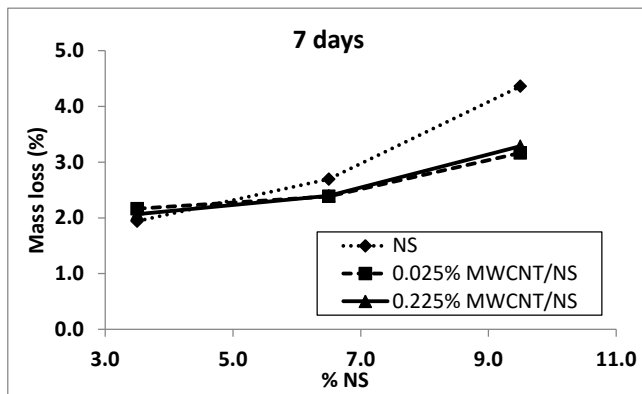
Figure 4. Mass loss results at 230 °C for Ca(OH)$_2$ pastes blended with MWCNT and NS after 7 days of hydration.
Source: The authors

For the ANOVA, the null hypothesis is that the factors MWCNT and NS do not have any effect over the mass loss of the pastes, and the alternative hypothesis is that the factors do have a significant effect over the mass loss. For a significance of $\alpha = 0.05$ and an associate confidence $\beta = 95\%$, it can be said that if a P-value is lower than α, the null hypothesis is rejected. For all the studied ages, the P-value of the factor MWCNT is always higher than 0.05 (accepting the

Figure 5. Mass loss results at 230 °C for Ca(OH)$_2$ pastes blended with MWCNT and NS after 21 days of hydration.
Source: The authors

Table 4.
ANOVA of the mass loss results at 230 °C of MWCNT/NS/Ca(OH)$_2$ pastes.

Age (days)	P-Value MWCNT	P-Value NS	R^2
1	0.505	0.042	81.06%
3	0.083	0.014	90.74%
7	0.397	0.018	87.46%
21	0.635	0.006	92.72%

Source: The authors

null hypothesis), and the P-Value of the factor NS is lower that 0.005 (rejecting the null hypothesis). This indicates that, with a 95% confidence for the all the tested ages, the addition of MWCNT did not have an effect on the mass loss of the pastes, and all the effects identified are from NS's pozzolanic activity. The experiments presented an average R^2 of 88%, which means that the statistical model was capable of accounting for 88% of the variability of the results.

To verify the correct sampling and statistical analysis of the results, the residue plots were verified for the assumptions of normality, constant variance, and independence of the response variable. These plots are not presented here due to space issues. For all the ages, it was concluded that the statistical model was valid because the plot of normal probability was linear, the histogram was Gauss shaped, and the residue versus order and adjustment plots did not show a tendency.

3.3. Super plasticizer demand

To guarantee that the measured strengths of all mortars were comparable, the flow of each mortar was standardized to a 110 ± 5 % value using a superplasticizer (SP), according to the ASTM 1437 standard. The SP demand of each mortar is presented in Table 5. It was found that the samples with NS presented a higher SP demand due to the high specific surface area of the nanoparticles. Mortars with MWCNT did not present a significant SP demand when compared to the plain cement sample because they already contained SP as a dispersing agent. Mortars blended with NS and MWCNT presented a SP demand similar to that of mortars blended only with NS. This indicates that the additive demand of NS is dominant over the presence of SP in MWCNT aqueous dispersion.

Table 5.
Superplasticizer demand of mortars blended with nanoparticles

Sample	SP (%)	Flow (%)
Cto	0.15	110.9
Cto/7.3%NS	1.32	105.0
Cto/0.0115%MWCNT	0.15	106.5
Cto/0.0115%MWCNT/7.3%NS	1.17	110.5

Source: The authors

Table 6.
ANOVA results for the results obtained for the flexural strength of mortars.

Factor	P-Value			
	1 day	3 days	7 days	28 days
MWCNT	0.640	0.642	0.004	0.221
NS	0.031	0.344	0.565	0.861
Interaction	0.935	0.068	0.030	0.199

Source: The authors

3.4. Mechanical strength

The four samples' flexural strength results for the four testing ages are presented in Fig. 6. Each strength result is presented with error bars of one standard deviation. These results were statistically studied using ANOVA. A summary of the P-values obtained for each factor and its interaction are presented in Table 6. For this given ANOVA, the null hypothesis is that the factors MWCNT and NS do not have any effect over the flexural strength of the mortars. The alternative hypothesis is that the factors MWCNT and NS do have a significant effect over the flexural strength of the mortars.

For a significance $\alpha = 0.05$, i.e. with a confidence $\beta = 95\%$, it can be said that if the obtained P-value is smaller than α, the hull hypothesis is rejected. This means that the MWCNT did not have any statistically significant effect on the flexural strength at any testing age, while NS only presented an improvement after 1 day of hydration. The MWCNT-NS interaction was statistically significant only after 7 days of hydration. Also, from Fig. 6 it can be seen that its effect was decreasing rather than increasing the flexural strength of the mortar.

The statistical analysis indicates that during the first day of hydration the MWCNT accelerated the pozzolanic activity of NS, which generated an improvement in the flexural strength by working as nucleation points. However, after 3 days of hydration, the improvement disappeared, and after 7 days it there was a decrease in strength. This can be explained by the reagglomeration process of the MWCNT [9-10], which inhibited the activity of the NS by decreasing its available surface area to react. After 7 and 28 days, not only the pozzolanic activity of NS was completely inhibited, but also the hydration process of cement was affected. This decreased the overall strength of the matrix.

Figure 6. Flexural strength of mortars blended with nanoparticles measured at 1, 3, 7, and 28 days.
Source: The authors
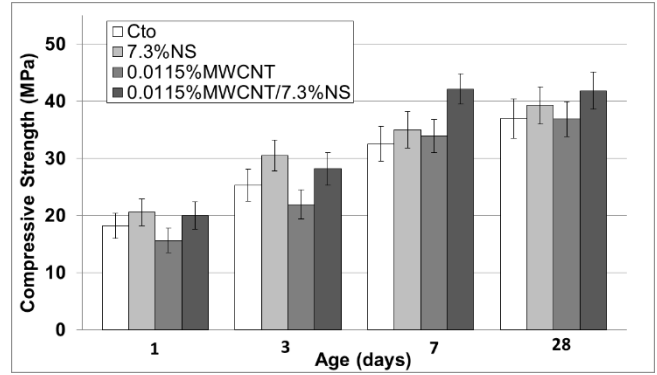
Figure 7. Compressive strength of mortars blended with nanoparticles measured at 1, 3, 7, and 28 days.
Source: The authors

Table 7.
ANOVA results for the results obtained for the compressive strength of mortars.

Factor	P-Value			
	1 day	3 days	7 days	28 days
MWCNT	0.095	0.149	0.385	0.685
NS	0.001	0.007	0.160	0.230
Interaction	0.270	0.804	0.860	0.655

Source: The authors

Compressive strength results for the 4 samples at the 4 testing ages are presented in Fig. 7. As with the flexural strength, all results are presented with error bars of one standard deviation. The statistical analysis was carried out with the same null and alternative hypotheses used for the flexural strength results. A summary of the P-values obtained for each factor and their interaction are presented in Table 7.

The addition of MWCNT did not have any statistically significant effect on the compressive strength of the samples at any testing age while NS presented a significant improvement only after 1 and 3 days of hydration. Again, this behavior can be explained by the reagglomeration process of the MWCNT. The pozzolanic activity of NS was probably inhibited by the MWCNT agglomeration after 7 and 28 days of hydration.

4. General discussion

From the results, it was verified that only NS presented a significant $Ca(OH)_2$ fixation that translated into a higher C-S-H production; MWCNT had no effect. This confirms the fact that MWCNT have no chemical affinity with the hydration products of cement [14], and any mechanical

improvement caused by the MWCNT is a consequence of physical effects such as nucleation or load distribution through a "bridge effect" [4].

Only during early ages were flexural and compressive strengths of the mortars improved by the combination of nanoparticles. At later ages the presence of the combination of nanoparticles was even found to be detrimental. These results are in agreement with previous reports that found that a predominance of a reagglomeration process of MWCNT inhibited the pozzolanic activity of NS and hindered the hydration reaction of cement at ages greater than 24 hours of curing [9-10]. This reagglomeration can account for the decrease in mechanical properties at later ages. The expected positive effects of the MWCNT on the mechanical properties of the cement matrix may have occurred before 1 day of hydration and were not detected in the experiments carried out due to a rapid reagglomeration of the MWCNT.

5. Conclusions

- NS particles maintained their pozzolanic activity up to the 7th day of hydration in $Ca(OH)_2$ pastes while MWCNT did not have any effect on it at any age.
- Mortars with MWCNT did not present a SP demand higher than that of plain cement ones. This is consequence of the presence of SP in the MWCNT as a dispersing agent.
- The addition of MWCNT did not have any significant effect on the compressive strength at any curing age while NS presented a significant improvement after 1 and 3 days of hydration.
- MWCNT did not present positive effects on the mechanical strength of mortars. After 3 days of hydration, they inhibited the pozzolanic activity of NS due to their reagglomeration process, which translated into a decrease in mechanical strength.

References

[1] Tobón, J.I., Payá, J.J., Borrachero, M.V. and Restrepo, O.J., Mineralogical evolution of Portland cement blended with silica nanoparticles and its effect on mechanical strength. Construction and Building Materials, 36, pp. 736-742, 2012. DOI: 10.1016/j.conbuildmat.2012.06.043

[2] Sanchez, F. and Sobolev, K., Nanotechnology in concrete – A review. Construction and Building Materials, 24(11), pp. 2060-2071, 2010. DOI: 10.1016/j.conbuildmat.2010.03.014

[3] Hewlett, P., Lea´s Chemistry of Cement and Concrete, 4th edition, Oxford: Elsevier Science & Technology Books, 2004.

[4] Li, G., Wang, P. and Zhao. X., Mechanical behavior and microstructure of cement composites incorporating surface-treated multi-walled carbon nanotubes. Carbon, 43(6), pp. 1239-1245, 2005. DOI: 10.1016/j.carbon.2004.12.017

[5] Yazdanbakhsh, A., Grasley, Z., Tyson, B. and Abu Al-Rub. R.K., Dispersion quantification of inclusions in composites. Composites Part A: Applied Science and Manufacturing, 42(1), pp. 75-83, 2011. DOI: 10.1016/j.compositesa.2010.10.005

[6] Shih, J., Chang, T. and Hsiao, T., Effect of nanosilica on characterization of Portland cement composite. Materials Science and Engineering: A, 424, pp. 266-274, 2006. DOI: 10.1016/j.msea.2006.03.010

[7] Srinivasan, S., Barbhuiya, S., Charan, D. and Pandey, S.P., Characterising cement–superplasticiser interaction using zeta potential measurements. Construction and Building Materials, 24(12), pp. 2517-2521, 2010. DOI: 10.1016/j.conbuildmat.2010.06.005

[8] Morsy, M.S., Alsayed, S.H. and Aqel, M., Hybrid effect of carbon nanotube and nano-clay on physico-mechanical properties of cement mortar. Construction and Building Materials, 25(1), pp. 145-149, 2011. DOI: 10.1016/j.conbuildmat.2010.06.046

[9] Mendoza, O., Sierra, G. and Tobón, J.I., Influence of super plasticizer and $Ca(OH)_2$ on the stability of functionalized multi-walled carbon nanotubes dispersions for cement composites applications. Construction and Building Materials, 47, pp. 771-778, 2013. DOI: 10.1016/j.conbuildmat.2013.05.100

[10] Mendoza, O., Sierra, G. and Tobón, J.I., Effect of the reagglomeration process of multi-walled carbon nanotubes dispersions on the early activity of nanosilica in cement composites. Construction and Building Materials, 54, pp. 550-557, 2014. DOI: 10.1016/j.conbuildmat.2013.12.084

[11] Rausch, J., Zhuang, RC. and Mäder, E., Surfactant assisted dispersion of functionalized multi-walled carbon nanotubes in aqueous media. Composites Part A: Applied Science and Manufacturing, 41(9), pp, 1038-1046, 2010. DOI: 10.1016/j.compositesa.2010.03.007

[12] Mendoza, O. and Tobón, J.I., An alternative thermal method for identification of pozzolanic activity in $Ca(OH)_2$/pozzolan pastes. Journal of Thermal Analysis and Calorimetry, 114, pp 589-596, 2013. DOI: 10.1007/s10973-013-2973-y

[13] Frias, M., Villar-Cociña, E., Saánchez, M.I. and Valencia-Morales. E., The effect that different pozzolanic activity methods has on the kinetic constants of the pozzolanic reaction in sugar cane straw-clay ash/lime systems: Application of a kinetic–diffusive model. Cement and Concrete Research, 35(11), pp. 2137-2142, 2005. DOI: 10.1016/j.cemconres.2005.07.005

[14] Makar, J.M. and Chan, G.W., Growth of cement hydration products on single-walled carbon nanotubes. Journal of the American Ceramic Society, 92(6), pp. 1303-1310, 2009. DOI: 10.1111/j.1551-2916.2009.03055.x

O.A. Mendoza-Reales, received a BSc. in Civil Engineering in 2009 and an MSc in Materials Engineering in 2013, both from the Universidad Nacional de Colombia. Medellin, Colombia. Currently he is a doctoral candidate in the Civil Engineering program at the Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. His research interests include: chemistry of cement, cement hydration, nanotechnology, pozzolanic materials, carbon nanotubes, oil well cementing pastes and environmental impact of cement based materials.
ORCID: 0000-0002-4241-1321

G. Sierra-Gallego, received a BSc. in Chemical Engineering in 1991, an MSc in Chemical Sciences in 2003 both from the Universidad de Antioquia, Colombia and a PhD in 2007 from Poitiers University. Currently, he is a Professor in the Materials and Minerals Department, Facultad de Minas, Universidad Nacional de Colombia. His research interests include: materials sciences, synthesis of carbon nanotubes, heterogeneous catalytic processes, biofuels and energetic evaluation of byproducts.
ORCID: 0000-0001-7981-2041

J.I. Tobón, received a BSc. in Geological Engineering in 1992, an MSc in Engineering in 2003, and a PhD in Science and Technology of Materials in 2011, all from the Universidad Nacional de Colombia, Medellin, Colombia. From 1992 to 1995, he worked for different companies in mining and oil; from 1995 to 1999 he worked for Cementos Argos S.A., at the same time from 1993 to 1999 he worked for the Universidad Nacional de Colombia as a part-time Professor, and since 1999 he has worked for the Universidad Nacional de Colombia on a full-time basis. Currently, he is a full professor in the Materials and Minerals Department, Facultad de Minas, Universidad Nacional de Colombia. His research interests include: industrial application of minerals and rocks, chemistry and mineralogy of cements, nanotechnology in construction materials, alternative cementitious materials, and high performance cements and concretes.
ORCID: 0000-0002-1451-1309
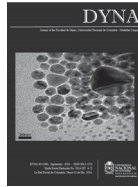
Influence of solid and liquid antioxidants on the formation of space charge in the XLPE insulation of medium voltage cables

Idalberto Tamayo-Ávila [a], José Manuel Nieto-Jalil [b], Jorge René Viteri-Moya [c] & Jenny Paulina Jácome-Jácome [a]

[a] Faculty of Engineering Sciences Industries, Equinoctial Technological University, Occidental Campus, Quito, Ecuador. idalberto.tamayo@ute.edu.ec
[b] Mechatronics Engineering Department, Monterrey Institute of Technology and Higher Education, Sonora Norte Campus, Mexico. jnietoj@itesm.mx
[c] Faculty of Chemical Engineering, Central University of Ecuador, Quito, Ecuador. jviteri@uce.edu.ec

Abstract
In this work, the influence of antioxidants in solid and liquid states was studied on the formation of space charge when cross-linked polyethylene (XLPE) was insulated in medium voltage cables that were thermally treated for up to 168 hours at a temperature of 120 °C. The results, which used the pulsed electro-acoustic technique (PEA) with a voltage of 120 kV applied to the cable show the use of solid antioxidants in an increased formation of bulk charge in the XLPE in comparison to those who have used liquid antioxidant. The PEA measures are consistent with those obtained by the thermally stimulated depolarization currents (TSDC) technique, which also have a greater depolarizing current in samples that use a solid antioxidant.

Keywords: Space charge, cross-linked polyethylene, solid and liquid antioxidant, Pulsed Electro-acoustic (PEA), Thermally Stimulated Depolarization Currents (TSDC).

Influencia del antioxidante sólido y líquido sobre la formación de carga de espacio en el aislamiento de XLPE de cables de media tensión

Resumen
En este trabajo se estudió la influencia del antioxidante en estados sólido y líquido en la formación de carga de espacio en el aislamiento de polietileno reticulado (XLPE) en cables de media tensión que fueron tratados térmicamente hasta 168 horas a la temperatura de 120 ºC. Los resultados mediante la técnica del Pulso Electroacústico (PEA) para una tensión de 120 KV aplicada al cable muestran que el uso de antioxidante sólido da como resultado una mayor formación de carga volumétrica en el XLPE comparado con los que se han usado antioxidante líquido. Las medidas PEA son coherentes con las obtenidas por la técnica de las Corrientes de Despolarización Estimuladas Térmicamente (TSDC) que también evidencian mayor corriente de despolarización en las muestras con antioxidante sólido.

Palabras clave: Carga de espacio, polietileno reticulado, antioxidantes sólido y líquido, Pulso Electroacústico (PEA), Corrientes de Despolarización Estimuladas Térmicamente (TSDC).

1. Introduction

The phenomenon of space charge is gaining more importance in the industry every day. It is also becoming more important for researchers as it helps them understand the space charge generating mechanisms or level of applied voltage, monitor the production process of insulating materials and study the possible deficiencies in the behavior of high voltage insulators. The term space charge describes the electrical condition of a poor conductor material that maintains electrical charges that are relatively immobile. The study of the effect of space charge on the properties of insulating materials is becoming more important due to its influence on the aging processes and dielectric rupture, which are limitations in their performance [1,2].

It is well known that having a space charge in insulating
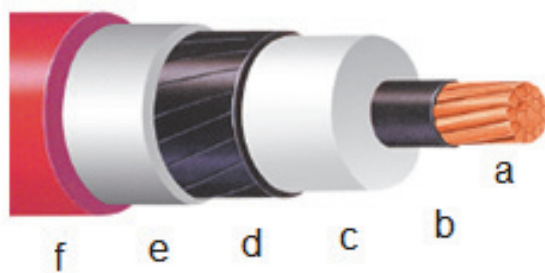
Figure 1. Medium voltage cables
Source: Adapted from internet http://www.viakon.com/pdf/categorias/10.pdf

materials increases the electric field several times in relation to the applied field. This contributes to the increase in conductivity and triggers the dielectric rupture process that eventually decreases the cable performance and its service life.

A great deal of research work has addressed the causes of space charge formation in XLPE cable insulation. The phenomena in which space charge formation in materials have important effects include temperature [3,4], acetophenones [5], the by-products from the cross-linking reaction [6], antioxidants [7], as well as dielectric strengths, arborescences and aging. During extrusion of medium voltage cables, polyethylene insulation is subjected to stresses that cause cracks in the material [8] contributing formation and charge conduction.

Taking into account all of the above, it is important to know all the factors that contribute to space charge formation in order to reduce its presence in cable insulation and thus optimize it to extend its service life and improve its performance. Currently, dielectric materials doped with nanoparticles that reduce the formation of space charge are being researched [9,10].

2. Medium voltage cables

The medium voltage cables in this study (Fig. 1) are obtained through a process of vulcanization in which temperatures of up to 220ºC can be reached. It consists of a hard aluminum conductor (a), two semiconductor layers (b and d) that form interfaces with the XLPE insulation (c), a copper metallic shield (e) and a PVC sheath (f). In general, these cables are used in primary distribution networks. They work within a 10-20 kV voltage range and during their operation they can reach temperatures up to 90 ºC.

3. Pulsed electro-acoustic technique (PEA)

Space charge measurement is one of the most important variables to be studied in order to discover the factors that degrade insulation and also for the purpose of improving the quality of insulating materials. There are many techniques used to study space charge measurement that have emerged since the end of the last century and the beginning of this one. For this study, we are focusing on the Pulsed Electro-acoustic (PEA) technique. It consists of applying an electric pulse of a controlled frequency and amplitude to the dielectric material. The facilities used are shown in Fig. 2.

This technique consists of applying a controlled frequency electric pulse (pulse time = 40ns) and amplitude (400V) to the dielectric material using a pulse generator (1). The objective of this experiment is to find out the different samples of a disturbance.

The samples are placed between two electrodes in the cell that is to be measured (5) in a high voltage DC (2), at 120 kV. The consequence is that the internal electric field is modified which causes an electromagnetic interaction and creates an elastic wave. This is proportional to the charge that propagates to a piezoelectric transducer, transforming itself into a voltage signal. From the voltage signal profile load space with a deconvolution software is obtained.

The amplitude of the acoustic wave in the material is recorded in an oscilloscope (3) and this is related with the quantity and its distribution in the sample volume. The cable samples studied were conveniently prepared so that the pulsed electroacoustic (PEA) could be applied with high fields. The measurements were undertaken in the Dielectric Materials Physics Laboratory (DILAB) at Barcelona Tech.

4. Thermally Stimulated Depolarization Current Technique (TSDC)

This technique is based on the combined effect of electric field and temperature to polarize the material and freeze the polarization mechanisms while samples are being cooled (Fig. 3). Next, the material is depolarized by increasing the temperature, which leads to a depolarization electric current. Integrating the curve $I = f(T)$ gives a measure of the amount of charge in the volume of the material.

The methodology used for the polarization and depolarization is described below:

a) The sample is heated up to the temperature of polarization T_p (1)

b) The polarizing field E_p is applied with high voltage DC (2) while the temperature remains constant.

c) The sample is cooled at a constant velocity (V_e) until the temperature T_0 (non-isothermal polarization) (6).

d) The electric field is eliminated, which short-circuits the sample through an electrometer at temperature T_0.

e) The sample is heated, usually at a constant velocity (V_c), generating thermo-stimulated depolarization while the current detected by the electrometer (4) is recorded according to the temperature in the PC (5).

Figure 2. Facilities for the measurements using PEA
Source: The authors.
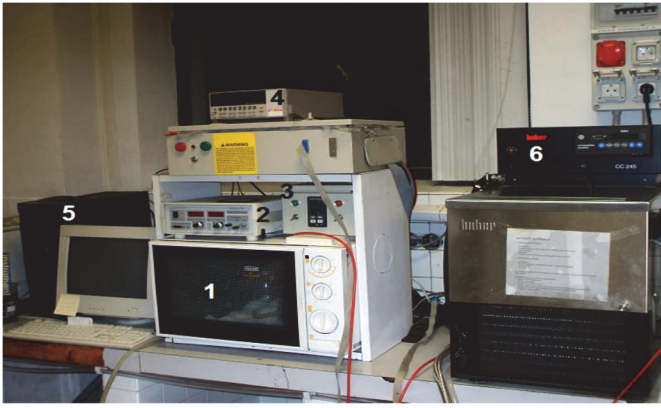
Figure 3. Facilities for the measurements using TSDC.
Source: The authors.

Figure 5.XLPE with liquid antioxidant 0 hrs at 120 ºC.
Source: The authors.

5. Results using the PEA technique

Figs. 4 and 5 show the PEA measurements on the cables that were not thermally treated. By undertaking a comparison, we are able to see that in the sample with liquid antioxidant a positive charge profile distribution can be observed that is close to the outer insulation surface (cathode). This is not observed in the cable with a solid antioxidant. If you refer to the charge density values in the Table 1 for 0 hours, it can be seen that the sample with liquid antioxidant charge density is higher in the cathode, the anode and the volume profile.

When the samples are treated for 8 hours at 120ºC, we can observe a similar charge distribution on both insulations near the cathode and anode (Figs. 6 and 7). However in the case of the solid antioxidant, two small charge distributions are formed close to the outer electrode. In the one with the liquid antioxidant, the initial distribution splits in two to produce a lesser value. The charge in the cathode is small for the sample with liquid antioxidants (values for 8 hours), as can be noted in Table 1.

Figure 6. XLPE with solid antioxidant at 120ºC for 8 hrs.
Source: The authors.

Finally, when the samples are treated for 168 hours at 120ºC, it can clearly be observed in Figs. 8 and 9 that, in the XLPE solid antioxidant charge distributions are formed all the way to the center of the insulation. However, in the XLPE with liquid antioxidant, the charge in the volume distributions decreases and a better behavior in terms of space charge is achieved. Also, the charge on the inner cathode continues to decrease. This can be seen in Table 1 charge values for 168 hours

6. Results using the TSDC technique

When comparing the TSDC measurements, note the area under each curve in Figs. 10 and 11. It is clear that the sample with liquid antioxidant and no thermal treatment (0 h) has

Figure 4.XLPE with solid antioxidant at 120 ºC for 0 hrs.
Source: The authors.
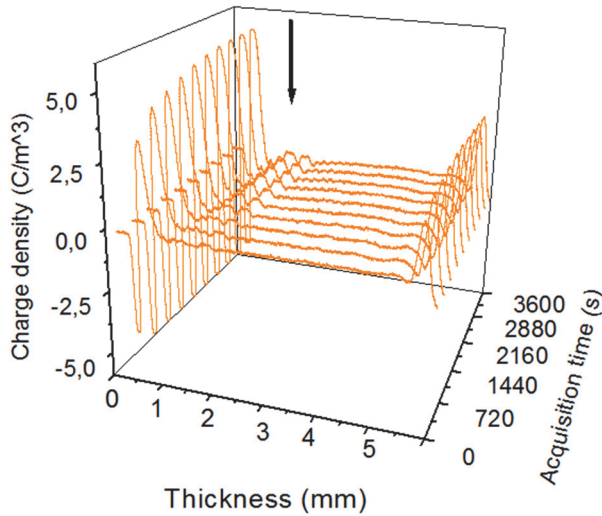
Figure 7. XLPE with liquid antioxidant at 120ºC for 8 hrs.
Source: The authors.
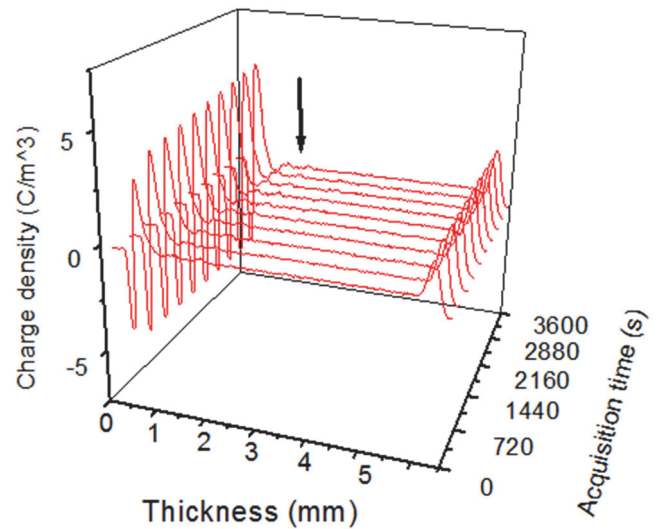
Figure 9. PEA in XLPE liquid antioxidant 168 hrs at 120ºC
Source: The authors.
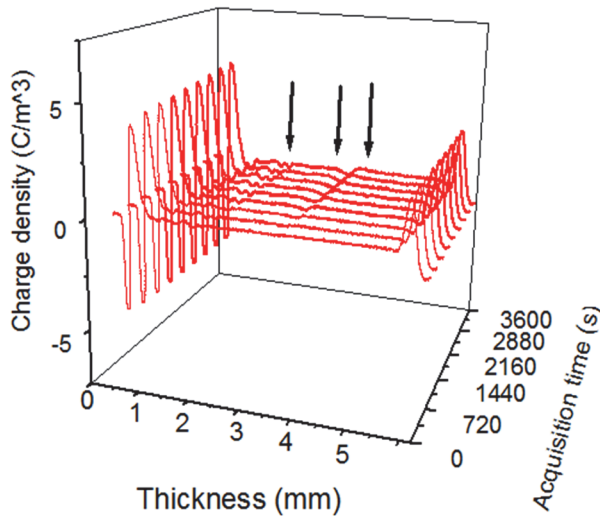
Figure 8. XLPE with solid antioxidant at 120ºC for 168 hr
Source: The authors.

Table 1.

Treatment To 120⁰ C	Outer Electrode (anode)	Positive Profile interface	Volume Profiles	Inner Electrode (cathode)
XLPE with Antioxidants	Density C/m³	Density C/m³	Density C/m³	Density C/m³
Solid (0 hours)	-9.87	5.86	Not	1.63
Liquid (0 hours)	-10.17	7.27	2.48	2.78
Solid (8 hours)	-5.12	5.11	0.56 0.10	2.28
Liquid (8 hours)	-5.20	5.09	0.45 -1.47	1.82
Solid (168 hours)	-5.12	4.79	0.07 0.07 0.16 -0.80	2.15
Liquid (168 hours)	-5.12	4.78	Not	1.64

Source: The authors.

more accumulated charge than the sample with solid antioxidant. However, if the treatment time is increased at a temperature of 120 ºC on both samples, the results is different behavior. In the case of the sample with a solid antioxidant, the accumulated charge increases from 0 to 168 hours. However, in the one with the liquid antioxidant, the charge decreases.

In Table 2 we can observe the changes occurring in each case. These were calculated by obtaining the area under the curve in Figs. 10 and 11.

Although the liquid antioxidant initially generates more charge, after the thermal treatment it produces a better behavior since it tends to decrease the accumulated charge. This result may be caused by the fast decomposition of the liquid antioxidant, which produces more by-products that contribute to the formation of more charge. However, with

Figure 10.TSDC in XLPE with solid antioxidant
Source: The authors.

Figure 11. TSDC in XLPE with liquid antioxidant
Source: The authors.

Table 2.

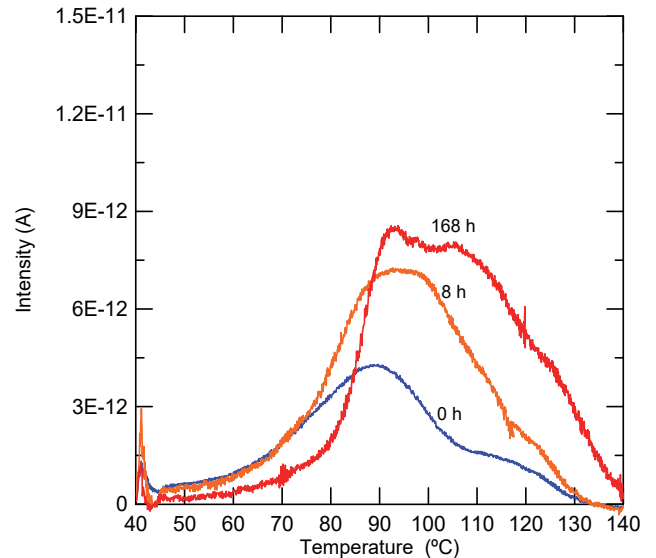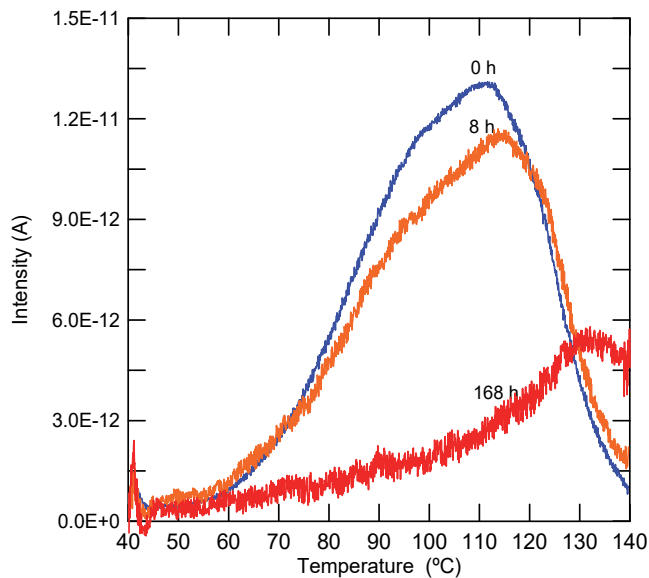| Treatment Time at 120ºC (hours)	Solid Antioxidants Charge (nC)	Liquid Antioxidants Charge (nC)
0	5.09	17.73
8	8.42	16.30
168	10.45	6.17

Source: The authors.

the treatment they vaporize and are released from the XLPE, leaving a more homogeneous structure with less defects and less charge. We suggest either using the liquid antioxidant and increasing the crosslinking temperature during the process or having previous treatment to guarantee a thermal cleaning of the by-products for the crosslinking reaction. This will guarantee a lower amount of space charge in the volume of the XLPE and in the XLPE-semiconductor interface.

7. Conclusions

The use of a solid antioxidant in the XLPE insulation of the medium voltage cables studied contributes to the formation of a greater amount of space charge. This could be due to the fact that the solid antioxidant does not totally decompose in the cross-linking process, and the residues create defects in the inner structure. These act as a charge trapping and accumulation center.

The results obtained by using the TSDC technique are in line with those obtained from the PEA technique. Both cases show that the samples with solid antioxidants accumulate more space charge if they are heat treated for a longer time at 120 ºC.

To understand the complex form in which space charge in the cable insulation evolves, there is a need to carry out systematic studies that involve thermal and electric aging. It should be noted that space charge is a degrading factor for insulating materials that can foster other factors and that can also be unfavorable for insulation as it can cause partial discharges and dielectric ruptures. As we have demonstrated in this study performing a simple analysis on the samples with no aging can lead to contradictory interpretations.

References

[1] Chen, G., Space charge and its impact on DC breakdown of polymeric materials. *18th International Symposium on High Voltage Engineering.* Seoul, KR, pp. 686-691, 2013.

[2] Sekii, Y., Suzuki, H., Noguchi, K. and Maeno, T., The negative heterocharge generation mechanism in polymeric dielectrics. IEEE Conf. Electr. Insul. Dielectr. Phenomena (CEIDP), pp. 404-408, 2007.

[3] Fu, F., Chen, G., Dissado, L. and Fothergill, J.C., Influence of thermal treatment and residues on space charge accumulation in XLPE for DC Power Cable Application, IEEE Trans. Dielectr. Electr. Insul., 14, pp. 53-64, 2007. DOI: 10.1109/TDEI.2007.302872

[4] Tamayo, I., Belana, J., Cañadas, J.C., Mudarra, M., Diego, J.A. and Sellarès, J., Thermally stimulated depolarization currents of crosslinked polyethylene relaxations in the fusion range of temperatures. J. Polym Sci. Part B: Polym. Phys., 41, pp. 1412-1421, 2003. DOI: 10.1002/polb.10489

[5] Doi, T., Tanaka, Y. and Takada, T., Measurement of space charge distribution in acetophenone coated low-density polyethylene. IEEE Annual Report of CEIDP, pp.32-35, 1997. DOI: 1109/CEIDP.1997.634552

[6] Maeno, Y., Hirai, N., Ohki, Y., Tanaka, T., Okashita, M. and Maeno, T., Effect of cross-linking byproducts on space charge formation in cross-linked polyethylene. IEEE Trans. Dielectr. Electr. Insul., 12, pp. 90-97, 2005.

[7] Sekii, Y., Taya, A. and Maeno, T., Effect of antioxidants on space charge generation in cross-linked polyethylene and EPR. IEEE Conf. Electr. Insul. Dielectr. Phenomena (CEIDP), pp. 133-137, 2006.

[8] Lavoie, F.L., Bueno, B.deS. y Lodi, P.C., Evaluación de la fisuración bajo tensión de geo membranas pos ensayos acaelerados. DYNA, 81(183), pp. 215-220, 2014

[9] Ohki, Y., Ishimoto, K., Kanegae, E., Tanaka, T., Sekiguchi, Y., Murata, Y. and Reddy, C.C., Suppression of packet-like space charge formation in LDPE by the addition of magnesia nanofillers. Properties and Applications of Dielectrics Materials. ICPADM 2009, pp. 9-14. 2009. DOI: 10.1109/ICPADM.2009.5252266

[10] Wang, X., Zepeng, L., Kai, W., Chen, X., Demin, T. and Dissado, L.A., Study of the factors that suppress space charge accumulation in LDPE nanocomposites. IEEE Transactions on Dielectrics and Electrical Insulation 21(4), pp. 1670-1679. 2014. DOI: 10.1109/TDEI.2014.004292

I. Tamayo-Ávila, gained a PhD in Applied Physics and Science Simulation from the Polythechnic University of Catalonia, Spain, in 2002. He has a BSc.degree in Physics and Astronomy from ISPH de Holguín, Cuba, in 1997. He worked as a professor of Physics at the Universidad de Holguín and as a researcher in Spectrum Analysis of Atomic Emission with Laser in Steel (1990-1998). He was an Associate Professor at the Polythechnic University of Catalonia (2005-2012) and a member of the DILAB group. He has worked on several research projects that have been published in scientific journals. He is currently a Professor in the Faculty of Engineering Sciences Industries Faculty at the Equinoctial Technological University in Quito, Ecuador.
ORCID ID: 0000-0001-8752-8235

J.M. Nieto-Jalil, has a PhD in Physics from the University of Sonora, Mexico, in 2006. He is a full time Professor and Director of Mechatronics at the Monterrey Institute of Technology and higher Education, Sonora Norte Campus (since 2004). He has been a columnist in the Expreso Newspaper, and the State of Sonora since 2010 and also in the Ingenio Magazine. He is

an evaluator in the National Register of Scientific and Technological Institutions and Enterprises of Mexico and a member of the FOMIX Evaluation Committee of Sonora as well as a technology Firms Consultant. He is also a researcher in Mechatronics and Nanotechnology areas.
ORCID ID: 0000-0001-9908-2621

J.R. Viteri-Moya, has a PhD in Technical Sciences from the University of Matanzas, Cuba. He has been a Production Manager at EDESA-Quito and EDESA VENCERÁMICA-Venezuela, a Plant Manager at Foundation Mariana de Jesus, a director of the Editorial Team at the Enfoque UTE Magazine, a representative for CONESUP at the Environmental Observatory in Quito, a member of the Technical Advisory Board at the Ecuadorian Institute of Standardization-INEN, a member of the Ecuadorian Certification Committee S.G.S., and an external evaluator for CEAACES. He is currently a Professor in the engineering Sciences Faculty at the Central University of Ecuador.
ORCID ID: 0000-0001-6444-301X

J.P. Jácome-Jácome, has a MSc. in Computer Engineering and Systems from the Autonomous University of the Andes and she has an undergraduate degree in Computer Science from Central University of Ecuador. She has a PGDip in Auditing Information Systems, a PGDip in Curriculum Design, a MSc. in Education Management, Sp. in Higher Education. She worked as an Instructor at the ICAM-Quito (Municipal Training Institute of Quito) and has also as a teacher at the Equinoctial Technological University in the Faculty of Engineering Sciences Industries since 2011.
ORCID ID: 0000-0003-0314-7246

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN

FACULTAD DE MINAS

Área Curricular de Ingeniería
Eléctrica e Ingeniería de Control

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Eléctrica

Mayor información:

E-mail: ingelcontro_med@unal.edu.co
Teléfono: (57-4) 425 52 64

Method for determining total losses in distribution transformers from the percentage of carbon silicon steel sheet

María Gabriela Mago-Ramos [a]; Luis Vallés [b]; Jhon Olaya [c]; Martha Zequera [d] & Jhon Vera-Vera [e]

[a,e] Vicerrectoría de Investigación, Universidad ECCI, Bogotá, Colombia. mmagor@ecci.edu.co, jverav@ecci.edu.co
[b] Doctorado en Ingeniería, Universidad de Carabobo, Valencia, Venezuela. Le.vallesd@gmail.com
[c] Universidad Nacional de Colombia, Facultad de Ingeniería, Bogotá, Colombia. jjolayaf@unal.edu.co
[d] Facultad de Ingeniería, Universidad Pontificia Javeriana de Bogotá, Bogotá, Colombia.mzequera@javeriana.edu.co

Abstract
Transformers are electrical devices that are responsible for the distribution and management of electrical power. In Latin America the recycling of silicon steel sheet transformer is very common and therefore its construction is very important because we must consider the losses that may occur by eddy currents and magnetic hysteresis especially when the team has operated under load condition. This paper proposes a novel method to determine the total losses in the single-phase distribution transformers Pole, analyzing the percentage of carbon silicon steel sheet and chemical characteristics. This diagnostic checks the iron losses in proportion to the square of the frequency induction and also applying the formula Steinmentz order to perform diagnosis and failure prediction. Although the results have approximation errors are useful for diagnosis and reuse of the silicon steel sheet, since changes in the microstructure (grain size in the nanometer scale) or ratio are significant quality B-H The contribution to the manufacturers and the industry in general is exegetical, because if they are considered the values set the standard for equipment operating under this condition, the service conditions are improved distribution networks and consequently, reliability and / or maintenance in electric utilities.

Keywords: Total losses, distribution transformer, silicon steel sheet.

Método para determinar las pérdidas totales en transformadores de distribución a partir del porcentaje de carbono de la chapa de acero al silicio

Resumen
Los transformadores son dispositivos eléctricos que se encargan de la distribución y manejo de la corriente eléctrica. En América Latina el reciclaje de la chapa de acero al silicio de los transformadores es muy común y por esta razón su construcción es de gran importancia debido hay que considerar las pérdidas que se pueden presentar por corrientes parásitas e histéresis magnética sobre todo cuando el equipo ha operado bajo condición de carga. Este artículo propone un novedoso método para determinar las pérdidas totales en los transformadores de distribución monofásicos de poste, analizando el porcentaje de carbono de la chapa de acero al silicio y sus características químicas. Este diagnóstico verifica las pérdidas en el hierro de forma proporcional al cuadrado de la inducción y frecuencia aplicando también la fórmula de Steinmentz con el fin de realizar diagnóstico y predicción de fallas. Los resultados obtenidos aunque presentan errores de aproximación son de gran utilidad para el diagnóstico y la reutilización de la chapa de acero al silicio, ya que los cambios en la micro estructura (tamaño de grano en escala nanométrica) o relación entre la calidad B-H son significativos, el aporte a los fabricantes y la industria en general es exegético, porque si son considerados los valores establecidos en la norma para los equipos operando bajo esta condición, se mejoran las condiciones del servicio en las redes de distribución y en consecuencia, la confiabilidad y/o mantenimientos en las empresas del sector eléctrico.

Palabras claves: Pérdidas totales, transformador de distribución, chapa de acero al silicio.

1. Introducción

Este artículo tiene como objetivo fundamental determinar las pérdidas totales en el núcleo de los transformadores de distribución monofásicos de poste debido a las Corrientes de Foucault "*que se producen en cualquier material conductor cuando se encuentra sometido a una variación de flujo magnético*". Se conocen también como pérdidas por corrientes parásitas y dependerán del material del que esté constituido el núcleo magnético y debido a la Histéresis Magnética, cuyo fenómeno "*se produce cuando la imantación de los materiales ferromagnéticos no solo depende del valor del flujo sino también de los estados magnéticos anteriores, la inducción magnética dependerá del tipo de chapa*". A partir del porcentaje de carbono obtenido de la composición química de la chapa de acero al silicio [1,2].

Hay que considerar las propiedades físicas y magnéticas para evaluar las pérdidas totales que establece la norma [3] para equipos operando bajo carga, de manera que, se comprueben en forma experimental cambios importantes en las pérdidas por Histéresis Magnética y Corrientes de Foucault a partir de la magnetización utilizando aparatos especiales [4]. En los procesos de fabricación debe tomarse en cuenta el tratamiento térmico para garantizar la durabilidad ante condiciones de falla, es por ello que [5, 6] se debe verificar la microestructura, cristalografía y propiedades en el proceso de laminado que es fundamental para este tipo de material. En el mismo orden de ideas, hay que incorporar recocidos y pruebas de tensión donde se evalúe la anisotropía y energía magneto cristalina a partir de la textura obtenida [7,8]. Las modelaciones en las condiciones de diseño [9,10,11] evidencian los efectos del coeficiente de oscilación del material mediante el análisis del impacto estructural, al igual que los cambios en el borde de corte [12]. Lo indicado anteriormente puede articularse para evaluar propiedades físicas y magnéticas en forma experimental. Otro comportamiento correspondiente a la condición de falla versus las pérdidas en el núcleo que ha sido objeto de estudio, es la polarización y frecuencia de magnetización versus la inducción circular teniendo en cuenta el efecto de la piel [13] que en investigaciones futuras permitiría construir una ecuación de difusión electromagnética a través del espesor de la muestra. Es conveniente considerar la dirección de la magnetización a fin de comprender el comportamiento de la chapa de acero al silicio [14, 15], el grado de acero eléctrico, la orientación de grano y las características microestructurales para las pruebas experimentales que estos requieran [16,17], incluso en la actualidad hay investigadores que están valorando las pérdidas en el núcleo con precisión [18] utilizando un método de aproximación para cuatro materiales magnéticos típicos de uso frecuente en electrónica de potencia: con 3% de acero al silicio de grano orientado, 6,5% de chapa de acero de silicio, material amorfo, y materiales nanocristalinos magnetizados con una tensión rectangular simétrica y asimétrica cuyas deducciones experimentales indican, que la componente de corriente continua de la intensidad de campo magnético causa errores en la medición, que para el caso propuesto en este artículo
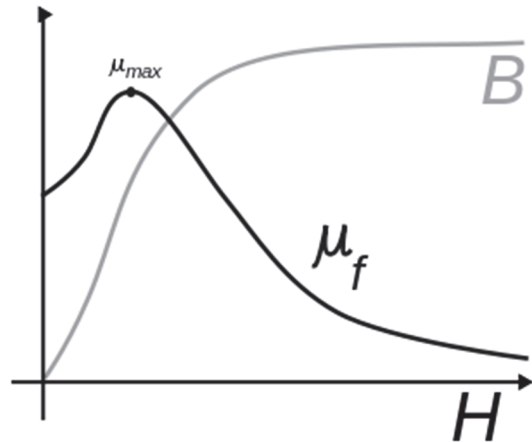
Figura. 1. Efecto de la Saturación Magnética.
Fuente: M. Mago, 2013

corresponde, a la aplicación de un método numérico que también presenta esta condición de error, sin embargo, la importancia radica en que si el material es reutilizado por las empresas que reparan o fabrican estos equipos de acuerdo al nivel de pérdidas totales obtenido, pueden decidir si es factible o no realizar este proceso para los transformadores de distribución. Aunque no se ha considerado en esta investigación es necesario tener en cuenta en investigaciones futuras, la eficiencia energética y el impacto ambiental para las pérdidas producidas en el interior del núcleo, cómo estas afectan el medio ambiente [19,20]. Las empresas fabricantes y los operadores del sistema eléctrico deben trabajar en políticas conjuntas para mejorar esta condición.

Las propiedades físicas del material como Resistencia Eléctrica, Conductibilidad Térmica, Saturación Magnética, Fuerza Coercitiva, Inducción Remanente y Permeabilidad Magnética son condiciones fundamentales para determinar las características de la aleación en estudio y por consiguiente las pérdidas en el núcleo [21]. Dentro de esas propiedades físicas la Saturación Magnética "*es un efecto que se observa en algunos materiales magnéticos, y se caracteriza como el estado alcanzado cuando cualquier incremento posterior en un campo de magnetización externo H no provoca un aumento en la magnetización del material. Esto demuestra que el campo magnético total B tiende a estabilizarse*".

Con esta propiedad física, se determina la Densidad de Flujo Magnético $B_{máx}$ y por consiguiente, las pérdidas totales.

2. Metodología

Se recolectaron muestras de equipos fallados procedentes de diferentes regiones de Colombia (Antioquia, Pacífico, Cundinamarca, Bogotá, y Caldas) las cuales corresponden a la población del estudio (Tabla 1), a las cuales se les realiza análisis químico para determinar el porcentaje de carbono con el equipo Spectrovac, Marca Baird (Tabla 2). Esta evaluación incluye dos (02) mediciones y el valor promedio verificando los contenidos de Hierro (Fe), Carbono (C), Manganeso (Mn), Fosforo (P), Azufre (S), Silicio (Si), Cobre (Cu), Niquel (Ni), Cromo (Cr), Vanadio (V),

Molibdeno(Mo), Wolframio o Tusgteno (W), Cobalto (Co), Titanio (Ti), Estaño(Sn), Aluminio (Al), Niobio(Nb), Boro (B), Plomo (Pb), Magnesio (Mg). El porcentaje de carbono obtenido y comparado con el valor promedio es el referente utilizado para determinar las pérdidas totales del núcleo de los transformadores fallados.

Es necesario verificar las características de la aleación utilizada en cuanto a espesor, tolerancia y porcentaje de silicio, a fin de comparar las pérdidas de potencia que establece el fabricante respecto a energía/peso (W/kg) (Tabla 3).

Tabla 1.
Población del Estudio: selección definitiva

Departamento de Procedencia del Transformador	Tipo de Falla	Cantidad
Caldas	Térmica	1
	Arco Eléctrico	1
	Descarga Parcial	1
Antioquia	Térmica	1
	Arco Eléctrico	2
	Descarga Parcial	1
Pacífico	Térmica	1
	Arco Eléctrico	1
	Descarga Parcial	1
Cundinamarca	Térmica	1
Bogotá	Térmica	1
	Arco Eléctrico	3
	Descarga Parcial	1
TOTAL		16

Fuente: M. Mago, 2011.

Tabla 2.
Análisis Químico

Regiones	Fallas	% Carbono
REGION ANTIOQUIA	TERMICA	0,193
	DESCARGA PARCIAL	0,069
	ARCO ELECTRICO	0,010
REGION CALDAS	TERMICA	0,067
	ARCO ELECTRICO	0,040
	ARCO ELECTRICO/TERMICA	0,019
	DESCARGA PARCIAL	0,014
REGION CUNDINAMARCA	TERMICA	0,031
REGION BOGOTA	ARCO ELECTRICO/TERMICA	0,105
	ARCO ELECTRICO	0,093
	DESCARGA PARCIAL	0,101
	TERMICA	0,1735
REGION PACIFICO	TERMICA/ARCO ELECTRICO	0,131
	ARCO ELECTRICO	0,002
	DESCARGA PARCIAL	0,033
PATRON DE REFERENCIA		0,033

Fuente: M. Mago, 2012.

Tabla 3.
Características de la aleación.

Espesor	Tolerancia	Aleación % SI	1 Tesla (10⁴ Gauss) W/kg
0,5	0,10	0,5 – 1	2,90
0,5	0,10	2,5	2,3
0,35	0,10	2,5	1,7
0,35	0,10	4	1,3
0,35	0,10	4,5	1,2
0,35	0,10	4,5	0,9

Fuente: M. Mago, 2012

Finalmente, se aplican métodos numéricos con el programa Matlab©, construyendo una *Recta de Regresión y=Ax+B que mejor se ajusta en el sentido de los mínimos cuadrados a los N datos $(X_1,Y_1),....,(Xn,Yn)$,* obteniendo el valor de *Saturación Magnética* [22].

A continuación, se aplica un ejemplo de cálculo para la muestra procedente del Departamento de Antioquia:

Donde:
X= esta serie corresponde al porcentaje de carbono de la aleación
Y= esta serie corresponde a la Saturación Magnética

* Tipo de Falla: Térmica. Causa: Sobrecarga.
>> X=[0 0.1 0.193 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2];
>> Y=[2000 1900 1850 1800 1700 1600 1450 1300 1100 900 700 500 300 0];
>> lsline(X,Y)

ans = -1.6463e+003
>> sum(X) ans = 7.9910
>> sum(Y) ans = 17100
>> G=((-7.9910*-1.6463e+003)+17100)/(14) = 2.1611e+003
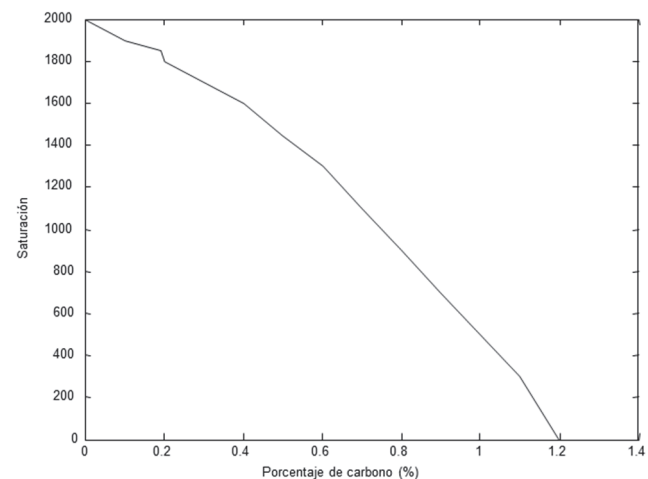>> y=(-1.6463e+003*(0.193))+2.1611e+003 = 1.8434e+003

Figura 2. Cálculo de la saturación magnética. Región Antioquia. Falla Térmica.
Fuente: M. Mago, 2013

El valor obtenido de 1.8434e+003 de Saturación Magnética, corresponde a la densidad de flujo máximo, el cual será utilizado para el determinar las pérdidas totales tomando en consideración el peso del equipo, que para un transformador de capacidad 5 kVA es de 75_kg.

Cálculo de pérdidas totales: pérdidas por corrientes de Foucault y pérdidas por Histéresis Magnética.

En un transformador se producen pérdidas por corrientes de Foucault *(Pf)*, pérdidas por histéresis *(Ph)* y en el cobre del bobinado *(Pcu)*

Pérdidas por corrientes de Foucault: se obtienen a partir de la siguiente expresión de la eq (1):

$$P_f = \frac{2,2\, f^2\, \beta m\grave{a}x^2\, \Delta^2}{10^{11}} \tag{1}$$

Donde:
P_f = pérdidas por corriente de Foucault (W/kg).
f = frecuencia en Hz.
Bmáx = inducción máxima en Gauss.
Δ = espesor de la chapa en mm.

Pérdidas por Histéresis Magnética: se obtiene aplicando la fórmula de Steinmentz indicada a continuación en la eq (2):

$$P_h = K_F\, f.\beta m\grave{a}x^n \tag{2}$$

Donde:
P_h = pérdidas por Histéresis Magnética (W/kg).
K_h = coeficiente de cada material.
f = frecuencia en Hz.
Bmáx = inducción máxima en Tesla.
n = 2 para β›1 Tesla y 1,6 para β‹1 Tesla

Pérdidas Totales: las pérdidas de potencia en el hierro *(PFE)* o en el núcleo magnético son la suma correspondiente a las pérdidas por Foucault *(Pf)* e histéresis *(Ph)* tal y como, se indica a continuación en la eq (3):

Pérdidas Totales en el Núcleo (Pt) $= P_t = (P_f) + (P_h)$ (3)

La ecuación final utilizada corresponde a la siguiente expresión, eq (4) [23]:

$$B_{m\grave{a}x} = \sqrt{\left(\frac{10^{11}\, P_t}{10^{11}\, K_h\, f + 2{,}2\, f^2\, \Delta^2}\right)} \tag{4}$$

3. Resultados

En las Tablas 4, 5, 6 y 7 se muestran los resultados obtenidos los cuales han sido comparados de forma adicional con las pruebas magnéticas:

4. Discusión de resultados

En las Tablas 4, y 5 se muestran los cálculos de pérdidas por *Corrientes de Foucault*, e *Histéresis Magnética* obtenidos de acuerdo al método propuesto, indicando en las mismas regiones de procedencia, tipos de fallas, valores obtenidos

Tabla 4.
Pérdidas por Corrientes de Foucault.

Región Antioquia	Tipo de falla	% de Carbono	Pruebas magnéticas
Muestra 1	Térmica	1,3608019	1,00939608
Muestra 2	Descarga parcial	1,3986951	1,00939608
Muestra 3	Arco eléctrico	1,4163174	1,00939608
Región Caldas			
Muestra 5	Térmica	1,39874171	1,00939608
Muestra 6	Arco eléctrico	1,4071184	1,01137626
Muestra 7	Arco eléctrico	1,41359885	1,00939608
Muestra 8	Descarga parcial	1,41500452	1,01335838
Región Pacífico			
Muestra 9	Térmica	1,88154258	1,44404568
Muestra 10	Arco eléctrico	1,418616	1,01335838
Muestra 11	Descarga parcial	1,408848	1,01335838
Región Cundinamarca			
Muestra 12	Térmica	1,653187	1,21267626
Región Bogotá			
Muestra 13	Arc eléctrico/Térmica	2,7915874	2,24450046
Muestra 14	Arc eléctrico	1,3913426	1,01137626
Muestra 16	Descarga parcial	1,3830362	1,01137626
Muestra 17	Térmica	1,3658764	1,01137626
Patrón Muestra 18		2,82052799	2,24155008

Fuente: M. Mago, 2012

Tabla 5.
Pérdidas por Histéresis Magnética.

Región Antioquia	Tipo de falla	% de Carbono	Pruebas magnéticas
Muestra 1	Térmica	0,16831935	0,13554054
Muestra 2	Descarga parcial	0,17299878	0,13554054
Muestra 3	Arco eléctrico	0,17518707	0,13554054
Región Caldas			
Muestra 5	Térmica	0,17300454	0,13554054
Muestra 6	Arco eléctrico	0,17404062	0,1355676
Muestra 7	Arco eléctrico	0,17484216	0,13554054
Muestra 8	Descarga parcial	0,17501602	0,13562171
Región Pacífico			
Muestra 9	Térmica	0,23272017	0,19504854
Muestra 10	Arco eléctrico	0,17546266	0,13559465
Muestra 11	Descarga parcial	0,17425457	0,13562171
Región Cundinamarca			
Muestra 12	Térmica	0,20447577	0,16388504
Región Bogotá			
Muestra 13	Arc eléctrico/Térmica	0,34527983	0,30460110
Muestra 14	Arc eléctrico	0,17208937	0,1355676
Muestra 16	Descarga parcial	0,17106199	0,13554054
Muestra 17	Térmica	0,16893957	0,13556760
Patrón Muestra 18		0,34885937	0,30456054

Fuente: M. Mago , 2012

Tabla 6.
Pérdidas Totales

Región Antioquia	Tipo de falla	% de Carbono	Pruebas magnéticas
Muestra 1	Térmica	114,688596	85,8702465
Muestra 2	Descarga parcial	117,877042	85,8702465
Muestra 3	Arco eléctrico	119,368088	85,8702465
Región Caldas			
Muestra 5	Térmica	117,880969	85,8702465
Muestra 6	Arco eléctrico	118,586927	86,020789
Muestra 7	Arco eléctrico	119,133076	85,8702465
Muestra 8	Descarga parcial	119,25154	86,173507
Región Pacífico			
Muestra 9	Térmica	158,569707	122,932067
Muestra 10	Arco eléctrico	119,555869	86,1714773
Muestra 11	Descarga parcial	118,732706	86,1735068
Región Cundinamarca			
Muestra 12	Térmica	139,324676	103,242098
Región Bogotá			
Muestra 13	Arc eléctrico/Térmica	235,265044	191,182617
Muestra 14	Arc eléctrico	117,257369	86,020789
Muestra 16	Descarga parcial	116,557365	86,018759
Muestra 17	Térmica	115,111196	86,020789
Patrón			
Muestra 18		131,828066	98,5467465

Fuente: M. Mago ,2012

Tabla 7.
Pérdidas versus tamaño del cristal (grano) de acuerdo al porcentaje de Carbono

Región	Pérdidas totales	Tamaño del Cristal	Error aproximado
Antioquia	1,59157447	5,05	1,421048634
Caldas	1,58115902	0	1,411749125
Pacífico	1,59407866	41,06	1,423284518
Bogotá	1,563432	54,36	1,395921429
Patrón	3,16938739	13,6	2,829810170

Fuente: M. Mago, 2012

de acuerdo al porcentaje de carbono y las pruebas magnéticas. De estos resultados los más aproximados a los establecidos por el fabricante Acesita (Brasil) hoy APERAM, corresponden a las muestras 1, 5 y 17 procedentes de las Regiones Antioquia, Caldas y Bogotá debido a fallas térmicas, y la muestra 11 procedente de la Región Pacífico debido a falla por descarga parcial. En contraparte a los resultados obtenidos, al aplicar pruebas magnéticas se aprecia un margen de error entre cinco (5) y ocho (8) porciento. Estos datos sirven de referencia a los expertos que reparan o fabrican núcleos para estos equipos.

En la Tabla 6 se hace el cálculo de las pérdidas totales con las mismas condiciones anteriores, sin embargo, se incluye el valor del peso del equipo para los resultados y se compara igualmente con la norma que establece el fabricante para las pérdidas bajo condición de carga. El margen de error obtenido en los cálculos para el porcentaje de carbono oscila entre veintisiete con cuarenta y dos por ciento (27,42%) y sesenta y uno con cuatro por ciento (61,4%), mientras que al

realizar pruebas magnéticas es de cinco con cuarenta y uno por ciento (5,41%) a doce con cuarenta y dos por ciento (12,42%). A pesar de estos errores de aproximación los resultados siguen siendo de suma importancia para las empresas responsables de los procesos de reparación, ya que pueden verificar si la chapa de acero al silicio está en condiciones de ser o no reutilizada.

En la Tabla 7 se muestran resultados obtenidos de forma específica para la falla debido al arco eléctrico, regiones de procedencia y pérdidas. Estos valores fueron comparados con el tamaño de cristal (grano) en escala nanométrica y el error de aproximación. Los valores obtenidos son representativos en cuanto al tamaño de grano para todas las muestras, ya que las mismas están fuera de norma lo que indica posibles cambios en la estructura física del material no importando la región de procedencia. Al utilizar este método en la determinación de las pérdidas totales deben considerarse estos errores en los cálculos prácticos.

5. Conclusiones

Se ha generado un nuevo conocimiento que utiliza un método para determinar las pérdidas totales del núcleo de los transformadores de distribución monofásicos de poste a través de las propiedades físicas de la chapa de acero al silicio.

Los métodos numéricos utilizados originan errores en la medición, sin embargo, por razones de inmediatez para las empresas que reparan o fabrican estos equipos este método resulta bastante útil, siempre y cuando, *"se considere dicho error"*.

Se evitan gastos innecesarios de reparaciones o mantenimientos de estos equipos si previamente, se determinan las pérdidas totales antes de reutilizar la chapa de acero al silicio, sobre todo si al aplicar este método, se comprueba que la misma no posee condiciones apropiadas establecidas en la norma, con lo cual, se aumenta la confiabilidad en las redes de distribución de energía eléctrica.

6. Recomendaciones

Establecer normas técnicas para que este método pueda ser aplicado por las empresas encargadas de reparar o fabricar transformadores de distribución, de acuerdo a lo establecido por los entes gubernamentales y regulatorios correspondientes.

Bibliografia

[1] E. STAFF DEL M.I.T. Circuitos Magnéticos y Transformadores. Editorial Reverte. Argentina. 1981. Disponible en: https://es.scribd.com/doc/99522748/Circuitos-Magneticos-y-Transformadores-MIT-Staff-Edit-Reverte-En-Espanol.

[2] Oda, Y., Toda, H., Shiga, N., Kasai, S. and Hiratani, T., Effect of Si content on iron loss of electrical steel sheet under compressive stress. IEEE Transactions on Fundamentals and Materials. 134(3), pp. 148-153. 2014. DOI: 10.1109/TMAG.2013.2290321.

[3] Norma RA7-60. Valores de pérdidas equipos bajo condición de carga. [En línea]. Disponible en: http://www.epm.com.co/site/Portals/0/Users/001/01/1/RA7-060.pdf. Consultado: 20-06-12 hora: 10:30 am.

[4] Miyamoto, M., Matsuo, T. and Nakamura, T., Measurement of vector hysteretic property of silicon steel sheets at liquid nitrogen temperature. Przeglad Elektrotechniczny. [Online]. 87(9B), pp. 111-114, 2011. Available at: https://www.researchgate.net/publication/265531441_Measurement _of_Vector_Hysteretic_Property_of_Silicon_Steel_Sheets_at_Liqui d_Nitrogen_Temperature

[5] Zhao, J.-W., Zhang, W.-K. and Miao, X., Effect of percentage reduction in two-stage cold rolling on microstructure, texture and magnetic properties of high grade non-oriented silicon steel sheets. Kang Tieh/Iron and Steel. [Online]. 47(11), pp. 69-72, 2012. Available at: https://www.researchgate.net/publication/287026490_Effect_of_per centage_reduction_in_two-stage_cold_rolling_on_microstructure_texture_and_magnetic_prope rties_of_high_grade_non-oriented_silicon_steel_sheets

[6] Pampa, G., Richard, R., Chromik, A.M., Knight, S.G. and Wakade, C., Effect of metallurgical factors on the bulk magnetic properties of non-oriented electrical steels. Journal of Magnetism and Magnetic Materials, 356, pp. 42-51, 2014. DOI: 10.1016/j.jmmm.2013.12.052.

[7] Ali, S., Mohammad, R.T., Hossein, E. and Jerzy, A.S., Effect of deformation route and intermediate annealing on magnetic anisotropy and magnetic properties of a 1 wt% Si non-oriented electrical steel. Journal of Magnetism and Magnetic Materials, 385, pp. 331-338, 2015. DOI: 10.1016/j.jmmm.2015.03.026.

[8] Pluta, W.A., Loss components in electrical steel with Goss texture. International Symposium on Electrodynamic and Mechatronic Systems, SELM 2013 – Proceedings, Article number 6562993, pp. 87-88. Zawiercie; Poland. 2013. DOI: 10.1109/SELM.2013.6562993.

[9] Magdaleno-Adamea, S., Olivares-Galvanb, J.C., Kulkarnic, S.V. and Escarela-Perezb, R., Analysis of slots in horizontal plates of T-beams in shell-form power transformers. Electric Power Systems Research, 101, pp. 88-95, 2013. DOI: 10.1016/j.epsr.2013.03.007.

[10] Sha, Y.H., Sun, C., Zhang, F., Patel, D., Chen, X., Kalidindi, S.R., Zuo, L., Strong cube recrystallization texture in silicon steel by twin-roll casting process. Acta Materialia, 76, pp. 106-117, 2014. DOI: 10.1016/j.actamat.2014.05.020.

[11] Paltanea, V.M., Paltanea, G., Gavrila, H. and Peter, I., The influence of punching and laser cutting technologies on the magnetic properties of non-oriented silicon iron steels. International Symposium on Fundamentals of Electrical Engineering, ISFEE. University Politehnica of Bucharest, Bucharest; Romania. 2014. DOI: 10.1109/ISFEE.2014.7050611.

[12] Pluta, W.A., Loss components in electrical steel with Goss texture. International Symposium on Electrodynamic and Mechatronic Systems, SELM 2013 – Proceedings, [Online]. Article number 6562993, pp. 87-88. Zawiercie; Poland. 2013. Available at: http://toc.proceedings.com/18784webtoc.pdf

[13] Appino, C., De La Barrière, O., Beatrice, C., Fiorillo, F. and Ragusa, C., Rotational magnetic losses in nonoriented Fe-Si and Fe-Co laminations up to the kilohertz range. IEEE Transactions on Magnetics. 50(11), Article number 6971426. Istituto Nazionale di Ricerca Metrologica, Department of Electromagnetics, Torino, Italy. 2014. DOI: 10.1109/TMAG.2014.2325968.

[14] Pluta, W., Directional properties of loss components in electrical steel sheets. International Journal of Applied Electromagnetics and Mechanics, 44(3-4), pp. 379-385, 2014. DOI: 10.3233/JAE-141800.

[15] Pluta, W.A., Influence of anisotropy on specific loss components in grain oriented electrical steel. Solid State Phenomena. 214, pp. 138-142. International Symposium on Electrodynamic and Mechatronic Systems, 2014. DOI: 10.4028/www.scientific.net/SSP.214.138.

[16] Pluta, W.A., Some properties of factors of specific total loss components in electrical steel. IEEE Transactions on Magnetics Electrical, 46(2), pp. 322-325, 2010. DOI: 10.1109/TMAG.2009.2033559.

[17] Füzer, J., Birčáková, Z., Zeleňáková, A., Hrubovčák, P., Kollár, P., Predmerský, M. and Huňady, J., Investigation of total losses of non-oriented electrical steels. IEEE Transactions on Magnetics Electrical, [Online]. 118(5), pp. 1018-1019, 2010. Available at: http://przyrbwn.icm.edu.pl/APP/PDF/118/a118z5p124.pdf.

[18] Hatakeyama, T. and Onda, K., Core, loss estimation of various materials magnetized with the symmetrical/asymmetrical rectangular voltage. IEEE Transactions on Power Electronics, 29(12), pp. 6628-6635, 2014. DOI: 10.1109/TPEL.2014.2306755.

[19] Pavlos, S.G., Environmental cost of distribution transformer losses. Applied Energy, 88, pp. 3146-3155, 2011. DOI: 10.1016/j.apenergy.2010.12.021.

[20] Trianni, A., Cagno, E. and De Donatis, A., A framework to characterize energy efficiency measures. Applied Energy, 118, pp. 207-220, 2014. DOI: 10.1016/j.apenergy.2013.12.042.

[21] Lajtin, Y., Metalografía y tratamiento térmico de los metales. 3ra Ed. Editorial MIR. Moscú. URSS. [En línea]. 1983. Disponible en: http://www.casadellibro.com/libro-metalografia-y-tratamiento-termico-de-los-metales-3-ed/2910009902282/968544

[22] Mathews, J. y Fink, K., Métodos numéricos con Matlab. Editorial Prentice Hall. 3ra Ed. Madrid, España. 2005.

[23] Mago, M., Sistema estocástico unificado para la predicción confiable de fallas por condiciones ambientales en transformadores de distribución. Tesis Doctoral, Área de Estudios de Posgrado, Universidad de Carabobo, Valencia, Venezuela, [En línea]. 2014. Disponible en: www.riuc.bc.uc.edu.ve/bitstream/123456789/1930/1/mmago.pdf.

[24] Forero, A., Laboratorio de Metales. Facultad de Ingeniería. Universidad Nacional de Colombia. Bogotá, Colombia. 1999.

[25] Reed-Hill, R., Principios de metalurgia física. Editorial C.E.C.S.A. 6ta Ed., México, 1976.

[26] Shackelford, J., Ciencia de materiales para ingenieros. Ed. Prentice Hall. 4ta Ed., México. 1998.

[27] Smith, W., Fundamentos de la ciencia e ingeniería de los materiales. Ed. Mc Graw Hill. 4ta Ed. España.2006.

M.G. Mago-Ramos, is an associate professor of the Postgraduate in the Universidad ECCI, Bogotá, Colombia. She received the MSc. degree in Industrial Engineering 1994 and Electrical Engineering 2011.He received the Dr. degree in Engineering in 2014 from the Universidad of Carabobo in Venezuela. His research interests include: simulation, modeling and forecasting in energy markets; nonlinear time-series analysis and forecasting using statistical and computational intelligence techniques; and optimization using metaheuristics.
ORCID: 0000-0001-7250-111X

L. Vallès, is a titular professor of the University of Carabobo, Venezuela. Dr. of Science from the University of Massachusetts. National Chiao Tung University Postdoctoral, Hsinchu, Taiwan. Today is dedicated to projects and consulting in the industrial sector.
ORCID: 0000-0002-3456-4846

J.J. Olaya, is a associate professor at the Departamento de Ingeniería y Mecatrónica in the Universidad Nacional de Colombia, Bogotá, Colombia. He conducts research in the general area of development and applications of thin films deposited by plasma assisted techniques, corrosion and wear. He received his PhD in 2005 from the Universidad Nacional Autonoma de México, Mexico.
ORCID: 0000-0002-6942-4907,

M. Zequera, is a titular professor at the Departamento de Ingeniería in the Universidad Pontificia Javeriana de Bogotá, Bogotá, Colombia. He conducts research in the general area of Bioengineering. Is Dr. in Bioengineering from the UK University in United Kingdom.
ORCID: 0000-0001-7315-5909

J.E. Vera-Vera, is an associate professor at the Departamento de Ingeniería y Mecatrónica in the Universidad ECCI, Bogotá, Colombia. He conducts research in the general area of power electronics and alternative energy. He received his Msc in 2013 from the Pontificia Universidad Javeriana, Bogotá, Colombia.
ORCID: 0000-0002-0237-1257
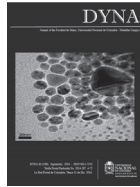
Methodology for hydrodynamic model selection. Case study: spatial variability of the thermal structure in the Riogrande II tropical reservoir, Colombia

Gabriel Betancur-Pérez [a], Francisco Mauricio Toro-Botero [a] & Andrés Gómez-Giraldo [a]

[a] Facultad de Minas, Universidad Nacional de Colombia – Sede Medellín, Medellín, Colombia. gabetancp@unal.edu.co, fmtoro@unal.edu.co, eagomezgi@unal.edu.co

Abstract
This article presents a methodology for a guided selection of a hydrodynamic model to study physical processes in a reservoir. It is based on both qualitative and quantitative criteria that are applied in order to study the spatial variability of the thermal structure in the Riogrande II tropical reservoir. The method consists of three main stages: i) definition and search, ii) pre-selection, and iii) selection. As a result, the DELFT3D and MIKE3 models were implemented, and the simulated temperature profiles were compared to field data that was collected between March 23rd and 27th, 2010. ELCOM was defined as a reference model, and the time series of the thermocline depth, the mixing layer depth, and its average temperature were compared with the selected models. The final selection was supported by using three different statistical parameters: the index of agreement (d_1), the MAE, and the RMSE. The quantitative analysis showed that the DELFT3D model behaves better than MIKE 3.

Keywords: methodology, model selection, three-dimensional hydrodynamic model, tropical reservoir, thermal structure, ELCOM, DELFT3D, MIKE3.

Metodología para la selección de modelos hidrodinámicos – Caso de aplicación: variabilidad espacial de la estructura térmica en el embalse tropical Riogrande II, Colombia

Resumen
Este artículo propone una metodología para la selección guiada de un modelo hidrodinámico para el estudio de los procesos físicos en un embalse, fundamentada en criterios cualitativos y cuantitativos, la cual fue aplicada para estudiar la variabilidad espacial de la estructura térmica en el embalse tropical de Riogrande II. La metodología considera tres etapas: i) definición y búsqueda, ii) preselección y iii) selección. Como resultado los modelos DELFT3D y MIKE3 fueron implementados y se compararon los perfiles de temperatura simulados con aquellos medidos entre el 23 y el 27 de marzo de 2010. ELCOM fue definido como modelo de referencia para comparar las series de tiempo de la profundidad media de la termoclina, la profundidad media de la capa de mezcla y su temperatura media. Para apoyar la selección final se utilizó tres estadísticos: i) el índice de concordancia (d_1), ii) el MAE y iii) RMSE, cuyos resultados indicaron que el modelo DELFT3D es para este caso la mejor alternativa.

Palabras clave: metodología de selección, modelos hidrodinámicos tridimensionales, embalse tropical, estructura térmica, ELCOM, DELFT3D, MIKE 3.

1. Introduction

The physical, climatic and hydrological conditions of rivers, lakes, reservoirs, estuaries and coastal zones, cover a complex dynamic of interaction and feedback between parameters of the representative processes. Given the limited information available on atmospheric variables, such as solar radiation, humidity, air temperature, etc., and on the variables

associated to the water body itself, such as free surface, flow rate, temperature, salinity, turbidity, conductivity, etc., the use of complex numerical models are required to understand the behavior of these variables and their interactions.

Moreover, the application of hydrodynamic models with one, two, or three dimensions to study the mass transport [1], temperature [2] and salinity variations -in addition to the velocity and circulation patterns [3] in different water bodies and geographic locations- has increased over recent years. However, in many cases the choice of the model to be used for a specific application is not clear, given the number of options available in the market. This makes the selection of the proper model (tool) a topic that needs to be carefully analyzed [4]. Selection of the proper model for a given application is neither simple nor straightforward since there are several factors involved in the process, such as the model accessibility (public, restricted or commercial), the budget available to resolve the problem, the personal involved, the available information to calibrate and validate the model (field measurements of key variables such as temperature, salinity, speed, current, etc.), and the time required to solve the problem.

As such, several authors have presented their views on the selection of a hydrodynamic model, considering topics such as the cost-complexity and the spatial and temporal dimension that are closely related to the phenomenon [5]. However, there is no systematic method that helps the user to make a good selection of a hydrodynamic model.

This article presents a methodology to perform a guided selection of a hydrodynamic model, based on qualitative criteria (multi-objective analysis) and quantitative tools (statistical analysis of model results). To illustrate the methodology, the spatial variability of the tropical reservoir Riogrande II's thermal structure, during a dry season, was defined as the case study.

2. Suggested methodology for model selection
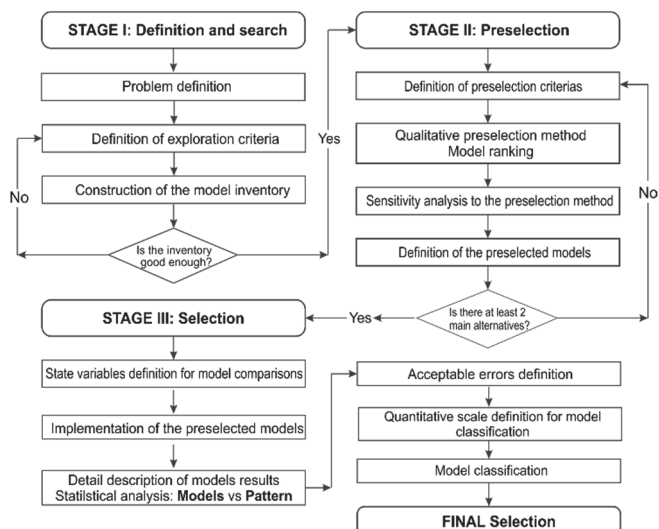
Figure 1. Flowchart of the suggested methodology for model selection
Source: Prepared by the authors

The suggested methodology considers three stages, the first two being of a descriptive and qualitative nature, while the third is quantitative. Fig. 1 shows a flowchart, which is discussed in the following sections.

2.1. Stage I: Problem definition and search for models

During the first stage, the physical problem to be studied must be clearly defined (*process definition*) in a mathematical sense by writing the equations (usually conservation laws) that describe the phenomenon under consideration. Then, some *exploration criteria* are defined and a *model inventory* should be made that must account for the state of the art of available models. Based on the knowledge gained during this search process, the model user defines if the inventory is robust or not. If it is, the model selection process continues to the next stage, if not, the decision maker will reappraise the exploration criteria and start a new search.

2.2. Stage II: Model preselection

In the second stage, it is necessary to define *preselection criteria* that are more restrictive than those used in the previous stage, and that will be later used to rank the pre-selected models through the implementation of a multi-purpose analysis [6]. Next, it is necessary to perform a *sensitivity analysis* for the selected multipurpose analysis method in order to assess whether the obtained solution (rank) is robust. If so, a threshold for the number of numerical models to be used in the next stage is defined and the process continues. Otherwise, the preselection criteria should be reevaluated in order to achieve the previously defined threshold.

2.3. Stage III: Model selection

This is the quantitative stage of the suggested methodology and it requires the numerical models that passed the previous stages to be implemented. In this final stage, it is necessary to clearly define the process control variables and select those that are going to be used to compare the results. To perform this stage, a threshold for the model errors (with respect to a pattern that can be another model or a series of field measurements) is defined, and the model's performance is evaluated by using a statistical tool such as the Mean Square Error (MSE), the Root Square Mean Error (RMSE), or the index of agreement (d_1). Each model's performance is quantified according to a pre-defined scale to rank it and to take the decision as to whether a model is selected.

3. Methods and application

3.1. Case of study: Riogrande II tropical reservoir

The Riogrande II reservoir is a multi-purpose (water and energy supply) reservoir located north of Medellín, Colombia, and is located between the coordinates 75°32'W and 75°26'W and 6°33'N and 6°28'N (Fig. 2). The main
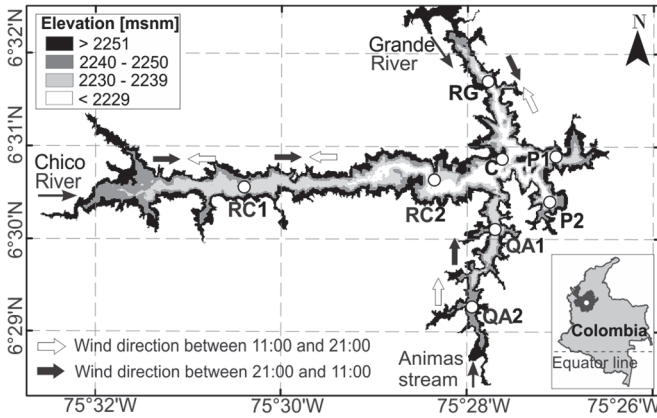
Figure 2. Location, bathymetry, and wind circulation pattern of Riogrande II reservoir. (○) Locations of the measured water body temperature profiles
Source: Prepared by the authors.

Figure 3. Typical temperature profiles measured in the Riogrande II reservoir between March 23rd and 27th, 2010
Source: Prepared by the authors.

tributaries are the Chico River ($Q_{mean} = 5m^3/s$) to the east, the Grande River ($Q_{mean} = 15m^3/s$) to the north and the Las Animas stream that has a negligible discharge to the south. The maximum water depth is 42.5m near the dam (near point P1 in Fig. 2), and the water storage volume is 240Mm^3 at 2270 m.a.s.l.[7]. The intake to the hydroelectric plant and to the water supply plant is located at point QA2 in Fig. 2.

This three-branch canyon valley reservoir, which according to [8] and [9], has a spatial and temporal variable wind field with a daily cycle (shown in Fig. 2. This includes a space and time variable wind field in the numerical models, which is important because it is an internal wave generator that interacts with the river plumes and generates fast vertical mixing [10].

Numerical simulations made by [9], as well as field measurements during different climatological seasons (rainy and dry seasons), suggest that the Grande River water moves in the reservoir like an intrusive plume from its entrance and then splits in the confluence (C in Fig. 2) into two branches. One travels upstream along the Chico River branch, which affects its plume dynamics, and the second travels south to the intake (QA2 in Fig. 2). The dynamics of these two plumes are mainly affected by the climatological seasons and the river´s water temperature [9].

3.2. Field Measurements for model calibration and validation

In order to facilitate the analysis of the model's results, the reservoir was divided into five zones with similar physical behavior that was based on the following field measurements:

- Zone 1: ChicoRiver Branch (RC1 y RC2)
- Zone 2: Confluence (C)
- Zone 3: Grande River Branch (RG)
- Zone 4: Las Animas stream (QA1 y QA2)
- Zone 5: Dam (P1 y P2)

In order to compare the different models' behavior that passed the second stage of the methodology the models' results were compared to those obtained from a field campaign undertaken between March 23rd and 27th, 2010. These data are described in the following paragraphs.
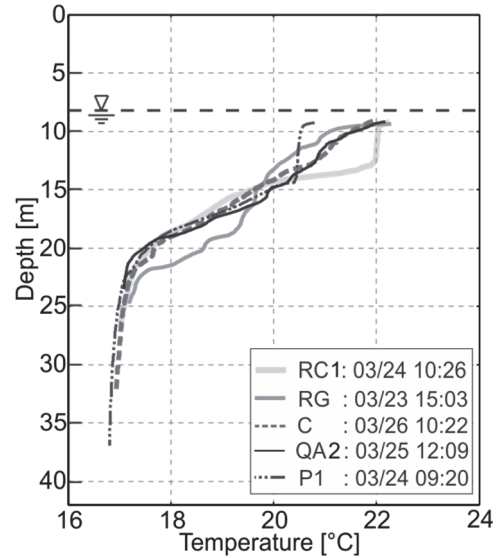
3.2.1. Field Measurements

Temperature profiles were measured between March 23rd and 27th, 2010 in the control stations shown in Fig. 2. Fig. 3 shows their typical behavior. The records of the reservoir level were used to assess the mass balance of the system that was reproduced by the implemented models.

Previous reservoir simulations [9] indicate that the temperature profiles in station RG are influenced by the Grande River. They consider that there is no well-defined mixing zone such as that shown in RC1 or P1 profiles (Fig. 3).

In the case of the QA2 profile, despite the selective withdrawal influence [8], it is possible that the reservoir shallowness in this area and the Grande River movement from its mouth to the intake introduce a mixing condition that reduces the temperature gradients. It shows a temperature profile similar to that in the RG station. In both cases, further research and field data are required to evaluate and analyze movement of the plumes and their influence on the reservoir's temperature conditions. Further analysis is also necessary to explain the hydraulic control produced by the Grande River on the Chico River branch. This analysis is outside the scope of this article.

3.2.2. The selection of a pattern for comparison

Those models that passed the two qualitative stages of the methodology will be implemented for this case study. In order to quantify their behavior, their simulated temperature profiles will be compared at selected points and zones of the reservoir using measurements and a reference model (in the methodology a comparison with a pattern, measurements, or another calibrated and validated model, is required). In our case study, results from a calibrated and validated model [9] were chosen to take advantage of the high space and time resolution obtained from the model. This was not available from the measurements. The reference model selected was

ELCOM [11], developed by the University of Western Australia, which was calibrated for the field campaign during March 23rd to 27th, 2010. ELCOM is a model that has been developed and progressively improved to capture the temperature gradients and the thermal structure in estuaries, lakes and reservoirs. ELCOM incorporates a special mixing model to account for the stratification processes [12], and, therefore, constitutes a reference for other models that seek to represent the thermal stratification in lakes and reservoirs.

ELCOM's results were used to extract time series of temperature at the stations indicated in Fig. 2. Based on these time series, the average depth of the mixing layer and the mean thermocline depth was estimated using the methodology proposed by [13]. Additionally, the average temperature of the mixed layer was calculated once the average depth of the mixed layer had been derived. These three variables will be used to test the models' behavior.

3.3. Multi-purpose analysis methods

The implementation of the suggested methodology requires, during the qualitative stages, the use of a multicriteria analysis method to quickly reduce the potentially suitable models' alternatives for this case study. In this project, the ELECTRE IV [14] method was used, because unlike other multicriteria techniques, such as the PROMETHE method [15] or AHP [16], ELECTRE IV does not use weights to perform the classification process. Instead, a table of metrics is used to define the values for each evaluation criteria, and every alternative forms an *impact matrix*. Two types of sorting are then performed: *ascending* and *descending*. Tie breaker rules are applied to finally rank the alternatives [6].

3.4. Statistical tools used for model behavior evaluation

Several techniques for quantitative evaluation of a model's behavior have been widely explored in atmospheric sciences [17]. Their adaptations to the coastal engineering, undertaken by [18] and [19], are the most important contributions to the numerical model behavior analysis. However, the introduction of the *index of agreement,* undertaken by [20], is a useful tool to compare and analyze the model results. The statistical parameters used in this paper for the model performance evaluation are presented below.

The most used statistical parameter in model behavior analysis is the Root Mean Square Error (RMSE), given by equation (1) as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2} \qquad (1)$$

where y_i are the reference values, x_i are the model predictions, and N is the number of data of the time series used in the comparison.

A robust measurement of the prediction's precision is the Mean Absolute Error (MAE) because it is not influenced by the time series' extreme values [21]:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - x_i| \qquad (2)$$

A third and fourth statistical parameter used in the study are the indexes of agreement d$_1$ and d$_2$, suggested by [20,21] for $\gamma = 1$ y 2, respectively:

$$d_\gamma = 1 - \frac{\sum_{i=1}^{N}(|y_i - x_i|)^\gamma}{\sum_{i=1}^{N}(|y_i - \bar{x}| + |x_i - \bar{x}|)^\gamma} \qquad (3)$$

As is the case for the RMSE, the d$_2$ index is highly influenced by the differences squared when they are large. For this reason, in this paper, only the d$_1$ index was used to evaluate the models results.

4. Suggested methodology step by step

4.1. Stage I

The problem under study is the space variation of the Riogrande II reservoir's thermal structure that is under the influence of the wind field during the dry climatological season. The time window to compare the models is between March 23rd and March 27th, 2010.

The search criteria to obtain the list of potential models to be used to study the stated problem are:

A Three Dimensional Hydrodynamic Model. The interaction that takes place between the different physical processes in a reservoir, such as the time-space variation of the water body temperature under the force of a variable wind field and the dynamics of the river plumes and their interaction makes it necessary to use a 3D hydrodynamic model to correctly approach the stated problem. This makes 1D and 2D models unsuitable. Additionally, the model should have implemented heat fluxes equations with the atmosphere and heat transport in the water body. The turbulent flow of the river plumes in their mixing layer requires an appropriate turbulence closure scheme. A long inventory list of numerical models analyzed in this stage is presented in Table 1.

Sediment transport module: This module was considered because, for future work, this will be a research field. However, no models with sediment transport were implemented in this study and only the hydrodynamics were considered.

4.2. Stage II: Pre-selection

For this stage, the selection criteria were grouped into three categories:

General: Including global characteristics. In this application the following were considered:
- Availability
- Graphical user interface
- Documentation
- Technical support

Flexibility: Configuration options available for the user to properly represent the specific conditions of the phenomenon

being considered as well as the representation of the computational domain. In this application, the following were considered:

- Domain horizontal discretization
- Domain Vertical discretization
- Turbulence closure schemes
- Non-hydrostatic module

Heat fluxes: Given the influence these fluxes have on the lake or reservoir's thermal structure, the following were considered:

- Formulae to estimate the latent and sensible heat fluxes
- Transport equation for temperature
- Vertical mixing model

Sediment Transport: This characteristic of the numerical model was considered as an option for future work because the sediment transport process requires a very well defined flow field. The following criteria were considered for the qualitative classification:

- Type of sediment transport process
- Type of sediments
- Bottom morphology evolution

4.2.1. ELECTRE IV method

Once the pre-selection criteria have been defined, the implementation of the ELECTRE IV method follows. To do this, two valuation scales (metrics) were used for the criteria analysis: one with integer numbers between 1 and 10 and a second one with odd numbers between 1 and 9. Using these two metrics, an *impact matrix* is built with all previously selected criteria. Then, following [14], the *reference relationships* were estimated by comparing the different alternatives, in groups of two, for each selection criteria. The values that were estimated for the impact matrix and following the preference rules are defined as follows:

Table1.
General Characteristics of models found under the search criteria

Model name[a]	ADC	D3D	ESD	EFDC	FVCOM	MIKE	MOHID	ROMS	SSIIM	TEL
Model Availability[b]	P	P	PR	P	PR	Co	PR	PR	P	P
Graphic Interface	± Co	+ R	-	± Co	± Co	+ Co	+ Co	±	±	±
Fully Documented[c]	±	+	±	±	±	+	+	±	±	+
Technical support and help[d]	-	+	±	±	+	+	+	+	±	+
Last Model update	2004	05/2013	02/2002	2007	08/2013	2013	2009	04/2013	09/2011	08/2013
Operating System and Programing Language [e]	W, F	W/L, F	L, F	W, F	W/L, F	W, F	W, F	L, FOR	W, C	W/L, F
Numerical Scheme[f]	FE	FD	FD	FD	FV	FD, FE	FV	FD	FV	FE
Coord. horizontal plane[g]	Ca	Ca, Cu	Ca	Ca, Cu	Ca, Ce	Ca, Ce	Ca, Cu	Ca, Cu	Ca	Ca, Ce
Coord. Vertical plane[h]	σ	Z, σ	σ	S, σ	S	H, σ, S	H, σ, S	S	Z	Z, H
Type of computational grid[i]	E, U	E, U	E	E	U	E, U	U	E	U	E, U
Turbulence closure models[j]	-	C, A, κ-l, κ-ε	MY	MY	κ-ε, MY	C, A, κ-ε	A, κ-l, κ-ε	C, A, κ-l, κ-ε, κ-ω, MY	κ-ε, κ-ω	C, κ-ε
Non-hydrostatic approach	-	+	-	-	+	+	-	-	-	+
Temp. Transport equation	-	+	+	+	+	+	+	+	+	+
Heat fluxes equations	-	+	+	+	+	+	+	+	±	-
Vertical mixing model [k]	-	-	-	-	-	-	-	-	-	-
Sediment transport module	+	+	+	+	+	+	+	+	+	+
Sediment types[l]	Co	Co, Nco	Co, Nc	Co, Nco, Ds	Co, Nco	Co, Nco	Co, Nco	Nco	Nco, Ds	Nc, Ds
Type of sediment load[m]	Cs	Cs, Cf	Cs	Cs, Cf	Cs, Cf	Cs, Cf	Cs	Ct	Cs, Cf	Cs, Cf
Morphologic evolution	-	+	-	+	+	+	-	+	-	+

Notes: +: included; ±: partially included; -: not included;

[a]ADC: Advanced Circulation Model (ADCIRC), D3D: Open Delft3D-FLOW, ESD: Estuarine, Coastal and Ocean Modeling System with Sediments (ECOMSED), TEL: Telemac Modeling System (TELEMAC).

[b] P: public domain; R: Restricted; Co: Commercial.

[c]Considers: User manual, Scientific manual, tutorials and examples.

[d]Considers: Direct technical support, internet discussion-help forum.

[e] W: Windows; L: Linux; Languages: Fortran: F; C.

[f]DF: Finite Difference; EF: Finite Elements; VF: Finite Volume.

[g]Ca: Cartesian; Cu: curve; Ce: spherical

[h]Z: Z coord.; σ: sigmacoord..; H: hybrid(Z-σ) coord..; S: General sigmacoord.

[i]E: structured; U: Non-structured.

[j]C: Constant Eddy Viscosity; A: Algebraic; κl: One equation; κ-ε:kappa-epsilon, κ-ω: kappa-omega; MY: Mellor-Yamada.

k Vertical mixing model developed by [12] and implemented in ELCOM.

[l]Co: cohesive; Nco: non-cohesive; Ds: sediment distribution [m]Cs: suspension load; Cf: bottom load; Ct: total load.

Source: Prepared by the authors

- *Over Relaxed classification*: alternative *i* **OverRelaxed Classifies** alternative *j* if the number of criteria for which alternative *i* is preferred over alternative *j* is at least twice the number of criteria alternative *j* is preferred over alternative *i*.
- *Weak Relaxed classification*: alternative *iWeakly Relaxed Classifies* alternative *j* if there is at least one criteria by which alternative *i* is strictly preferred over alternative *j* and there is no criteria by which alternative *j* is strictly preferred over alternative *i*. Also, there is the additional condition that the *Strong Over Classification* is not satisfied.
- *Strong Classification:* Alternative *i* is **Strongly Classified** over alternative *j* if there is at least one criteria by which alternative *i* is strictly preferred over alternative *j* and there is not a criteria by which alternative *j* is strictly preferred over alternative *i*. Also, if the number of criteria by which alternative *j* is weakly preferred over alternative *i* is smaller or equal to the number of criteria by which alternative *i* is preferred (strictly or weakly) over alternative *j*.

Finally, based on the classifications just defined for each alternative, two lists of alternatives, organized in ascending and descending order are prepared. The final ordered list of alternatives is obtained as the intersection of these two previous lists.

Table 2 shows the final ordered list obtained by following the previously described methodology and for each one of the two previously mentioned evaluation metrics. The numbers appearing in the columns Classification 1 and Classification 2 refer to the model number used in the inventory table (Table 1): alternative 1 refers to the ADCIRC model and alternative 10 refers to the TELEMAC 3D model.

The ELECTRE IV method shows that the DELFT 3D model is the dominant alternative because both classifications placed it in first position on the list. The MIKE 3 and the EFCD 3D models in Classification 1 shared the second position, but in Classification 2, the MIKE3 model only appears in the second position. For this reason, the MIKE3 model was finally classified in the second position. Based on these results, models DELFT 3D and MIKE 3 are chosen for the next stage in the methodology.

4.3. Stage III: Selection

4.3.1.   Variables used for comparison

Table 2.
List of alternatives for the ELECTRE IV method

Order	Classification 1	Classification 2
1	2	2
2	4 y 6	6
3	-	4
4	10	10
5	5	5
6	8	8
7	9	7
8	7	1 y 9
9	1	-
10	3	3

Source: Prepared by the authors

Based on the field measurements and on the time series obtained from the reference model (ELCOM), the variables used to evaluate the behavior of the preselected models are:
- Temperature profiles
- Water level variations in the reservoir
- Mean depth of the mixing layer
- Mean Temperature of the mixing layer
- Mean depth of the thermocline

It is very important that the list of models selected for this stage captures some of the most relevant components of the physical process under study. In this case, the thermal structure of the water body, under the influence of two forcing factors such as the wind field and the river plumes, is well represented by the above-mentioned variables. The depth of the mixing layer and that of the thermocline are important variables because they are a key part of the possible internal waves' appearance in the reservoir [23] as well as being important for its ecological evolution. The mean temperature of the mixing layer is important because it is in this zone that the energy exchange between the water body and the atmosphere occurs [12].

4.3.2.   Implementation of the pre-selected models.

DELFT3D is a 3D hydrodynamic model whose governing equations are solved numerically by the Finite Difference technique. The model has been used to simulate flow circulation patterns, heat transport, water quality and bottom morphology under several forcing factors such as wind field and tidal waves with a wide field of applications in coastal zones and estuaries [24]. Since 2010, DELFT3D has been an open code for the hydrodynamic (FLOW), Morphodynamic (MOR), and waves (WAVE) modules. In the year 2013, the water quality module (DELWAQ) was released as an open code. In our application, graphic interphase version 4.00.02 and code 5.00.10.2136 of January 22nd, 2013 were used for the model implementation of the Riogrande II reservoir.

MIKE3 is the 3D version of the MIKE 21 model developed by the Danish Hydraulic Institute (DHI). It has been widely used in coastal and estuarine applications. The model includes several modules coupled to a graphic interphase. To simulate the thermal structure of the Riogrande II reservoir, the 2012 flexible grid version was used. This version uses the non-structured grid with a finite element discretization in the horizontal plane and a hybrid σ – Z coordinate discretization in the vertical plane [25].

Both models were implemented following the reference model (ELCOM) as closely as possible. The reference model was included in this analysis (selection) because it was fully calibrated and validated with field measurements and will be used to generate data not directly available from the measurements. Table 3 shows the general configuration for these three models, in which some differences in the time interval are clear; these are differences due to space discretization. Fig. 4 shows the river discharges and outflows from the reservoir and the atmospheric forcing factors used for all the three models.

4.3.3.   Model results and behavior analysis of the models
Error limits

Temperature profiles shown in Fig. 3, typical for tropical zones, differ from those for extra tropical regions. In extra
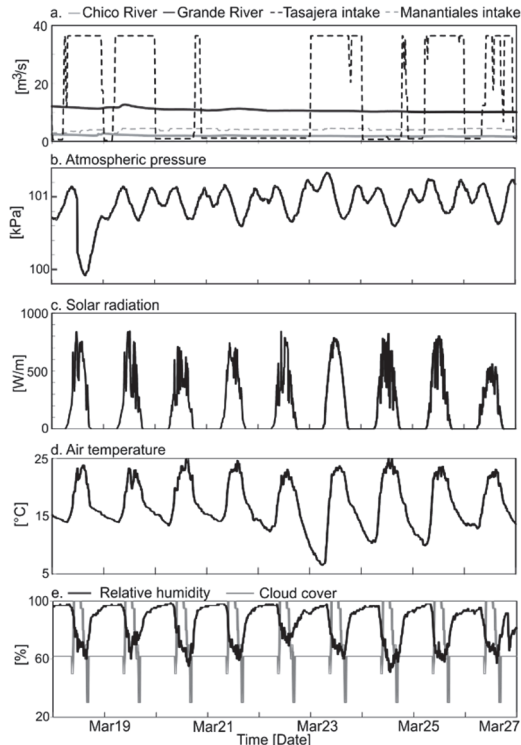
Figure 4. Time series of (a) rivers and outflow discharges (b) atmospheric pressure, (c) Solar radiation, (d) air temperature (and e) cloud cover and air relative humidity
Source: Prepared by the authors

Table 3.
General configuration for the three models (Prepared by the authors).

Variable	ELCOM	Delft3D	MIKE 3
Total simulation time after the "model warming up period"	9.5 days	9.5 days	9.5 days
Computation time interval, Δt	30 s.	15 s.	Hydro = 0.25 s Trans. = 0.5 s
CPU time	3 hours	12 hours	26 hours (Hydro) 36 hours (Tracer)
Horizontal grid	229 rows 136 columns	229 rows 136 columns	2896 elements 2177 nodes
Vertical grid	40 variable height layers	40 variable height layers	1 σ layer, and 69 uniform Z layers
Initial conditions	Water levels and Temp. profiles	Water levels and Temp. field	Water levels and Temp. field
Climatology	Time series	Time series	Time series
Wind field	Time series by zones	Wind field	Wind field
Inflows and outflows	Time series: discharges and temperature	Time series: discharges and temperature	Time series: discharges and temperature

Source: Prepared by the authors

tropical reservoirs typical temperature differences between the surface and bottom of the reservoir may

reach values of 20°C [26] with an error of 10% (2°C), which may be reasonable for simulated results. In tropical reservoirs and lakes, the vertical temperature gradients (Figure 3) are much milder than those in extra tropical zones and there are temperature differences between the surface and the bottom in the order of 5°C or 6°C all year long, with a daily cycle. This makes the value of 2°C a very high tolerance error for these cases. By keeping the 10% value for the tolerance error, an equivalent of ± 0.5°C was used in this work as an acceptable limit for the RMSE and MAE errors.

Reservoir free Surface level

Measured and simulated free surface water levels are shown in Fig. 5. It is clear from the figure that all three models closely follow the measured tendency. The differences found between the measured simulated values, a maximum of 20 cm for March 23rd, are due to deficiencies in the measured inflows and outflows time series.

The corresponding MAE and RMSE values, that are close to zero (Table 4), indicate the good agreement with measurements. Although, the very similar values obtained for all three models do not allow us to draw any conclusion about which models have a better behavior with respect to this variable. The same behavior can be noticed for the index of agreement, d_1; although, MIKE3 is the one that behaves the poorest. This variable is not good enough for model classification because, as will be shown later, major differences in the parameters are found for the models: differences that are not important for this variable (all models sufficiently capture the free surface water level).

Temperature Profiles

A comparison between simulated and measured temperature profiles is shown in Fig. 6, in which the profiles correspond to those showing the maximum difference between the simulated and computed ones. In zone 1, ELCOM and DELFT3D closely follow the measured profile for which corresponding values for d_1 of 0.93 and 0.91 were found with MAE corresponding values of 0.25°C and

Table 4.
Free water level statistics computed for the three models

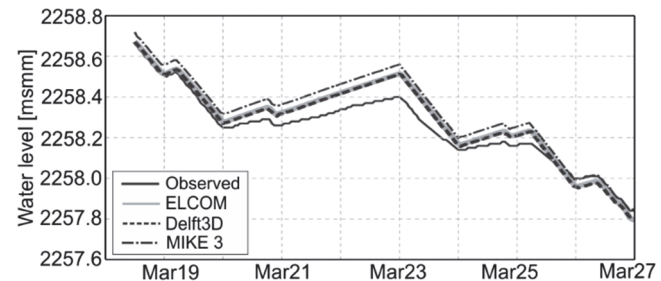| Statistic parameter/Model	ELCOM	Delft3D	MIKE 3
RMSE [m]	0.06	0.06	0.09
MAE [m]	0.05	0.05	0.08
d_1	0.82	0.84	0.74

Source: Prepared by the authors

Figure 5. Simulated and measured free surface levels for the Riogrande II reservoir
Source: Prepared by the authors

0.30°C. MIKE3 is less precise: in the hypolimnium, the model closely follows the measurements, but as the free water surface is approached the model departs from measurements and even shows an unstable profile. This behavior leads to a d_1 value of 0.83 and an MAE value of 0.48°C (see Fig. 7 and Fig. 8 for more details).

In zone 2, measurements do not show the near surface mixing layer (Fi. 6), which generates a constant temperature vertical gradient that is not sufficiently captured by any model. It instead shows a mixing layer that is about 5m deep. In this zone, ELCOM and DELFT3D produce a d_1 value higher than 0.9 and RMSE values that are smaller than 0.5°C (Fig. 9). MIKE3 shows a value of 0.86 for d_1 and a value of 0.61°C for RMSE.

In zone 3, near the entrance of Grande River, the absence of the surface mixing layer is clear (Fig. 6) and there is a constant temperature vertical gradient. The vertical mixing produced at the river's entrance is responsible for these types of profiles in addition to the wind action in the shallow parts of the reservoir. Values of the index of agreement are presented in Fig. 7, and it is seen that MIKE3 has the smallest values for several profiles in the zone. Analyzing the MAE and RMSE parameters (Figs. 8 and 9), once again MIKE3 has poorer behavior with respect to ELCOM and DELFT3D. Similar behavior is noticed in zone 4 where the effect of the hydro-plant intake is relevant.

Finally, in zone 5, ELCOM and DELFT3D behave in a similar way to in zone 1, with values close to unity for the index of agreement (Fig. 7) and values for the MAE and RMSE that are close to zero (Figs. 8 and 9). In this zone, MIKE3 also has a poorer behavior than the other two models.

These results suggest that the MAE parameter shows better behavioral models. For example, in zone 3 the index of agreement indicates that all three models have poor behavior without being able to differentiate the best one (Fig. 8). The RMSE parameter, on the other hand, demonstrates the better behavior of the ELCOM models, but it is unable to show the differences between DELFT3D and MIKE3. The MAE parameter, however, did show the difference between the three models.
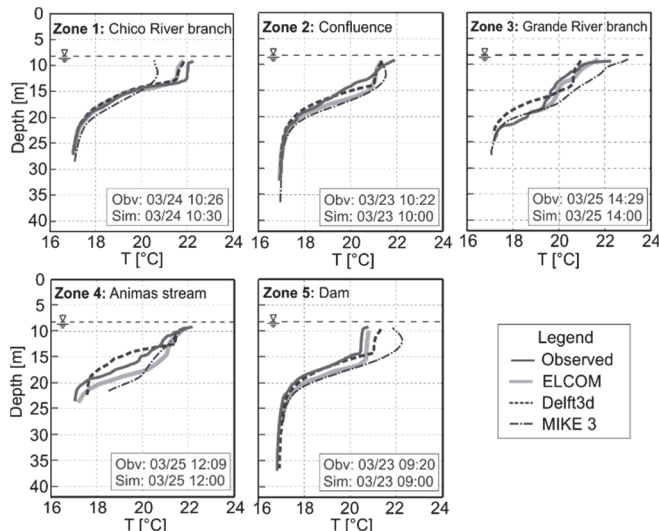
Figure 6. Observed and simulated temperature profiles in all 5 zones of the Riogrande II reservoir
Source: Prepared by the authors
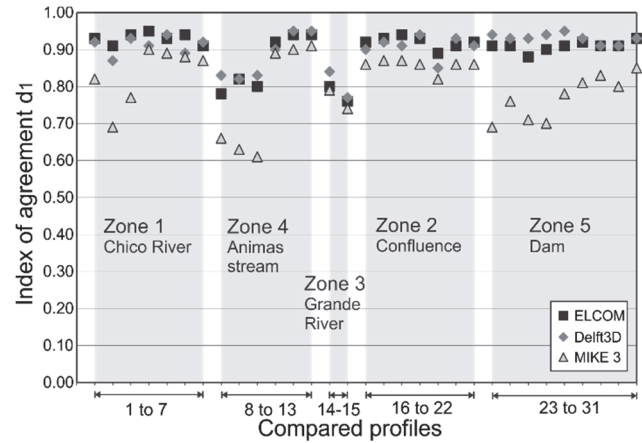
Figure 7. Index of agreement (d_1) computed for several temperature profiles in the five zones of the Riogrande II reservoir
Source: Prepared by the authors
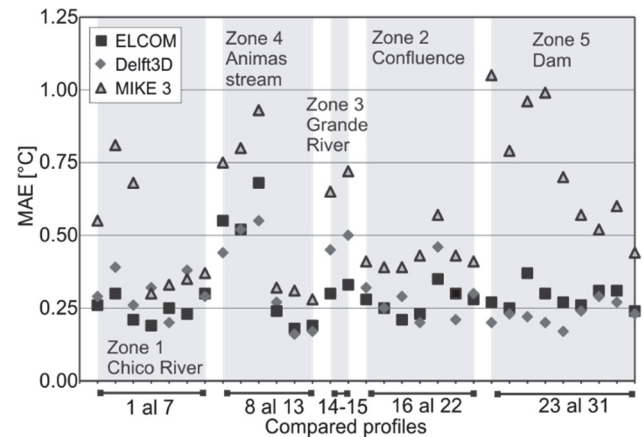
Figure 8. MAE parameters computed for several temperature profiles in the five zones of the Riogrande II reservoir
Source: Prepared by the authors

Time series

Three variables of physical interest for the thermal structure of the reservoir are selected for analysis: the thickness of the mixing layer, the mean temperature of this layer, and the depth of the thermocline [13]. The time series of these variables are shown in Fig. 9 for monitoring stations RC1 and QA2.

The thickness of the mixing layer at RC1 (Fig. 10a) is captured by DELFT3D and MIKE3 for the four-day computed period. DELFT3D closely follows the reference model (ELCOM) in both the shape of the series and in the maximum and minimum values. MIKE3, on the other hand, captured the peaks but the shape of the time series looks different from that of the reference model. It showed a constant peak for a period of six hours, which is not in the reference time series. At QA2, both models show more differences with the reference model, although DELFT3D seems to be closer to it. The index of agreement computed for this variable and for several monitoring points in the reservoir is shown in Fig. 8, in which it is clear that DELFT3D better follows the reference model (higher d_1 values than MIKE3).

Time series for the mean temperature of the mixing layer

for the monitoring points RC1 and QA2 are shown in Fig. 9b. Mean temperature computed from both models monitoring station RC1 follow the shape of the reference model, but the MIKE3 results are about 0.75°C cooler for the reported period than the reference time series. The ones for DELFT3D are about 0.5°C warmer than the reference time series. For monitoring station QA2, DELFT3D continues with the same behavior, but the temperature difference is now between 1.2°C and 0.5°C warmer. MIKE3 maintains the shape of the daily cycle but is now sometimes warmer and sometimes cooler than the reference. The statistical parameters (d_1, MAE and RMSE) confirm that DELFT3D behaves more similarly to the reference model than MIKE3 does. The behavior of MIKE3 can be explained, with respect to this variable, due to the calibration undertaken to the vertical dispersion coefficient (0.0025). This, for some higher values, produced stable temperature profiles near the surface, resulting in a thicker mixing layer and a deeper thermocline position. For smaller values it generated unstable temperature profiles near the surface with cooler temperature.

This behavior is reflected in the statistical parameters in Fig. 11. The behavior of DELFT3D, generating warmer mean temperatures for the mixing layer, may be related to the parametrization for the heat fluxes included in the model as well as to the wind shear stress on the surface, although a detailed study is needed to confirm this.

Finally, the analysis undertaken to the thermocline depth shows large differences between the reference time series and the computed ones in both monitoring stations (Fig. 10c). It confirms the values of the statistical parameters computed for

Figure 10. Statistical parameters for the time series analysis. (♦) Thickness of the mixing layer, (▲) Mean temperature of the mixing layer, and (■) Depth of the thermocline
Source: Prepared by the authors

several stations (Fig. 11). It is important to notice that the methodology the reference model, ELCOM, has implemented a special mixing model differing from those implemented in DELFT3D and MIKE3.

4.3.4. Models classification and final selection

ELCOM was taken as a reference model. The results were used to generate temperature time series that are not available from direct field measurements. This model was not considered for classification purposes due to its proven capacity to simulate thermal stratification and heat fluxes dynamics in lakes and estuaries [23]. It is also widely used to simulate similar cases around the world [27].

Based on the ranges defined by Table 5, the models' performance were estimated for several points in the reservoir zones defined in section 3. Fig. 11 shows that DELFT3D performs better than MIKE3, and its results are similar to ELCOM.

Table 5.
Suggested Scale values for Final Classification

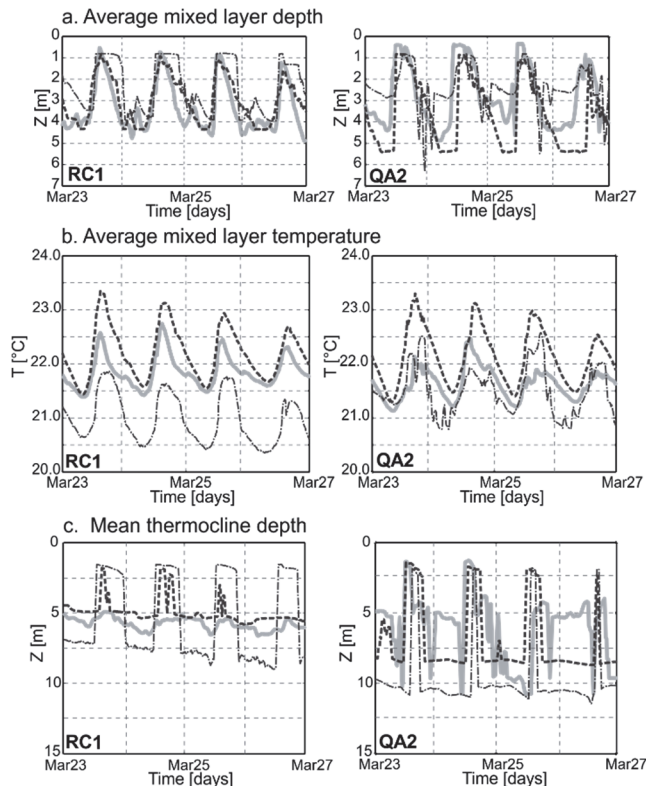| Range	Model performance
$d_1 \geq 0.90$	High
$0.85 \geq d_1 < 0.90$	Mean-High
$0.80 \geq d_1 < 0.85$	Mean
$0.75 \geq d_1 < 0.80$	Mean-Low
$d_1 < 0.75$	Low

Source: Prepared by the author

Figure 9. Models time series (——) ELCOM, (- - -) Delft3D and (— • —) MIKE 3
Sourde: Prepared by the authors

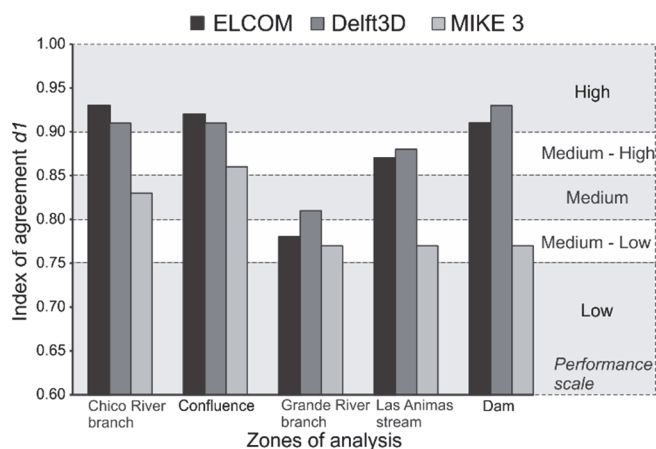Figure 11. Final model classification based on d_1
Source: Prepared by the authors

5. Discusion

On the general methodology

The qualitative stages of the methodology are a fundamental part of the selection process because, for practical reasons, it is not possible to implement the long list of available models in the literature. During these stages, it is possible to make a quick review of the alternatives and create a long inventory of possible models that can be used for future work. However, the subjective character of the methods used in these stages may lead to a skewed decision by choosing a model that may not be the best technical option.

This subjective character, which is associated with the definition of the selection criteria and with the implementation of the multi-purpose decision analysis, such as the ELECTRE IV method, may be reduced by looking for methodologies that reduce the subjectivity of the definition of the selection criteria and their evaluation scales. One example is the AHP method [16], which is a technique that involves more than one decision maker and includes a panel of experts that participates in the final decision.

Regarding the quantitative stage of the methodology, the index of agreement for statistical parameters has the advantage over the other two. This is because it is non-dimensional and can be used in a general sense as well as being compared to other models in similar cases. However, care must be taken when using it because it represents a result in a statistical sense. This depends on the set of data used to make comparisons and it is the modeler´s decision to use it instead of the physical results. In this case, the d_1 provides a coherent interpretation of the results and that is why it was selected as the leading parameter for decision-making.

The Statistical parameters

The comparison results showed that the MAE parameter is much more robust, in its physical interpretation, than RMSE, as suggested by [21] and later reported by [28]. In the same sense, the index of agreement is the best parameter to use, given its direct relation to MAE. However, despite the acceptance of this parameter as the leading parameter for selecting and classifying the behavior of hydrodynamic

models, there are other alternatives reported in the literature like the "skill scores", reported by [29], who proposed a different scale for the errors. Also [18] suggested the use of the ARMAE (the square root of the modified MAE by including the errors in the measurements) parameter to compare and classify hydrodynamic models.

6. Conclusions

The presented methodology approximates a guided selection of a model (hydrodynamic model in the case study presented here). It is based on qualitative as well as on quantitative criteria that systematically filter the list of possible candidates to simulate of a given process. In this study, the results showed that DELFT3D is the best-fitting model to simulate the thermal structure of a reservoir. The ELCOM model was not included in the selection process because of its role as a reference model.

The analysis of the performance results shows that MAE is a better alternative than RMSE to compare the models' results because MAE clearly distinguishes its behavior in each one of the compared zones. As a complement, the index of agreement, d_1, is a very useful tool to classify and compare models' performance due to its dimensionless nature, and allows a direct comparison between results.

Acknowledgments

References

[1] Rueda, F.J. and MacIntyre, S., Modelling the fate and transport of negatively buoyant storm–river water in small multi-basin lakes, Environ. Model. Softw., 25(1), pp. 146-157, 2010.

[2] Román-Botero, R., Gómez-Giraldo, A. y Toro-Botero, M., Efecto estacional de los afluentes en la estructura térmica de un pequeño embalse neotropical, La Fe – Colombia. DYNA, 80(177), pp. 152-161, 2013.

[3] Beletsky, D., Schwab, D. and McCormick, M., Modeling the 1998–2003 summer circulation and thermal structure in Lake Michigan, J. Geophys. Res., 111(C10), p. C10010, 2006.

[4] Osorio, A., Peláez-Zapata, D., Guerrero-Gallego, J., Álvarez-Silva, O., Osorio-Cano, J., Toro-Botero, F.M. y Giraldo, A., Hidrodinámica aplicada a la gestión y la conservación de ecosistemas marinos y costeros: Isla Gorgona , Océano Pacífico Colombiano, Rev. Biol. Trop., 62, pp. 133-147, 2014.

[5] Papanicolaou, A.N.T., Elhakeem, M., Krallis, G., Prakash, S. and Edinger, J., Sediment transport modeling review — Current, J. Hydraul. Eng., 2008.

[6] Smith, R., Mesa, O., Dyner, I. Jaramillo, P., Poveda, G. y Valencia, D. Decisiones con múltiples objetivos, Segunda ed. Universidad Nacional de Colombia, Sede Medellín, 2000.

[7] INTEGRAL-EEPPM, Proyecto hidroelectrico Rio-Grande II. Estudio de factibilidad: Informe Final, 1982.

[8] Marín, L.M., Modelamiento de la estructura térmica en embalses tropicales. Aplicación al embalse de Riogrande II, Tesis. Universidad Nacional de Colombia, Sede Medellín, Colombia, 2008.

[9] Franco, J.D., Modelación de la estructura térmica de un embalse ramificado mediante el análisis de los procesos físicos gobernantes. Aplicación al embalse multipropósito Riogrande II, Tesis. Universidad Nacional de Colombia, Sede Medellín, Colombia, 2011.

[10] Morillo, S., Imberger, J., Asce, M., Antenucci, J.P. and Woods, P.F., Influence of wind and lake morphometry on the interaction between two rivers entering a stratified lake, 134(11), pp. 1579-1590, 2009.

[11] Hodges, B. and Dallimore, C., Estuary, Lake and coastal ocean model: ELCOM v2.2 Science Manual, 2012.

[12] Spigel, R.H., Imberger, J. and Rayner, K.N., Modeling the diurnal mixed layer, Limnol. Oceanogr., 31(3), pp. 533-556, 1986.

[13] Read, J.S., Hamilton, D.P., Jones, I.D., Muraoka, K., Winslow, L.A., Kroiss, R., Wu, C.H. and Gaiser, E., Derivation of lake mixing and stratification indices from high-resolution lake buoy data, Environ. Model. Softw., 26(11), pp. 1325-1336, 2011.

[14] Roy, B., The outranking approach and the foundations of ELECTRE methods, Theory Decis., 31, pp. 49-73, 1991.

[15] Behzadian, M., Kazemzadeh, R.B., Albadvi, A. and Aghdasi, M., PROMETHEE: A comprehensive literature review on methodologies and applications, Eur. J. Oper. Res., 200(1), pp. 198-215, 2010.

[16] Saaty, T. and Vargas, L., Models, methods, concepts & applications of the Analytic Hierarchy Process, Second ed. Springer, 2012.

[17] Hetland, R.D. and DiMarco, S.F., Skill assessment of a hydrodynamic model of circulation over the Texas–Louisiana continental shelf, Ocean Model., 43–44, pp. 64-76, 2012.

[18] Sutherland, J., Peet, A.H. and Soulsby, R.L., Evaluating the performance of morphological models, Coast. Eng., 51(8–9), pp. 917-939, 2004.

[19] Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S. Robson, B., Seppelt, R., Voinov, A., Fath, B.D. and Andreassian, V., Characterising performance of environmental models, Environ. Model. Softw, 40, pp. 1-20, 2013.

[20] Willmott, C.J., On the validation of models.pdf, Phys. Gepgraphy, 2(2), pp. 184-194, 1981.

[21] Willmott C.J. and Matsuura, K., Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30, pp. 79-82, 2005.

[22] Willmott, C.J., Ackleson, S.G., Davis,R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J. and Rowe, C.M., Statistics for the evaluation and comparison of models, J. Geophys. Res., 90(C5), 8995 P., 1985.

[23] Hodges, B.R., Imberger, J., Laval, B. and Appt, J., Modeling the hydrodynamics of stratified lakes, in Hydroinformatics 2000 Conference, 2000, July, pp. 23-27.

[24] Deltares, Delft3D-FLOW - Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. Deltares, Rotterdamseweg, The Netherlands, pp. 1-706, 2013.

[25] D.H.I. DHI, Mike 3 Flow Model FM - Hydrodynamic Module. DHI, pp. 1-130, 2012.

[26] Boehrer B. and Schultze, M., Stratification of lakes, Rev. Geophys., 446, pp. 1-27, 2008.

[27] Lee, H., Chung, S., Ryu, I. and Choi, J., Three-dimensional modeling of thermal stratification of a deep and dendritic reservoir using ELCOM model, J. Hydro-environment Res., 7(2), pp. 124-133, 2013.

[28] Winter, C., On the evaluation of sediment transport models in tidal environments, Sediment. Geol., 202(3), pp. 562-571, 2007.

[29] van Rijn, L., Walstra, D.J., Grasmeijer, B., Sutherland, J., Pan, S. and Sierra, J., The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based Profile models, Coast. Eng., 47(3), pp. 295-327, 2003.

G.A. Betancur-Pérez, received the BSc. Eng in Civil Engineering in 2010 and the MSc. degree in Water Resources in 2014 both from the Universidad Nacional de Colombia, Medellin, Colombia. Currently, he works as full time engineer in the Centro de Ciencia y Tecnología de Antioquia (www.cta.org.co). His research interests include hydrodynamic modeling, physical limnology and water resources management.
ORCID:orcid.org/0000-0002-8923-4414

F.M. Toro-Botero, received his BSc. Eng in Civil Engineering in 1981 from the Universidad Nacional de Colombia. Medellin, Colombia. The MSc. degree in Engineering in 1985 and the PhD degree in Hidroscience and Engineering in 1994, both of them from the University of Mississippi, USA. Currently, he is a Full Professor in the Geosciences and Environmental Department, Facultad de Minas of the Universidad Nacional de Colombia, Medellín, Colombia. His research interests include hydrodynamics, fluid mechanics, numerical methods, water quality simulation and hydrodynamic modeling.
ORCID: 0000-0002-4763-7469

E.A. Gómez-Giraldo, received his BSc. Eng in Civil Engineering in 1995 and the MSc. degree in Water Resources in 1998 both from the Universidad Nacional de Colombia. Medellin, Colombia. His PhD in Environmental Engineering in 1997 from the University of Western Australia. Currently, he is an assistant Professor in the Geosciences and Environmental Department, Facultad de Minas, of the Universidad Nacional de Colombia, Medellín, Colombia. His research interests include physical limnology and hydrodynamic simulation.
ORCID: 0000-0001-7103-9

Facile one-pot synthesis of uniform silver nanoparticles and growth mechanism

Daniel Ramirez & Franklin Jaramillo

Centro de Investigación, Innovación y Desarrollo de Materiales-CIDEMAT, Universidad de Antioquia - UdeA, Medellín, Colombia.
estiben.ramirez@udea.edu.co, franklin.jaramillo@udea.edu.co

Abstract

Size controlled silver nanoparticles were obtained via chemical reduction using one-pot synthesis. Differently from other reported methods for silver nanoparticles, 1-octanol was used as both solvent and reduction agent, oleylamine and oleic acid acted as capping agents and silver nitrate was used as the metal precursor. Ultraviolet-visible and Raman spectroscopy were used to monitor the in situ growth of the nanoparticles and to corroborate the oxidation of the alcohol to caprylic acid. X-ray diffraction (XRD) and transmission electron microscopy (TEM) served to find the size and shape of the nanoparticles. It was found that the temperature used and the reagents proportions were appropriated to produce silver nanoparticles. A growth mechanism was proposed including the formation of silver carboxylates as an intermediate step of the reaction. As a systematic use of oleic acid, we could observe that a higher concentration of this capping agent led to smaller and more homogenous nanoparticles, less than 5nm in size.

Keywords: silver nanoparticles, nanoparticles growth, capping agents.

Fácil síntesis en un paso y mecanismo de formación de nanopartículas de plata

Resumen

En este trabajo se obtuvieron nanoparticulas de plata por la ruta de reducción química usando síntesis en un paso. Diferente a otros métodos reportados para nanoparticulas de plata, se usó 1-octanol como solvente y agente reductor, oleilamina y ácido oleico como agentes estabilizantes, mientras que el precursor del metal fue nitrato de plata. El crecimiento de las nanopartículas fue monitoreado por espectroscopia de absorción ultravioleta-visible y Raman, lo cual corroboró la oxidación del alcohol a ácido caprilico. Los ensayos de difracción de rayos X y microscopia electrónica de transmisión permitieron conocer la forma y el tamaño de las nanopartículas. Se encontró que la temperatura y los reactivos empleados fueron apropiados para producir nanopartículas de plata. Se formaron carboxilatos de plata como un paso intermedio en la reacción. Finalmente, el incremento en la concentración de ácido oleico permitió la obtención de nanopartículas de plata de menos de 5nm.

Palabras clave: nanopartículas de plata, crecimiento, agentes estabilizantes.

1. Introduction

Silver nanoparticles are very attractive due to their remarkable size and shape-dependent electronic and optical properties [1,2]. Hence, they have been used in a broad range of fields which include catalysis, photonics and surface-enhanced Raman scattering (SERS) [3-10]. It has also been demonstrated that they have highly effective bactericide properties related to the continuous release of silver ions, and therefore they have become one of the most commercialized nanotechnology products in the area of health care; i.e., for bandages, clothing, and cosmetics [11].

The absorption spectrum of silver nanoparticles is sharp and strong, a feature which is ascribed to the so called surface

plasmons, which consists in the excitation of collective electron oscillations, in response to an electromagnetic field [12]. The optical properties of silver nanoparticles are dictated by the morphological parameters, such as size and shape; hence it is crucial to be able to control these parameters [13].

Many strategies have been developed for the preparation of silver nanoparticles. Mehta S.K. et al [14] prepared silver nanoparticles with controlled shapes and size in homogeneous aqueous solutions using silver nitrate and different saccharides in a micellar media with sodium dodecyl sulfate as surfactant. More complex methods have also been developed, like the use of gamma radiation known as the radiolytic reduction method, in which an aqueous solution of the silver precursor (commonly silver nitrate) is irradiated with gamma rays in order to generate reactive species capable of reducing silver ions Ag+ into zero-oxidation state silver atoms (Ag^0) [15,16]. Tri-sodium citrate, tannic acid and silver nitrate have also been used [17], but some other methods require the usage of uncommon silver tetradecanoate as a precursor reagent [18]. One of the most common reducing agents for the synthesis of nanoparticles is 1,2-hexadecanediol [19,20], and other less expensive capping and reducing agents like liquid paraffin [21] have also been used. In this case, 1-octanol, which is a very common and available solvent, is explored both as a solvent and a reducing agent.

2. Experimental methods

In order to fulfill the requirements of a standard, high purity (>99.8%) silver nitrate and 1-octanol were purchased from Merck, Oleic Acid (90%) from Alfa Easer and technical grade oleylamine (70%) from Sigma-Aldrich. All these reagents were used without any further purification.

2.1. Synthesis of silver nanoparticles

Silver nanoparticles were obtained by a facile method via one-pot synthesis. The reagents used are described in Table 1. Changes were made to the quantity of Oleic acid rather than the quantity of oleylamine or 1-octanol because of this reagent's lower cost. In order to synthesize nanoparticles with uniform sizes, a "hot injection" rather than "heating up" method was adopted as proposed by Chen Y. et al [22]. 1-octanol, oleylamine and oleic acid were added in a 100 mL two-neck flask, and the mixture was heated to 180°C under magnetic stirring. Upon reaching the set temperature, AgNO$_3$ was added and the reaction continued and at the same time aliquots were taken at different times and cooled in an ice bath in order to follow the kinetic by Raman and UV-vis absorbance measurements. The resulting dark brown solution was cooled down to room temperature, and the product was precipitated by adding ethanol to the solution and collected by centrifugation at 5000 rpm for 5min, which was further purified by washing with ethanol 2–3 times. Final products were re-dispersed in hexane for later use and analysis.

Table 1.
Reagents used in the synthesis.

Nomenclature	AgNO$_3$ (mg)	Oleic acid (ml)	Oleylamine (ml)	1-Octanol (ml)
Ag1	203	2.0	2.5	30
Ag2	203	3.0	2.5	30
Ag3	203	4.0	2.5	30

Source: The authors

2.2. UV–Vis absorption spectrophotometry

The UV-vis absorption spectra were taken at room temperature on a UV-vis spectrophotometer Cary 100 from Varian Inc, with a variable wavelength between 300 and 800 nm using a glass cuvette with an optical path of 1 cm.

2.3. X-ray diffraction (XRD)

The X-ray powder diffraction (XRD) pattern was recorded using a PANalytical X-ray diffractometer with Cu Kα radiation (1.5406Å). Samples for measurement were prepared by dropping silver colloids (dispersion in hexane) on the Mylar grid and allowing them to dry at room temperature. The scan step size was 0.0525, ranging from 20 to 90° with a time per step of 64s.

2.4. Micro-Raman spectroscopy

Raman spectra of the samples were recorded using a Horiba Yvonjobin dispersive micro-Raman spectrophotometer. A 785nm laser for the excitation radiation was used. All the spectra were collected in the range 3700–100cm^{-1}.

2.5. Transmission electron microscopy (TEM)

Transmission electron microscopic images were obtained on a FEI Tecnai TEM. Samples for the TEM were dispersed in hexane and deposited on an amorphous carbon film-coated copper grid followed by natural evaporation at room temperature.

3. Results

Fig. 1 shows a picture of the resulting nanoparticles dispersed in hexane after being purified. The brown color is very characteristic of the colloidal dispersion. Visual monitoring was carried out for 2 months and no color change or sedimentation was observed, indicating that the particles were very stable in this solvent.

3.1. Growth process of silver nanoparticles: UV-Vis and Raman spectra

2 mL aliquots of the reacting solution were deposited into test tubes at different times (2, 10, 30, 60 and 120min) and then immediately cooled in ice water to stop the reaction. To carry on the optical spectra of the nanodispersions, 20μL of each aliquot were diluted into 3mL of ethanol.

Figure 1. Picture of silver nanoparticles dispersed in hexane.
Source: The authors

The evolution of the growing silver nanoparticles was monitored by their change of the UV-visible spectrum. Fig. 2a shows the resultant absorbance spectra for Ag2. A very characteristic absorption peak could be observed around 460nm. We can see that the absorbance increased rapidly in just 2 minutes, indicating a higher concentration of silver nanoparticles; after 2 hours the change was not significant, because all particles were already formed. A blue shift was found while the reaction occurred, which is consistent with other reports [23]. As the position of the peak is very dependent on the nanoparticles' size [21] the 460nm peak is not the real one associated to the plasmon resonance of the particles, because they were not well dispersed in the samples taken from the reacting solution (aggregates absorb at higher wavelengths). Lower peak positions were obtained after purification and dispersion in hexane (See Fig. 5).

The two well-defined processes associated with the nucleation and growth of the particles can be observed in Fig. 2b. These are similar to those proposed by Lamer and Dinegar [24] who studied the variation of the solute concentration as a function of time. Nucleation implies an increase in the number of scattering centers (number of particles) for a given system, and therefore gives an increase in the scattered intensity. In contrast, the growth of particles is associated with a decrease of the scattered intensity since the observation window corresponds to the diffraction of smaller particles that are disappearing during the growth process or dissolution of the unstable nucleus. It is consistent with the mechanism of the reduction of Ag^+ ions and the association of Ag^0 atoms to produce metallic Ag particles as proposed by several authors [14,25,26].

Raman measurements were also carried out in order to understand the evolution of the reaction. It was not possible to find a clear reported position band for the absorption of silver nanoparticles in the literature, but it is known that silver presents a lattice vibrational mode between 50 and 300cm^{-1} depending on whether it is present as oxide, nitrate, chloride or as some other compound [26]. Nevertheless, as a reference we decided to use a 99.99% pure silver wire to compare with
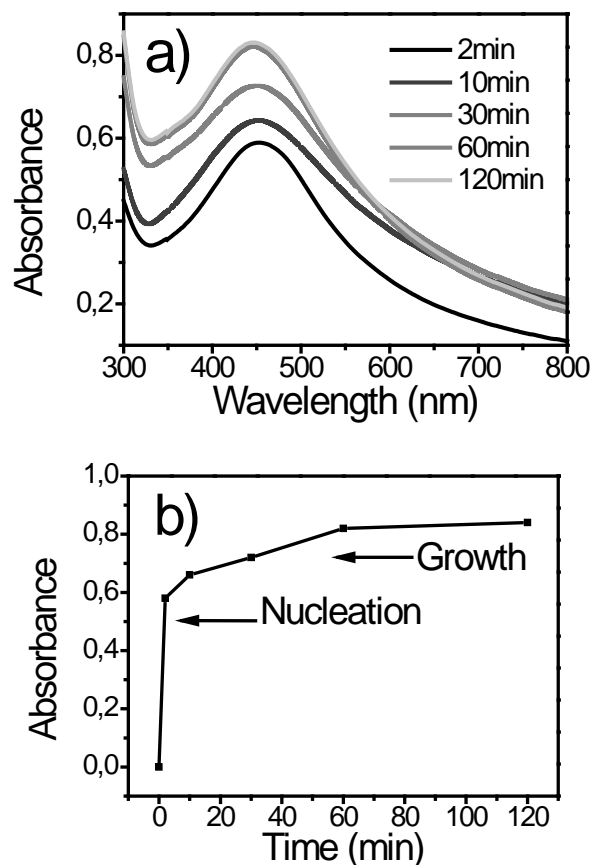
Figure 2. UV-Vis absorption spectra of the Ag2 reaction mixture.
Source: The authors

Table 2.
Raman assignment.

Raman shift (cm^{-1})	Assignment [27, 28]
244	Ag lattice vibrational mode of silver standard
270	Ag lattice vibrational modes of silver nanoparticles
1302	C-O stretching
1388	COO stretching
1438	CH$_2$ scissoring
1658	C=O stretching

Source: The authors

the position for the metallic silver. The observed bands were found to be 244cm^{-1} for the silver wire, while 270cm^{-1} for the growing silver nanoparticles. Table 2 shows the signals assignation.

The Raman spectra of the reaction mixture are plotted in Fig. 3. This behavior was attributed to an intermediate step, in which silver carboxylates could be formed; this is thought to be because the symmetric strong COO stretch band is usually seen at 1450-1360cm^{-1} [28]. The symmetric stretching vibrations of C=O groups could be identified at 1658cm^{-1}; from 2 to 10 minutes its intensity remained almost the same, but later, the peak became more intense, indicating that more C=O groups were formed due to the oxidation of 1-octanol to octanoic acid (Caprylic acid). Oleylamine can
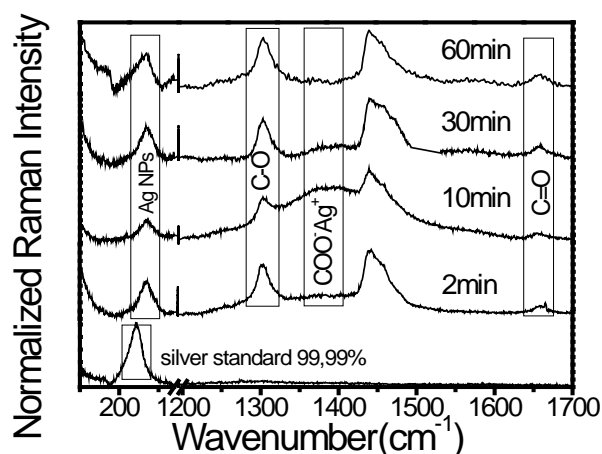
Figure 3. Spectra of the Ag2 reaction mixture.
Source: The authors
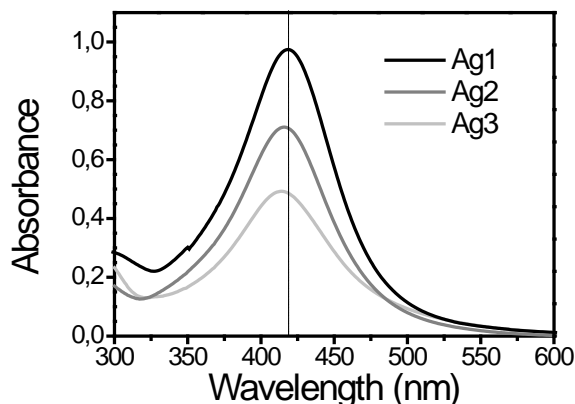
Figure 5. UV-Vis absorption spectra of silver nanoparticles in hexane after purification.
Source: The authors
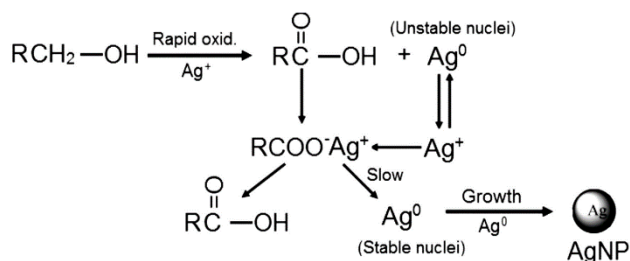
Figure 4. Schematic illustration of the proposed formation mechanism of silver nanoparticles obtained by 1-octanol as reducing agent.
Source: The authors

Figure 6. X-ray diffraction patterns of Ag1 and Ag3 silver nanoparticles
Source: The authors

also act as reducing agent and undergo metal-ion-induced oxidation to nitriles [29], but it was not possible to observe changes because the content of Oleylamine was relatively low.

Fig. 4 shows the proposed formation mechanism of silver nanoparticles according to the results observed by Raman. The formation of intermediated unstable silver nanoparticles occurs very rapidly (less than two minutes) as observed in the absorption spectra. These nanoparticles are not large enough to grow and these unstable nuclei can be partially dissolved, and therefore as shown in the Raman results, the peak intensity attributed to the already formed nanoparticles decreased while the silver carboxylate increased (see Fig. 3 at 10 minutes) [24]. As the silver carboxylate formation is an intermediate step, and as the salts are not very stable at 180 °C, they provide the Ag+ ions, which are later reduced to form stable nuclei for the nanoparticles. This means that the formed carboxylates tend to disappear (as seen in the Raman spectrum at 30 minutes) and the intensity associated to the formed nanoparticles has to increase.

Fig. 5 shows the resulting absorbance spectra for the three sets of nanoparticles synthesized in this work. If the particle size becomes comparable to or smaller than the mean free path of the conduction band electrons or "free" electrons, the collisions of the electrons with the particle surface becomes important and the effective mean free path is less than that existing in bulk materials. This usually results in broadening and blue-shift of the plasmon band for particles smaller than

about 10 nm [30]. There was a slightly blue shift in the maximum absorbance peak from Ag1 to Ag3. These results suggest that smaller nanoparticles were obtained by increasing the oleic acid concentration, probably because of the steric effects; the growing particles are rapidly capped or stabilized by the molecules of the acid and therefore the opportunity for a particle to interact with others becomes lower.

3.2. Particle size: XRD and TEM

The X-ray diffraction patterns of Ag1 and Ag3 samples are shown in Fig. 6. The peaks of Ag1 perfectly match the face-centered cubic (fcc) structure of the bulk silver, with the broad peaks around $2\theta = 38°$, $44°$, $65°$, and $78°$ corresponding to (111), (200), (220) and (311) lattice planes, respectively.

The average crystallite size of both samples were calculated over the (111) reflection plane using the classical Scherrer Eq.1 [31]. k is the Scherrer constant (k = 0.89) λ is the wavelength of the X-ray, β is the FWHM of the peak and θ is half of the Bragg angle.

$$D = \frac{k\lambda}{\beta \cos\theta} \tag{1}$$

The calculated crystallite size was found to be 5.97nm and 2.63nm for Ag1 and Ag3, respectively. These results are consistent with the absorbance spectra of Fig. 5.

Fig.s 7 and 8 show TEM images of the samples. The corresponding histogram of the particles size distribution for the respective samples is presented along with the TEM images. Ag1 (Fig. 7) ranged from 4 to 71nm and Ag3 from 2.1 to 4.7nm, the latter being more uniform and smaller than Ag1, with an average particle size of 3.8±0.5nm.

A lower concentration of oleic acid resulted in larger crystallites and nanoparticles because it was probably not enough to completely cover the particles' surface and stabilize; therefore, some non-regular shapes (elongated spheres and prisms) were obtained.

According to Fig. 8, there was an important improvement in the distribution of the particles indicating that a simple change in the capping agent's concentration allows for controlling the nanoparticles' shape and size and that can lead to the formation of almost monodispersed nanoparticles

4. Conclusions

Well dispersed silver nanoparticles with controllable size and shape were prepared by reducing silver nitrate with 1-octanol in the presence of oleic acid and oleylamine as capping agents. The nucleation and growth processes could be recognized by UV-vis
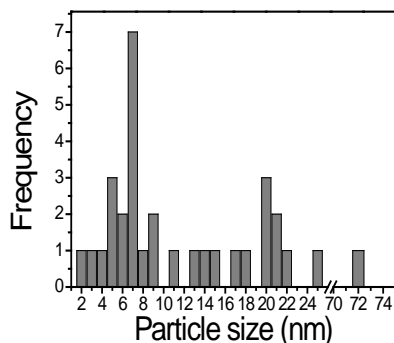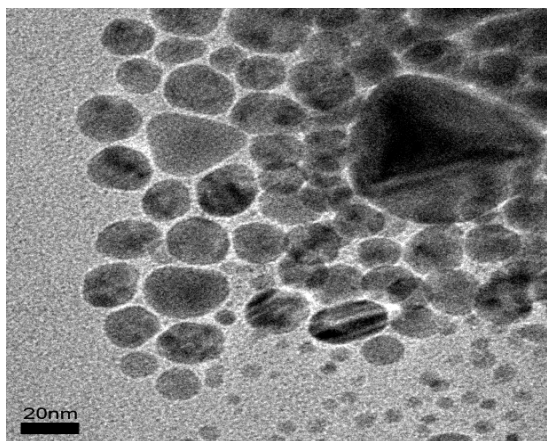
Figure 7. TEM image (top) and the corresponding particle size distribution (bottom) of Ag1.
Source: The authors
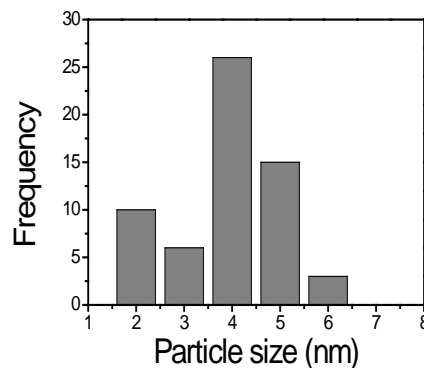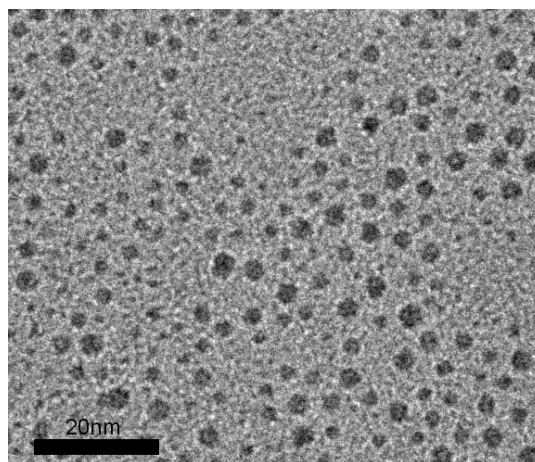
Figure 8. TEM image (top) and the corresponding particle size distribution (bottom) of Ag3.
Source: The authors

spectrophotometry. A mechanism for the formation of the silver nanoparticles was proposed according to the Raman measurements, which allowed us to identify that during the reduction of silver ions and oxidation of 1-octanol to caprylic acid, an intermediate step occurred, in which silver carboxylates were formed. Finally, oleic acid limited further aggregation and particle size of silver nanoparticles and fully stabilized the dispersed silver nanoparticles in solution, which served to obtain smaller and more homogenous spherical nanoparticles.

Acknowledgements

References

[1]   Tsuboi, A., Nakamura, K. and Kobayashi, N., Chromatic control of multicolor electrochromic device with localized surface plasmon resonance of silver nanoparticles by voltage-step method. Solar Energy Materials and Solar Cells, 145, pp 16-25, 2016. DOI: 10.1016/j.solmat.2015.07.034

[2]   Xiangheng, N., Libo, S., Jianming, P., Fengxian, Q., Yongsheng, Y., Hongli, Z., and Minbo, L., Modulating the assembly of sputtered silver nanoparticles on screen-printed carbon electrodes for hydrogen peroxide electroreduction:

Effect of the surface coverage. Electrochimica Acta, 199, pp 187-193, 2016. DOI: 10.1016/j.electacta.2016.03.100

[3] Steffan, M., Jakob, A., Claus, P. and Lang, H., Silica supported silver nanoparticles from a silver (I) carboxylate: Highly active catalyst for regioselective hydrogenation. Catalysis Communications, 10, pp. 437-441, 2009. DOI: 10.1016/j.catcom.2008.10.003

[4] Kalfagiannis, N., Karagiannidis, P.G., Pitsalidis, C., Panagiotopoulos, N.T., Gravalidis, C., Kassavetis, S., Patsalasc, P. and Logothetidisa, S., Plasmonic silver nanoparticles for improved organic solar cells. Solar Energy Materials and Solar Cells, 104, pp 165-174, 2012. DOI: 10.1016/j.solmat.2012.05.018

[5] Pei, J., Tao, J., Zhou, Y., Dong, Q., Liu, Z., Li, Z, Chen, F., Zhang, J., Xu, W. and Tian, W., Efficiency enhancement of polymer solar cells by incorporating a self-assembled layer of silver nanodisks. Solar Energy Materials and Solar Cells, 95, pp. 3281-3286, 2011. DOI: 10.1016/j.solmat.2011.07.007

[6] Kang, Y., Si, M., Zhu, Y., Miao, L. and Xu, G., Surface-enhanced Raman scattering (SERS) spectra of hemoglobin of mouse and rabbit with self-assembled nano-silver film. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 108, pp. 177-180, 2013. DOI: 10.1016/j.saa.2013.01.098

[7] Vanamudan, A. and Sudhakar, P.P., Biopolymer capped silver nanoparticles with potential for multifaceted applications. International Journal of Biological Macromolecules, 86, pp. 262-268, 2016. DOI: 10.1016/j.ijbiomac.2016.01.056

[8] Bi, L., Dong, J., Xie, W., Lu, W., Tong, W., Tao, L. and Qian, W., Bimetallic gold–silver nanoplate array as a highly active SERS substrate for detection of streptavidin/biotin assemblies. Analytica Chimica Acta, 805, pp. 95-100, 2013. DOI: 10.1016/j.aca.2013.10.045

[9] Geetha, K., Umadevi, M., Sathe, G.V., Vanelle, P., Terme, T. and Khoumeri, O., Orientation of 1,4-dimethoxy-3–bromomethylanthracence-9,10-dione on silver nanoparticles: SERS studies. Journal of Molecular Structure, 1059, pp. 87-93, 2014. DOI: 10.1016/j.molstruc.2013.11.013

[10] Raza, A. and Saha, B., In situ silver nanoparticles synthesis in agarose film supported on filter paper and its application as highly efficient SERS test stripes. Forensic Science International, 237, 2014.

[11] Chernousova, S. and Epple, M., Silver as antibacterial agent: Ion, nanoparticle, and metal. Angewandte Chemie International Edition, 52, pp. 1636-1653, 2013. DOI: 10.1002/anie.201205923

[12] Chatre, A., Solasa, P., Sakle, S., Thaokar, R. and Mehra, A., Color and surface plasmon effects in nanoparticle systems: Case of silver nanoparticles prepared by microemulsion route. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 404, pp. 83-92, 2012. DOI: 10.1016/J.COLSURFA.2012.04.016

[13] An, W., Zhu, T. and Zhu, Q., Numerical investigation of radiative properties and surface plasmon resonance of silver nanorod dimers on a substrate. Journal of Quantitative Spectroscopy and Radiative Transfer, 132, pp. 28-35, 2014. DOI: 10.1016/j.jqsrt.2013.01.013

[14] Mehta, S.K., Chaudhary, S. and Gradzielski, M., Time dependence of nucleation and growth of silver nanoparticles generated by sugar reduction in micellar media. Journal of Colloid and Interface Science, 343, pp. 447-453, 2010. DOI: 0.1016/j.jcis.2009.11.053

[15] Saion, E., Gharibshahi, E. and Naghavi, K., Size-Controlled and optical properties of monodispersed silver nanoparticles synthesized by the radiolytic reduction method. International Journal of Molecular Sciences, 14, pp. 7880-7896, 2013. DOI: 10.1021/la0600245

[16] Shin, H.S., Yang, H.J., Kim, S.B. and Lee, M.S., Mechanism of growth of colloidal silver nanoparticles stabilized by polyvinyl pyrrolidone in γ-irradiated silver nitrate solution. Journal of Colloid and Interface Science, 274, pp. 89-94, 2004. DOI: 10.1016/J.JCIS.2004.02.084

[17] Dadosh, T. Synthesis of uniform silver nanoparticles with a controllable size. Materials Letters, 63, pp. 2236-2238, 2009. DOI: 10.1016/J.MATLET.2009.07.042

[18] Yamamoto, M., Kashiwagi, Y. and Nakamoto, M., Size-controlled synthesis of monodispersed silver nanoparticles capped by long-chain alkyl carboxylates from silver carboxylate and tertiary amine. Langmuir, 22, pp. 8581-8586, 2006. DOI: 10.1021/LA0600245

[19] Chiang, I.C., Chen, Y.T. and Chen, D.H., Synthesis of NiAu colloidal nanocrystals with kinetically tunable properties. Journal of Alloys and Compounds, 468, pp. 237-245, 2009. DOI: 10.1016/J.JALLCOM.2008.01.063

[20] Li, Y., Zhang, X.L., Qiu, R. and Kang, Y.S., Synthesis and investigation of SmCo5 magnetic nanoparticles. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 313, pp. 621-624, 2008. DOI: 10.1016/J.COLSURFA.2007.04.150

[21] Chen, M., Feng, Y.G., Wang, X., Li, T.C., Zhang, J.Y., and Qian, D.J., Silver nanoparticles capped by oleylamine: Formation, growth, and self-organization. Langmuir, 23, pp. 5296-5304, 2007. DOI: 10.1021/LA700553D

[22] Chen, Y., Gao, N. and Jiang, J., Surface matters: Enhanced bactericidal property of core–shell Ag–Fe2O3 nanostructures to their heteromer counterparts from one-pot synthesis. Small, 9, pp. 3242-3246, 2013. DOI: 10.1002/SMLL.201300543

[23] Jana, N.R., Sau, T.K. and Pal, T., Growing small silver particle as redox catalyst. The Journal of Physical Chemistry B, 103, pp. 115-121, 1998. DOI:10.1021/JP982731F

[24] Lamer, V.K. and Dinegar, R.H., Theory, production and mechanism of formation of monodispersed hydrosols. Journal of the American Chemical Society, 72, pp. 847-854, 1950. DOI: 10.1021/ja01167a001

[25] Goia, D.V., Preparation and formation mechanisms of uniform metallic particles in homogeneous solutions. Journal of Materials Chemistry, 14, pp. 451-458, 2004. DOI: 10.1039/B311076A

[26] Harada, M, Inada, Y. and Nomura, M., In situ time-resolved XAFS analysis of silver particle formation by photoreduction in polymer solutions. Journal of Colloid and Interface Science, 337, pp. 427-438, 2009. DOI: 10.1016/j.jcis.2009.05.035

[27] Martina, I., Wiesinger, R., Jembrih, D. and Schreiner, M., Micro-raman characterisation of silver corrosion products: Instrumental set up and reference database. E-Preservation Science: Morana RTD [Online]. pp. 1-8, 2012. Available at: http://www.morana-rtd.com/e-preservationscience/2012/Martina-05-03-2012.pdf

[28] Lin-Vien, D., Colthup, N. B., Fateley, W.G. and Grasselli, J.G., CHAPTER 9 - Compounds containing the carbonyl group. In: Lin-Vien, D, Colthup, NB, Fateley, WG, Grasselli, JG, eds. The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules. San Diego: Academic Press, 1991, pp. 117-154.

[29] Mourdikoudis, S. and Liz-Marzán, L.M., Oleylamine in nanoparticle synthesis. Chemistry of Materials, 25, pp. 1465-1476, 2013. DOI: 10.1021/cm4000476

[30] Zhang, J.Z., Optical Properties and Spectroscopy of Nanomaterials. Cahpter 7 - Optical Properties and Spectroscopy of Nanomaterials: World Scientific, 2009, pp. 205-235.

[31] Scherrer, P., Bestimmung der Grösse und der inneren Struktur von Kolloidteilchen mittels Röntgenstrahlen. [Online]. pp. 98-100, 1918. Available at: http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002505045&IDDOC=63709

D. Ramirez, is currently a PhD student in Materials Engineering at Universidad de Antioquia, Medellín, Colombia, where he also completed his BSc. Eng in Materials Engineering in 2014. His research interests include nanothecnology and nanostructured solar cells. He is currently working on synthesis and characterization of nanoparticles, nanocomposites and novel semiconductors for thin film photovoltaic applications in The Center of Research, Innovation and Development of Materials – CIDEMAT at Universidad de Antioquia.
ORCID: 0000-0003-2630-7628

F. Jaramillo, completed his BSc. Eng in Chemical Engineering at Universidad de Antioquia, Medellín, Colombia, in 2001 and, in 2005, a PhD in Chemistry at The University of Manchester, USA. He is the current director of the solar cells lab EPM-UdeA, part of the Nanotechnology Regional Iniciative at Ruta N-Medellín, and a member of the national advisory council in Nanoscience and Nanotechnology - Red NanoColombia. His research interest areas include nanotechnology, nanostructured solar cells, novel semiconductors, energy materials, nanocomposites, and functional polymers from renewable resources. He is currently an associated professor at Materials Engineering Department and member of The Center of Research, Innovation and Development of Materials – CIDEMAT at Universidad de Antioquia.
ORCID: 0000-0003-1722-5487

Adsorption and catalytic oxidation of asphaltenes in fumed silica nanoparticles: Effect of the surface acidity

Camilo Andrés Franco-Ariza, Juan David Guzmán-Calle & Farid Bernardo Cortés-Correa

Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. caafrancoar@unal.edu.co, jdguzmanc@unal.edu.co, fbcortes@unal.edu.co

Abstract

This study aims to evaluate the effect of surface acidity of fumed silica nanoparticles in adsorption and subsequent thermal cracking of Colombian asphaltenes. The acidities of the surfaces were established through Temperature Programed Desorption (TPD) experiments. The adsorption equilibrium of asphaltenes was determined using a static batch method, and the data obtained was fitted using the Langmuir model, the Freundlich model and the SLE Model. Asphaltenes catalytic oxidation experiments were conducted, and it was found that this process was surface nature dependent. In all cases, the temperature of asphaltenes oxidation was reduced regarding the virgin asphaltene sample. The effective activation energies were estimated with the iso-conversional OFW method. This energy was found to be related to adsorption affinity and asphaltenes self-association on nanoparticles surface.

Keywords: asphaltenes; adsorption isotherms; thermal cracking; nanoparticles; superficial modification.

Adsorción y oxidación catalítica de asfaltenos en nanopartículas de sílice fumárica: Efecto de la acidez superficial

Resumen

Este trabajo busca evaluar el efecto de la acides superficial de nanopartículas de sílice fumárica en la adsorción y craqueo térmico de asfaltenos colombianos. La acidez de las superficies fue determinada a través de pruebas TPD. Los experimentos de adsorción fueron elaborados utilizando un método por lotes y los datos obtenidos fueron ajustados al modelo de Langmuir, al modelo de Freundlich y al modelo SLE. Se llevó a cabo la oxidación catalítica de los asfaltenos y se encontró que este proceso era dependiente de la superficie en que se llevaba a cabo. En todos los casos, la temperatura de oxidación de los asfaltenos fue reducida en consideración con los asfaltenos vírgenes. Las energías de activación fueron estimadas con el método isoconversional OFW. Se encontró que esta energía está relacionada con la afinidad del proceso adsortivo y la auto-asociación de los asfaltenos en la superficie de la nanopartícula

Palabras clave: asfaltenos; isotermas de adsorción; craqueo térmico; nanopartículas; modificación superficial.

1. Introduction

According to a projection made by the Organization of the Petroleum Exporting Countries (OPEC), a growth in energetic global demand of at least 50% is expected over the next decades [1]. Demand will increase from 267.6 million barrels of oil equivalent per day (mboe/d) in 2013 to 399.4 mboe/d in 2040 [1]. In this year it is also expected that 53% of this demand will be supplied by hydrocarbon industry [1]. In this scenario, it is worthwhile noting that Extra-Heavy Oil (EHO) reserves correspond to 32% of world oil reserves [2]; therefore, it is necessary to investigate about this kind of crude oil could as an alternative to supplying the energy demand. EHO have a lot of heavy hydrocarbon compounds such as asphaltenes that reduce the American Petroleum Institute gravity (°API) and drastically increase oil viscosity [3,4]. Therefore, asphaltenes are defined as the heaviest fraction of the crude oil soluble in aromatics hydrocarbons, but it is insoluble in alkanes such as n-pentane

and n-heptane [5, 6]. Generally, it has a polyaromatic core that is attached to alkyl chains and heteroatoms like O, N, S, Ni, V and Fe [7-10]. Some problems associated with these compounds are the reduction of oil mobility at reservoir conditions, changes in the reservoir wettability, and their deposition over refining equipment [11]. Typically, asphaltenes are removed by solvent injection, vapor injection, or mechanically [12-14], but these techniques, in most the cases, can lead to re-deposition and then a lower Energy Return On Investment (EROI) [15]. In summary, the presence of asphaltenes can affect production, transportation and the refinery processes, involving a large amount of capital and operational costs because conventional recovery techniques cannot be used effectively [16].

In this sense, new cost-effective and environmentally friendly technologies that enhance EHO from unconventional resources with lower operational and capital costs are a priority. As such, nanoparticles have recently become an area of research that is attractive for the oil industry [5, 16-26]. The adsorption and subsequent thermal decomposition of asphaltenes onto surfaces of nanoparticles was first introduced by Nassar and colleagues [5,16,17,26-31]. In his earlier study [5], he investigated the effect of different metal oxides in the adsorption process and post-oxidation, pyrolysis or gasification of asphaltenes. He found that these processes are adsorbent specific. Recently, at the Universidad Nacional de Colombia, several studies have been conducted on functionalized nanoparticles for asphaltenes adsorption and post-decomposition [19,20,23,25].

A in-depth and extensive investigation into different key variables, such as asphaltenes' chemical structure and nanoparticles functionalities will provide insight into the mechanism and behavior of nanoparticles' efficiency as catalysts and inhibitors. This will result in enhanced oil recovery techniques and quality improvements. For this reason, this paper is the continuation of our work, and its primary objective is to modify the fumed silica nanoparticles surface in order to obtain different acidities and then evaluate their effect on the adsorption and catalytic oxidation of asphaltenes. Also, a correlation between the calculated effective activation energies and the Langmuir model, the Freundlich model, and the SLE model parameters are presented to provide a better understanding of the role of the adsorption process in asphaltenes catalytic decomposition.

2. Materials and methods

2.1. Materials

A Colombian crude oil with 7.9°API, a viscosity of 595688 cP at 25°C, and approximately 13 wt% of asphaltenes content is used as the asphaltenes source. These were extracted by n-Heptane (99%, Sigma-Aldrich, St. Louis, MO). Toluene (99.5%, MerkK GaG, Germany) was used to prepare heavy oil model solutions. Fumed silica nanoparticles, HCl (37%), and NH_4OH (28%) were used for superficial modifications and were purchased from Sigma-Aldrich (St. Louis, MO). To determine superficial acidity, 10% NH_3/He and He (99.9%) gasses were provided by Linde (Medellín, Colombia).

2.2. Methods

2.2.1. Asphaltenes extraction

Solid n-C_7 asphaltene was extracted from the crude oil by adding an excess amount of n-heptane following a standard procedure described in previous studies [5, 19]. 40 mL of n-heptane are added to the crude oil per every gram of it that was used. The mixture is sonicated for 2 h at 25°C and then stirred at 300 rpm for 20 h. Samples are further centrifuged at 5000 rpm for 15 min. The precipitated is filtered through 8 μm Whatman filter paper and washed with n-heptane until the color of the asphaltenes became shiny black. Finally, the asphaltenes obtained were homogenized and fined using a mortar and left to dry in a vacuum oven at room temperature (RT) for 12 h.

2.2.2. Nanoparticles surface modification

The nanoparticles' surface is modified to neutral, basic or acidic by adding the nanoparticles to aqueous solutions with different pH values of 3, 7 and 10. The pH is adjusted by adding aliquots of HCl or NaOH at 0.1N. The mixture is sonicated at (RT) for 2 hours. Then, the nanoparticles are separated and washed with deionized water until the pH value of the residue remains constant. Finally, the product is dried at 120°C for 6 hours to eliminate any humidity.

2.2.3. Particle size and surface area measurements

The sizes of the nanoparticles were obtained using an XPert PRO MPD X-ray diffractometer (PANalytical, Almelo, Netherlands) with Cu Kα radiation operating at 60 kV and 40 mA with a θ/2θ goniometer. The mean nanoparticle diameter (crystallite size) was determined by applying the Scherrer equation to the principal diffraction peak. Additionally, the surface areas (S_{BET}) were measured using a ChemBET 3000 (Quantachrome Instruments, Florida, USA) by employing the Brunauer-Emmet-Teller (BET) method [32,33]. This was undertaken by degassing the samples overnight at 140°C under a N_2 flow and after performing nitrogen adsorption-desorption at -196 °C.

2.2.4. Temperature Programmed Desorption

NH_3 temperature programmed desorption (TPD) measurements were carried out with 0.1 g samples at a flow rate of 80 mL/min. Before the TPD measurements were made, nanomaterials were pretreated in a flow of Helium at 110°C for 3 h. Then, samples were treated with 10% NH_3/He at 100°C for 1 h ensuring NH_3 saturation and purged with He for 1 h. TPD runs were carried out between 100°C to 900°C at a heating rate of 10°C/min. The online ChemBet 3000 (Quantachrome Instruments, Florida, USA) with a Thermal Conductivity Detector (TCD) was used to monitor the desorbed NH_3.

2.2.5. Equilibrium adsorption isotherms

Model solutions for the batch adsorption experiments are prepared by dissolving a desired amount of the obtained

asphaltenes in toluene. All samples were made from a stock solution that was prepared at a maximum concentration (C_{max}) of 5000 mg/L without asphaltenes precipitation. The initial concentration of asphaltene solutions used in the adsorption experiments ranged from 500 mg/L to C_{max}. Before the experiments were undertaken, a calibration curve of absorbance against concentration was constructed at a wavelength of 295 nm [34], using a Genesys 10S UV-VIS spectrophotometer (Thermo Scientific, Waltham, MA). Nanoparticles are added to the solutions in a relation of 100mg/10 mL. The solutions are stirred at 300 rpm for 24 h at 25°C to ensure that they were in equilibrium. The amount adsorbed q in units of mg of asphaltenes / m^2 of nanoparticles surface area is estimated according to eq. (1):

$$q = \frac{C_0 - C_E}{A} V \tag{1}$$

where C_0 (mg/L) and C_E (mg/L) are the initial concentration of asphaltenes in the solution and the equilibrium concentrations of asphaltenes in the supernatant, respectively; V (L) the solution volume and A (m^2) the nanoparticles' dry surface areas.

2.2.6. Thermogravimetric analysis

After the adsorption experiments, the nanoparticles with asphaltenes adsorbed are extracted from the solutions by centrifugation for 15 min at 4000 rpm using a Hermle Z 306 Universal Centrifuge (Labnet, NJ) and dried overnight in a vacuum oven. Thermogravimetric analysis is performed on a fixed amount of asphaltenes that adsorbed 0.2 mg/m^2 using a Q50 Thermogravimetric Analyzer (TGA) made by TA Instruments, Inc. (New Castle, DE). The TGA analyzer is coupled to an IRAffinity-1 FTIR device (Shimadzu, Japan) that is equipped with a gas cell to analyze the outcome gasses. The samples (nanoparticles with asphaltene adsorbed and virgin nanoparticles) are heated in an air atmosphere from 30 to 1000°C at the following different heating rates: 5, 10, and 20°C/min. The air-flow was kept at a constant 100 mL/min during the experiment. It is worthwhile mentioning that the sample mass was kept low at approximately 5 mg to avoid any diffusion limitations [19, 35, 36].

2.3. Modeling

The experimental data obtained for adsorption isotherms were modeled using two commonly used adsorption models: the Langmuir model and the Freundlich model [37, 38]. Likewise, the experimental results were modeled using a novel Solid-Liquid Equilibrium model [39]. For the kinetic oxidation of the asphaltenes on nanoparticles, the Ozawa-Flyn-Wall (OFW) method was applied. This allows the reaction kinetics to be described and the effective activation energy to be estimated [40].

2.3.1. The Langmuir model

The Langmuir model has been widely used since it was published in 1916 to correlate experimental data on equilibrium adsorption [37]. In order to do this, monolayer adsorption takes place on a homogeneous surface and it was derived by taking into account the fact that equilibrium is obtained when the rates of adsorption and desorption are equivalent. The Langmuir equation can be expressed as follows:

$$N_{ads} = N_{ads,max} \left(\frac{K_L C_E}{1 + K_L C_E} \right) \tag{2}$$

where $N_{ads,max}$ (mg/m^2) is the amount of asphaltenes adsorbed onto the nanoparticles, C_E (mg/L) is the equilibrium concentration of asphaltenes in the supernatant, K_L (L/mg) is the Langmuir equilibrium adsorption constant related to the affinity of binding sites, and $N_{ads,max}$ (mg/m^2) is defined as the monolayer saturation capacity. The latter represents the maximum amount of asphaltenes adsorbed per unit of nanoparticle surface area for complete monolayer coverage.

2.3.2. The Freundlich model

The Freundlich approach was created in 1906. It is an empirical expression that represents the isothermal variation of the adsorption of a quantity of mass, which is adsorbed by unit of surface area (of solid adsorbent) at equilibrium concentration [38]:

$$N_{ads} = K_F C_E^{1/n} \tag{3}$$

where K_F ((mg/m^2)/(L/mg)$^{1/n}$)) is the Freundlich constant related to the adsorption capacity, and $1/n$ (unitless) is the adsorption intensity factor.

2.3.3. The Solid-Liquid Equilibrium (SLE) model

Recently, Montoya et al. [39] proposed a three-parameter model to describe the adsorption isotherms of asphaltene by non-porous materials. The model is based on a theoretical framework of adsorption by self-associated asphaltenes on solid surfaces. The SLE model equation is expressed as follows:

$$C_E = \frac{\psi H}{1 + K\psi} \exp\left(\frac{\psi}{q_m \cdot A}\right) \tag{4}$$

with

$$\psi = \frac{-1 + \sqrt{1 + 4K\xi}}{2K} \tag{5}$$

and

$$\xi = \left(\frac{q_m \cdot q}{q_m - q}\right) A \tag{6}$$

where q (mg/m^2) is the amount of asphaltenes adsorbed by the nanoparticle surface, q_m (mg/m^2) is the maximum adsorption capacity of the nanoparticles, A (m^2/mg) is the measured surface area per mass of nanoparticles and C_E (mg/g) is the equilibrium concentration of asphaltenes in the supernatant. K (g/g) is the reaction constant related to the asphaltenes' degree of association on the nanoparticles surface, and H (mg/g) is the Henry's law constant linked to the asphaltenes' preference for being in the liquid phase or

the adsorbed phase [39]. The *H*, *K* and q_m SLE parameters were estimated by model fitting.

For the three adsorption models, the correlation coefficient (R^2) and a nonlinear chi-square (χ^2) analysis were used to estimate the goodness of fit using the Solvent feature that is part of the Excel package [41].

2.3.4. OFW model

After TGA experiments were performed at different heating rates, the OFW method can be used to estimate the effective activation energies [40,42]. The Ozawa-Flynn-Wall method (OFW) gave us an estimate of the effective activation energies, assuming that for a constant reaction conversion, the reaction rate, eq. (7), is a function of the temperature and the state.

$$\frac{d\alpha}{dt} = K_\alpha \exp\left(-\frac{E_\alpha}{RT}\right) f(\alpha) \tag{7}$$

where K_α (1/s) is the pre-exponential factor, E_α (kJ/mol) is the effective activation energy for a constant conversion, R (J/mol·K) is the ideal gas constant, T (K) is the reaction temperature and α is the reaction conversion described by the next equation:

$$\alpha = \frac{m_0 - m_T}{m_0 - m_f} \tag{8}$$

with m_0 as the initial mass of the sample; m_f the final mass of the sample and m_T the mass at a given temperature. Eq. (9) can be obtained when the heating rate is defined as $\beta = dT/dt$ and integrating. Using the Doyle approximation [43], the effective activation energy can by estimated from the slope of the linear fit from the plot of log (β) against 1/T, according to eq. (10)[35]:

$$g(\alpha) = \int_0^\alpha \frac{d\alpha}{f(\alpha)} = \int_0^T \frac{K_\alpha \exp(-E_\alpha/RT)}{\beta} dT \tag{9}$$

$$log(\beta) = log\left(\frac{K_\alpha E_\alpha}{R g(\alpha)}\right) - 2{,}315 - 0{,}4567 \frac{E_\alpha}{RT} \tag{10}$$

3. Results and discussion

3.1. Characterization of nanoparticles

Silica nanoparticules are usually synthetized through Sol-Gel method [44,45]. As mentioned above, commercial silica nanoparticles were used. After the superficial modification process that was described above, the acidic silica nanoparticles (AS), the neutral (NS) and basic silica nanoparticles (BS) show the same particle diameter (dp); this can be seen in Table 1. It can be seen that the superficial modification apparently has no effect on the size of the nanoparticle. On the other hand, the surface area (S_{BET}) decreased when the particle became more acidic. This was determined with total acidity that is related to the amount of NH_3 adsorbed by each nanoparticle. In this case, it can be noticed that superficial treatments were effective because AS is the most acidic of three materials.

Table 1.
Nanoparticles characterization.

Nanoparticle	$d_{P\text{-}50}$ (nm)	S_{BET} (m²/g)	Total acidity (µmol/g)
AS	7	237.9	2847
NS	7	171.8	2156
BS	7	138.1	1285

Source: The authors
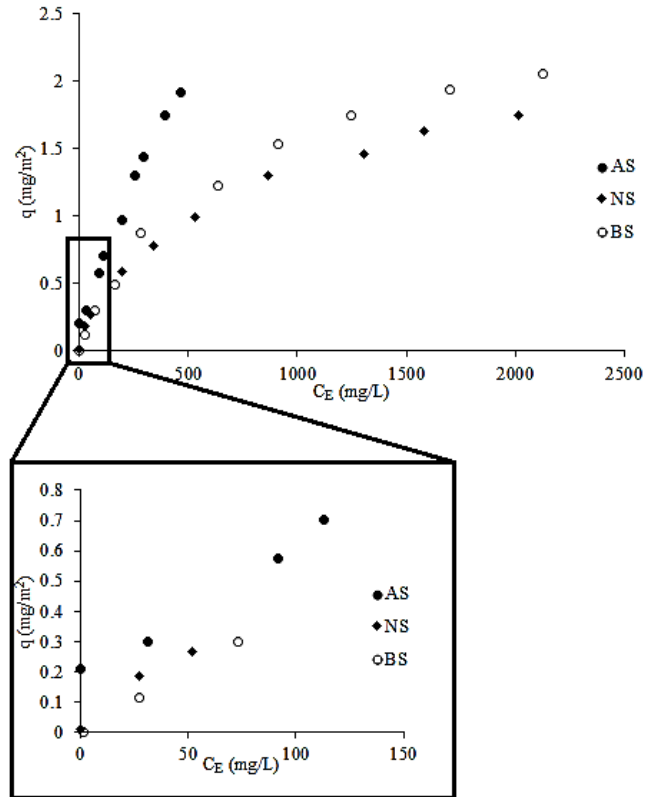
Figure 1. a) Adsorption isotherms of asphaltenes by fumed silica nanoparticles with different surface acidities at 25°C and b) zoom on Henry's region.
Source: The authors

3.2. Asphaltenes adsorption isotherms

Fig. 1a shows the obtained adsorption isotherms of asphaltenes by AS, NS and BS that were constructed at 25°C on surface area basis. It can be observed that the trend of the amount adsorbed at CE of 464 mg/L, where the isotherms can be compared, follows the order AS > BS > NS, which indicates that the surface is selective to determined functional groups that are present in the asphaltenes structure.

Also, Fig. 1b shows that in the Henry's region, the adsorption affinity is also surface specific. As the Langmuir model K_L is related to the affinity of binding sites, larger values implied more adsorption affinity [17,19,20]. In the empirical model proposed by Freundlich, this affinity cannot be seen so clearly. Hence, frequently K_F is taken as a rough indicator of the adsorption capacity and *1/n* is the adsorption intensity factor. A larger K_F value suggests greater adsorption capacity, and a lower *1/n* indicates stronger adsorption strength [17,19]. Furthermore, the reciprocal of the parameter

Table 2.
Langmuir, Freundlich and SLE model parameters of asphaltenes adsorption by AS, NS and BS at 25°C.

	Parameter	AS	NS	BS
Langmuir model	$N_{ads,max}$ (mg/m^2)	4.43	2.22	2.71
	K_L(L/mg) x 10^{-3}	1,602	1.580	1.409
	R^2	0.986	0.991	0.998
	X^2	9.858	0.209	0.027
Freundlich model	K_F ((mg/m^2)/(L/mg)$^{1/n}$))	0.025	0.047	0.043
	$1/n$ (unitless)	0.71	0.48	0.51
	R^2	0.989	0.996	0.988
	X^2	2.156	0.018	0.113
SLE model	H (mg/g)	0.15	0.25	0.28
	K (g/g) x 10^{-4}	2.45	1.75	1.45
	q_m (mg/m^2)	7.75	3.49	4.58
	R^2	0.99	0.99	0.99
	X^2	1.30	1.43	0.58

Source: The authors

H in the SLE model is a clear indicator of the adsorption affinity. A smaller H value implies greater adsorption affinity [21,39]. In this regard, we can see in Table 2 that the adjusted parameters for the Langmuir and SLE models indicate higher adsorption affinity with the most acidic nature of the surface of the nanoparticle. Also, the SLE model tells us that the trend followed by the K parameter is opposite to the one observed for the H parameter, meaning that BS nanoparticles lead to a lower degree of asphaltenes self-association over its surface. Freundlich model's parameters do not show a direct relationship with the acidic nature of the surface.

3.3. Catalytic Oxidation of Asphaltenes

The TGA experiments for asphaltenes oxidation were conducted at three different heating rates: 5, 10 and 20°C/min with an oxygen flow at 100 cm3/min. Fig. 2a-d shows the conversion for oxidation of a) virgin asphaltenes and asphaltenes in the presence of b) AS, c) NS and d) BS. For all systems, it can be observed that the percentage of conversion decreases as the heating rate increases.

To show a comparison, Fig. 3 shows the conversion of asphaltenes in the presence and absence of the evaluated nanoparticles at a fixed heating rate of 10°C/min. It is observed that for a fixed value of conversion, the temperature increases in the order of BS < NS < AS < virgin asphaltenes. This confirms the catalytic effect of the nanoparticles and shows that catalytic oxidation of the asphaltenes is specific to the adsorbent surface conditions. It can also be noted that the catalytic effect of the surface is related to the adsorption affinity and asphaltenes self-association. As lower signifies adsorption affinity (higher value of H at SLE model) and as lower signifies asphaltenes self-association (lower value of K at SLE model), higher is the catalytic behavior of the surface.

Fig. 4 shows the rate of mass loss for virgin asphaltenes oxidation and asphaltenes oxidation in the presence of AS, NS and BS. It can be seen in Fig. 4 that asphaltenes oxidation starts before for those asphaltenes adsorbed onto the nanoparticles, in comparison with the virgin asphaltenes. It is worthwhile mentioning that the virgin nanoparticles were also submitted for TGA analysis. For the virgin asphaltenes,

Figure 2. Conversion of a) virgin asphaltenes and asphaltenes in the presence of b) AS, c) NS and d) BS at different heating rates.
Source: The authors

Figure 3. Conversion for oxidation of asphaltenes in the presence and absence of AS, NS and BS nanoparticles at a heating rate of 10°C/min.
Source: The authors
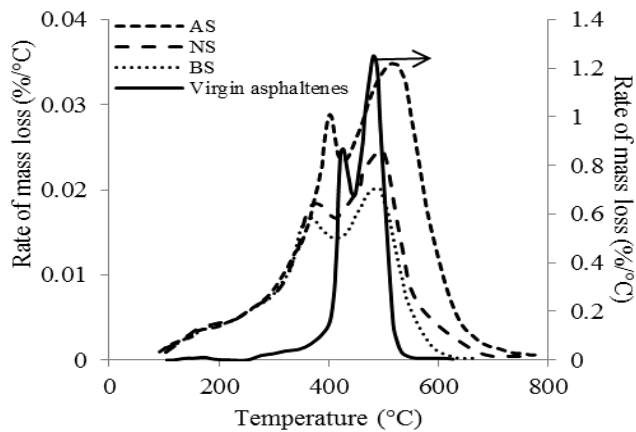
Figure 4. Rate of mass loss for oxidation of asphaltenes in presence and absence of AS, NS and BS nanoparticles at a heating rate of 10°C/min.
Source: The authors

two peaks are observed at approximately 422 and 479 °C. The first peak temperature is considerably reduced by the nanoparticles at 38, 52 and 67°C. This indicates that the catalytic activity of the BS is higher than that of the other two nanoparticles. These results agree with the estimated values of the H and K parameters of the SLE model.

The evolution of gas production was also evaluated using an FTIR device coupled with the TGA analyzer. Figs. 5a-d show the evolution of CO, CO_2, CH_4 and other light hydrocarbon production to oxidize a) virgin asphaltenes and asphaltenes in the presence of b) AS, c) NS and d) BS at a fixed heating rate of 10°C/min. It is worthwhile mentioning that the results are normalized based on the signal with the highest intensity that corresponds to the CO_2 production. In all cases, it is observed that the major production corresponds to the CO_2. Except for the system with BS, the trend of effluent production increases in the order of CO < CH_4 < Hydrocarbons < CO_2. However, for the BS, the trend followed is CH_4 < Hydrocarbons < CO < CO_2.

3.4. Estimation of activation energies

The results for the effective activation energies (EAE) are shown in Fig. 6. The asphaltenes follow an increasing trend

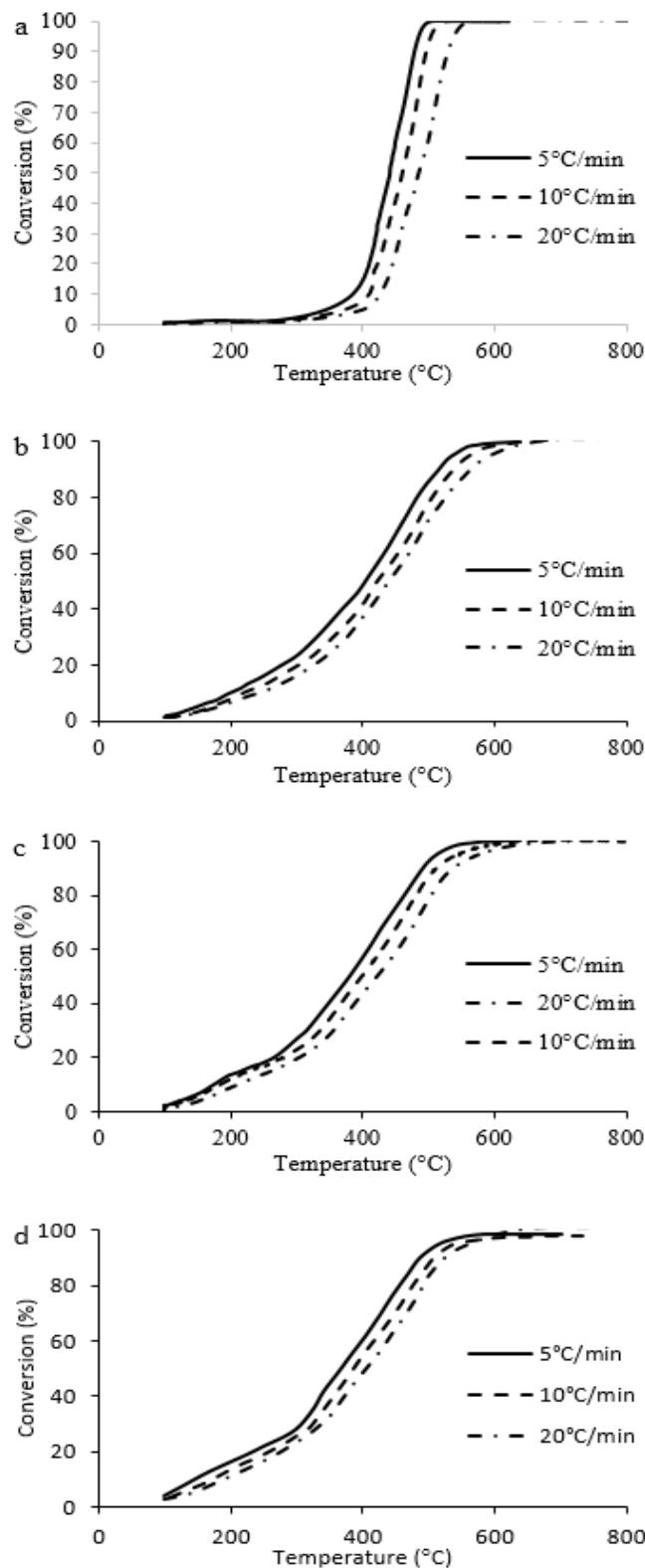Figure 5. Evolution of CO, CO2, CH4 and other hydrocarbon production for oxidation of a) virgin asphaltenes and asphaltenes in the presence of b) AS, c) NS and d) BS at a heating rate of 10°C/min.
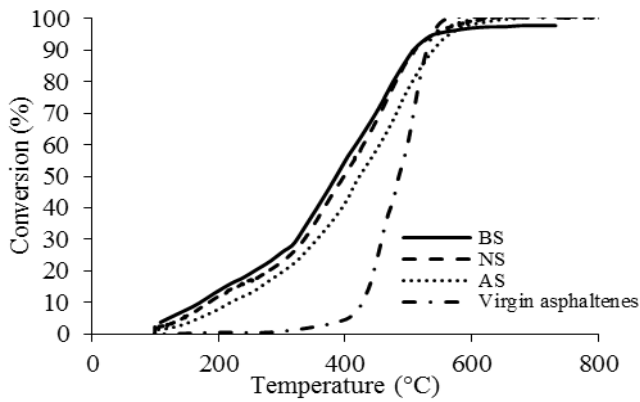Source: The authors

Figure 6. Estimated effective activation energies for asphaltenes oxidation in the presence and absence of AS, NS and BS nanoparticles.
Source: The authors

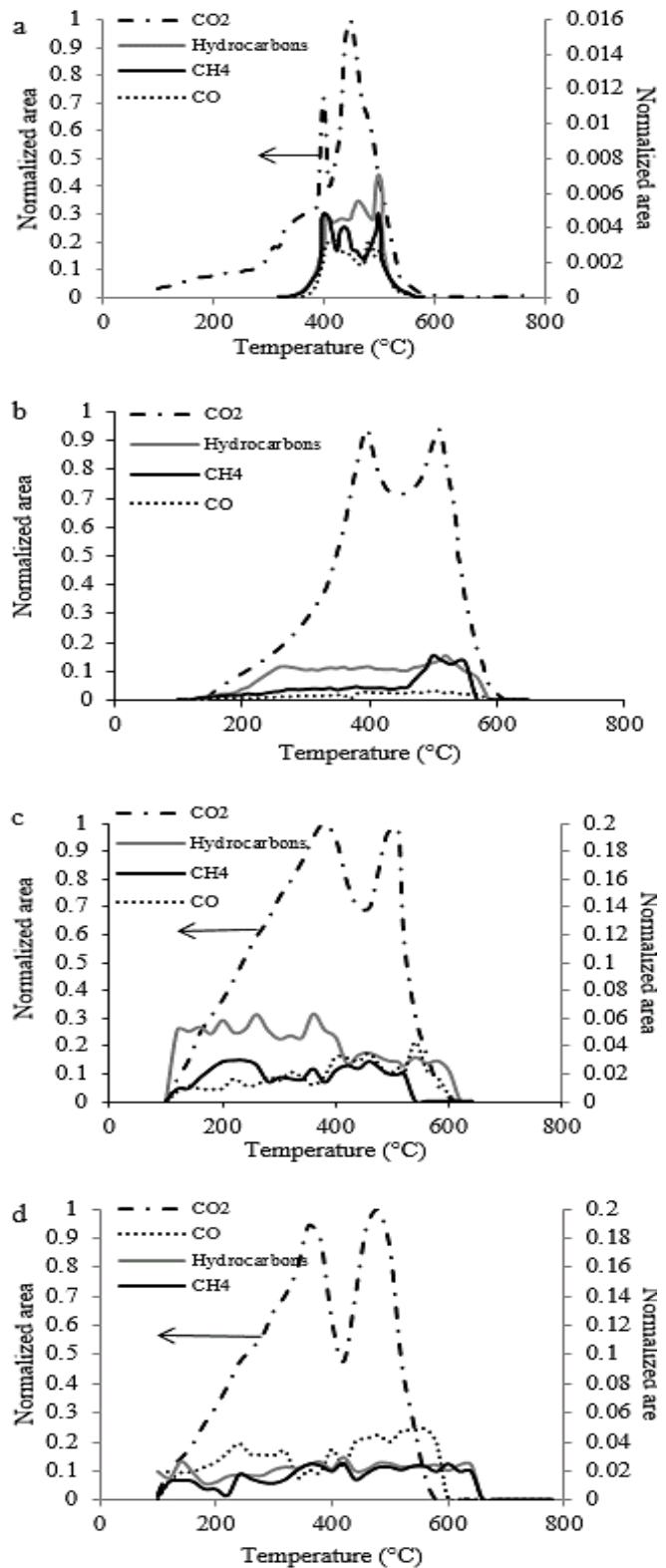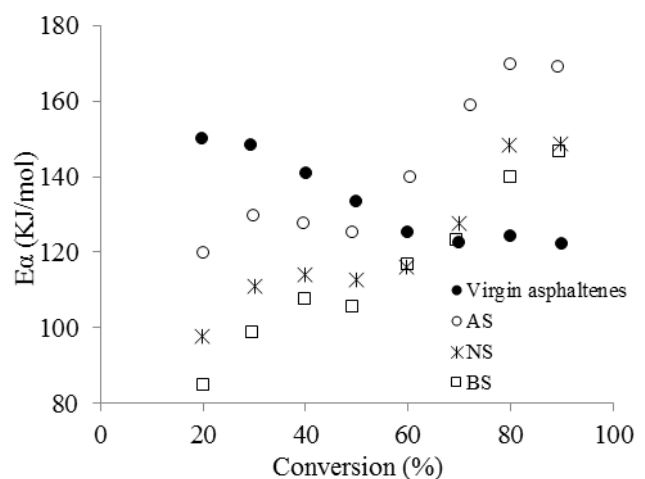in the presence of the nanoparticles, which increase as the asphaltenes conversion increases. This could be due to addition reactions that occur in the process. However, for the virgin asphaltenes, the trend observed is the opposite. Both behaviors are indicative of the fact that asphaltene oxidation in the presence and absence of the nanoparticles is not a single-step process and it involves more than one mechanism. Fig. 6 shows that higher EAE are obtained for AS, followed by NS, and the lower values are obtained for BS. Also, the pathway followed by the EAE is similar for the three nanoparticles, indicating that the reaction order is possibly similar for the three catalysts.

Fig. 7 shows the correlation of parameters H and K from the SLE model and K_L from the Langmuir model. These have the estimated EAE for the following percentages of conversion: 20, 50 and 80%. It can be observed that as the adsorption affinity increases (i.e. the H parameter decreases in the SLE model and the K_L parameter increases in the Langmuir model), the EAE increases. This indicates that an over attachment of the asphaltenes in the catalysts surface would lead to a decrease in the catalytic activity of the nanoparticles. However, the correlation observed for the K parameter indicates that as the degree of asphaltenes self-association increases the catalytic activity of the nanoparticles decreases. This is also shown in Fig. 8 where the evolution of CO is correlated with the K parameter for different degrees of conversion. It is observed that as the degree of asphaltenes self-association decreases, the production of CO increases. This could be due to bigger asphaltenes aggregates tending to hide some active sites that are released as asphaltenes are oxidized. This leads the CO to be chemisorbed over the nanoparticle surfaces to form CO_2.

4. Conclusions

The surface modification process effectively modified the acidity of nanoparticle surfaces, as can be seen in TPD experiments. Through batch adsorption experiments, the isotherms of asphaltenes adsorption onto fumed silica with different surface acidities were successfully constructed. The

Langmuir model, the Freundlich model, and the SLE model have an excellent adjustment for the experimental results of the adsorption isotherms. The first and the third of these models, allow us to conclude that the acidic surface creates a greater affinity for the asphaltenes in the adsorption process, but it leads to higher EAE and to disfavor catalytic activity. Also, if the asphaltenes self-association is higher on the nanoparticle surface, represented in parameter K of the SLE model, the catalytic activity is lower. In this sense, the BS showed a better performance in the catalytic oxidation of asphaltenes. Also, it was possible to detect the effluent gasses in the oxidation process by using a coupled FTIR device, which showed that they are surface nature dependent.
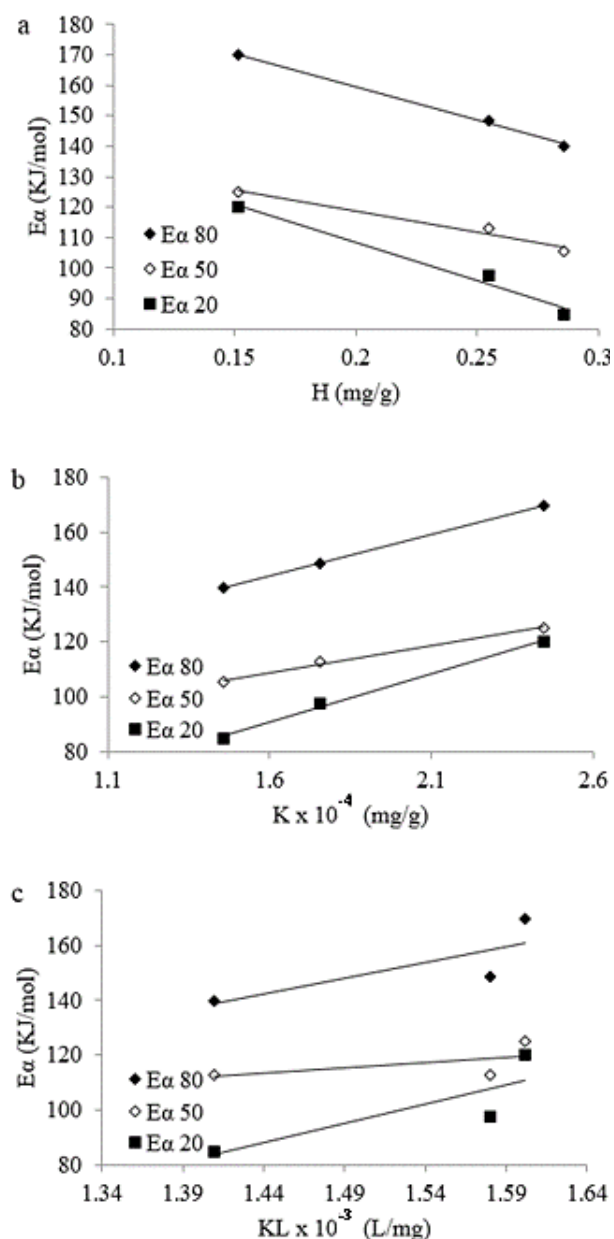

Figure 7. Correlation of the a) H and b) K SLE parameters and c) KL Langmuir parameter, with the estimated effective activation energies for percentages of conversion of 20, 50 and 80%.
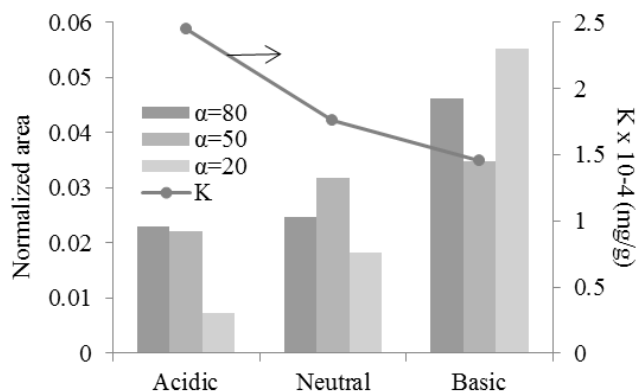Source: The authors

Figure 8. Correlation of the CO production and the K parameter of the SLE model at percentages of conversion of 20, 50 and 80%.
Source: The authors

References

[1]   OPEC, 2015 World Oil Outlook, Vienna, Austria, OPEC Secretariat, 2015, pp. 29-88.

[2]   IEA, World Energy Outlook 2012, Paris, France, International Energy Agency, 2012, pp. 81-120.

[3]   Ghanavati, M., Shojaei, M.-J. and Ramazani, A., Effects of asphaltene content and temperature on viscosity of Iranian heavy crude oil: experimental and modeling study. Energy & Fuels, 27(12), pp. 7217-7232, 2013. DOI: 10.1021/ef400776h.

[4]   Leontaritis, K., Amaefule, J. and Charles, R., A systematic approach for the prevention and treatment of formation damage caused by asphaltene deposition. SPE Production & Facilities, 9(03), pp. 157-164, 1994. DOI: 10.2118/23810-PA.

[5]   Nassar, N.N., Asphaltene adsorption onto alumina nanoparticles: kinetics and thermodynamic studies. Energy & Fuels, 24(8), pp. 4116-4122, 2010. DOI: 10.1021/ef100458g.

[6]   Fergoug, T. and Bouhadda, Y., Determination of Hassi Messaoud asphaltene aromatic structure from 1 H & 13 C NMR analysis. Fuel, 115, pp. 521-526, 2014. DOI: 10.1016/j.fuel.2013.07.055.

[7]   Acevedo, S., Castro, A., Negrin, J.G., Fernández, A., Escobar, G., Piscitelli, V., Delolme, F. and Dessalces, G., Relations between asphaltene structures and their physical and chemical properties: The rosary-type structure. Energy & fuels, 21(4), pp. 2165-2175, 2007. DOI: 10.1021/ef070089v.

[8]   Groenzin, H. and Mullins, O.C., Asphaltene molecular size and structure. The Journal of Physical Chemistry A, 103(50), pp. 11237-11245, 1999. DOI: 10.1021/jp992609w.

[9]   Mullins, O.C., The asphaltenes. Annual Review of Analytical Chemistry, 4, pp. 393-418, 2011. DOI: 10.1146/annurev-anchem-061010-113849.

[10] Mullins, O.C., Sabbah, H., Eyssautier, J.l., Pomerantz, A.E., Barré, L., Andrews, A.B., Ruiz-Morales, Y., Mostowfi, F., McFarlane, R. and Goual, L., Advances in asphaltene science and the Yen–Mullins model. Energy & Fuels, 26(7), pp. 3986-4003, 2012. DOI: 10.1021/ef300185p.

[11] Adams, J.J., Asphaltene adsorption, a literature review. Energy & Fuels, 28(5), pp. 2831-2856, 2014. DOI: 10.1021/ef500282p.

[12] Cenegy, L.M., Survey of successful world-wide asphaltene inhibitor treatments in oil production fields. in SPE Annual Technical Conference and Exhibition (2001, New Orleans, Louisiana). Society of Petroleum Engineers, New Orleans, Louisiana, 2001, P. 7. DOI: 10.2118/71542-MS.

[13] Allenson, S.J. and Walsh, M.A., A novel way to treat asphaltene deposition problems found in oil production. in SPE international symposium on oilfield chemistry (1997, Houston, Texas). Society of Petroleum Engineers, Houston, Texas, 1997, pp. 699-702. DOI: 10.2118/37286-MS.

[14] Torres, C.A., Treint, F., Alonso, C.I., Milne, A. and Lecomte, A., Asphaltenes pipeline cleanout: A horizontal challenge for coiled tubing. in SPE/ICoTA coiled tubing conference and exhibition (2005, The Woodlands, Texas). Society of Petroleum Engineers, The Woodlands, Texas, 2005, P. 19. DOI: 10.2118/93272-MS.

[15] Zabala, R., Acuna, H.M., Cortes, F., Patino, J.E., Cespedes-Chavarro, C., Mora, E., Botero, O.F. and Guarin, L., Application and evaluation of a nanofluid containing nanoparticles for asphaltenes inhibition in well CPSXL4. in OTC Brasil (2013, Rio de Janeiro, Brazil). Offshore Technology Conference, Rio de Janeiro, Brazil, 2013, P. 14. DOI: 10.4043/24310-MS.

[16] Nassar, N.N., Hassan, A. and Pereira-Almao, P., Application of nanotechnology for heavy oil upgrading: catalytic steam gasification/cracking of asphaltenes. Energy & Fuels, 25(4), pp. 1566-1570, 2011. DOI: 10.1021/ef2001772.

[17] Nassar, N.N., Hassan, A. and Pereira-Almao, P. Effect of surface acidity and basicity of aluminas on asphaltene adsorption and oxidation. Journal of colloid and interface science, 360 (1), pp 233-238, 2011. DOI: 10.1016/j.jcis.2011.04.056.

[18] Franco, C.A., Nassar, N.N. and Cortés, F.B., Removal of oil from oil-in-saltwater emulsions by adsorption onto nano-alumina functionalized with petroleum vacuum residue. Journal of colloid and interface science, 433, pp. 58-67, 2014. DOI: 10.1016/j.jcis.2014.07.011.

[19] Franco, C.A., Montoya, T., Nassar, N.N., Pereira-Almao, P. and Cortés, F.B., Adsorption and subsequent oxidation of colombian asphaltenes onto nickel and/or palladium oxide supported on fumed silica nanoparticles. Energy & Fuels, 27(12), pp. 7336-7347, 2013. DOI: 10.1021/ef4018543.

[20] Franco, C.A., Nassar, N.N., Ruiz, M.A., Pereira-Almao, P. and Cortés, F.B., Nanoparticles for inhibition of asphaltenes damage: adsorption study and displacement test on porous media. Energy & Fuels, 27(6), pp. 2899-2907, 2013. DOI: 10.1021/ef4000825.

[21] Franco, C.A., Nassar, N.N., Montoya, T., Ruíz, M.A. and Cortés, F.B., Influence of asphaltene aggregation on the adsorption and catalytic behavior of nanoparticles. Energy & Fuels, 29(3), pp. 1610-1621, 2015. DOI: 10.1021/ef502786e.

[22] Hamedi-Shokrlu, Y. and Babadagli, T., In-situ upgrading of heavy oil/bitumen during steam injection by use of metal nanoparticles: A study on in-situ catalysis and catalyst transportation. SPE Reservoir Evaluation & Engineering, 16(03), pp. 333-344, 2013. DOI: 10.2118/146661-PA.

[23] Cortés, F.B., Mejía, J.M., Ruiz, M.A., Benjumea, P. and Riffel, D.B., Sorption of asphaltenes onto nanoparticles of nickel oxide supported on nanoparticulated silica gel. Energy & Fuels, 26(3), pp. 1725-1730, 2012. DOI: 10.1021/ef201658c.

[24] Hashemi, R., Nassar, N.N. and Pereira-Almao, P., Transport behavior of multimetallic ultradispersed nanoparticles in an oil-sands-packed bed column at a high temperature and pressure. Energy & Fuels, 26(3), pp. 1645-1655, 2012. DOI: 10.1021/ef201939f.

[25] Nassar, N.N., Hassan, A. and Vitale, G., Comparing kinetics and mechanism of adsorption and thermo-oxidative decomposition of Athabasca asphaltenes onto TiO 2, ZrO 2, and CeO 2 nanoparticles. Applied Catalysis A: General, 484, pp. 161-171, 2014. DOI: 10.1016/j.apcata.2014.07.017.

[26] Nassar, N.N., Hassan, A. and Pereira-Almao, P., Effect of the particle size on asphaltene adsorption and catalytic oxidation onto alumina particles. Energy & Fuels, 25(9), pp. 3961-3965, 2011. DOI: 10.1021/ef2008387.

[27] Nassar, N.N., Hassan, A. and Pereira-Almao, P., Clarifying the catalytic role of NiO nanoparticles in the oxidation of asphaltenes. Applied Catalysis A: General, 462(463), pp. 116-120, 2013. DOI: 10.1016/j.apcata.2013.04.033.

[28] Hassan, A., Lopez-Linares, F., Nassar, N.N., Carbognani-Arambarri, L. and Pereira-Almao, P., Development of a support for a NiO catalyst for selective adsorption and post-adsorption catalytic steam gasification of thermally converted asphaltenes. Catalysis Today, 207(0), pp. 112-118, 2013. DOI: 10.1016/j.cattod.2012.05.010.

[29] Nassar, N.N., Hassan, A., Carbognani, L., Lopez-Linares, F. and Pereira-Almao, P., Iron oxide nanoparticles for rapid adsorption and enhanced catalytic oxidation of thermally cracked asphaltenes. Fuel, 95, pp. 257-262, 2012. DOI: 10.1016/j.fuel.2011.09.022.

[30] Nassar, N.N., Hassan, A. and Pereira-Almao, P., Comparative oxidation of adsorbed asphaltenes onto transition metal oxide

nanoparticles. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 384(1-3), pp. 145-149, 2011. DOI: 10.1016/j.colsurfa.2011.03.049.

[31] Nassar, N.N., Hassan, A. and Pereira-Almao, P., Metal oxide nanoparticles for asphaltene adsorption and oxidation. Energy & Fuels, 25(3), pp. 1017-1023, 2011. DOI: 10.1021/ef101230g.

[32] Brunauer, S., Emmett, P.H. and Teller, E., Adsorption of gases in multimolecular layers. Journal of the American Chemical Society, 60(2), pp. 309-319, 1938.

[33] Rouquerol, F., Rouquerol, J. and Sing, K.S.W., Adsorption by powders and porous solids; principles, Methodology and Applications, London, U.K., Academic Press, 1999.

[34] Goncalves, S., Castillo, J., Fernandez, A. and Hung, J., Absorbance and fluorescence spectroscopy on the aggregation behavior of asphaltene–toluene solutions. Fuel, 83(13), pp. 1823-1828, 2004.

[35] Nassar, N.N., Hassan, A., Luna, G. and Pereira-Almao, P., Comparative study on thermal cracking of Athabasca bitumen. Journal of thermal analysis and calorimetry, 114(2), pp. 465-472, 2013.

[36] Mateus, F.A.D., Chaves, A., Maradei, M.P., Fuentes, D.A., Guzman, A. and Picon, H.J., Kinetic analysis of the termal decomposition of colombian vacuum residua by termogravimetry. Ingeniería e Investigación, 35(3), pp. 19, 2015. DOI: 10.15446/ing.investig.v35n3.49498.

[37] Langmuir, I., The constitution and fundamental properties of solids and liquids. Part I. Solids. Journal of the American Chemical Society, 38(11), pp. 2221-2295, 1916.

[38] Freundlich, H., Over the adsorption in solution. J. Phys. Chem, 57(385), pp. 385-470, 1906.

[39] Montoya, T., Coral, D., Franco, C.A., Nassar, N.N. and Cortés, F.B., A novel solid–liquid equilibrium model for describing the adsorption of associating asphaltene molecules onto solid surfaces based on the "Chemical Theory". Energy & Fuels, 28(8), pp. 4963-4975, 2014.

[40] Ozawa, T., A new method of analyzing thermogravimetric data. Bulletin of the chemical society of Japan, 38(11), pp. 1881-1886, 1965.

[41] Montgomery, D.C. and Runger, G.C., Applied statistics and probability for engineers, John Wiley & Sons, 2010.

[42] Flynn, J.H. and Wall, L.A., A quick, direct method for the determination of activation energy from thermogravimetric data. Journal of Polymer Science Part B: Polymer Letters, 4(5), pp. 323-328, 1966.

[43] Doyle, C., Synthesis and evaluation of thermally stable polymers. II. Polymer evaluation. Appl Polym Sci, 5, pp. 285-292, 1961.

[44] Fajardo, C.A.G. y Castellanos, F.J.S., Síntesis de catalizadores de Fe-Mo soportados sobre sílice para la oxidación selectiva de metano hasta formaldehído. Ingeniería e Investigación, 29(1), pp. 53-59, 2009.

[45] Stöber, W., Fink, A. and Bohn, E., Controlled growth of monodisperse silica spheres in the micron size range. Journal of colloid and interface science, 26(1), pp. 62-69, 1968. DOI: 10.1016/0021-9797(68)90272-5.

C.A. Franco-Ariza, is a BSc. in Petroleum Engineer in 2012 and obtained a PhD degree in Energy Systems in 2015, both from the Universidad Nacional de Colombia, Medellín, Colombia. He has since dedicated his professional and research activities to nanotechnology applied to the oil & gas industry in different areas such as formation damage, heavy oil upgrading, EOR/IOR processes and wastewater remediation. Currently he is a researcher in the Chemistry and Petroleum Engineering Department, Facultad de Minas, Universidad Nacional de Colombia. Dr. Franco is the author of more than 20 published articles and presentations, and more than 10 International Conferences and Seminars.
ORCID: 0000-0002-6886-8338

J.D. Guzmán-Calle, is a BSc. an Economist in 2012 and as Petroleum Engineer in 2015 from Universidad de Antioquia and Universidad Nacional de Colombia, respectively, both in Medellín, Colombia. For the past three years he has been dedicated to research activities relating to development and application of nanotechnology in the oil and gas industry, specifically in different areas such as formation damage, heavy oil upgrading, fracking fluids, EOR/IOR processes and wastewater remediation. Currently he is pursuing a MSc degree in Petroleum Engineering, Facultad de Minas, Universidad Nacional de Colombia.
ORCID: 0000-0002-3144-7337

F.B. Cortés-Correa, is a BSc. in Chemical Engineer in 2004 from the Universidad Nacional de Colombia, Medellín, Colombia. Later, he graduated with an MSc. in Chemical Engineering and a PhD. in 2006 and 2009, respectively. Currently, he is an associate professor at the Department of Chemical and Petroleum Engineering in the Universidad Nacional de Colombia, Medellín, Colombia, and has been since 2011. He is the author of more than 40 published articles
ORCID: 0000-0003-1207-3859

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN

FACULTAD DE MINAS

Área Curricular de Ingeniería
Química e Ingeniería de Petróleos

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Química
Maestría en Ingeniería - Ingeniería de Petróleos
Doctorado en Ingeniería - Sistemas Energéticos

Mayor información:

E-mail: qcaypet_med@unal.edu.co
Teléfono: (57-4) 425 5317

Development of a simulation model as a decision support system for sugarcane supply

Claudia C. Bocanegra-Herrera [a] & Carlos Julio Vidal [b]

[a] Departamento I.O.P., Universidad Santiago de Cali, Cali, Colombia, claudia.bocanegra00@usc.edu.co
[b] Escuela de Ingeniería Industrial, Universidad del Valle, Cali, Colombia, carlos.vidal@correounivalle.edu.co

Abstract

This research presents a discrete-event simulation model of cane supply as a decision support system for managers in a sugarcane mill. The research considers harvesting, transporting, and unloading cane at the mill yard, including the time windows, inherent uncertainty in the system, and queues of all operations. The model was implemented in C++, and the *Qt Creator* development environment was used to generate a graphical interface. We found that fifty percent of the time, the trailers are waiting; we also assessed the impact of mill downtime on the unloaded cane in the mill yard. Additional environments were also evaluated: rainy period, harvesting in remote places, and an alternative configuration at the unloading area that improves the process efficiency.

Keywords: Discrete-event simulation; sugarcane supply; decision support system.

Desarrollo de un modelo de simulación como un sistema de soporte de decisiones para el abastecimiento de caña de azúcar

Resumen

Esta investigación presenta un modelo de simulación de eventos discretos del sistema de abastecimiento de caña como sistema de soporte a las decisiones para los administradores en los ingenios azucareros. Comprende las operaciones de cosecha, transporte y descarga de la caña en el patio del ingenio, además tiene en cuenta ventanas de tiempo, la incertidumbre inherente del sistema y considera las colas de todas las operaciones. El modelo fue implementado en C++ utilizando el entorno de desarrollo *Qt Creator* para generar una interfaz gráfica. Se encontró que el 50% del tiempo las tractomulas se encuentran en espera, además se evaluó el impacto de los paros de molienda en la cantidad de caña descargada. Otros ambientes fueron también evaluados: época lluviosa, cosecha en campos lejanos y una configuración alternativa en la zona de descarga de caña que mejora la eficiencia del proceso.

Palabras clave: Simulación por eventos discretos; abastecimiento de caña de azúcar; sistema de soporte a las decisiones.

1. Introduction

Sugarcane supply systems at mills are complex due to the high uncertainty in the system and the interrelationships among processes. There is a central administration of a sugarcane mill that coordinates the management and logistics of the supply operations. This coordination is achieved through communication with stakeholders using information systems. Historical average values are used for decision making and compensate for the differences created by the system uncertainty with buffers, whose sizes are based on the experience of decision makers. Uncertainties can cause delays in cane deliveries and transportation overcosts. The primary sources of uncertainty are weather, cane milling rate, design of fields, traffic on roads, and equipment damage.

Power [1] defined a decision support system (DSS) as an interactive computer-based system that helps people use computer communications, data, knowledge, and models to solve problems and make decisions. A DSS based on a discrete event simulation (DES) answers questions concerning what-if scenarios in a statistically valid manner [2] and considers uncertainty, system dynamics, and their interrelationships. Lejars et al. [3] implemented a DSS based on a simulation model that improved the transparency of

revenue for both millers and growers. A DSS has been used to improve harvesting and transportation processes [4-8]. The authors in [9] presented a web-based tool that can support decisions for biomass multi-crop production systems on the strategic, tactical, and operational levels.

Specifically for the sugarcane industry, South African and Reunion Island have developed several DESs in Arena, Excel, and a simulation tool called MAGI for decision support of cane supply management [10-15]. For a Mexican sugar plantation, researchers developed a DES to analyze the utilization of machinery and personnel [16]. The researchers in [17] applied a simulation model to reduce the number of mechanical harvesting machines, and the researchers in [18] implemented a simulation model in several sugarcane mills in Australia to decrease harvesting and transportation costs. Jannoni and Morabito [19] developed a DES in Arena focused on the reception area in a Brazilian mill to decrease the time vehicles spent in queues. For the Moroccan sugar industry, the researchers in [20] developed a simulation in SIMUL8; their results showed that the number of transportation units could be reduced by 50% while maintaining the same efficiency. The researchers in [21] developed a simulation model to analyze the impact of the freight and lead time on sugarcane supply profit in a mill. An object-oriented simulation model was built in [22] as an integral part of a DSS to advise farmers with making in-field operational decisions and machinery dimensioning. Other models exist that are different from a simulation such as DSS for sugarcane logistics; these models include algorithms for hierarchical planning of sugarcane harvests [23], a sugar cane operational planning model using mixed-integer programming and adaptive genetic algorithms [24], and a multi-choice goal programming model for sugarcane harvest scheduling, which was proposed in [25].

The structures of sugar mills in these countries are different from those in Colombia in terms of harvest time, capacity, type of equipment, and farm layout. This research describes a DSS based on a DES created with free software called COMDES (Colombian Mill's Discrete Event Simulation) for logistics supply cane of a sugar mill. The system considers time windows, cane download settings, problems at the farm, and the effect of distance, among other factors. To demonstrate the importance of this technique as a decision-making tool for configurations (tactical decisions), policies (strategic decisions), and the effect of uncontrollable variables in the system, in section 2, the logistics of sugarcane mill supply is presented. Next, in section 3, the simulation model is described. Section 4 presents the results and a discussion, and section 5 presents the main conclusions.

2. Supply logistics of sugarcane mills

The sugar cane industry in Colombia is located in the Andean valley of the Cauca River; its climatic conditions allow harvesting throughout the year. Harvesting is performed by an *alce* (set of operators and machines used to harvest) allocated to a *frente* (current field being harvested). Harvest can be done manually or mechanically and uses two types of trucks: tractor with a chaser bin in the farm and trailer-truck (TT) for road transportation. The latter uses two

train configurations (tractor unit + semitrailer + 3 trailers) and a *cabezote* (tractor unit). Fig. 1 shows a trailer-truck. Although part of the harvest and transport of sugarcane is done by outsourcing, the sugar mills have a central command of operations, where decisions are made by experts using information systems, such as SIAGRI, and communication with stakeholders using radiotelephones and cellphones. The experts use buffers to deal with randomness.

In this case, the average distance between the cane field and sugarcane mills is 25 km. The harvesting and processing of sugarcane is performed continuously every day for 24 hours.

Fig. 2 illustrates the basic operations of a sugarcane supply system from the transport unit point of view. The process between the *frente* and the sugarcane mill is as follows: "harvesting" occurs when sugarcane stalks are cut and thrown onto the tractor, and "loading" occurs when the cane in the tractor is loaded onto the train; "transporting" occurs when the train is full or when there is a shift change. The "weighing" operation describes the arrival of every train to the sugarcane mill, and the "sampling" operation describes its inspection. Then, the train goes to the "unloading" area, and the cane is delivered. In the "fuel" operation, the trailer
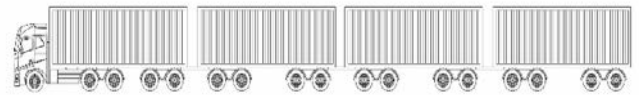

Figure 1. Trailer-truck configuration.
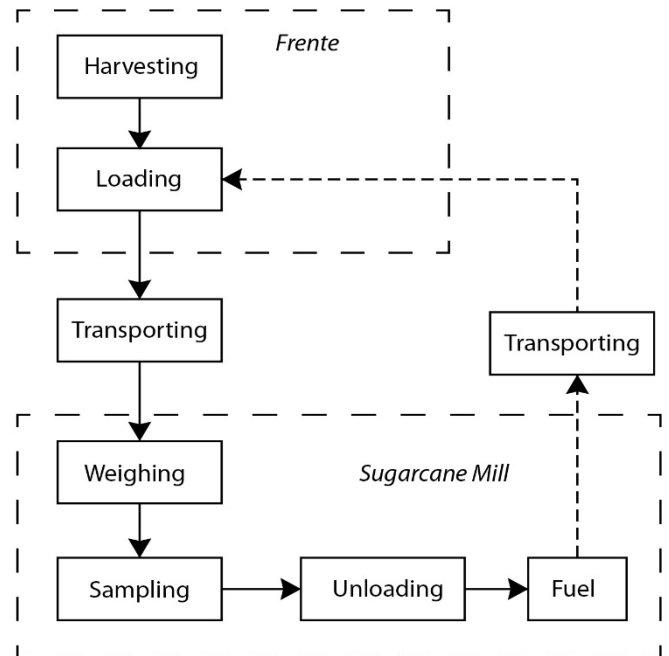Source: The authors.

Figure 2. Scheme of sugarcane supply system.
Source: The authors.

truck fills up with fuel and returns to the field as "empty transporting". The simulation model described in the next section highlights the queues, related random rates, and service times at each stage.

3. Simulation of sugarcane supply system

3.1. Model building and assumptions

The model was built in three steps: conceptual model, implementation of the simulation, and validation.

The model conceptualization defines the scope and level of detail; the main assumption is that the only scarce resources are *cabezotes* and mechanical harvestings. The rates and service times exhibit a random behavior that can be described by a statistical distribution. Figs. 3 and 4 show the flowchart for events in the sugarcane mill and the *frentes* that occur within a shift. However, the simulation also takes into account shift changes and equipment setups.

The model was implemented in C++ as an object-oriented model and used the Qt Creator 2.72 development environment to generate the graphical interface. The model was structured into 25 events that represent each of the outcomes necessary to describe the actual system; each event has its own parameters

Figure 3. Flowchart of sugarcane mill's events.
Source: The authors.

Figure 4. Flowchart of *frentes'* events.
Source: The authors.

and variables. To quantify the parameters, historical sugar mill data found in literature [26] and data collected through interviews with staff from different mills were used. Most data are given as random distributions to imitate the actual behavior.

The system components are accessed through a menu structure that is easy to use (see Fig. 5), which allows data manipulation, such as the number of tractors in service, number of *alces*, distributions rates, and working speeds; the simulation output is displayed and can be exported to a spreadsheet format. The simulated system is nonterminating, i.e., it reaches steady state using minutes as the unit of time. For verification and validation, both the conceptual model and the translation to the computer were made from simple to complex.

Throughout the translation process, the proper behavior and its resemblance to the real system were constantly checked using ASSERTs, message lists, visual presentation, and a response variables list. Fig. 6 shows the input values and the performance measures of the simulation.

3.2. Decision rules

Several routines were made inside events, which considered the inherent uncertainty of the system; the following highlights the most important ones:

- Trains to assign to an *alce* per shift: This routine calculates the minimum number of trains required according to eq. (1)-(5). The allocation of tractor-trailers to an *alce* is done at the beginning of the shift in the zone of *cabezotes* that represent the driver change; there are two rules: begin shift, by which trailer-trucks in the zone of *cabezotes* are allocated to every *alce* and ending a trip, by which the current *alce* is no longer required; then, the train can be allocated to another *alce*.
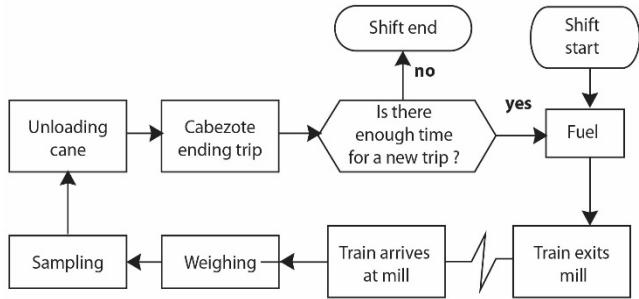
$$\textbf{Cane to haverst per shift} \left[\tfrac{ton}{frente \cdot shift}\right] =$$
$$Min(Total\ Cane\ per\ shift, Alce\ Capacity, Milling\ Capacity) \quad (1)$$
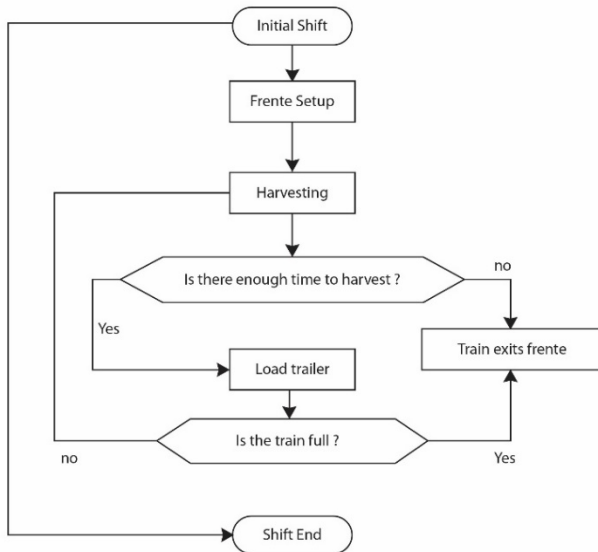
$$\textbf{Number of trips per shift} \left[\tfrac{train}{shift}\right] =$$
$$round.up\left(\tfrac{Cane\ to\ harvest\ per\ shift}{Train\ Capacity}\right) \quad (2)$$

$$\textbf{Cycle time}\ [h] = \tfrac{Distance\ mill\ to\ frente}{Speed\ of\ empty\ train} + \tfrac{Distance\ frente\ to\ mill}{Speed\ of\ full\ train}$$
$$+ Time\ in\ mill + \ Time\ in\ frente \quad (3)$$

$$\textbf{Number of a train trips per shift} \left[\tfrac{1}{Shift}\right] =$$
$$round.down\left(\tfrac{Work\ hours\ per\ shift}{Cycle\ time}\right) \quad (4)$$

$$\textbf{Number of Trains}\ [train] =$$
$$round.up\left(\tfrac{Number\ of\ trips\ per\ shift}{Number\ of\ a\ train\ trips\ per\ shift}\right) \quad (5)$$

- Working time for transporting: This is used to decide when a train that is unloading cane can make another trip in the same shift:

if current time + cycle time < ending time of shift *then*
Return to *frente*
Else
End shift for train
End if

Figure 5. Collage of snapshot of COMDES.
Source: The authors.

Figure 6. General model with inputs and outputs.
Source: The authors.

- Time between successive trains: When a shift starts, tractor-trailers are allocated to every *alce*. Immediately, a train is directed to the field, and then, the other tractor-trailers leave with a departure time described by eq. (6).

$$Departure\ time = Tractortrailer\ in\ field[min] = \frac{Maximun\ capacity\ of\ train}{Harvest\ Rate * Number\ of\ harvestings} \quad (6)$$

- Mill shutdown: When the amount of cane is not enough, the mill is interrupted and restarts when there is a certain amount of sugarcane in the unloading area.

4. Results and discussion

4.1. Current system

The maximum capacity of milling considered was 10,000 t of sugarcane per day with two tandems of mills, 11 alces, and 34 trailer-trucks. In the yard, trains must wait until the cane is unloaded in the mills to continue the route. The simulation warm-up period is the time taken to reach steady state, which was calculated by Welch's method. The system rapidly reaches steady state in 48 hours, and the shift change subsequently forces the transportation and alces to return to the initial state.

The experiments were executed with 12 runs (each run with a different random seed), and each run had 120 batches of 12 h, which results in a total of 1,440 simulation results for every trial and scenario. Table 1 summarizes the results; there is a low variability in the measures except for shutdown time and the number of shutdowns because in a few shifts, the mill breaks down randomly, which is due to the occasional downtime of the mill caused by a random delay in sugarcane delivery.

The maximum amount of cane that can be delivered is approximately 5,520 t per shift. There is a 1:1 ratio between the transport time and the time in queue; in other words, a tractor-trailer spends half its time waiting. It is important to focus on decreasing downtime waiting to improve transport eff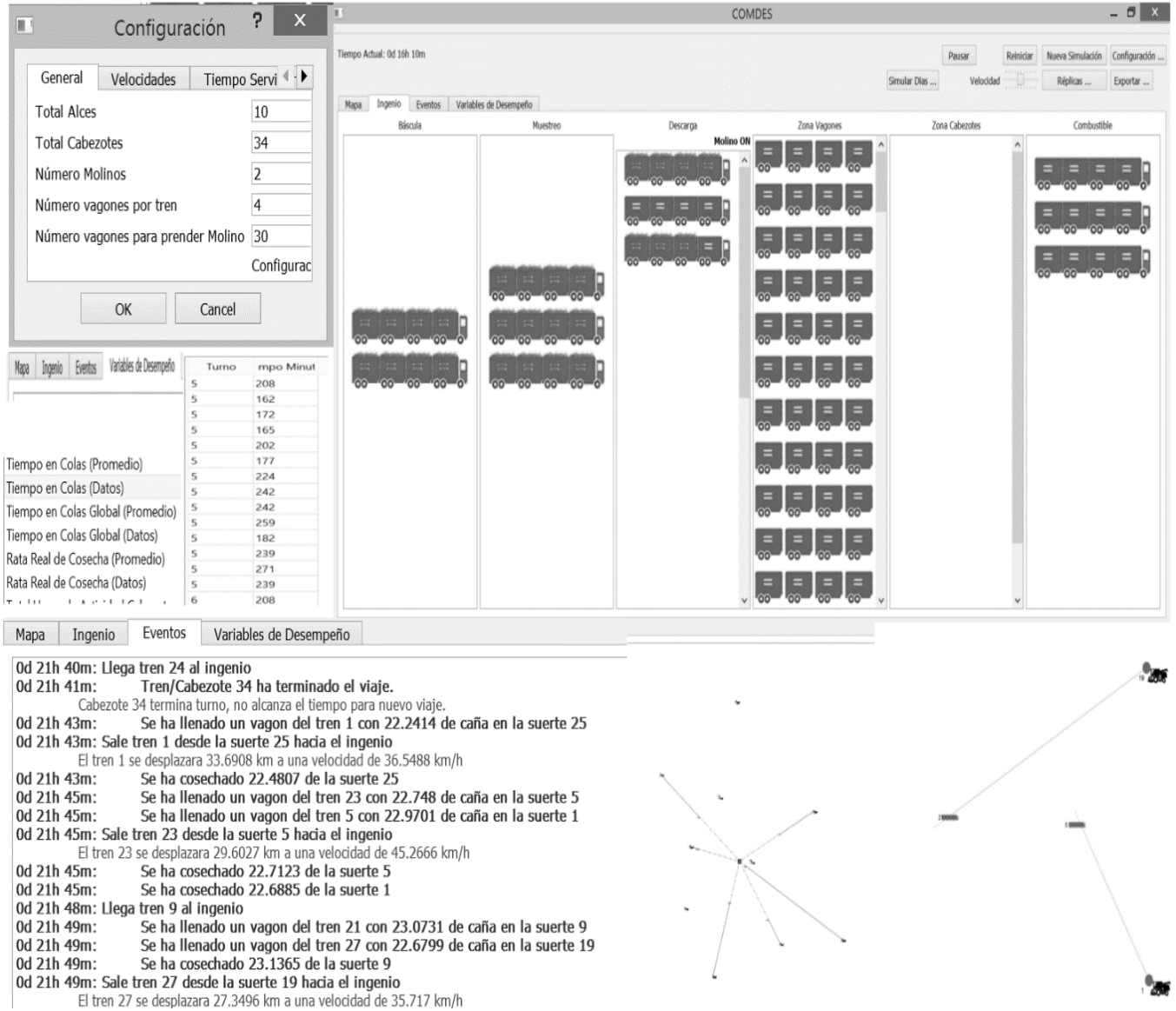iciency. Furthermore, because of the need for trailer-trucks drivers to be at the starting place before the shift change, we observed that transportation cannot be fully used at all times.

In this research, we suggest four scenarios to demonstrate the impact of the use of simulation techniques by sugarcane mill managers for analysis and decision making in the supply system.

Scenario 1. *Environmental*: Rainy period. This scenario considers the random (uncontrollable) factors in operations during rainy periods, recreates the decreased train speed due to wet roads and the decrease in the harvesting rate due to the risk of soil compaction.

Table 1.
Summary of results of original scenario.

Measure	Units	Expected Value ($\overline{X}$)	Coefficient of Variation (CV)
Cane unloading	t / shift	5,509	0.014
Shutdown time	min	1.504	6.380
Numbers of shutdowns		0.024	6.355
Average time in queue	min	168.0	0.032
Average cycle time	min	335.4	0.017
Max queue in *frente*		2.923	0.092

Source: The authors.

4.2. Analysis of scenarios

Scenario 2. *Policy*: All *frentes* are located away from the sugarcane mill. In many occasions, due to the need to achieve daily production targets and the lack of strategic and tactical planning, sugarcane mills with mature cane are at distances far away.

Scenario 3. *Setting:* Change of layout in the unloading area. This situation considers that all trailer-trucks release their cane-filled trailers by unloading cane, and the empty trailers return to the trailers area to continue travel. The loaded trailers will be pulled by a tractor.

Table 2 summarizes the findings. In Scenario 1, it was necessary to adjust a single mill in use, which is a common practice in rainy periods. The amount of unloaded cane decreases by 57%; however, the milling operation is never shut down at any time. The percentage of trains waiting is 65%, which increases the cycle time because of the slower load capacity, which reduces the amount of performed and results in slower milling. Due to the weather, it is important to design new plans for operational decisions to address new *frentes*, which is a task that is easier done using COMDES by constructing a hierarchical plan.

In Scenario 2, the amount of unloaded cane decreases by 15% due to the increase in the distances of the *alce* with respect to the sugarcane mill; also, the cycle time increases to 46 minutes, and thus, more tractor trailers are required. In this case, a mill shutdown always occurs per shift, and the expected value of shutdown time increases to 110 minutes. The layout of Scenario 3 has a better performance because it increases the amount of unloaded cane and decreases the time in queues (36% from the queue time). Furthermore, the queue time decreases, the travel time is reduced, and the mill does not stop; therefore, the amount of unloaded cane increases by 0.2% and reaches maximum capacity. An economic and financial evaluation is required to assess if less tractor-units and drivers compensate for the increase in the number of trailers and tractors.

Table 2.
Summary of performance operations of the scenarios.

Measure	Units	Scenario 1		Scenario 2		Scenario 3	
		$\overline{X}$	CV	$\overline{X}$	CV	$\overline{X}$	CV
Cane unloading	t/shift	2,385	0.005	4,663	0.012	5,520	0.001
Shutdown time	minutes	0.000	-	111.7	0.066	0.000	-
Numbers of shutdowns		0.000	-	1.000	0.000	0.000	-
Average time in queue	minutes	529.1	0.010	180.4	0.020	80.80	0.062
Average cycle time	minutes	819.5	0.006	380.9	0.009	223.0	0.023
Max queue in *frente*		2.000	0.000	2.138	0.162	2.930	0.087

Source: The authors.

5. Conclusions

When making decisions, one of the primary problems of sugarcane mills in Colombia is not considering the effect of process capability given time windows, uncertainty, and the interactions of system elements. Moreover, due to the large number of variables that change over time and their unclear interactions, it is important to use a DSS, such as COMDES, which uses discrete-event simulations and is adapted to the features of Colombian sugarcane mills. We highlight the fact that the DSS was created using free software and allowed the realization of different types of analyses, both visible and transparent, and improved operational performance of the sugarcane supply chain. COMDES can help make decisions, analyze decisions made by managers, and measure the consequences in a matter of seconds.

Scenario 1 showed the results of difficult environmental conditions by rainy weather (random factors). Scenario 2 presented the effects of a bad policy, in which *frentes* are located far away from the mill. Scenario 3 showed the impact of a new setting in the yard; however, economic and financial evaluations are required to complete the analysis.

Although the simulation does not use an optimizer, its internal algorithm explores different factors and can help determine near optimal system configurations under the considered performance measures. As future work, it is important to use business software for decision making that can be adapted to the simulation with real time data. This model can be used to study the effect of disturbances, such as transportation breakdowns (terminal simulation) and thus, is recommended for future studies.

Acknowledgements

References

[1] Power, D.J., Decision support systems: Concepts and resources for managers, Westport: Greenwood publishing Group, 2002.

[2] Cochran, J.K. and Chen, M.S., An integrated multicomputer DSS design for transport planning using embedded computer simulation and database tool, Decision Support Systems, 7, pp. 87-97, 1991.

[3] Lejars, C., Auzoux, S., Siegmund B. and Letourmy P., Implementing sugarcane quality-based payment systems using a decision support system, Computers and Electronics in Agriculture, 70, pp. 225-233, 2010.

[4] Nilsson, D., Dynamic simulation of straw harvesting systems: Influence of climatic, geographical and biological factors on performance and costs, Journal of Agricultural Engineering Research, 76, pp. 27-36, 2000. DOI: 10.1006/jaer.1999.0456

[5] Sokhansanj, S., Kumar, A. and Turhollow, A.F., Development and implementation of integrated biomass supply analysis and logistics model (IBSAL), Biomass and Bioenergy 30, pp. 838-847, 2006. DOI: 10.1016/j.biombioe.2006.04.004

[6] Kumar, A. and Sokhansanj, S., Switchgrass (Panicum vigratum, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model, Bioresource Technology, 98, pp. 1033-1044, 2007. DOI: 10.1016/j.biortech.2006.04.027

[7] Parthanadee, P. and Buddhakulsomsiri, J., Simulation modeling and analysis for production scheduling using real-time dispatching rules: A case study in canned fruit industry, Computers and Electronics in Agriculture, 70, pp. 245-255, 2010. DOI: 10.1016/j.compag.2009.11.002

[8] Mobini, M., Sowlati, T. and Sokhansanj, S., Forest biomass supply logistics for a power plant using the discrete-event simulation approach, Applied Energy, 88, pp. 1241-1250, 2011. DOI: 10.1016/j.apenergy.2010.10.016

[9] Busato, P. and Berruto, R., A based-web tool for biomass production systems, Biosystem Engineering Special Issue: Operations Management, 120, pp. 102-116, 2013. DOI: 10.1016/j.biosystemseng.2013.09.002

[10] Barnes, A., Meyer, E. and Schmidt, E., Evaluation of methods to reduce harvest-to-crush delays using a simulation model, Proc. S. Afr. Sug. Technol. Ass., 74, pp. 25-28, 2000.

[11] Guilleman, E., Le Gal, P.Y., Meyer, E. and Schmidt, E., Assessing the potential for improving mill area profitability by modifying cane supply and harvest scheduling: A South African study, Proc. S. Afr. Sug. Technol. Ass., 7, pp. 566-579, 2003.

[12] Le Gal, P.Y., Lejars C. and Auzoux, S., Magi: A simulation tool to address cane supply chain management. Proceedings of the South African Sugar Technologists' Association. Durban, South African, 77, 200, pp. 553-565, 2003.

[13] Lejars, C., Le Gal, P.Y. and Auzoux, S., A decision support approach for cane supply management within a sugar mill area, Computers and Electronics in Agriculture, 60, pp. 239-249, 2008. DOI: 10.1016/j.compag.2007.08.008

[14] Le Gal, P.Y., Lyne, P.W.L., Meyer E. and Soler, L.G., Impact of sugarcane supply scheduling on mill sugar production: A South African case study, Agricultural Systems, 96, pp. 64-74, 2008. DOI: 10.1016/j.agsy.2007.05.006

[15] Le Gal, P.Y., Le Masson, J., Bezuidenhout, C.N. and Lagrange, L.F., Coupled modelling of sugarcane supply planning and logistics as a management tool, Computer and Electronics in Agriculture, 68, pp. 168-177, 2009. DOI: 10.1016/j.compag.2009.05.006

[16] Arjona, E., Perez, V., Bueno, G. and Salazar, L., An interactive simulation environment for evaluating managerial decisions in a sugarcane plantation in Mexico, ARPN Journal of Agricultural and Biological Science, 9, pp. 24-37, 2014.

[17] Higgins, A., Antony, G., Sandell, G., Davies, I., Prestwidge, D. and Andrew, B., A framework for integrating a complex harvesting and transport system for sugar production, Agricultural Systems, 82, pp. 99-115, 2004. DOI: 10.1016/j.agsy.2003.12.004

[18] Higgins, A. and Davies, I., A simulation model for capacity planning in sugarcane transport, Computers and Electronics in Agriculture, 47, pp. 85-102, 2005. DOI: 10.1016/j.compag.2004.10.006

[19] Iannoni, A.P. and Morabito, R., A discrete simulation analysis of a logistics supply system, Transportation Research Part E: Logistics and Transportation Review, 42, pp. 191-210, 2006. DOI: 10.1016/j.tre.2004.10.002

[20] Jorio, R., Legendre, B., Gautz, L. and Abdellaouir, R., Incorporation of sugarcane harvesting and transport variables into a sugarcane harvest and transport scheduling model. Proc. S. Afr. Sug. Technol. Ass., 80, pp. 71-85, 2006.

[21] Rangel, A.J.J., Prado, A.C. and Rangel, L.R., A simulation model to evaluate sugarcane supply systems. Proceedings of the 2010 Winter Simulation Conference, Baltimore, USA, pp. 2114-2125, 2010. DOI: 10.1109/WSC.2010.5678861

[22] Hameed, I.A., Bochtis, D.D., Sørensen, C.G. and Vougioukas, S., An object-oriented model for simulating agricultural in-field machinery activities, Computer and Electronics in Agriculture, 81, pp. 24-32, 2012.

[23] Tischer, I. y Carrion, A., La planificación jerárquica y su aplicación a la cosecha de la caña de azúcar en Colombia, Ingeniería y Competitividad, 4, pp. 42-52, 2003.

[24] Neungmatcha, W., Sethanan, K., Gen, M. and Theerakulpisut, S., Adaptive genetic algorithm for solving sugarcane loading stations with multi-facility services problem, Computers and Electronics in Agriculture, 98, pp. 85-99, 2013. DOI: 10.1016/j.compag.2013.07.016

[25] Da Silva, A.F., Silva, F.A. and Dias, E.X., Addressing uncertainty in sugarcane harvest planning through a revise multi-choice goal

programming model, Applied mathematical modelling, 39(18), pp. 5540-5558, 2015. DOI: 10.1016/j.apm.2015.01.007

[26] Amú, L.G., Modelo de simulación y optimización para la gestión logística del sistema de abastecimiento de caña en un ingenio sucro-alcoholero colombiano, Tesis de Grado, Departamento de Ingeniería Industrial, Universidad del Valle, Cali, Colombia, 2011.

C.C. Bocanegra-Herrera, received her BSc. Eng in Chemical Engineering in 2009 with a Sp. in Logistics in 2011 and an MSc. in Engineering in 2014, all of them from the Universidad del Valle, Cali, Colombia. She currently works at the Universidad Santiago de Cali in Cali, Colombia. Her research interests include the following: Discrete-event simulation; supply chain optimization; and lean manufacturing.
ORCID: 0000-0002-8593-8879

C.J. Vidal received his BSc. Eng in Mechanical Engineering in 1983 and his MSc. in Industrial and Systems Engineering in 1990, both from the Universidad del Valle, Cali, Colombia. Afterwards, he traveled to the U.S. and received an MSc. and a PhD. in Industrial Engineering in 1998, both from the Georgia Institute of Technology, Atlanta, Georgia, USA. He is currently a professor at the School of Industrial Engineering at the Universidad del Valle. His research interests include the following: Applied optimization and operations research; logistics; supply chain optimization; and inventory systems.
ORCID: 0000-0002-4774-9591

Algorithm for detection of overlapped red blood cells in microscopic images of blood smears

Miguel Fabián Romero-Rondón [a], Laura Melissa Sanabria-Rosas [a], Lola Xiomara Bautista-Rozo [a]
& Alfonso Mendoza-Castellanos [b]

[a] Escuela de Ingeniería de Sistemas e Informática, Universidad Industrial de Santander, Colombia. miguelfromeror@gmail.com, melissasanabriarosas@gmail.com, lxbautis@uis.edu.co
[b]Biosys Ltda, Colombia. almecas99@yahoo.com

Abstract

The hemogram is one of the most requested medical tests as it presents details about the three cell series in the blood: red series, white series and platelet series. To make some diagnostics, the specialist must undertake the test manually, observing the blood cells under the microscope, which implies a great physical effort. In order to facilitate this work, different digital image processing techniques to detect and classify red blood cells have been proposed. However, a common problem is the presence of overlapped cells, which generate various flaws in the analysis. Therefore, the implementation of an algorithm to address the problem of red blood cells overlapped in cellular smear images is proposed in order to support the clinician in the visual reading process. The method was tested with 50 images in which the indices of sensitivity and specificity were calculated, and the effectiveness of the algorithm developed was shown.

Keywords: digital image processing; hematology; Hough; k-means; overlap; red blood cells; watershed.

Algoritmo para la detección de glóbulos rojos superpuestos en imágenes microscópicas de extendidos de sangre periférica

Resumen

El hemograma es uno de los exámenes médicos más solicitados, ya que ofrece información detallada sobre las tres líneas celulares presentes en la sangre: la serie roja, la blanca y la plaquetaria. Para emitir algunos diagnósticos el especialista debe hacerlo de forma manual, observando en el microscopio las células sanguíneas, lo que implica mayor esfuerzo. Con el propósito de facilitar este trabajo, se han propuesto diferentes técnicas de procesamiento digital de imágenes para la detección y clasificación de glóbulos rojos, pero se ha encontrado un problema muy común que es la presencia de células superpuestas, lo cual genera diversos errores en los resultados. Por esta razón, se propone la implementación de un algoritmo que permita abordar el problema de superposición de glóbulos rojos en imágenes de frotis celular, con el fin de dar soporte al especialista en el proceso de lectura visual. El método fue probado con 50 imágenes, con las cuales se calcularon los índices de especificidad y sensibilidad, mostrando la efectividad del algoritmo desarrollado.

Palabras clave: Procesamiento digital de imágenes; hematología; hough; k-means; superposición; glóbulos rojos; watershed.

1. Introduction

Blood is composed of a liquid part known as plasma and formed elements such as red blood cells (RBC), white blood cells (WBC) and platelets [1]. In order for it to be analyzed, a study called a hemogram is undertaken, which counts the number of figurative elements in a certain volume of blood. It also allows specialists to confirm or assume various diseases according to the alterations presented in the count [2].

Currently, this test is performed automatically using specialized equipment and calibrated according to the ranges specified by the manufacturer [3,4]. When the counting is outside

this range, the specialists analyze the state of maturation, staining characteristics, the content of granules, inclusions and cellular forms. The observation of cells from a blood smear using a microscope then becomes necessary, which is a tedious process for the individual [5].

In order to facilitate this work and make it more efficient a wide range of algorithms in digital image processing have been developed to capture, represent, analyze and classify objects in microscopic images. Moreover, in automation tasks such as segmentation or classification of elements in an image, it has been found that a common problem is the presence of overlapped cells [6].

Different methods to address the problem of overlapped cells have been proposed in several studies. In [7-12], watershed transformation is used in which over-segmentation problems occur; however, better results can be obtained if sign marking is employed. Therefore, most researchers recommend performing suitable image preprocessing or pre-selection of objects in order to appropriately choose the benchmarks.

In other studies [13-17], separation was performed using techniques based on the objects edge. These take into account its concavity or convexity in order to find the points of maximum curvature and then join them by using a triangulation technique. The errors in these methods can be observed when the cells have an irregular edge.

One way to do the groups separation is to use the morphological operation of erosion, as shown in [18-21]; however, this method is only effective when the convexity of the groups is deep enough. Some alternatives to solve this problem is to erode certain parts of the cell, but this is very limited due to the sensitivity of the parameters, and the computational cost is also high.

Another alternative is proposed by [22-26], who use the Hough transform to detect circular shapes and thus segment images of RBC in peripheral blood smears. This takes advantage of circular shape of erythrocytes.

This article describes the design of an algorithm that uses morphological operations, the watershed method, the Hough transform and the clustering method of k-means to detect overlapped RBC. This will, in turn, smooth out the curve of their edge with Bézier curves. Finally, we will present the specificity and sensitivity rates found in test scores from a group of 50 microscopic images of blood cells.

2. Algorithm design

The proposed algorithm consists of three stages, as shown in Fig. 1. The first stage is the pre-processing stage, in which the removal of white blood cells and platelets is undertaken.

The second step is segmentation, in which the separation of groups of cells to subsequently detect the centers of each is made. Finally, there is a post-processing stage, in which the edges of all RBC appearing in the image are plotted.

2.1. Preprocessing

2.1.1. White blood cell removal

In order to highlight the white blood cells, image preprocessing is performed by converting the image from an
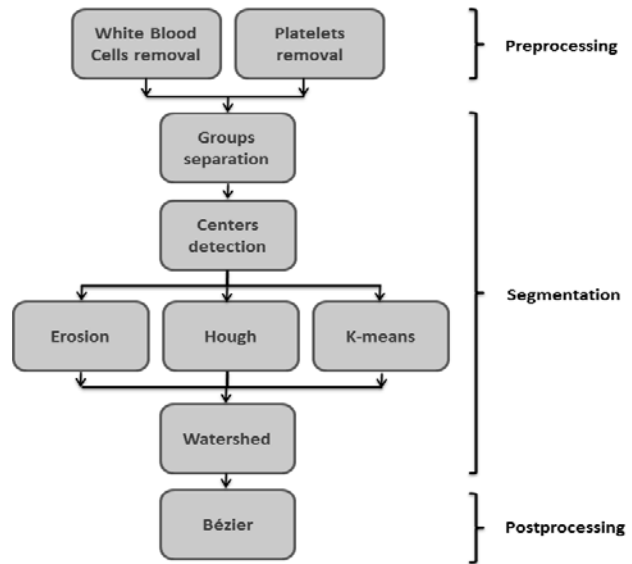
Figure 1. Flow diagram of the proposed algorithm.
Source: The authors.

RGB color model to an HSV model. The S channel is chosen because it has a higher contrast, which means that platelets and white blood cell nuclei can be clearly distinguished, as is shown in Fig. 2 (b). To expedite the process the image is binarized using the Otsu's method [27] in order to create a black and white image containing the nuclei of white blood cells and platelets (see Fig. 2 (c)).

As can be observed, some noise and the WBC cytoplasm still remain on the image. Therefore, the small objects are removed using a morphological opening; the result is shown in Fig. 2 (d). Subsequently, the image is dilated with a disc-shaped structuring element in order to remove part of the WBC cytoplasm. This result can be seen in Fig. 2 (e).

Finally, white blood cell removal on the original image is shown in Fig. 2 (f).

2.1.2. Platelets removal

As seen in the previous section, the extraction of channel S and binarization also make it possible to distinguish platelets. Therefore these steps are repeated to remove them. Subsequently, a subtraction between the binarized image of S channel and binarization of the original image is performed.

Despite this process, some noise and platelets still remain, and as such, objects with a small area are removed by employing a morphological opening operation with a disc-shaped structuring element.

2.2. Segmentation

2.2.1. Group Segmentations

In order to improve the image contrast, a background template is obtained using a morphological opening over the entire image. It is then subtracted from the grayscale image.
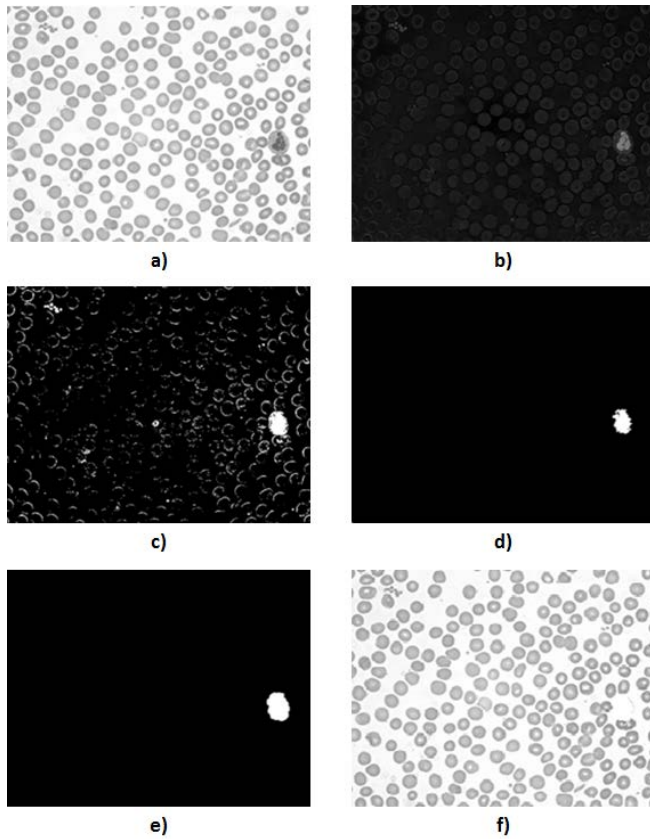
Figure 2. (a) Original image. (b) S channel of original image. (c) S channel Binarization. (d) Removal of small objects. (e) Morphologic dilation. (f) Result of white blood cell removal.
Source: The authors.

If the full image after background removal is binarized (objects take the value 1 and background 0), the black and white image presented in Fig. 3 (b) is obtained. As a result of this binarization, holes can be seen inside the cells due to lighting and the concave shape of RBC. Thus, these dark spaces are filled using a technique that employs morphological dilation to close the RBC centers. The result of this operation can be seen in Fig. 3 (c).

The noise that is still present in Fig. 3 (c) is removed by using a morphological opening with a disc-shaped structuring element. A morphological closing is then made on the image to close the RBC that were opened in the binarization step due to illumination on its surface; the result is presented in Fig. 3 (d).

Considering that the average circular area of a RBC is 35 micrometers, a sketch can be plotted for the distribution curves of the areas for the individual. There will be two overlapped and three or more overlapped RBC, as shown in Fig. 4.

Misclassification can occur at the intersection of probability distribution curves. To overcome these errors, the following measures are ensured:

- In the detection of a single cell or two overlapped cells, the convex hull is used. If it is equal to the cell edge, the connected component is classified as a single cell; otherwise it is classified as a group of two cells.

- In the classification of groups of two or three cells, the Hough transform is used to detect the number of circles present. If the result differs with minimal difference in the number of cells detected according to the area, Hough determines the amount taken to be true. But, if the difference is greater than 1, the number of cells defined by the area is maintained. This procedure is also useful to classify groups of more than three cells.

Based on this idea, we can classify according to the area of the connected components found in the image in order to gain an understanding of the amount of cells that could be in each one. The result of this classification is shown in Fig. 5; individual cells, groups of two cells, and groups of three or more cells are highlighted in the image.

2.2.2. Centers detection

The watershed transformation is an image segmentation tool that is based on mathematical morphology. The image can be considered to be a topographic relief where the gray
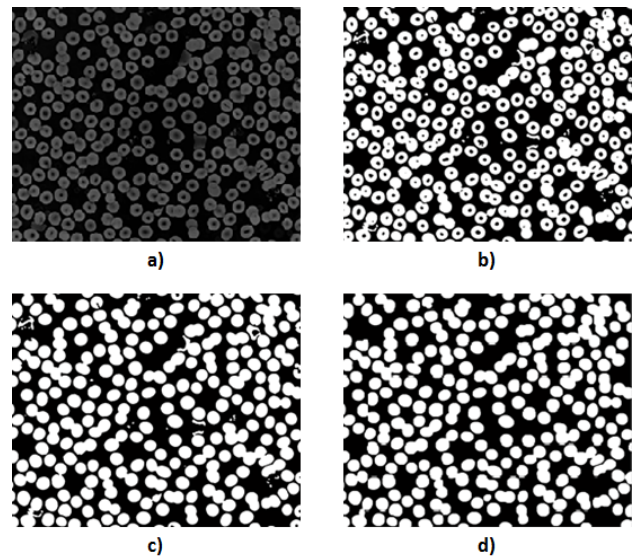
Figure 3. (a) Background removal. (b) Black and white image. (c) Filling the centers of red blood cells. (d) Result of opening and closing operations.
Source: The authors.
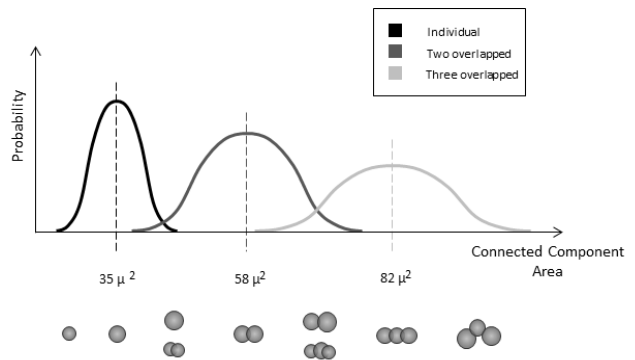
Figure 4. Probability distribution sketch depending on the blood area, which belongs to one of the classification sets (Single, Two cells, three cells).
Source: The authors.
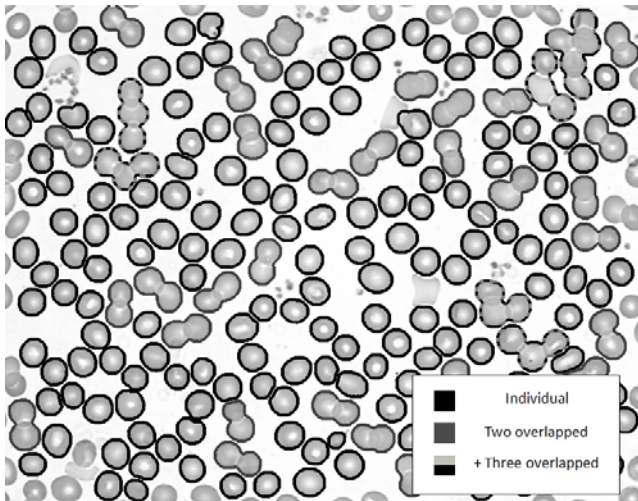
Figure 5. Group separation according to their area.
Source: The authors.

Figure 6. (a) Connected component. (b) Erosion to separate the connected component. (c) Watershed using results from erosion as markers.
Source: The authors.

level of each pixel corresponds to a physical elevation. When the landscape is immersed in a lake with holes pierced in local minima, catchment basins fill up with water, starting at these local minima. At points where water coming from different basins meets, dams are built. As a result, the landscape is partitioned into regions or basins separated by dams; these are called watershed lines [28].

The main drawback of this method is the over-segmentation due to the presence of many local minima. To decrease the effect of severe over-segmentation, marker-controlled watershed transformations can be used [29].

For this reason, in this step, different approaches are used in order to define the suitable markers in watershed transform.

Using erosion: After classifying the cell groups we need to find out the centroid of each cell that defines the connected component. As a first attempt to discover this center, successive erosions are used. In order to do this, an image such as Fig. 6(a) is created for each connected component that was not classified as an individual [30].

First, erosions are made iteratively until the connected component is separated. Then, in order to look for the line that allows the cells to be separated we had to find the components' center of mass that were in Fig. 6 (b) in order to use them as markers in the watershed method.

Fig. 6 (c) shows the centroids and watershed lines that are drawn on the connected component.

Using Hough: The problem that arises is that, after erosion, some cells are not separated; however, their morphology was taken to be an advantage (typically RBC have a circular shape) using the Hough transform.

The Hough transform is applicable to any function of the form $g(v, c) = 0$, where v is a vector of coordinates and c is a vector of coefficients. In Cartesian coordinates, the equation for a circle is given by:

$$(x - a)^2 + (y - b)^2 = r^2 \qquad (1)$$
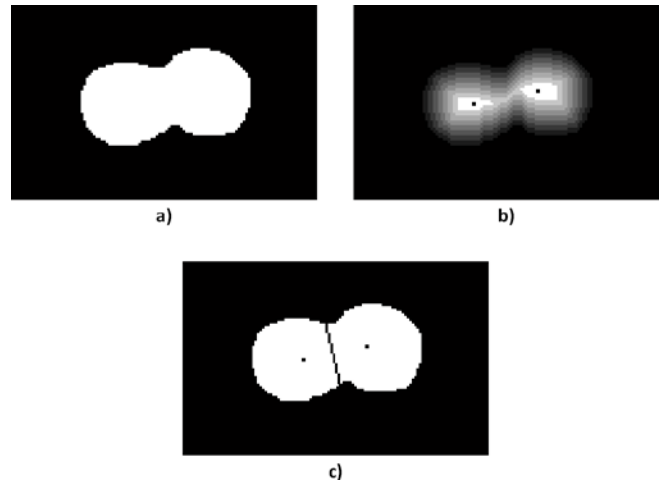
To detect a circle we need to search parameter triplets (a, b, r) in a three dimensional parameter space. The edge pixels are used as the input point for the Hough transform circle detection.

If the edge pixel lies on a circle, the locus for the parameters of that circle is a right circular cone surface. This means that each point (x, y) produces a cone surface in parameter space.

If a set of edge pixels in a cell image are lying on the circle with parameters (a_0, b_0, r_0), the resultant loci of each point will pass the same point (a_0, b_0, r_0) in parameter space. Thus, many right circular cone surfaces will intersect at a common point [23].

Below, Fig. 7 shows the centers of the circles detected by Hough in groups of two and three cells. Subsequently, using these centers as markers, watershed finds the overlapping lines.

Using K-means: After undertaking the steps above, there are still unseparated cells; many of these have an irregular shape, so it is important to use a method that does not depend on the morphology of the object.

The third technique used is k-means, it is the clustering algorithm used to determine the natural spectral groupings present in a data set. This receives as an input the number of clusters to be located in the data. The algorithm then arbitrarily seeds or locates that number of cluster centers in multidimensional measurement space. Each pixel in the image

Figure 7. Centers detected by Hough in groups of two and three cells.
Source: The authors.
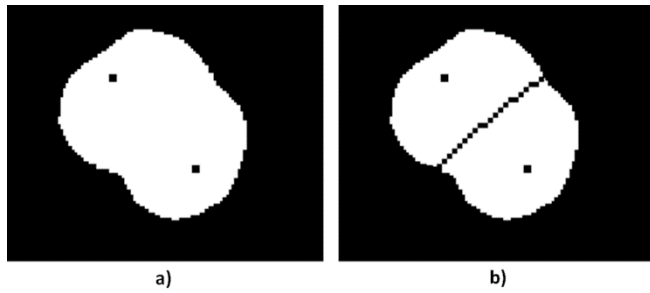
Figure 8. (a) Result of k-means. (b) Watershed using k-means results as markers.
Source: The authors.

Figure 9. (a) Cells with two endpoints on its edge. (b) Cells with more than two endpoints on its edge.
Source: The authors.

is then assigned to the cluster whose arbitrary mean vector is closest. The procedure continues until there is no significant change in the location of class mean vectors between successive iterations of the algorithm [31].

The set of points that belong to the edge is passed as a parameter to a k-means algorithm, and it also approximates the center of each cell, as shown in Fig. 8 (a). Again, using the result of k-means as markers, watershed detects overlapping lines. This result is presented in Fig. 8 (b).

2.3. Post-Processing

The purpose of this stage is to draw the cell edge in the overlapping area. First, each cell is separated using the images that come from each of the above steps. Their edge is found and then the endpoints of that edge are obtained. In this step, there are two types of cells; the first is shown in Fig. 9 (a), in which there are only two endpoints.

In this case, Bézier is used to soften the missing edge of the cell. The following are used as control points: 50 points taken from the edge, including 25 points behind each endpoint, as shown in Fig. 10 (b); and an arc which is formed using the two endpoints and the cell's center of mass(see Fig. 10 (c)). The result, presented in Fig. 10 (d), is a smooth curve of the cell edge.

The second type is shown in Fig. 9 (b), in which a cell has more than two endpoints on its edge. For these types of cells, Bézier is not used since it presents several problems: the execution time increases, there are fewer control points and the edge of these objects is very irregular. Therefore, the distance of each point on the edge to the center of mass of the cell was measured, and an average radius was selected to draw a circle in order to define the edge of the cell.

3. Results and discussion

A computer with an Intel Core I3 processor was used for the test. The processing time per image was, on average, 25 seconds. 50 images of 1076 x 840 were selected, in which the three types of blood cells (white, red and platelets) are present; although, some of them are blurred and have noisy spots.

Table 1 shows the rates of specificity and sensitivity to detect groups of two and three cells, as these are the most common in the observation area of peripheral blood smears. The numbers were estimated by comparing with manual counting results.
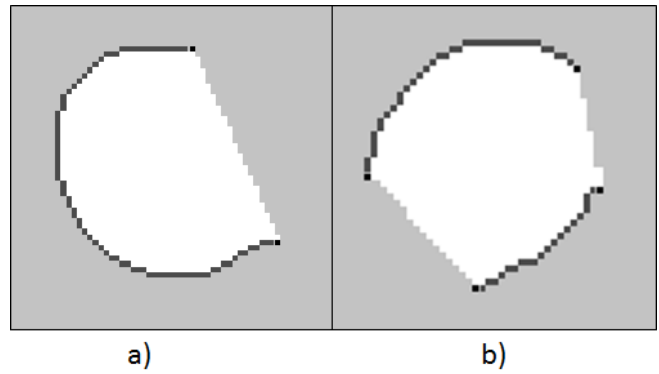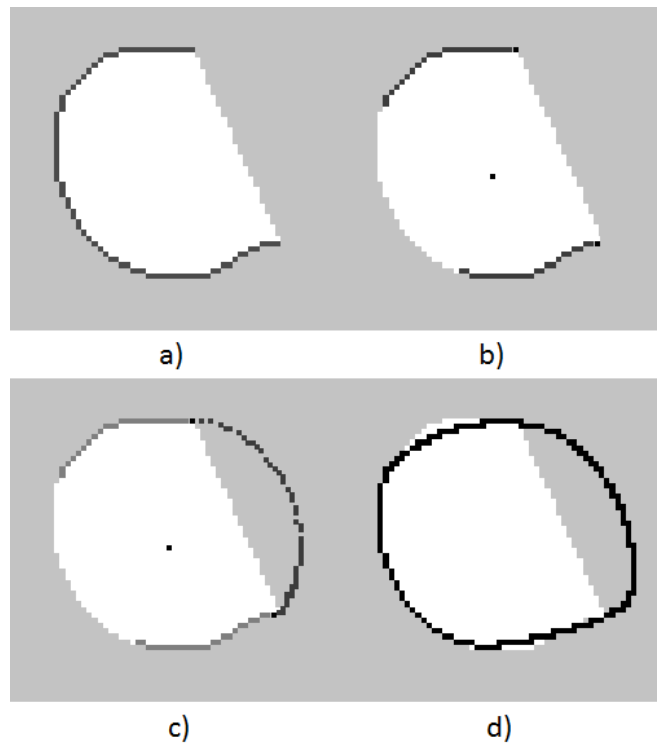
Figure 10. (a) Initial edge of the cell (b) Previous points to the endpoints (c) Arc using the final points. (d) Edge of a cell smoothed with Bezier. Source: The authors.

Table 1.
Rates result.

Group	Sensitivity [%]	Specificity [%]
1 cell	99,90	96,70
2 cells	96,19	99,79
3 cells	99,04	99,90

Source: The authors.

The sensitivity of the algorithm quantifies its ability to correctly identify the groups of two and three cells. In other words, it is the proportion of the true positives (TP) that are correctly identified by the algorithm, which is given by

$$Sensitivity = \frac{TP}{TP+FN} \quad , \qquad (2)$$

where FN represents the false negatives, the groups that were not detected. Moreover, the specificity is the algorithm's ability to not identify as groups of two or three cells the groups that do not meet these characteristics. In other words, the proportion of the true negatives (TN) is given by (3), where FP represents the false positives: the groups that did not have two or three cells and were detected as such [32].

$$Specificity = \frac{TN}{FP + TN} \qquad (3)$$

The final output looks like the one in Fig. 11, which shows that the algorithm also manages to classify groups of more than three cells. In addition, Fig. 12 shows comparative results for three different algorithms: a watershed approach, an algorithm using the Hough transform, and our method. The Hough-based algorithm was superior to the watershed approach, but it still could not identify cells with no circular shape. In contrast, our algorithm performs better in cells with an irregular shape, and does not present over-segmentation as the watershed-based algorithm does.

However, Fig. 13 shows some errors in our algorithm. When there is a platelet over the RBC, the preprocessing step deforms the cell, and when the size of the cells is out of the average range, the area classification of the groups is flawed.
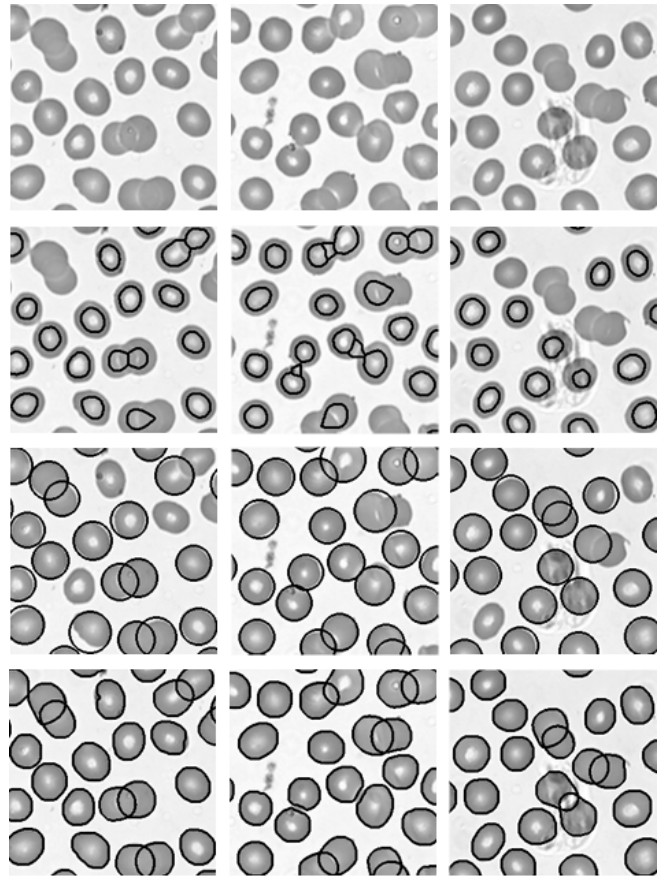
Figure 12. Comparison with other techniques. The original image, watershed results, Hough results, and our method results are in the first, second, third and fourth rows, respectively.
Source: The authors.

Figure 11. (a) Original image. (b) Result of edge detection of all the cells present in the image.
Source: The authors.
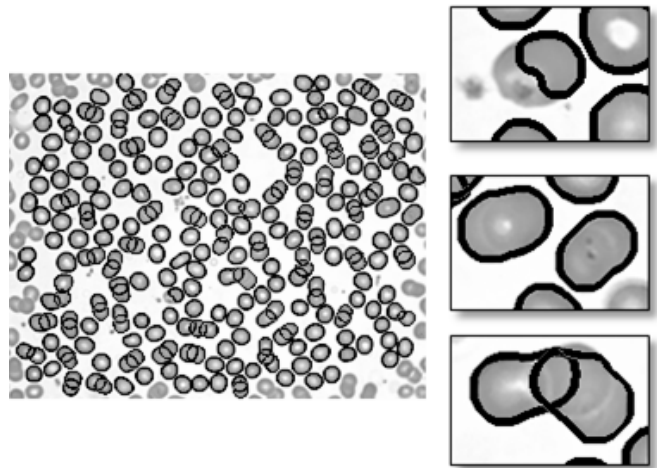
Figure 13. Algorithm errors.
Source: The authors.

Changes in illumination and contrast also significantly affect the results. The light intensity of the microscope Olympus CX-31 was set at 5.7 volts, and the aperture iris diaphragm was set to 80% of the aperture of the objective. For these specific values, in the thresholding step of the nuclei of white blood cells, the value obtained by the Otsu's method increased by 1.5. If the values of illumination and contrast vary, the scaling value should be changed accordingly in order to prevent the result being affected.

In the Red Blood Cell Segmentation step, as the Wright's stain was used, the Green channel of the RGB input image is thresholded as the variations of tone are best appreciated in this channel. If the staining method changes, the segmentation step can give different results.

4. Conclusions

- The illumination of the microscope, the staining of the plate, and properly performing the extent are very important in achieving a correct segmentation.
- It is important that the images are taken in the observation area that is used by the specialist. This is because in this area the shape of RBC allows Hough transform to offer better results since the cells in other regions of the plate are usually deformed.
- Using Watershed, Hough and K-means separately does not equate to employing an effective technique because each one of them has its own limitations. Therefore, an algorithm that takes advantage of each of these techniques to get better results was developed.
- The specificity and sensitivity indices obtained show that the algorithm developed for detection and classification of overlapped red blood cells was highly accurate.

Acknowledgments

References

[1] Rogers, K., Blood: Physiology and Circulation. New York: Britannica Educational, 2010

[2] Bijlani, R., Fundamentals physiology a textbook for nursing students. New Delhi: Jaypee Brothers Publishers, 2001.

[3] Estridge, B., Basic Medical Laboratory Techniques. Clifton Park: Cengage Learning, 2000.

[4] Rodak, B., Hematology: Clinical principles and applications. Missouri: Elsevier Health Sciences, 2007.
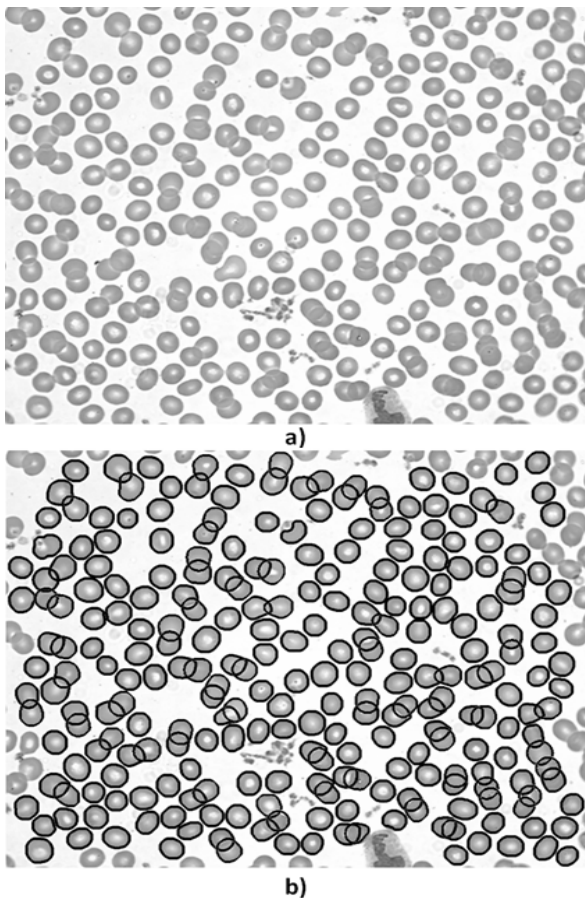
[5] Ruiz, G., Fundamentos de hematología. Buenos Aires: Ed. Médica Panamericana, 1994.

[6] Kothari, S., Chaudry, Q. and Wang, M., Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques, Biomedical Imaging: From Nano to Macro, 2009. pp. 795-798. DOI: 10.1109/ISBI.2009.5193169

[7] Cloppet, F. and Boucher, A., Segmentation of complex nucleus configurations in biological images. Pattern Recognition Letters, 31(8), pp. 755-761, 2010. DOI: 10.1016/j.patrec.2010.01.022

[8] Sharif, J., et al., Red blood cell segmentation using masking and watershed algorithm: A preliminary study, Biomedical Engineering (ICoBE), 2012. pp. 258-262. DOI: 10.1109/ICoBE.2012.6179016

[9] Veta, M., et al., Marker-controlled watershed segmentation of nuclei in H&E stained breast cancer biopsy images, Biomedical Imaging: From Nano to Macro, 2011. pp. 618-621. DOI: 10.1109/ISBI.2011.5872483

[10] Xu, S., Liu, H. and Song, E., Marker-controlled watershed for lesion segmentation in mammograms. Journal of Digital Imaging, 24(5), pp. 754-763, 2011. DOI: 10.1007/s10278-011-9365-2

[11] Di Cataldo, S. et al.. Automated segmentation of tissue images for computerized IHC analysis. Computer Methods and Programs in Biomedicine, 100(1), pp. 1-15, 2010. DOI: 10.1016/j.cmpb.2010.02.002

[12] Karvelis, P., Likas, A. and Fotiadis, D., Identifying touching and overlapping chromosomes using the watershed transform and gradient paths. Pattern Recognition Letters, 31(16), pp. 2474-2488, 2010. DOI: 10.1016/j.patrec.2010.08.002

[13] Wen, Q., Chang, H. and Parvin, B., A Delaunay triangulation approach for segmenting clumps of nuclei, Biomedical Imaging: From Nano to Macro, 2009. pp. 9-12. DOI: 10.1109/ISBI.2009.5192970

[14] Kumar, S. et al., A rule-based approach for robust clump splitting. Pattern Recognition, 39(6), pp. 1088-1098, 2006. DOI: 10.1016/j.patcog.2005.11.014

[15] Wang, H., Zhang, H. and Ray, N., Clump splitting via bottleneck detection and shape classification. Pattern Recognition, 45(7), pp. 2780-2787, 2012. DOI: 10.1016/j.patcog.2011.12.020

[16] Berge, H. et al., Improved red blood cell counting in thin blood smears, Biomedical Imaging: From Nano to Macro, 2011. pp. 204-207. DOI: 10.1109/ISBI.2011.5872388

[17] Latorre, A. et al., Segmentation of neuronal nuclei based on clump splitting and a two-step binarization of images. Expert Systems with Applications, 40(16), pp. 6521-6530, 2013. DOI: 10.1016/j.eswa.2013.06.010

[18] Di Ruberto, C., Dempster, A., Khan, S. and Jarra, B., Analysis of infected blood cell images using morphological operators. Image and Vision Computing, 20, pp. 133-146, 2002. DOI: 10.1016/S0262-8856(01)00092-0

[19] Buggenthin, F. et al., An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. BMC Bioinformatics, 14(297), pp. 1-12, 2013. DOI: 10.1186/1471-2105-14-297

[20] Prasad, K. et al., Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images. Journal of digital imaging, 25, pp. 542-549, 2012. DOI: 10.1007/s10278-011-9442-6

[21] Amit-Kunar, P., Tembhare, P. and Pote, C., Enhanced identification of malarial infected objects using otsu algorithm from thin smear digital images. International Journal of Latest Research in Science and Technology. 1, pp. 159-163, 2012.

[22] Priyankara, G.P.M. et al., An extensible computer vision application for blood cell recognition and analysis. Thesis, Department of Computer Science and Engineering, University of Moratuwa, Sri Lanka, 2006.

[23] Guan, P.P. and Hong, Y., Blood cell image segmentation based on the Hough transform and fuzzy curve tracing. Machine Learning and Cybernetics (ICMLC), 2011. pp. 1696-1701. DOI: 10.1109/ICMLC.2011.6016961

[24] Prasad, D.K., Leung, M.K. and Cho, S.Y., Edge curvature and convexity based ellipse detection method. Pattern Recognition, 45(9), pp. 3204-3221, 2012. DOI: 10.1016/j.patcog.2012.02.014

[25] Mahmood, N. et al., Blood cells extraction using color based segmentation technique, An International Journal IJLBPR, 2(2), 2013. DOI: 10.1016/j.patcog.2012.02.014

[26] Ramesh, N., Salama, M. and Tasdizen, T., Segmentation of haematopoeitic cells in bone marrow using circle detection and splitting techniques, 9th IEEE International Symposium on Biomedical Imaging (ISBI), 2012. pp. 206-209. DOI: 10.1109/ISBI.2012.6235520

[27] Otsu, N., A threshold selection method from gray-level histograms. Automatica, 11, pp. 23-27, 1975.

[28] Jayaraman. Digital image processing. Image segmentation. Nueva Delhi: Tata McGraw-Hill Education, 2011.

[29] Gomez, W., Leija, L., Alvarenga, A., Infantosi, A. and Pereira, W. Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation. Med Phys, 37, pp. 82-95, 2010. DOI: 10.1118/1.3265959

[30] Huang, J., An improved algorithm of overlapping cell division, Intelligent Computing and Integrated Systems (ICISS), 2010. pp. 687-691. DOI: 10.1109/ICISS.2010.5655507

[31] Chitade, A.Z. and Katiyar, S.K., Colour based image segmentation using k-means clustering. International Journal of Engineering Science and Technology 2(10), pp. 5319-5325, 2010.

[32] Vianney, J., Rosales, A. and Gallegos, F., Computer-aided diagnosis of brain tumors using image enhancement and fuzzy logic. DYNA, 81(183), pp. 148-157, 2014. DOI: 10.15446/dyna. v81n183.36838

M.F. Romero-Rondón, Received a BSc. Eng in Systems Engineering in 2014 from the Universidad Industrial de Santander, Bucaramanga, Colombia. He is currently a graduate student at the University of Nice-Sophia Antipolis, France. His current research interests include bioinformatics, biomedical signal and image analysis, telemedicine, and computer vision.
ORCID: 0000-0001-6199-8722

L.M. Sanabria-Rosas, received a BSc. Eng in Systems Engineering in 2014 from the Universidad Industrial de Santander, Bucaramanga, Colombia. She has worked in the biomedicine and telemedicine fields. She is currently a student on the International Masters in Computer Science at the Université Nice-Sophia Antipolis.
ORCID: 0000-0003-4345-1074

L.X. Bautista-Rozo, received a BSc. Eng. in Systems Engineering in 2003 from the Universidad Industrial de Santander, Bucaramanga, Colombia and an MSc. in Computer Engineering in 2007 from the University of Puerto Rico, Mayaguez Campus. She was the director of the Biomedical Engineering (GIIB) Research Group from 2011 to 2013. She is currently a PhD candidate at the Ecole Doctorale des Sciences et des Technologies de l'Information et de la Communication at the Université Nice-Sophia Antipolis.
ORCID: 0000-0002-3853-007X

A. Mendoza-Castellanos, received a BSc. in Computer Science from New York Institute of Technology, an MSc in Automation, Robotics and Informatics from the Université Franche Comté. From 1980 to 2011 he was Full Professor in the Universidad Industrial de Santander. He was the founder and director of the Biomedical Engineering (GIIB) Research Group. He is currently the general manager of the company Biosys LTDA.
ORCID: 0000-0002-3550-9638

An online algorithm for the container stacking problem

Roberto Guerra-Olivares [a], Neale R. Smith [a] & Rosa G. González-Ramírez [b]

[a] Centro de Calidad y Manufactura, Tecnológico de Monterrey, Monterrey, México roberto.guerra@itesm.mx nsmith@itesm.mx
[b] Facultad de Ingeniería y Ciencias Aplicadas, Universidad de Los Andes, Santiago de Chile, Chile, rgonzalez@uandes.cl

Abstract
Efficient cargo handling is a key element for a maritime port to compete and provide good service levels to its users. The performance of a port is related to ship-turnaround, which is conditioned by the ships loading and unloading operational efficiency. At the yard, containers are temporarily stacked in order to later either load them onto a ship or dispatch them to external users. Stacking has a strong impact on ships' service times. This paper proposes a container stacking policy, considering the particular characteristics of a container terminal in Chile. In order to measure the performance of the procedure, an upper bound for the number of re-handles of containers is estimated as a function of the block's capacity. Numerical results are provided in comparison to an upper bound, and a good performance by the proposed procedure is demonstrated.

Keywords: Container Terminal, Yard Management, Container stacking problem.

Un algoritmo en línea para el problema de apilamiento de contenedores

Resumen
El manejo eficiente de carga es un elemento clave para un puerto marítimo pueda competir y proveer adecuados niveles de servicio a sus usuarios. El desempeño de un puerto depende del tiempo de permanencia de la nave, que está condicionado por la eficiencia en las operaciones de carga y descarga de las naves. En el patio, los contenedores son almacenados temporalmente para ser cargados a la nave o despachados a los usuarios externos con un alto impacto en los tiempos de atención de las naves. Este artículo propone una política para stacking de contenedores, considerando las características particulares de un terminal de contenedores en Chile. Para medir el desempeño de este procedimiento, se propone una cota superior para el número de despejes de un contenedor en función de la capacidad de los bloques. Se presentan resultados numéricos en comparación con la cota superior, mostrando un buen desempeño del procedimiento propuesto.

Palabras clave: Terminal de Contenedores, Gestión del Patio, Problema de Stacking de Contenedores

1. Introduction and literature review

Containerization has lowered shipping costs and significantly increased productivity of operations related to international trade. Port terminals play a crucial role as intermodal interfaces and act as a linking node with other inland transport modes. International trade has been increasing over recent years, and with the increasing number of container shipments competition between port terminals has also increased. This means that there has been a greater demand for better service levels and value-added services to the users.

The competitiveness of a container seaport strongly relies on the service time of the ships and, hence, the minimization of the time a ship is at the berth is an overall objective with respect to terminal operations.

In order to guarantee low service times for the ships, efficient cargo handling is required, especially considering that a higher number of containers are received by the terminal due to the tendency of increasing ship sizes. The previous have motivated research and development of methods and tools to support decisions related to operations

management in port terminals. Steenken et al. [1] and Stahlbock and Voß [2] provide a comprehensive survey of the state of the art of operations at container terminals as well as some models and methods that have been proposed in the literature to optimize import and export operations.

Port terminals consist of three main areas: the quay or seaside interface, the yard, and the gate or landside interface. The quay is the area where ships are berthed and quay cranes are used to transfer containers from and to the ships. There may also be some ships that have their own cranes to transfer cargo. The gate is the interface to landside, where the containers are received or dispatched to the external users by trucks, trains or barges. The yard act as a buffer for the terminal, and it is the place where containers are temporarily stored during the interval of time in which they are received and loaded on the ship (export containers), or they are unloaded from the ship and dispatched to external carriers (import containers).

The arrangement of containers within the yard clearly influences the operational continuity of the quay cranes, and hence, the operational efficiency of the port terminal. In this work we consider the problem of assigning storage space, particularly for export containers that arrive at the yard, with the aim of enhancing quay cranes productivity. In order to do this, we propose a heuristic procedure that assigns storage space to containers based on minimizing container relocation. This takes place when containers need to be moved so as a container can be reached that will be loaded onto a ship or dispatched to an external carrier.

International trade operations in Chile have shown a high level of dynamism: evidence from the current integration with world markets with a significant growth in the volume of trade over the past decade. Maritime transport is the most significant mode of transportation in Chile; more than 90% of the cargo is transferred through maritime port terminals. Chile is the economy that has the highest number of containers per capita in Latin America, and it is ranks as the 4th economy that has the highest amount of container traffic in Latin America and the Caribbean region [3]. In [4], the evolution of the port system in LAC is presented. Chile shows high level of dynamism with the emergence of secondary ports and a greater geographical spread of ports towards the south of the country.

The Port System in Chile can be classified into three main regions: northern, central, and southern. The central region of Chile contains more than 60% of the volume transferred in Chile, the ports of Valparaiso and San Antonio being the most significant in terms of TEUs. According to the port ranking provided by UN-ECLAC [5], in 2013 Chile presented a 6% positive increase of container movements in contrast to other Latin American countries that had a decreasing rate. The Port of San Antonio is positioned in 12th place in the ranking with 1,196,844 TEUs, and the Port of Valparaiso is in 16th place with 910,780 TEUs. Both ports share the same hinterland, which is comprised of the Metropolitan Region of Santiago, the Fifth and Sixth Regions of Chile, to a smaller extent from the Fourth and Seventh regions, as well as the Cuyo Region in Argentina.

The research presented in this paper is motivated by the current operations of the San Antonio International Terminal

(STI for its acronym in Spanish). STI is the main terminal in the port of San Antonio, and one of the main containerized terminals in the country.

1.1. Literature review

Research into container terminal operations has attracted the attention of academics since the 1990s. Up until now, more than 6,000 articles related to container handling have been published in different international journals. The work that deals with either automated or non-automated container terminals can be distinguished in the literature, and most of the work found addresses tactical and operational planning decisions.

On a strategic decision-making level, we can find several contributions related to the location of a facility in the global supply chain network. For instance, Osorio-Ramírez et al. [6] present an evolutionary algorithm to locate container yards for a 3PL provider.

Several authors have addressed related problems in terms of yard management operations. One of the earliest works is by Kim [7], who estimated the number of moves to retrieve one container, both analytically and by simulation. Kim and Kim [8] extended previous study and incorporated yard cranes allocation decisions. Kim and Bae [9] present a model to minimize the number of export containers to be moved in the shortest possible traveling distance, considering a remarshaling strategy (containers are relocated in a dedicated space, near to the place in which the ship will be berthed). This paper does not consider a remarshaling strategy. However, the heuristic procedure proposed could be implemented for both a direct and remarshaling of containers storage space policy.

Kim and Kim [10] consider the storage space allocation problem for import containers and analyze cases with a constant, cyclic and dynamic arrival rate of import containers; the main objective is to minimize the expected number of re-handles. This work differs mainly in that we consider export containers. Kim et al. [11] use dynamic programming to determine the storage locations of export containers, grouping containers by weight. The aim of the procedure is to reduce the total number of re-handling or relocations. Previous work mainly differs in that the policy that we propose is based on a known loading sequence of containers that are to be loaded based on an estimated stowage plan.

Zhang et al. [12] formulated the storage space allocation problem (SSAP) for the Container Terminals´ storage yards. They first determine the number of containers to be placed at each block in the yard, and aim to balance the workload among blocks. They then define the number of containers associated with each vessel in terms of the total number of containers in each block and each period; the objective is to minimize the total distance traveled. The exact location of containers at the block is not defined, which is the main difference between previous research and this paper.

Bazzazi et al. [13] extend the model proposed in [12] in order to consider different types of containers. However, the authors only consider import containers, and they propose a genetic algorithm to solve the problem. Park et al. [14] present a stacking policy for incoming containers at an automated terminal. They propose an online search algorithm

that dynamically adjusts and optimizes a stacking policy by generating variants of stacking policies and evaluates them while they are actually being applied, in order to determine stacking positions. The main difference with respect to this work is that we do not consider an automated terminal, and we assume that yard cranes are reach-stackers. The equipment employed at the yard significantly influences the policies that will be implemented to assign space to containers and potential container relocations.

Tapia et al. [15] present a mathematical model for the storage space allocation problem that extends upon the basic ideas proposed in [12], but goes further in considering objective functions, assigning close locations in the yard to containers that belong to the same group or segregation of containers. Ries et al. [16] propose an online algorithm to help stack of containers at the yard, based on a fuzzy logic framework. In this work, we address the same problem but we propose a heuristic algorithm in which the rules are defined based on crispy values rather than fuzzy logic rules. Another important difference in the strategy employed herein, is that we consider the current practices of the Port Terminal in which they assign storage space to containers in a two phase approach. In this approach, blocks are pre-assigned to a block, prior to the beginning of each shift, and the exact location (bay and tier) of each container is assigned in real time. Another related work is presented by Valdés-González et al. [17], which proposes a fuzzy-based strategy for the container stacking problem. It specifically considers the case study of the Port of Valparaíso in Chile. The main difference with this paper is the methodology used, which is similar to that used in [16]. Another important difference is that the Port of San Antonio and Valparaíso are very different in terms of the logistics operations; the Port of Valparaiso has a pre-terminal where all trucks are directed, and also, the port terminals have more control on the sequence of truck arrivals to their gates. A complementary problem related to the relocation of export containers incurred during the ship´s loading operations is presented by Guerra-Olivares et al. [18].

The remainder of the manuscript is organized as follows: Section 2 presents a description of the problem and practical concerns. Section 3 presents the policy description, and an estimation of an upper bound is presented in section 4. Section 5 presents numerical experiments and results. Conclusions and recommendations for future research are given in section 6.

2. Problem description and practical considerations

The container yard serves as a temporary buffer for intermodal transportation systems. The export containers must be stored in the yard for the period of time between the container drop off at the yard by external trucks and the time when the container is loaded onto the ship. Based on STI current operations, the typical time window for the stacking of export containers is around 72 hours prior to the arrival of the ship. In general, containers arrive during that time window interval in a random sequence, that is unless the terminal has implemented an appointment system, which is not the case for the STI at present.

We assume that the yard is operated by RTGs (Rubber Tyred Gantry) cranes, although in practice, only certain sectors of the yard use this handling equipment. The yard manager should decide where to allocate those containers at the yard so that a continuous flow of containers to the ship may be guaranteed in order to ensure good service times. The need for a good pre-specified space allocation policy for the containers is fundamental for the overall efficiency of the port. During the service time of a ship, it is desirable to have a continuous flow of containers between the yard and the quay.

Particularly for export operations, which are the focus of this work, it is desirable that the rate at which the containers are retrieved from the yard be equal to the rate at which the quay crane loads containers onto the vessel. If the container retrieval rate is slower, the quay crane will have idle time, and the service time of the ship will be increased. Despite the random arrivals of the containers to the port, the port has advanced knowledge of at least a preliminary version of the loading sequence of containers to the vessel. Hence, a basic assumption that we consider is that the stowage plan is a piece of input data and is already known, even if it is a preliminary version that may be updated once the ship has arrived at the terminal.

Typically, the container storage area of a port yard is constituted by blocks divided into Bays, Rows and Tiers according to the BAROTI system. A slot is defined as a BAROTI coordinate in which a TEU (Twenty-Foot Equivalent Unit) is stored. The blocks may be subdivided into sub-blocks of bays for space allocation purposes. A sub-block of bays is a set of adjacent bays. The containers are arranged based on common characteristics like destination port, weight, type, and size. A group of containers with the same characteristics is known as segregation. It is attractive to store all the containers with the same segregation in the same sub-block of bays or in consecutive ones as this may minimize the relocation or re-handling of containers during the loading operation.

The procedures to store export containers vary among port terminals. The flow of containers could be directly from the yard stack to the quay side during the loading process, or in other cases, the containers could be organized in another area (marshaling space) in which containers are sorted according to the sequence of the stowage plan. This minimizes the ship service time, but also increases the number of movements. In this work, we assume the current practice of STI in which the managers define a specific area for export containers to be stacked during the time window that has been defined prior to the arrival of the ship. There are, therefore, no pre-marshaling areas, and hence, containers are located directly to position in the yard from which they will later be retrieved to be loaded onto the ship.

3. Proposed policy description

Considering a two-stage procedure for the space assignment policy, we will focus on the online or real time assignment of space, assuming that the port manager has pre-assigned a block position for each container to be segregated in the yard. In [16], a mathematical model to support the pre-assignment of block positions to segregations is proposed. In this paper, we focus on the assignment of containers in real time to a pre-specified block, as we will further describe.

Table 1.
Example of a sorting of containers based on arriving time.

Container	1	2	3	4	5	6	7	8	9	10
Arriving time	12:05	12:15	12:45	13:20	13:22	13:58	14:11	14:12	15:01	15:37

Source: The authors

Table 2.
Example of a loading sequence of containers

Container	1	2	3	4	5	6	7	8	9	10
Loading sequence	3	2	5	1	10	7	6	4	9	8

Source: The authors
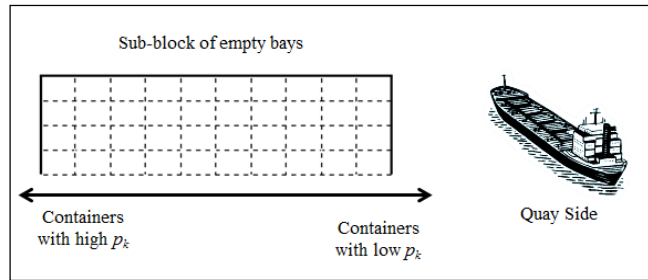
Figure 1. Using the horizontal dimensions of a block to sort the containers according to the loading sequence.
Source: The authors

3.1. Basic policy description and notation

The policy to allocate containers within the port yard consists of a set of criteria to assign a location to each arriving export container (BAROTI coordinate). We propose a heuristic procedure based on the corresponding ship´s stowage plan. Consider the particular instance of 10 containers with their arrival times, as shown in Table 1.

The sequence in which containers will be loaded to the ship is shown in the Table 2.

As we can observe from the Tables 1 and 2, the loading sequence of containers does not follow a FIFO (first in-first out) policy, and hence, the arrival order of containers does not define the sequence in which they are loaded. The horizontal dimension of the block can be used as an indicator of how early a specific container should be loaded to the ship, according to the stowage plan. For instance let p_k be the loading sequence according to the stowage plan of the kth container to arrive at the port. Therefore, a high p_k value indicates that the container will be loaded later than a container with a low p_k. Fig. 1 illustrates the sub-block space allocation depending on the containers´ loading sequence).

Thus, the main principle of the proposed policy is that containers to be first loaded on to the ship, should be allocated nearer to the right hand side of the block, and the containers to be loaded later should be allocated closer to the left hand side of the block.

3.1. Algorithm description

The following notation is defined:
- Let N be the number of available bays in the block,
- Let n be the number of sub-blocks into which N bays are subdivided. $1 \leq n \leq N$,
- Let N_j be the number of bays assigned to sub-block j, $\forall j \in \{1, 2, ..., n\}$,
- Let A be the total number of containers to be allocated,
- Let A_j be the number of containers assigned to the sub-block j, $\forall j \in \{1, 2, ..., n\}$,
- Let ran_j be the range of p_k stored in sub-block j, $\forall j \in \{1, 2, ..., n\}$,
- Let sb_k be the sub-block that stores container k, $\forall k \in \{1, 2, ..., A\}$,
- Let u_i be the number of containers stored at bay i $\forall i \in \{1, 2, ..., N\}$,
- Let Q be the maximum bay capacity, and
- Let B_j the set of opened bays in sub-block j, $\forall j \in \{1, 2, ..., n\}$.

3.2.1. Parameters Estimation

Some parameters should be determined prior to the execution of the algorithm. This is described by the following subsections:
a) Minimum bay capacity to store the containers. Due to random container arrivals the position of each container k must be identified according to its position in the stowage plan, p_k. Let d^* be the minimum bay capacity to store A containers in the yard. Assuming that an equal number of containers is allocated to each Bay, this may be computed in the way that is shown in the following equation:

$$d^* = \frac{A}{N} \qquad (1)$$

b) Number of bays that will conform each sub-block (Nj). The procedure to compute Nj is as follows; the number of sub-blocks (n) is an input parameter, so $\lfloor \frac{N}{n} \rfloor$ bays are allocated to each sub-block. To allocating all the bays of each sub-block, one extra bay should be allocated for the first mod (N, n) bays. For example, with N = 20 bays and n = 3 sub-blocks, each of the 3 sub-blocks will contain 6 bays. Since mod (20, 3) = 2, sub-blocks 1 and 2 will each be assigned one extra bay. The final bay allocation to each sub-block is shown in Table 3. The sub-blocks are sorted according to their proximity to a quay. In this example, sub-block number 1 is the closest to the quay while sub-block number 3 is the farthest from the quay.

c) Number of containers to be allocated to sub-block j (Aj). For each sub-block j, the product of the minimum bay capacity and the number of bays assigned to the sub-block j

Table 3
Numerical example of bay allocation to sub-blocks

Sub-block	1	2	3
Bay Allocation (N_j)	5	6	7
	6	6+1 = 7	8
		= 7	6

Source: The authors

gives the number of containers assigned to each sub-block, $A_j = N_j \cdot d*$. If this value is not an integer, it should be adjusted based on the following considerations:

- For the sub-blocks with the minimum the value of A_j will be equal to $\lfloor N_j \cdot d* \rfloor$.
- For the sub-blocks with the maximum N_j, the value of A_j will be equal to $\lceil N_j \cdot d* \rceil$.

Additional adjustment of A_j may be required if $\sum_{j=1}^{n} A_j \neq A$.

When the total number of containers assigned to all n sub-blocks is greater than A, the A_j for the sub-block with the maximum container allocation should decrease by $\sum_{j=1}^{n} A_j - A$ units.

- Equivalently, when the total number of containers assigned to all n sub-blocks is less than A, the A_j for the sub-block with the minimum container allocation should increase by

$A - \sum_{j=1}^{n} A_j$ units.

Once A_j is defined, parameter ran_j can be determined as follows:

- Let R_j be the accumulate value of A_j:

$$R_j = \sum_{x=1}^{j} A_x \quad \forall j \in \{1, 2, ..., n\}$$
(1)

- Sub-blocks with j = 1 store the containers with p_k in the range $ran_1 = [1, R_1]$.
- Sub-blocks with j = {2, 3, …, n} store containers with p_k in the range $ran_j = [R_{j-1} + 1, R_j]$

In order to illustrate previous procedure, consider the case with a value of $A = 50$ containers, $N = 20$ Bays, $d* = 2.5$ and $n = 3$ sub-blocks. After applying the previous steps, the final container allocation of each sub-block is presented in the Table 4.

Table 4
Numerical Example to assign containers to sub-blocks of bays

9	Sub-block (j)	10	1	11	2	12	3
13	(d*)(N_j)	14	2.5	15	2.5	16	2.5
		*7 = 17.5		*7 = 17.5		*6 = 15	
17	Container Allocation, A_j	18	18	19	17	20	15
21	R_j	22	18	23	35	24	50
25	ran_j	26	[1, 18]	27	[19, 35]	28	[36, 50]

Source: The authors

3.4. Heuristic description

A heuristic procedure is designed to assign containers to bays. The block of bays is subdivided into n sub-blocks and instances; different values of n were generated to examine the dependence of the number of re-handles with this parameter. The heuristic proceeds to fill the bays by allocating containers from the left to the right of the sub-block and from the back to the front of the sub-block. The algorithm assumes that the sub-blocks of bays are empty at the beginning and that each sub-block is filled up with containers according to ran_j. When a new container arrives it is allocated either in an opened sub-block bay or in a new empty sub-block bay. The containers are allocated so that the number of re-handles required to load them onto the ship are minimized.

The parameters that define an instance of the problem are the number of sub-blocks in which the block of bays will be subdivided (n), and the maximum tolerable difference between loading sequences in the stowage plan (Cr_{max}). Thus, one instance consists of the following parameters:

- Number of containers, A.
- Number of bays, N.
- Number of sub-blocks, n.
- Maximum capacity of the bay, Q.
- Maximum tolerable difference Cr_{max}.

Suppose the first container to arrive at the port yard has $p_k = p_1$ and is allocated in the upper left corner of the corresponding sub-block. Due to random arrivals, suppose that the second has $p_k = p_2$ with $p_2 < p_1$; so the second container could be located in front of the first. To retrieve both containers will not necessarily involve any re-handle movement. However, all the containers having p_k between p_2 and p_1 should be located somewhere else, and the total delayed time to load all the containers onto the ship will increase due to crane displacement to different sectors of the sub-block. To avoid the previous situation we estimate the difference between the last container allocated in each opened bay of the sub-block j and the container waiting to be allocated as shown in the following equation:

$$Cr_i = p_k - p_{k-1}^i \quad \forall i \in B_j$$
(2)

Where:
- p_{ik-1} is the loading sequence of the last container allocated to bay i,
- p_k is the loading sequence of the arriving container, and,
- B_j is the set of all opened bays in sub-block j.

A Cr_i difference should be computed between the loading sequence of the incoming container and the loading sequence of the last container assigned to each opened bay in the sub-block. We are interested only in negative values of Cr_i because they indicate that no re-handles will be required if the incoming container is allocated in bay i. The value of Cr_i is restricted to a maximum tolerable value, defined as Cr_{max}. If the value of Cr_i is greater to or equal than Cr_{max}, the container can be located in bay i.

The methodology can be summarized by the following series of steps:

1. Sort the containers according to their arrival order $k = \{1, 2, ..., A\}$.
2. Match the containers with the stowage plan and determine p_{k}.
3. Determine $sb_k \quad \forall k \in \{1, 2, ..., A\}$ by matching with ran_j.
4. Initialize $B_j = \emptyset \quad \forall j \in \{1, 2, ..., n\}$.
5. Initialize $u_i = 0 \quad \forall i \in \{1, 2, ..., N\}$.
6. Set $k = 1$.

7. If sb_k is empty, open the first bay of sb_k and go to step 8, otherwise go to step 9.

8. Allocate the container in the bay opened in step 7 and add the bay to B_{sbk}, then go to step 17.

9. Calculate $cr_i \forall i \in B_{sbk}$.

10. Determine the maximum negative element of cr_i. If no element of cr_i is negative go to step 12.

11. If the cr_i obtained in step 10 is greater or equal than Cr_{max}, allocate the container in the bay i and go to step 17. If not, proceed to step 12.

12. If possible, open a new bay in the sub-block and add the bay to B_{sbk}. If all bays of the sub-block are already opened, go to step 14.

13. Allocate the container in the block opened in step 12, and then go to step 17.

14. Calculate $cr_i \forall i \in B_{sbk}$ for any container with arrival order m, with $m = \{k+1, k+2, ..., A\}$, and $sb_m = sb_k$.

15. Reserve bay i when a negative value of cr_i is found.

16. Allocate the container k in the emptiest and unreserved bay of sb_k.

17. Set $k = k + 1$ and $u_i = u_i + 1$.

18. If $u_i = Q$, remove the bay i from B_{sbk}.

19. If $k \le A$ go to step 7, otherwise go to step 20.

20. End of the algorithm.

To illustrate the methodology, we provide a numerical example. Figure 2 shows a graphical representation of a block of 5 bays, each one has a maximum capacity of 5 containers. The bay number 5 has not yet been opened. There are 9 containers already in the block, and the tenth container to arrive to the port has a $p_k = 16$. Cr_i values for bays 1, 2, 3 and 4 should be calculated in order to determine in which bays the incoming container can be assigned.

The value of Cr for the bay 1 is calculated as follows: $Cr_1 = 16 - 8 = 8$. Table 5 shows the Cr_i differences of all opened bays. Some Cr_i differences are positive and others are negative. A Cr_i greater than 0, such as the Cr of bay numbers 1 and 3, indicates that if the container with $p_k = 16$ is allocated in any of those a re-handle movement will be required to retrieve the container.

According to the algorithm, the bays that are allowed to receive a container are those with a negative Cr_i value. Table 5 shows the Cr_i values for all opened bays. In this example, bays 2 and 5 are possible candidates to receive the container. The Cr_{max} is equal to -2, so bays with a value of $Cr = -1$ or $Cr = -2$ can be chosen. The algorithm will compare the Cr_i value of all candidates with the Cr_{max}, and the final decision will be to allocate the container in bay number 2.
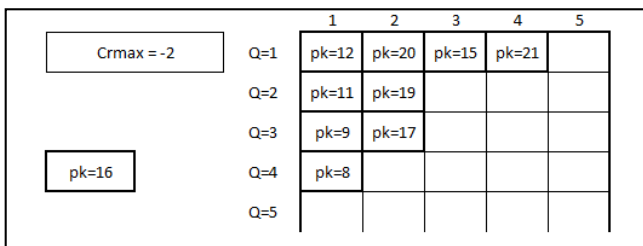
Table 5.
Cr_i calculation.

Bay	1	2	3	4
Cr	16-8	16-17	16-15	16-21
Cr	8	-1	1	-5

As can be observed in Table 5, a value of $Cr = -1$ implies that the containers have consecutive values of p_k, and in general, a negative value indicates that no re-handles are required to retrieve the pair of containers. The decision to open a new bay must be taken only when all the Cr_i differences are positive or there is no longer a negative Cr greater or equal than Cr_{max} for all opened bays in the block.

To open or use a new bay there must be at least one empty bay in the sub-block. When a container arrives, but it is not possible to open a new bay due to all bays being opened in the sub-block, the following methodology is performed:

- Inspect if there is any $Cr_i < 0$ for any container that has not yet arrived from the sub-block.
- When a $Cr_i < 0$ is found, reserve the bay for that container.
- Allocate the newcomer container in the emptiest and a bay that is not reserved in the sub-block.

Fig. 3 shows a numerical example of this situation. For the container with $p_k = 24$, the Cr_i vector is: $\{16, 7, 10, 3, 21\}$. Neither component of the vector is negative, so the algorithm will verify if there is a negative Cr_i for any not container that has not arrived.

4. Upper bound estimation

In this section we present an estimation of an upper bound on the maximum and an expected number of re-handles of a bay with capacity Q. In order to estimate the number of re-handles, we assume that containers are retrieved by a reach-stacker crane. The reach stacker crane can retrieve containers only from the front side of the stack, and the maximum stack height reached by the equipment is five. We assume that all the containers that need to be moved to retrieve the container are relocated in the same configuration as they were previously stacked in the block. Consider a bay with a capacity of 4 containers, as is illustrated in Figure 4.

Suppose that the sequence to retrieve the containers according to the stowage plan is $\{A, B, C, D\}$. Three re-handles (D, C, B) are required to retrieve container A. Once container A was retrieved, two additional re-handles are required to retrieve container B. Then, one additional re-handle should be made to retrieve container C. The last container to be retrieved does not require re-handles, so the total re-handles required to retrieve all the containers are: $3+2+1 = 6$. This configuration is the worst one possible due to the order of retrieving the containers being exactly the opposite of the order of the containers allocated in the bay.

Let Max_Rhand be the upper bound in the number of re-handles of a Bay with capacity Q. The maximum number of re-handles of a bay with capacity Q provides an upper bound (the worst case) for the number of re-handles that may be used to evaluate the effectiveness of the algorithm, which is expressed in the following equation:

		1	2	3	4	5
Crmax = -2	Q=1	pk=12	pk=20	pk=15	pk=21	
	Q=2	pk=11	pk=19			
	Q=3	pk=9	pk=17			
pk=16	Q=4	pk=8				
	Q=5					

Figure 2. Example of allocating an arriving container with pk = 16.

Figure 3 Example of a sub-block with all bays opened.
Source: The authors

Figure 4. Initial configuration of a bay of capacity 4.
Source: The authors.

$$Max\_Rhand = \sum_{i=1}^{Q-1} i = \frac{Q(Q-1)}{2} \qquad (3)$$

We may also use for comparison purposes the expected value of the number of re-handles for a bay with capacity Q. To de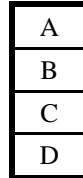fine this value we compute the distribution of the number of interferences and the number of permutations. For example, the same bay in the previous example with capacity 4 has 4! = 24 possible permutations. The minimum number of re-handles that can be obtained is 0 when the order of retrieving the containers is exactly the same as the order of the containers allocated in the bay. By generating the 24 permutations and the number of re-handles in each permutation, the corresponding distribution obtained is that shown in Table 6.

Table 6.
Distribution of re-handles of a bay with capacity 4

44	Re-handles	45	Number of permutations
45	0	46	1
47	1	48	3
49	2	50	5
51	3	52	6
53	4	54	5
55	5	56	3
57	6	58	1
59	Total	60	24

Source: The authors

Assuming an equal occurrence probability for each permutation, the expected value of the number of re-handles can be obtained by the weighted average as follows:

$$E[re-handles] = \frac{(0)(1)+(1)(3)+(2)(5)+(3)(6)+(4)(5)+(5)(3)+(6)(1)}{24} = 3 \qquad (4)$$

The previous procedure was applied to different values of the capacity of a bay in order to obtain the corresponding value of the maximum and expected number of re-handles, as shown in Table 7.

Table 7
Maximum re-handles and number of re-handles expected as a function of Q

61	Bay Capacity	62	Maximum Number of Re-handles	63	E[re-handles]
64	2	65	1	66	0.5
67	3	68	3	69	1.5
70	4	71	6	72	3
73	5	74	10	75	5
76	6	77	15	78	7.5
79	7	80	21	81	10.5
82	8	83	28	84	14
85	9	86	36	87	18
88	10	89	45	90	22.5
91	11	92	55	93	27.5
94	30	95	435	96	217.5

Source: The authors

A polynomial trend line was fitted to the data, and we obtained the following expression to determine the expected number of re-handles:

$$E[re-handles] = f(Q) = \frac{Q(Q-1)}{4} \qquad (5)$$

5. Numerical results and experimentation

5.1. Instance generation

We consider three types of instance sizes, according to the relationship between the number of containers and the total available slots to allocate the containers: Tight, Medium and Relaxed. We defined the values of the capacity of each bay as the product of the number of rows and tiers. Typical values based on the case study of STI are 6 and 5 respectively, which corresponds to a value of $Q = 30$. These are also the values considered for the instances generated. The number of sub-blocks could be any integer in the range of $(1, N)$, where the extreme cases are represented by the values of 1 and N when $n = 1$. The entire block is treated as the unique sub-block in the instance while if $n = N$ each single bay of the block is treated as a sub-block.

We define a range of values for each parameter that will later be described. Based on the historical data of the port terminal under consideration we assume that, on average, 90 trucks per hour arrive at the port terminal during the first and second shift in the day. We will consider therefore, that for a congested operational day, there is a maximum arrival of 1500 trucks. For an uncongested day, we will consider that 800 trucks arrive: thus, $A \in \{800, 1500\}$. The rest of the parameters are defined as: $N \in \{tight, medium, relaxed\}$; $n \in \{1, 3, 20, N\}$; $Q \in \{30\}$ and $Cr_{max} \in \{2, 4\}$.

By considering previous values for each parameter we generate a total of 2x3x4x2=48 types of instances. The values of N corresponding to tight, relaxed, and medium ranges are estimated as follows:

$$N_{tight} = \left\lceil \frac{A}{Q} \right\rceil \qquad (6)$$

$$N_{relaxed} = \left\lceil \frac{1.5 * A}{Q} \right\rceil \qquad (7)$$

$$N_{medium} = (N_{tight} + N_{relaxed}) / 2 \qquad (8)$$

5.2. Numerical experiments

The numerical experimentation was performed on an Intel Core i5 Computer with 2.50 GHz and 6.0 GB of RAM. The maximum number of re-handles in a bay with a capacity of 30 containers is defined as:

$$Max\_Rhand = \frac{(30)\ (29)}{2} = 435 \qquad (9)$$

While the expected value of *Re-handles* is defined as:

$$\mathrm{E}\left[Re\text{-}handles\right] = \frac{(30) \cdot (29)}{4} = 217.5 \qquad (10)$$

Considering the solution provided by the algorithm, we determined the number of re-handles per bay (*R_bay*) and compared this number to the upper bound that can be defined as the expected number of re-handles, as shown in the following equation:

$$GAP = \frac{\left|(R\_bay) - \left(\mathrm{E}\left[Re\text{-}handles\right]\right)\right|}{\mathrm{E}\left[Re\text{-}handles\right]} \times 100 \qquad (11)$$

The value of the gap obtained may be interpreted as the percentage below the expected value of Re-handles. The bigger the gap, the better the algorithm performs. Tables 8 and 9 show the results for the 800 size and 1500 size instances respectively. We consider a single replicate for the 48 types of instances, which are refereed based on the values of its parameters. For example, the "*800ctr27bays1subcrm2*" is an instance with 800 containers, 27 bays, 1 sub-block and Cr_{max} = 2. The tables present the total number of re-handles obtained by the solution, the number of re-handles per bay and the corresponding gaps and computation times. Gaps are expressed as a percentage value below the expected number of re-handles.

As can be observed in the tables, all instances found a number of re-handles that were less than the expected number of re-handles that was estimated as an upper bound. As we previously defined, positive gaps indicate a better performance of the heuristic procedure. From the above tables, the 800 size instances obtained an average gap of 38.48% and the 1500 size instances an average gap of 42.97%. This indicates that for both instance sizes the heuristic performs on average similar. Also, there are no big differences with respect to the size of the instance in terms of the number of containers is observed.

Moreover, for both sizes of containers, we observe that gaps increase proportionally as the instances are less restricted. Also, the best results are obtained when all the block of bays are treated as a single sub-block, e.g. with *n* = 1. Execution

Table 8.
Numerical Results for instances of 800 containers.

Instance	Total Re-handles	Re-handles / bay	GAP %below expected	Computational Time (seconds)
800ctr27bays1subcrm2	4742	175.63	19.25%	0.008
800ctr27bays1subcrm4	4742	175.63	19.25%	0.008
800ctr27bays3subcrm2	6135	227.22	4.47%	0.005
800ctr27bays3subcrm4	6135	227.22	4.47%	0.006
800ctr27bays20subcrm2	5758	213.26	1.95%	0.004
800ctr27bays20subcrm4	5758	213.26	1.95%	0.004
800ctr27bays27subcrm2	5713	211.59	2.72%	0.004
800ctr27bays27subcrm4	5713	211.59	2.72%	0.004
800ctr34bays1subcrm2	2415	71.03	67.34%	0.008
800ctr34bays1subcrm4	2415	71.03	67.34%	0.007
800ctr34bays3subcrm2	4509	132.62	39.03%	0.006
800ctr34bays3subcrm4	4480	131.76	39.42%	0.007
800ctr34bays20subcrm2	4853	142.74	34.37%	0.005
800ctr34bays20subcrm4	4853	142.74	34.37%	0.004
800ctr34bays34subcrm2	4328	127.29	41.47%	0.007
800ctr34bays34subcrm4	4328	127.29	41.47%	0.004
800ctr40bays1subcrm2	1377	34.43	84.17%	0.007
800ctr40bays1subcrm4	1370	34.25	84.25%	0.008
800ctr40bays3subcrm2	3315	82.88	61.90%	0.006
800ctr40bays3subcrm4	3342	83.55	61.59%	0.005
800ctr40bays20subcrm2	4317	107.93	50.38%	0.005
800ctr40bays20subcrm4	4309	107.73	50.47%	0.003
800ctr40bays40subcrm2	3948	98.70	54.62%	0.000
800ctr40bays40subcrm4	3948	98.70	54.62%	0.010

Source: The authors

Table 9.
Numerical Results for instances of 1500 containers.

Instance	Total Re-handles	Re-handles / bay	GAP (%below expected)	Computational Time (seconds)
1500ctr50bays1subcrm2	6950	139.00	36.09%	0.020
1500ctr50bays1subcrm4	7318	146.36	32.71%	0.020
1500ctr50bays3subcrm2	10725	214.50	1.38%	0.020
1500ctr50bays3subcrm4	10719	214.38	1.43%	0.010
1500ctr50bays20subcrm2	11158	223.16	2.60%	0.000
1500ctr50bays20subcrm4	11160	223.20	2.62%	0.010
1500ctr50bays50subcrm2	10959	219.18	0.77%	0.010
1500ctr50bays50subcrm4	10959	219.18	0.77%	0.010
1500ctr63bays1subcrm2	2945	46.75	78.51%	0.020
1500ctr63bays1subcrm4	3321	52.71	75.76%	0.020
1500ctr63bays3subcrm2	6052	96.06	55.83%	0.010
1500ctr63bays3subcrm4	6327	100.43	53.83%	0.010
1500ctr63bays20subcrm2	9374	148.79	31.59%	0.000
1500ctr63bays20subcrm4	9377	148.84	31.57%	0.010
1500ctr63bays63subcrm2	8448	134.10	38.35%	0.010
1500ctr63bays63subcrm4	8448	134.10	38.35%	0.010
1500ctr75bays1subcrm2	1705	22.73	89.55%	0.020
1500ctr75bays1subcrm4	1912	25.49	88.28%	0.010
1500ctr75bays3subcrm2	3778	50.37	76.84%	0.010
1500ctr75bays3subcrm4	3773	50.31	76.87%	0.010
1500ctr75bays20subcrm2	7742	103.23	52.54%	0.010
1500ctr75bays20subcrm4	7724	102.99	52.65%	0.010
1500ctr75bays75subcrm2	7142	95.23	56.22%	0.010
1500ctr75bays75subcrm4	7142	95.23	56.22%	0.010

Source: The authors

time for all the instances is less than 0.1 second, which is a very short computational time due to the polynomial characteristics of the algorithm that is proposed. Smaller computational times are observed for the 800 size instances, but even for the 1500 size these are quite small and adequate for an online procedure. Tables 10, 11 and 12 present results that are summarized by the type of instance: tight, medium, and relaxed, respectively.

As we can observe from previous tables, tight instances present the lower gaps, with an average gap of 8.45%, and a maximum gap of 36%, which shows that are the hardest instances to be solved and the number of re-handles is closer to the expected number computed as an upper bound. It is important to notice that for this type of instance, better results are obtained for the 1500 size instances than for the 800 size.

Table 10.
Numerical results for tight instances.

Instance	Total Re-handles	Re-handles / bay	GAP (%below expected)	Computational Time (seconds)
800ctr27bays1subcrm2	4742	175.63	19.25%	0.008
800ctr27bays1subcrm4	4742	175.63	19.25%	0.008
800ctr27bays3subcrm2	6135	227.22	4.47%	0.005
800ctr27bays3subcrm4	6135	227.22	4.47%	0.006
800ctr27bays20subcrm2	5758	213.26	1.95%	0.004
800ctr27bays20subcrm4	5758	213.26	1.95%	0.004
800ctr27bays27subcrm2	5713	211.59	2.72%	0.004
800ctr27bays27subcrm4	5713	211.59	2.72%	0.004
1500ctr50bays1subcrm2	6950	139.00	36.09%	0.020
1500ctr50bays1subcrm4	7318	146.36	32.71%	0.020
1500ctr50bays3subcrm2	10725	214.50	1.38%	0.020
1500ctr50bays3subcrm4	10719	214.38	1.43%	0.010
1500ctr50bays20subcrm2	11158	223.16	2.60%	0.000
1500ctr50bays20subcrm4	11160	223.20	2.62%	0.010
1500ctr50bays50subcrm2	10959	219.18	0.77%	0.010
1500ctr50bays50subcrm4	10959	219.18	0.77%	0.010

Source: The authors

Table 11.
Numerical results for medium instances

Instance	Total Re-handles	Re-handles / bay	GAP (%below expected)	Computational Time (seconds)
800ctr34bays1subcrm2	2415	71.03	67.34%	0.008
800ctr34bays1subcrm4	2415	71.03	67.34%	0.007
800ctr34bays3subcrm2	4509	132.62	39.03%	0.006
800ctr34bays3subcrm4	4480	131.76	39.42%	0.007
800ctr34bays20subcrm2	4853	142.74	34.37%	0.005
800ctr34bays20subcrm4	4853	142.74	34.37%	0.004
800ctr34bays34subcrm2	4328	127.29	41.47%	0.007
800ctr34bays34subcrm4	4328	127.29	41.47%	0.004
1500ctr63bays1subcrm2	2945	46.75	78.51%	0.020
1500ctr63bays1subcrm4	3321	52.71	75.76%	0.020
1500ctr63bays3subcrm2	6052	96.06	55.83%	0.010
1500ctr63bays3subcrm4	6327	100.43	53.83%	0.010
1500ctr63bays20subcrm2	9374	148.79	31.59%	0.000
1500ctr63bays20subcrm4	9377	148.84	31.57%	0.010
1500ctr63bays63subcrm2	8448	134.10	38.35%	0.010
1500ctr63bays63subcrm4	8448	134.10	38.35%	0.010

Source: The authors

Table 12.
Numerical results for relaxed instances

Instance	Total Re-handles	Re-handles / bay	GAP (%below expected)	Computational Time (seconds)
800ctr40bays1subcrm2	1377	34.43	84.17%	0.007
800ctr40bays1subcrm4	1370	34.25	84.25%	0.008
800ctr40bays3subcrm2	3315	82.88	61.90%	0.006
800ctr40bays3subcrm4	3342	83.55	61.59%	0.005
800ctr40bays20subcrm2	4317	107.93	50.38%	0.005
800ctr40bays20subcrm4	4309	107.73	50.47%	0.003
800ctr40bays40subcrm2	3948	98.70	54.62%	0.000
800ctr40bays40subcrm4	3948	98.70	54.62%	0.010
1500ctr75bays1subcrm2	1705	22.73	89.55%	0.020
1500ctr75bays1subcrm4	1912	25.49	88.28%	0.010
1500ctr75bays3subcrm2	3778	50.37	76.84%	0.010
1500ctr75bays3subcrm4	3773	50.31	76.87%	0.010
1500ctr75bays20subcrm2	7742	103.23	52.54%	0.010
1500ctr75bays20subcrm4	7724	102.99	52.65%	0.010
1500ctr75bays75subcrm2	7142	95.23	56.22%	0.010
1500ctr75bays75subcrm4	7142	95.23	56.22%	0.010

Source: The authors

In the case of the 800 tight instances, no difference on the results obtained for different values of Cr_{max} are observed. On the other hand, the relaxed instances show the best results (higher gaps), with an average value of 65%. Medium instances have an average gap of 48%, which is an intermediate value with respect to the relaxed and tight instances.

6. Conclusions and recommendations for further research

We have proposed a heuristic procedure for the allocation of storage space for export containers arriving at a Port Terminal. We specifically consider the case in which the terminal defines a specific area in which export containers are stacked during the time window that is fixed prior to the arrival of the ship. Containers that arrive with more anticipation are allocated to other spaces from which they are relocated to the stacking area during the time window period.

The aim of the procedure is to minimize the number of potential re-handles of containers once they are retrieved from the yard to be loaded onto the ship. The heuristic consists of a set of criteria to assign a location in a space within the yard based on the sequence in which the container will be loaded to the ship, and as is indicated in the stowage plan. We present computational results with a set of different sized instances and the restrictiveness that were generated.

An upper bound on the number of expected re-handles of containers as a function of the capacity of a bay was

determined in order to measure the performance of the heuristic procedure. Results show that the heuristic always finds lower values of the number of re-handles with respect to the upper bound. The instances that turned out to be more difficult are the tight type in which the gap with respect to the upper bound is not as high as for the rest of the instances. Computation times are very small with lower values of 0.1 seconds.

For further research, we propose that the algorithm to the import containers that are unloaded from the ship in extended in order to minimize the number of re-handles when they are dispatched to external trucks. We also consider an extension of the procedure in which the number of sub-blocks of bays (n) may be determined instead of considering it as a parameter to be another area of future research.

References

[1] Steenken, D., Voß, S. and Stahlbock, R., Container terminal operations and operations research—A classification and literature review. OR Spectrum vol. 26, pp. 3-49, 2004. DOI 10.1007/s00291-003-0157-z.

[2] Stahlbock, R., and Voß, S., Operations research at container terminals: A literature update. OR Spectrum, 30, pp. 1-52, 2008. DOI 10.1007/s00291-007-0100-9.

[3] Rodriguez, J.-P., The benefits of logistics investments: Opportunities for Latin America and the Caribbean. Technical notes, Inter-American Development Bank, 2012. IDB-TN-395.

[4] Wilsmeier, G., Monios, J. and Pérez-Salas, G., Port system evolution in Latin America and the Caribbean. Journal of Transport Geography, 39, pp. 208-221, 2014. DOI: 10.1016/j.jtrangeo.2014.07.007

[5] Doerr, O., Latin American and the Caribbean Container port throughput, Ranking 2013. Infrastructure Services Unit |NRID| ECLAC | United Nations, 2013

[6] Osorio-Ramírez, C., Arango-Serna, M.D., Adarme-Jaimes, W., Implementing an evolutionary algorithm for locating container yards of a 3PL provider. DYNA, 81(187), pp. 49-55, 2014. DOI: 10.15446/dyna. v81n187.40044.

[7] Kim, K.H., Evaluation of the number of re-handles in container yards. Computers and Industrial Engineering, 32(4), pp. 701-711, 1997. DOI 10.1016/S0360-8352(97)00024-7.

[8] Kim, K.H. and Kim, H.B., The optimal determination of the space requirement and the number of transfer cranes for import containers. Computers and Industrial Engineering, 35(3–4), pp. 427-430, 1998. DOI: 10.1016/S0360-8352(98)00125-9.

[9] Kim, K.H. ans Bae, J.W., Re-marshaling export containers in port container terminals. Computers and Industrial Engineering, 35(3–4), pp. 655-658, 1998. DOI: 10.1016/S0360-8352(98)00182-X.

[10] Kim, K.H. and Kim, H.B., Segregating space allocation models for container inventories in port container terminals. International Journal of Production Economics, 59, pp. 415-423, 1999. DOI: 10.1016/S0925-5273(98)00028-0

[11] Kim, K.H., Park, Y.M. and Ryu, K.-R., Deriving decision rules to locate export containers in container yards, European Journal of Operational Research, 124, pp. 89-101, 2000. DOI: 10.1016/S0377-2217(99)00116-2

[12] Zhang, C., Wan, Y.-W., Liu, J. and Linn, J.R., Dynamic crane deployment in container storage yards, Transportation Research Part B, 36, pp. 537-555, 2003. DOI: 10.1016/S0191-2615(01)00017-0

[13] Bazzazi, M., Safaei, N. and Javadian, N., A genetic algorithm to solve the storage space allocation problem in a container terminal. Computers & Industrial Engineering, 56, pp. 44-52, 2009. DOI: 10.1016/j.cie.2008.03.012

[14] Park, T., Choe, R., Kim, Y. H. and Ryu, K.R., Dynamic adjustment of container stacking policy in an automated container terminal, International Journal of Production Economics, 133, pp. 385-392, 2011. DOI: 10.1016/j.ijpe.2010.03.024

[15] Tapia, F., Covarrubias, R., Miranda, P. and González-Ramírez R.G., On the storage space allocation problem. Proceedings of 22nd International Conference on Production Research (ICPR), Brazil, 2013.

[16] Ries, J., González-Ramírez, R.G. and Miranda, P., A fuzzy logic model for the container stacking problem at the container terminal. Lecture Notes in Computers Sciences, LNCS 8760, pp. 93-111, 2014. DOI: 10.1007/978-3-319-11421-7_7

[17] Valdés-González, H., Reyes-Bozo, L., Vyhmeister, E., Salazar, J.L., Sepúlveda, P.E. and Mosca-Arestizábal, M., Container stacking revenue management system: A fuzzy-based strategy for Valparaiso port. DYNA 82(190), pp. 38-45. DOI: 10.15446/dyna.v82n190.42311

[18] Guerra-Olivares, R., González-Ramírez, R.G. and Smith, N., A heuristic procedure for the outbound container relocation problem during export loading operations. Mathematical Problems in Engineering, Article ID 201749, 2015, 13 P. DOI: 10.1155/2015/201749

R. Guerra-Olivares, is an Operations Analyst at Madisa-Cat in Monterrey, Mexico. He holds a Ph.D in Engineering Science from the Tecnológico de Monterrey, Monterrey Campus in Mexico. He holds a BSc. in Chemical Engineering from the Tecnológico de Monterrey and a MSc. in Energy Engineering from the same university. He also holds a MSc. degree in Industrial Engineering from the Pontificia Universidad Católica de Valparaíso, Chile. He has one book published in LAP LAMBERT Academic Publishing GmBH & Co. KG and several papers in proceedings of national and international seminars.
ORCID: orcid.org/0000-0001-6929-800X.

N.R. Smith, is an associate professor at the Quality and Manufacturing Center at the Tecnológico de Monterey, Campus Monterrey in Mexico where he currently directs the Doctoral Program in Engineering Science. He holds a BSc. in Industrial Engineering from the University of Arizona and MSc. and PhD degrees in Industrial Engineering from Georgia Tech. He has taught both undergraduate and graduate level courses in operations research, logistics and supply chain management. His research interests are in the areas of operations research and logistics and he has advised and co-advised both Master's and PhD theses. He has published in journals such as the International Journal of Production Research and the European Journal of Operational Research, among others. Before pursuing an academic career, he was a consultant and implemented logistics-related solutions for numerous companies.
ORCID: orcid.org/0000-0002-1477-8388

R.G. González-Ramírez, is a professor and researcher in the Faculty of Engineering and Applied Sciences at the Universidad de Los Andes, Chile. She holds a BSc. in Industrial Engineering from the Technologic Institute of Morelia, a MSc. in Industrial Engineering from Arizona State University, a MSc. in Quality and Productivity Systems and a PhD. in Engineering Sciences from the Tecnológico de Monterrey in Mexico. Her research includes port logistics and intermodal transport of cargo, supply chain management, trade facilitation and optimization techniques. She has been working in applied research projects and is currently an active researcher in the Program Digital and Collaborative Network of Ports in Latin America and the Caribbean leaded by the Economic System of Latin America and the Caribbean, SELA and the Latin America Development Bank CAF.
ORCID: orcid.org/0000-0001-8029-5432.

Marine safety: Asbestos in the maritime industry

José Ángel Fraguela-Formoso [a], José Luís Fernández-Soto [b], Pablo Fariñas-Alvariño [a]
& Luís Carral-Couce [a]

[a] E.P.S. Ferrol. University of A Coruña, Spain. E-mail: jafraguela@udc.es, pfarina@udc.es, lcarral@udc.es
b Germanischer Lloyd Spain, Ltd., Spain. E-mail: jlfernandezsoto@gmail.com

Abstract

Over the last decade the International Maritime Organization (IMO) has highlighted the negative effect for human health of exposure to asbestos. The process they undertook has culminated in the amendment of the relevant SOLAS regulations. Changes came into force on January 1st, 2011; in these changes the new installation of asbestos containing materials (ACMs) on ships has been forbidden. This work analyses the evolution of the international regulations related to this change. The involvement of the implied parts (ship owners, shipyards, etc.) in the fulfillment of these standards, as well as the presence and treatment of these materials onboard are also considered.

Keywords: safety; ships; asbestos; removal; rules.

Seguridad marítima: Asbestos en la industria marítima

Resumen

En la última década la Organización Marítima Internacional (OMI) ha puesto de relieve el efecto negativo para la salud humana de la exposición al amianto. El proceso ha culminado con la modificación de los reglamentos pertinentes del Convenio SOLAS. Los cambios que entraron en vigor el 1 de enero de 2011 han prohibido la instalación de materiales que contengan asbesto en los buques. En este artículo se presta atención a la evolución de las regulaciones internacionales relacionadas con este. A bordo también se consideran la participación de las partes implicadas (armadores, astilleros, etc…) en el cumplimiento de estas normas, así como la presencia y el tratamiento de estos materiales.

Palabras clave: seguridad; buques; amianto; eliminación; normas.

1. Introduction

World Health Assembly Resolution 58.22 from 2005 on cancer prevention and control [1] urged Member States to pay special attention to cancers for which avoidable exposure is a factor, particularly exposure to chemicals at the workplace.

Worker exposure to asbestos occurs through inhalation. The World Health Organization (WHO) estimates that over 107.000 people die each year due to asbestos-related diseases from occupational exposure.

In terms of the European Community Directive 2009/148/EC [2], "asbestos" is understood to be the following group of fibrous silicates, identified by their number in the register of the Chemical Abstract Service (CAS): actinolite asbestos (CAS No. 77536-66-4), grunerite asbestos (amosite) (CAS No. 12172-73-5), anthophyllite asbestos (CAS No. 77536-67-5) chrysotile (CAS No. 12001-29-5), crocidolite (CAS No. 12001-28-4), and tremolite asbestos (CAS No. 77536-68-6). Chrysotile -also known as white asbestos- is the most commonly used (up to 90%), followed by crocidolite (blue asbestos), and then amosite (brown asbestos). The wide range of commercial applications of these materials is due to their excellent tensile strength, low thermal conductivity and significant resistance to chemical attack.

The maritime industry is concerned as the risk of being affected by the harmful effects of these materials not only lies

with the people handling such materials during assembly, but also the crews that inhabit the ship during its operation. Even people from outside the ship (agents, inspectors, shipyard personnel, passengers, etc.) who occasionally visit facilities where such materials are present may be also affected.

To reduce risks of exposure, the International Maritime Organization (IMO), under SOLAS regulation II-1/3-5 [3], which has been in force since July 1st, 2002, prohibited any materials from being installed that containing asbestos on board a ship, except for some specific exceptional cases. This regulation must be applied to all materials used for the structure, machinery, electrical installations and equipment covered by the Convention. For ships built before that date, the presence of asbestos was allowed as long as its installation on board does not imply a risk for their crews' health. In addition, in the MSC/Circ.1045 [5] that was published the same year, a maintenance and supervision program of the ACM on a ship was established in order to minimize any person on board's exposure to asbestos fibers from.

Subsequently, successive amendments to the regulation [3] culminated in the prohibition, without exception, of the on board installation of asbestos containing materials; this has been in force since January 1st, 2011 [4].

However, despite this tightening of regulations, the IMO has identified serious contraventions, as stated in MSC.1/Circ. 1374 [6]. In some vessels certified as "asbestos free", dangerous materials have been found on board as a result of repairs having been carried out in shipyards or purchases of spare parts after the issuance of such certification.

Because of the serious risks to crewmembers, inspectors, and shipyard workers' health, every action must be taken to remove ACMs within a maximum of 3 years from the date found at which it was found on board. This task must be developed in close consultation with and be under the supervision of the Flag State concerned.

2. Impact of exposure to asbestos on health

Handling materials containing asbestos in assembly operations, cutting, maintaining, repairing and dismantling of facilities, etc., may rise asbestos fibers, which then remain suspended in the air. People present in the working area may then be at risk of inhaling these fibers.

The International Agency for Research on Cancer (IARC) has classified all asbestos fibers as "carcinogenic to humans". Asbestos is one of the most important occupational carcinogens, causing about half of all deaths from occupational cancer.

Asbestos is especially resistant to the internal defenses of the human body. Once inhaled and lodged inside the lungs, most fibers will not break up or dissolve, and they cannot be neutralized or removed. The most common asbestos related diseases are lung cancer, mesothelioma and asbestosis. Lung cancer is the major cause of death for workers exposed to inhaling asbestos. Its latency period is very long, as it manifests itself between 15 and 40 years after exposure. Mesothelioma is a malignant nodular cancer type in the membranes that line the lung cavity (80% in the pleura and 20% in the peritoneum). Generally, there is a latency period of 25 to 30 years. Asbestosis or pulmonary fibrosis is a thickening and scarring of lung tissue, which causes difficulty breathing. It usually has a latency period of 20 years.

There is no safe level of exposure to asbestos. The risk of occurrence of a disease associated with asbestos is related to the concentration of fibers in the air, exposure duration, exposure frequency, the size of inhaled fibers, and the elapsed time since the initial exposure. The higher the dose of exposure to asbestos, the greater the risk of asbestos related diseases. However, the amount of asbestos in a product is not necessarily related to an increased health risk.

According to the International Labour Office (ILO) [7], it is estimated that 4% of global gross domestic (incomings) product is lost in direct and indirect costs resulting from accidents, occupational diseases and deaths.

The dose-response related to asbestos exposure was not well known when this material was introduced into the industry. Thus, in Spain, in 1961 asbestosis was considered as the disease associated with exposure to asbestos; it was not recognized as carcinogen until 1978.

The regulation of occupational diseases is relatively recent. In Appendix III of [7], lung cancer and mesothelioma caused by asbestos appear as occupational diseases.

With European Commission Recommendation 2003/670/EC [8], the EU aims to improve understanding of asbestos-related diseases. Preventive guides for exposed workers are also given and it tries to relate the dose-response between exposure to a particular agent and disease. Annex I presents mesothelioma, pleura fibrotic diseases with respiratory restriction, and lung cancer, as well as diseases caused by inhaling asbestos dust. Annex II deals with lung cancer.

Directive 2009/148/EC [2] assumes that, in addition to the above, diseases such as bronchial carcinoma and gastro-intestinal carcinoma may also be associated with exposure to asbestos.

International medical research has concluded that there is no level of exposure to asbestos below which clinical effects do not occur. The risks associated with low levels of cumulative exposure are, so far, not well known. Since asbestos fibers do accumulate in the lungs, and since the risk of developing disease does increase as the cumulative dose increases, exposure to asbestos should be controlled or avoided whenever possible.

According to the International Ban Asbestos Secretariat (IBAS), the use of asbestos is forbidden in 54 countries throughout the world, most of them industrialized; however, it is still extensively used in emerging countries.

3. The influence of international standards on preventive asbestos actions

According to the American Industrial Hygiene Association (AIHA), Industrial Hygiene is a science and art devoted to the anticipation, recognition, evaluation, prevention, and control of those environmental factors or
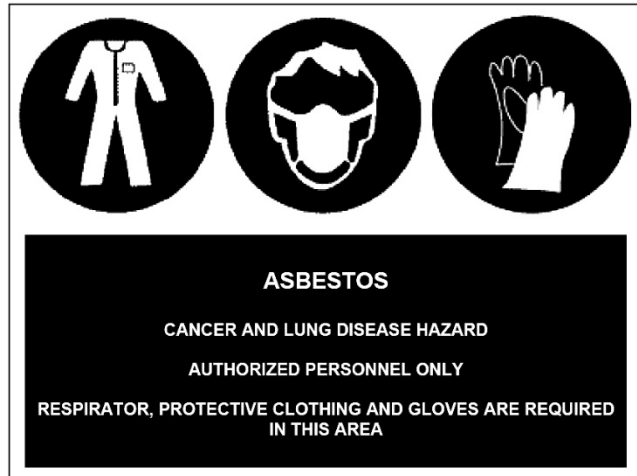
Figure 1. Asbestos warning signs
Source: The authors

stresses arising in or from the workplace. These may cause sickness, impaired health and well being, or create significant discomfort among workers or citizens in the community.

In the Industrial Hygiene field, preventive techniques are applied to the technical and human factors present in the workplace. However, those techniques can only be applied in accordance with up-to-date knowledge on the harmful effects of the materials used in a particular industrial process on people.

International laws consider aspects related to restrictions on the trading and use of certain dangerous substances [9-12], asbestos measurement methods and assessment in workplaces [13-15], analysis [14,16-18] as well as prevention and protection of workers exposed to this risk [19-25].

European Directive 2003/18/EC on the protection of workers from risks related to exposure to asbestos at work [24] contains the following elements: general provisions of employer obligations, exposure limits and prohibitions, evaluation and control of working environments, general organization and technical measures for prevention, personal hygiene, personal protection equipment for respiratory specific situations and certain activities, development and processing of working plans and training, information, consultation, participation, and monitoring of workers health.

This directive includes the maintenance, repair and removal of existing asbestos containing materials in equipment, units (such as ships), facilities and structures.

Worker exposure is allowed up to a concentration of airborne asbestos at or below the Permissible Exposure Level (PEL) of 0.1 fibers per cubic centimeter (f/cm^3), as an 8-hour time-weighted average (TWA).

Since January 1st, 2011, the new installation of materials containing asbestos has been prohibited for all ships covered by SOLAS Convention, and, consequently, the PEL has been limited to 0.0 f/cm^3 onboard vessels.

The present restricted exposure levels have been reached based on the results obtained by medical research into asbestos-related diseases, which has strongly recommended reducing exposure levels. EU countries have followed this tendency over the last 50 years. As a matter of comparative interest, in Spain, the original permissible exposure level in 1961 was 175 f/cm^3 for all varieties of asbestos. In 1982 it was reduced to 2.0 f/cm^3. In 1984 the PEL was once again reduced to 1.0 f/cm^3 for all varieties except for crocidolite or blue asbestos, the latter having been reduced to 0.0 f/cm^3 for. In 1993, the PEL was fixed at 0.6 f/cm^3 for chrysotile, 0.3 f/cm^3 for other varieties, and 0.0 f/cm^3 for crocidolite or blue asbestos. Although the PEL is now fixed at 0.1 f/cm^3 in a working environment, there are still an estimated significant number of between 3 and 4 asbestos-related deaths per 1000 workers.

4. Presence of ACMs on board ships

Due to the particular characteristics of asbestos, its most frequent applications on board ships are the following: bulkhead and deck insulation (blankets, panels and sprayed insulation); wall and ceiling panels (sandwich type); floating floors; floor tiles; cement, adhesive-like glue (e.g., mastics) and fillers; packing in pipe/cable penetrations; seals and sealing putty; gaskets (in pipe flanges and manhole covers of tanks); strings fastening insulation around valves in pipes; textile fabric in HVAC ducts as vibration dampener material; boiler insulation; steam pipe insulation; exhaust duct insulation; vanes used in rotary vane compressors; vanes used in rotary vane vacuum pumps; electrical cable material and fuses; friction material of brakes; paintings; fire blankets.

Since the largest amount of ACMs on board are installed as bulkhead/deck/pipe insulation, the most common places in which this material can be found is in the accommodation (or living area) and in the engine room.

Asbestos containing materials can be classified into two categories: friable and non-friable. The friable ACMs have their fibers weakly bound and, when dry, can be easily crumbled, pulverized and reduced to powder by hand pressure. When these microscopic mineral fibers are released into the air, human health can be negatively impacted via inhalation.

This type of asbestos is commonly used as insulation on exhaust pipes, as insulation in blanket forms on bulkheads and decks, as a rope for fastening isolation around valves, and as the filling material in wall and ceiling panels [Figs. 2 and 3]. The effect of typical vessel vibrations can cause the small fibers arising from asbestos or ACMs to remain suspended in the air.

In the following void spaces on ships, the content of fibers or asbestos dust in the air is likely to be high: the backside of partition walls in front of bulkheads; the space between partition walls; the space between ceilings and upper decks; the space between floating floors and decks; and inside escape trunks.
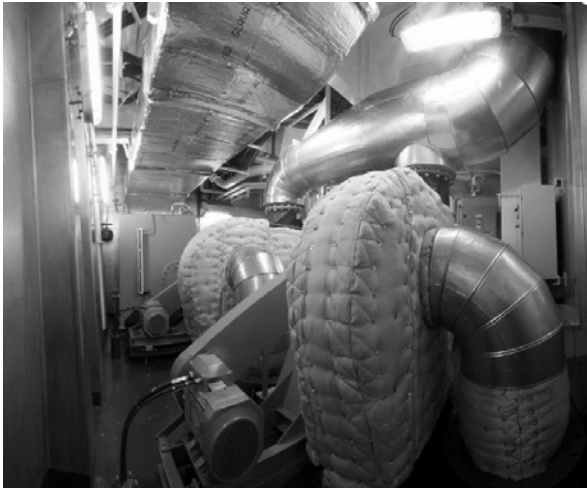
Figure 2. Insulation asbestos material covering equipment. On board arrangement.
Source: The authors

Figure 4. Floor covering.
Source: The authors

Figure 3. Lined bulkhead.
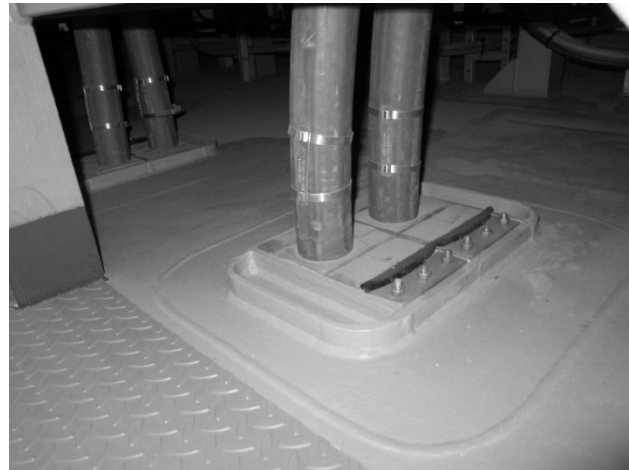Source: The authors

Figure 5. Sealed cables.
Source: The authors

As these spaces are not for regular crew access, they are usually not ventilated and friable ACMs with no cover are used in them. With ship vibrations, the fibers tend to release accumulating asbestos inside these spaces and create hazardous atmospheres that pose a serious risk when opening walls, ceilings or floors.

Moreover, non-friable ACMs are characterized by greater compactness of their fibers. Due to this, they cannot be simply pulverized by hand pressure. Hence, their level of danger to health is lower than for the friable, non-releasing fibers with the ship vibrations.

However, the non-friable asbestos can become friable and release fibers into the air when it are subjected to mechanical works such as sanding, cutting, grinding, drilling, especially when these are carried out with inadequate tools, or when they are burned or welded [Figs. 4 and 5].

This type of asbestos is often used as gaskets in pipe flanges and as manhole covers of tanks, in floor tiles (see Fig. 4), and paintings. The packing used in cable penetrations (see Fig. 5) and pipe penetrations is also non-friable.

5. Treatment of materials on board

In order to study the treatment and control of ACMs on board a ship, we will distinguish between two different cases:
I. Ships containing ACMs that do not breach the SOLAS regulation II-1/3-5 [3], either because they were constructed before July 1st, 2002 and their ACMs do not pose a risk to the crew's health, or because, having being built later, the ACMs are on the list of the exceptions allowed by the regulation issued on that date.
II. Vessels that contain ACMs on-board, are breaching the SOLAS regulations II-1/3-5. This means ships built on or after July 1st, 2002 and that contain ACMs not covered by the exceptions of this rule, or those ships built on or after January 1st, 2011 and that contain any type of ACMs on board, are clearly in contravention of [4].

In compliance with MSC/Circ.1045 [5], which was dated on May 28th, 2002, ships belonging to *group I* should implement in their Safety Management System (developed in compliance with the ISM Code) a maintenance and control

of ACMs program on board in order to minimize the exposure to asbestos of any person on board (passengers, crew, port stevedores, surveyors, etc.) while the vessel is "in service". The shipping company must identify and locate the presence of ACMs on board and make an assessment of their condition. This assessment is to be completed using an evaluation checklist shown in Circular Appendix 1 in Annex 1. This includes the materials accessibility, their protection characteristics, their degree of degradation, their exposure to shocks and vibration, and the presence of air currents in the area. Then, using Appendix 2 in Annex 1, a diagnosis of the conservation state of these materials shall be drawn up. Based on this diagnosis, the materials will be subject to a periodic check of the conservation state, to a monitoring of asbestos dust levels in the surrounding area, and to on board protection, repairing, or removal works.

When the material conservation or the asbestos fiber levels into the air make it appropriate, the following precautions are necessary: control (abatement actions) fiber release (friable asbestos) and maintain a safe health environment without resorting to remove onboard material. These actions include:

- **Enclose asbestos in some airtight material.** This could be a plastic wrapping taped over steam pipes, a new floor over an existing one, a steel plate put up over the exposed fibers of an insulation panel, etc.
- **Encapsulate asbestos in a material that will solidify and make them it friable.** This strategy only works when the surface of the ACM is slightly porous. A penetrating material should be used for the first layer. The finished material should be sprayed on rather than rolled or brushed to reduce the chances of altering the asbestos.

None of these methods should be used as a substitute for asbestos dismantling and removal in the case that it is in an advanced state of deterioration. They should be used only when asbestos is in a good state of repair.

Ships in *group II* are generally those that either, as a result of a repair made in a country where the asbestos is not prohibited, or an error in the provision of spare parts for the ship, have ACMs on board. The SOLAS regulation II-1/3-5 [4] is therefore broken. In this case, and according to MSC.1/Circ.1374 [6], the removal of such materials must be made within a period not exceeding 3 years from the date at which the infringement was detected. This action must be developed under the guidance and supervision of the Flag State.

6. Asbestos removal and recycling

Due to the danger in the direct contact with ACMs, asbestos dismantling and removal should be conducted by specialized companies that handling these materials, or by shipyard workers who have received special training.

Asbestos is listed in Annex VIII (List A) of the Basel Convention [26], and it is considered a hazardous waste for health, and therefore, should not be re-used or re-cycled. The potential health impacts associated with the use of ACMs are of such a serious nature that it is necessary to take extreme caution when drawing and handling these materials. Precautions include: protection of workers when removing

asbestos from the vessel; management of asbestos as waste; and measures to prevent asbestos from being used in other applications.

It is recommended that shipyards belonging to countries whose national laws do not address the above precautions implement a management plan for asbestos in their waste management plan. This plan will include an inventory of ACMs on board the vessel so that they can be located, quantified and identified before being removed (instead of extracting samples for analysis in a laboratory of all materials suspected of containing asbestos, it might be more feasible and economical to assume that such materials contain asbestos). In addition, this plan should include personal protective equipment (PPE) for workers removing the material and procedures for both removal and disposal. National legislation should determine the maximum permissible exposure levels (PEL).

Workers involved in the removal and management of ACMs will have to use certified personal protective equipment (PPE) with appropriate respiratory protective equipment (RPE) as well as protective clothing such as overalls, gloves and head coverings, face shield or vented goggles, and appropriate footwear. To comply with the rules of hygiene at work, the yard should have decontamination areas, equipped with a room to leave the equipment and work clothes, a shower area, and a dining room.

In asbestos extraction or removal works on a ship it is essential to keep the asbestos wet during such operations in order to avoid dispersion of fibers into the air. Thus, in this type of work, two people should be involved: one will ensure that asbestos is wet during removal, and the other one will be in charge of the extraction work.

The ACMs removed from a vessel will be introduced in airtight containers with lids and labeled for the transport to the disposal area. Generally, the ACMs are disposed of in landfills where they are buried underground.

7. Actions to be taken by the port state control inspectors (PSC) and classification society surveyors

MSC.1/Circ.1374 [6] underlines the importance of the proper training of inspectors/surveyors in the identification of ACMs on board ships. Asbestos can be found throughout ships from the top of the bridge to the bilge. Identifying the location and type of asbestos and assessing its condition involve qualified inspectors/surveyors. They should know the precautions to be taken in the presence of this material and should also be trained in taking samples.

As in many cases when there are suspicions that a particular material may contain asbestos (see Fig. 8), the visual identification of asbestos is not easy. If sampling and analysis of such material by experts cannot be conducted, both crew and inspectors/surveyors should consider the material to be an ACM in order to avoid possible health risks.

In the aforementioned Circular it is also recommended that inspectors make sure that whenever a new material is installed on board, it is delivered with a statement of compliance with a SOLAS regulation II-1/3-5, or similar. This may take the form of an "asbestos free declaration".

Table 1.
SWOT analysis.

Strengths	Weaknesses

- The Regulation MSC.1/Cir.1374 II-1/3-5 strongly contributes to global control of the presence of asbestos on board ships.
- Vessels certified as "free of asbestos" are guaranteed to have no risk of exposure to asbestos. This will appear in its Safety Management System.
- Avoid asbestos in new construction and repairs; this means controlling the risk in shipyards and for workers and surveyors, which implies a reduction and / or elimination of diseases resulting from the use of asbestos.
- Reducing exposure to asbestos is a participatory process in which all stakeholders play an important role. Demands will be made from shipyards, shipping and inspection bodies, and there will be criminal, civil and administrative liabilities that arise from illness and death of those exposed.
- There will be monitoring and verification by the MOU surveyors to make sure that these new preventive measures are being fulfilled.
	- For ship-owners, the removal of asbestos from the ship represents, in general, a high cost, and in many cases it will be impractical for the following reasons:
	   Such withdrawal must be conducted by a certified company and specialized and under the supervision of the ship's flag.
	   During this process, the ship must remain immobilized.
	- There is a lack of knowledge about the risks and consequences of the use of asbestos as well as lack of training of preventive measures to control the risks.
	- The IMO's actual sanctioning capacity.
	- For shipyards:
	   · It will be necessary to implement specific measures to control the risks of exposure to asbestos.
	   · They will need to implement the prevention of this risk in their management system, and the management of toxic and hazardous wastes.
	   · They will need to enable decontamination areas to be used exclusively for these exposed workers.
	- For the inspection services
	   · There will need to be training, awareness and personal protective measures.

OPPORTUNITIES
- For the ship-owners:
   · There is a new challenge in the implementation of these new requirements, in the Safety Management System, and identifying, locating and assessing risk.
   · Crews need to be trained and make aware of this risk.
-      These preventive measures may join the Voluntary IMO Member State Audit Scheme.
-      The safety management and safe operation of ships should be improved and be in line with the International Safety Management Code (ISM).
- Vessels and operators that meet these standards may incorporate this information in the European Quality Information System (EQUASIS).
	THREATS
	- In many countries:
	   · Asbestos is not prohibited.
	   · There is a lack of awareness and political will to legislate on the control of exposure to asbestos.
	- Port states that allow the use of asbestos in ground facilities will not sanction its use on ships.
	- There is a lack of awareness and willingness to implement the measures by shipyards and ship owners.
	- There are ship owners with substandard ships and flags of convenience.
	- The assumption of obligations by ship owners and shipyards is a drastic change in risk management.

Source: The authors

The following precautions should be adhered to if inspectors/surveyors have to undertake their work on a ship on which repairs or other work such as grinding, welding or blasting is being carried out, and consequently, friable and non-friable ACMs may be affected.

- Do not enter in those atmospheres that may contain asbestos fiber concentrations in suspension until they have been properly ventilated. To ensure that the level of fibers in the air is within the levels permitted by local laws, measurements will be taken by experts.
- The inspectors/surveyors will carry a particle filtering half mask and disposable overalls as PPE to protect themselves in case they walk in an area containing high levels of asbestos fibers. In this case, the PPE will be used to leave immediately the area.
- The above PPE should not be ever considered as protective equipment for inspectors/surveyors who intend to enter a contaminated area and stay there performing their inspection work. Entry into these spaces will be reserved for specialized companies.
- Sampling and analysis of ACMs will be left to experts. The analysis of a sample to detect whether it is an ACM or not should be performed in a laboratory.

8. Conclusions

Research in recent years on the health effects of workers' exposure to asbestos has shown that this is a carcinogenic material and, therefore, poses a serious risk to health. As most industrialized countries have becoming aware of this, they have regulated the use of this material, which has had drastic reductions in the permissible level of exposure (PEL). *This pattern of reduction in both the level of exposure and the use of these materials in the industry is the expected trend in the coming years.*

In the field of shipping, the recent appearance of ACMs on ships certified as "asbestos free" has made alarms bells ring out, pushing the IMO to take a decisive step with their prohibition of the new installation of ACMs on board all ships on January 1st, 2011. In the fulfillment of this new regulation ship-owners will play a key role. *Asbestos will have a significant entry path into the vessels through shipyard repairs or purchases of spare parts in countries that are not Member States of the IMO or whose national laws do not control the use of these materials.* Therefore, in these cases, it is the ship-owners responsibility to be alert and to prevent asbestos from accessing their ships.

Owners of vessels that contain asbestos (vessels belonging to group I of item 5) will have to identify their type and location, assess their condition, and comply with a maintenance and control program. The friable ACMs are the most dangerous and the tendency is to wrap them into other material or encapsulate them while they are in an acceptable condition and then remove them from the ship when they deteriorate. Thus, void spaces on the backside of partition walls and above ceilings are areas likely to contain high concentration of airborne fibers, this being a considerable potential risk.

Inspectors/surveyors obtaining proper qualifications to be able to identify asbestos is another crucial point in tackling the challenge of getting ships asbestos free. Once identified on board, a professional decision should be taken as to whether to remove, encapsulate, or leave the asbestos undisturbed. If samples are taken, they shall be subsequently analyzed in duly equipped laboratories.

Captains and officers must know and be concerned about the importance of the presence of asbestos on board. Crews and workers on board should take into account authorities' recommendations. It is the responsibility of captains and officers to inform crew and workers as they sometimes do not have complete information regarding asbestos.

Asbestos removal operations requires the preparation of specific working protocols, workers specially trained and informed of the risks they are exposed to, and finally, its treatment by managers trained in hazardous wastes. The flag State of the vessel surveyors must be involved in its removal.

Despite great advances over the last decade, the war against asbestos in the maritime sector will not be won until there is full global commitment for asbestos control from all stakeholders, especially from the governments of the least developed countries.

A SWOT analysis (Strengths, Weaknesses, Opportunities, and Threats) permits the strengths, weaknesses, threats and opportunities, of the implementation of the measures contained in Table 1 to be determined.

According to previous analysis, it is possible to determine the most appropriate strategies to reach a successful conclusion of the IMO rules.

Offensive strategies (strengths & opportunities): Assumption by ship-owners, shipyards, the MOU surveyors and operators, goodness and necessity of implementation of these preventive measures.

Defensive strategy (weaknesses & threats): with the conviction that these preventive measures will favor the health of crews, it will be possible to conduct risk assessments for exposure to asbestos, specific medical examinations in shipyard workers, crews and surveyors, and monitor the implementation of preventive measures.

Reorientation strategy (weaknesses & opportunities): the change in the provisions of the IMO should be used to manage these risks by all parties, and in particular by carriers with international information systems on improving the management of their vessels.

Survival Strategy (weaknesses & threats): the steps initiated by the IMO confirmed that there was a bad situation in terms of working conditions with asbestos. It is necessary to exploit the MSC.1/Cir.1374 II-1/3-5 and promote a change in the way of thinking, working procedures, and the use of new available materials that do not contain asbestos. Maritime organizations' sanction capabilities of IMO member states should be fully exploited.

References

[1] WHO, Summary Consensus Report of WHO Workshop on mechanisms of fibre carcinogenesis and assessment of chrysotile asbestos substitutes, Lyon, France. Geneve, World Health Organitation, 8-12 November.2005.

[2] European Union, Directive 2009/148/EC of the European Parlament and of the Council on the protection of workers from the risks related to exposure to asbestos at work, 30.11. 2009.

[3] SOLAS, regulation II-1/3-5, New installation of materials containing asbestos, July1st. 2002.

[4] SOLAS, regulation II-1/3-5, New installation of materials containing asbestos, January1st. 2011.

[5] International Maritime Organization (IMO). MSC/Circ.1045, Guidelines for maintenance and monitoring of on-board materials containing asbestos, 28.05.2002.

[6] International Maritime Organization (IMO), MSC.1/Circ.1374, Information on prohibiting the use of asbestos on board ships, 03.12.2010.

[7] International Labour Office Geneva, International Labour Conference, 90. to Session 2002, Report V (1), Recording and reporting of accidents work and occupational diseases and ILO list occupational diseases, Geneva. 2001

[8] European Union, Commission Recommendation 2003/670/EC of 19 September 2003 on the European list of occupational diseases [Notified under document number C (2003) 3297] (Text with EEA relevance), Official Journal L 238, 25.09.2003.

[9] European Union, Directive 76/769/EEC Council of 27 July 1976 on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations, OJ L 262, 27.9.1976.

[10] European Union, Council Directive 83/478/EEC of 19 September 1983 amending for the fifth time (asbestos) Directive 76/769/EEC on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations. OJ L 263, 24.9.1983.

[11] European Union, Commission Directive 91/659/EEC of 3 December 1991 adapting to technical progress Annex I to Council Directive 76/769/EEC on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations (asbestos), OJ L 363, 31.12.1991.

[12] European Union, Commission Directive 1999/77/EC of 26 July 1999 adapting to technical progress for the sixth time Annex I to Council Directive 76/769/EEC on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations (asbestos), Official Journal of the European Communities 6. 8. 1999

[13] Asbestos International Association (A.I.A.), Reference method for the determination of airborne asbestos fibre concentrations at workplaces by ligth microscopy (Membrane filter method). Recommended technical method number 1. 1982.

[14] Ministerio de Trabajo y Asuntos Sociales de España. MTA/MA-051/A04, Determinación de fibras de amianto y otras fibras en aire. Método del filtro de membrana. Microscopía óptica de contraste de fases. (Método multifibra).

[15] Ministerio de Trabajo y Asuntos Sociales de España. Criterio/Recomendación CR-02/2005. Medida fiable de las concentraciones de fibras de amianto en aire. Aplicación del método de toma de muestras y análisis MTA/MA-051/A04.

[16] National Institute for Occupational Safety and Health (NIOSH). Analysis method 7400. Asbestos fibres. 1984.

[17] Schlecht., P.C. and Shulman, S.A., Performance of asbestos fiber counting laboratories in the NIOSH proficiency analytical testing (PAT) program, American Industrial Hygiene Association Journal, 47(5), pp 259-269, 1986. DOI: 10.1080/15298668691389748

[18] Baron, P.A., Shulman, S.A., Evaluation of the magiscan image analyzer for asbestos fiber counting, American Industrial Hygiene Association Journal, 48, pp. 39-46,1987. DOI: 10.1080/15298668791384346

[19] European Union, Council Directive 83/477/EEC of 19 September 1983 on the protection of workers from the risks related to exposure to asbestos at work (second individual Directive within the meaning of Article 8 of Directive 80/1107/EEC), Official Journal L 263, 24.09.1983.

[20] European Union, Council Directive 90/394/CE, of 28 June 1990 on the protection of workers from risks related to exposure to carcinogens at work (Sixth individual Directive under paragraph 1 of Article 16 of Directive 89/391/EEC), Official Journal L 196, 26.07.1990.

[21] European Union, Council Directive 97/42/EC of 27 June 1997 amending for the first time Directive 90/394/EEC on the protection of workers from risks related to exposure to carcinogens at work (Sixth individual Directive under paragraph 1 of Article 16 of Directive 89/391/EEC), Official Journal L 179, 08.07.1997.

[22] European Union, Directive 1999/38/EC of 29 April 1999 the second time Directive 90/394/EEC on the protection of workers from exposure to carcinogens at work and extending it to mutagens, Official Journal L 138.OJ. 1.6.1999

[23] European Union, Council Directive 91/382/EEC of 25 June 1991 amending Directive 83/477/EEC on the protection of workers from the risks related to exposure to asbestos at work (second individual Directive within the meaning of Article 8 of Directive 80/1107/EEC), OJ L 206, p. 16–18. 29.7.1991

[24] European Union, Directive 2003/18/EC of the European Parliament and of the Council of 27 March 32003 amending Council Directive 83/477/EC on the protection of workers from the risks related to exposure to asbestos at work (Text with EEA relevance) OJ L 97, 15.4.2003

[25] International Labour Organisation (ILO). C162 Convention, concerning Safety in the Use of Asbestos. 1986.

[26] United Nations Environment Programme (UNEP). Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and their Disposal, Sixth meeting, Geneva, 9-13.12.2002.

L.M.Carral-Couce holds a PhD. in Naval Architecture and Marine Engineering from the Universidad Politécnica de Madrid (Technical Univ. of Madrid) and a MSc.in Maritime Law and Port Administration from the University of Coruna, Spain. He has worked in the former navy shipyard of Bazan (now Navantia) prior to entering the EPS-Ferrol at the University of Coruna, Spain, as a Professor. At the University he is part of the GII (a multidisciplinary research group) and he has promoted several Masters' degrees such as Design, Production and Inspection of Motor and Sailing Craft, and Transport & Logistics. He has also been Chairman of the Spanish Association of Naval Architects in the region of Galicia for several years.
ORCID 0000 – 0003 – 1109 – 1131

J.A. Fraguela-Formoso, PhD in Naval Architecture and Marine Engineer from the Polytechnic University of Madrid, is Professor at the University of A Coruña, Spain. Quality deputy director of the Polytechnic School of the University of A Coruña. Director Course at Second University Cycle of the University of A Coruña Degree in Management of Prevention, Quality and Environment, from 2000 to 2010. Creator, Organizer and Director 10 Masters in Prevention Labor risks of the University of a Coruña, two of them in Ibero America. Creator, Organizer and Manager 7 Masters in Integrated Quality Management, Environment and Occupational Health and Safety at the University of A Coruña. He has been head of the Services Occupational Health and Safety and Occupational Health at the National Military Bazán company Construcciones Navales S.A. Ferrol. Chairman of ENOGA (Association of Enxeñeiros Navais and Oceania Galicia), IPEN-SPAIN (Pan American Institute of Naval Engineering). Member of the Advisory Council of IPEN-AMERICAS. Member of the Editorial Board of IPEN digital JOURNAL (ISSN 1011-5951) of the Pan American Institute of Naval Architecture and Marine Engineering
ORCID 0000-0002-5308-387X

P. Fariñas-Alvariño, is PhD. Naval Architect and Marine Engineer from Universidade da Coruña, Spain, since 2004. He has been professor in Heat Engines, Fluid Mechanics and Shipbuilding. Currently he is professor of naval hydrodynamics and computational hydrodynamics. Develops research work in transport phenomena in fluids focused mainly to energy transport. In recent years he has developed own models for CFD study of nanotechnology applied to fluid mechanics for the study of suspensions.
ORCID 0000-0002-9598-5249

J.L. Fernández-Soto, is PhD. Naval Architect and Marine Engineer from the University of La Coruna, has developed its activity in the EN Bazán (Navantia) and is now surveyor of the company DNV - GL.
ORCID 0000 - 0002 - 9807 – 2007

UNIVERSIDAD **NACIONAL** DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Medio Ambiente

Oferta de Posgrados

Especialización en Aprovechamiento de Recursos Hidráulicos
Especialización en Gestión Ambiental
Maestría en Ingeniería Recursos Hidráulicos
Maestría en Medio Ambiente y Desarrollo
Doctorado en Ingeniería - Recursos Hidráulicos
Doctorado Interinstitucional en Ciencias del Mar

Mayor información:

E-mail: acia_med@unal.edu.co
Teléfono: (57-4) 425 5105
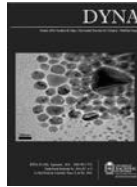
Contributions to the online fault diagnosis of interturn short circuit in three-phase induction motor by of means negative sequence components

Javier Ernesto Micolta [a], José Luis Oslinger [a] & Edinson Franco [a]

[a] Facultad de Ingeniería, Universidad del Valle, Cali, Colombia. micolta.javier@gmail.com, jose.oslinger@correounivalle.edu.co, edinson.franco@correounivalle.edu.co

Abstract
In this article the current and the negative sequence impedance as fault indicators inter-turn short-circuit of an induction motor squirrel cage is used. The effect of voltage unbalances and different load levels in these diagnostic parameters is also discussed, which are calculated from data obtained from simulations using the Finite Element Method (FEM) and laboratory experiments.

Keywords: Induction motor fault diagnosis, negative sequence current, negative sequence impedance, finite element method, voltage unbalance, CVUF factor.

Aportes al diagnóstico en línea de fallas de cortocircuito entre espiras de un motor de inducción mediante parámetros eléctricos de secuencia inversa

Resumen
En este artículo se emplea la corriente y la impedancia de secuencia negativa como indicadores de fallas de cortocircuito entre espiras de un motor de inducción tipo jaula de ardilla. También se analiza el efecto que tienen los desequilibrios de tensión y los diferentes niveles de carga en dichos parámetros de diagnóstico, los cuales se calcularon a partir de datos obtenidos de simulaciones por medio del Método de Elementos Finitos (MEF) y experimentos de laboratorio.

Palabras clave: Motor de inducción, diagnóstico de fallas, Corriente de secuencia negativa, impedancia de secuencia negativa, Método de elementos finitos, desequilibrio de tensión, factor CVUF.

1. Introducción

Las máquinas de inducción tipo jaula de ardilla constituyen más del 90 por ciento del total de las máquinas eléctricas rotativas [1], esto se debe a su bajo costo, robustez y bajo mantenimiento comparado a otro tipo de máquinas eléctricas rotativas, aun así, éstas no se encuentran exentas de fallas, las cuales causan paradas en los procesos de producción y por lo tanto pérdidas económicas.

Las fallas en el motor de inducción están distribuidas principalmente en el rotor, los rodamientos y el estator, donde éstas últimas constituyen unas de las causas de daño más común [2-9]. De ahí la importancia de emplear métodos de diagnóstico para este tipo de fallas.

Entre los diferentes métodos de diagnósticos de fallas estatóricas [10], Los métodos de la corriente y la impedancia de secuencia negativa tienen la ventaja de ser técnicas no invasivas y de bajo requerimiento computacional [11], y

aunque son confiables, se pueden presentar errores en el diagnóstico debido al efecto de los desequilibrios de tensión en los parámetros de secuencia inversa.

Diferentes autores han estudiado el efecto que tiene el factor desequilibrio de tensión en los parámetros de secuencia negativa, donde como resultado muestran las variaciones de dichos parámetros en función del factor VUF (factor de desequilibrio de tensión) [11-18] pero como muestran algunos estudios, existe una cantidad infinita de desequilibrios de tensión con igual factor. Por lo anterior, útilmente la comunidad científica está empleando un nuevo factor denominado CVUF (factor complejo de desequilibrio de tensión), el cual caracteriza el desequilibrio de tensión por magnitud y ángulo [19-24].

En este artículo se analizará el efecto de las variaciones de la magnitud y ángulo del factor CVUF en el módulo de la corriente y la impedancia de secuencia negativa del motor de inducción cuando éste se encuentra bajo fallas de cortocircuito entre espiras.

2. Cálculo de los parámetros de secuencia negativa

Las componentes de secuencia negativa se determinan por medio de la transformación de Fortescue, donde las tensiones y las corrientes de secuencia se obtienen de las tensiones y corrientes de fase por medio de las ecuaciones:

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \frac{1}{3} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{bmatrix} \cdot \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} = \frac{1}{3} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{bmatrix} \cdot \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (2)$$

Donde V_a, V_b, V_c, I_a, I_b, I_c corresponden a los valores de fase de tensión y corriente expresados en forma fasorial y V_0, V_1, V_2, I_0, I_1, I_2 son las componentes de secuencia, donde I_2 es empleada como indicador de falla [25].

Debido a que I_2 aparece cuando el motor de inducción presenta fallas y también cuando es alimentado con fuentes desequilibradas, es necesario dividir la corriente de secuencia negativa calculada en dos componentes:

$$I_2 = I_2^f + I_2^V \quad (3)$$

Donde I_2^f e I_2^V corresponde a la componente de la corriente de secuencia negativa generada por la condición de falla y a la componente de la corriente de secuencia negativa asociada al desequilibrio de tensión de la fuente de alimentación respectivamente. Para calcular I_2^V, se emplea el modelo matemático del motor de inducción sin falla alimentado con fuentes desequilibradas [13,17] o la impedancia de secuencia negativa, Z_2^s, del motor real sin falla, como lo muestra la siguiente ecuación [16].

$$I_2^V = \frac{V_2}{Z_2^s} \quad (4)$$

Para el cálculo de la impedancia de secuencia negativa, usada como parámetro de diagnóstico se emplea la siguiente expresión [26]:

$$Z_2 = \frac{V_2}{I_2} \quad (5)$$

3. Cálculo del desequilibrio de tensión

Para el cálculo del desequilibrio de tensión, se empleará el factor CVUF (factor complejo de desequilibrio de tensión) propuesto por Wang [19,20] el cual además del módulo, incluye la diferencia angular entre la tensión de secuencia negativa y positiva, como se muestra en la siguiente expresión:

$$CVUF = \frac{V_2 \angle \theta_2}{V_1 \angle \theta_1} \quad (6)$$

$$CVUF = VUF \angle \theta_V \quad (7)$$

Donde

V_1: tensión se secuencia positiva
V_2: tensión se secuencia negativa
θ_2: ángulo de la tensión de secuencia negativa
θ_1: ángulo de la tensión de secuencia positiva
θ_V: ángulo del CVUF, que establece que tan separados están los fasores de secuencia positiva y negativa.

De lo anterior se puede decir que el factor VUF por sí sólo no da información completa debido a que existe una cantidad infinita de desequilibrios de tensión con igual modulo.

La Fig. 1 muestra las tensiones de fase en función del ángulo θ_V para $V_1 = 127$ V y VUF = 3%. En ella se puede observar que hay un número infinito de distribución voltajes con igual módulo de desequilibrio de tensión, por lo tanto, para hacer una mejor caracterización es necesario considerar el ángulo que hay entre la tensión de secuencia negativa y la tensión de secuencia positiva.

4. Resultados de simulación y laboratorio

A continuación se muestran los resultados obtenidos empleando el método de corriente de secuencia negativa y el método de impedancia de secuencia negativa como herramientas para detectar fallas de cortocircuito entre espiras
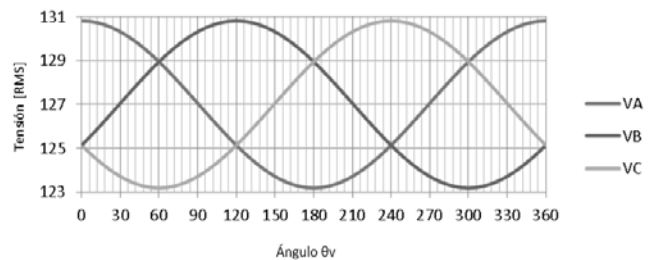

Figura 1. Distribución de tensiones en función de θ_V para un VUF = 3%.
Fuente: Los autores.

de un motor de inducción tipo jaula de ardilla. En los resultados se presentan datos de laboratorio y de simulación, donde estos últimos se obtuvieron empleando el método de elementos finitos (MEF) por medio del software Flux2D®.

En la Tabla 1, se muestran los datos del motor de inducción usado en el estudio, La Fig. 2, muestra el enmallado del motor de inducción en elementos finitos y la Fig. 3, muestra el motor real acoplado al generador de CC con el cual se aplicaron los diferentes niveles de carga.

Tabla 1.
Características del motor de inducción.

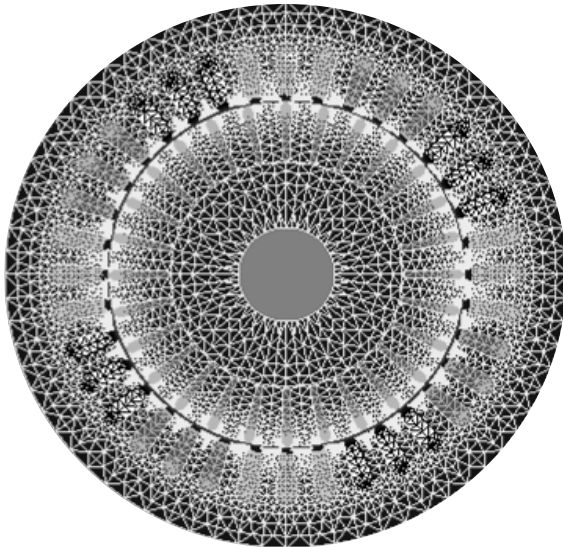|		
Marca	:	Westinghouse
Tipo	:	Asíncrono tipo jaula de ardilla NEMA B
Potencia	:	3 HP
Tensión	:	220/440 V
Corriente	:	4,2/8,4 A
Fases	:	3
Velocidad	:	1740 RPM
Frecuencia	:	60 Hz

Fuente: Los autores.

Figura 2. Modelo del motor en elementos finitos.
Fuente: Los autores.

Figura 3. Generador CC acoplado al motor de inducción.
Fuente: Los autores.

El motor bajo estudio fue acondicionado para introducir fallas de cortocircuito entre espiras, para lo cual se rebobinó y se derivaron terminales hacia el exterior de la máquina conformando una bornera.

Para registrar los datos de tensión y corriente, se empleó el equipo medidor de calidad de potencia Power Xplorer PX5. Con el fin de proteger el devanado de la máquina, se empleó una resistencia de falla de 0.14 Ω para realizar el cortocircuito entre espiras. La Fig. 4, muestra el banco de pruebas empleado en el laboratorio.

La Fig. 5, muestra los resultados experimentales y de simulación de la corriente y la impedancia de secuencia negativa en función del número de espiras en cortocircuito, de ella se puede observar que el módulo de I_2 aumenta conforme el número de espiras en cortocircuito se hace mayor, mientras ocurre lo contrario con Z_2. Lo anterior se debe a que el desequilibrio en el devanado del motor de inducción se hace mayor a medida que el grado de falla aumenta.

Para observar el efecto de las variaciones del ángulo θ_V en los parámetros de secuencia negativa, se realizaron simulaciones en elementos finitos del motor bajo diferentes espiras en cortocircuito y alimentado con un desequilibrio de tensión VUF = 3%, V1 = 127 V. Los resultados se muestran en la Fig. 6, en ellos se puede observar que para un mismo valor de VUF se pueden tener diferentes valores en los módulos de los parámetros de secuencia negativa según el número de espiras en cortocircuito y el ángulo del factor CVUF.

La Tabla 2, muestra las tensiones aplicadas al motor de inducción real y al modelo en MEF con 10 espiras en cortocircuito en la fase C bajo diferentes niveles de carga. Para obtener los diferentes valores del ángulo θ_V, la magnitud de tensión de fase más bajo debido al desequilibrio de tensión se rotó en cada una de las fases de motor, de esta manera la diferencia entre el voltaje de secuencia negativa y positiva
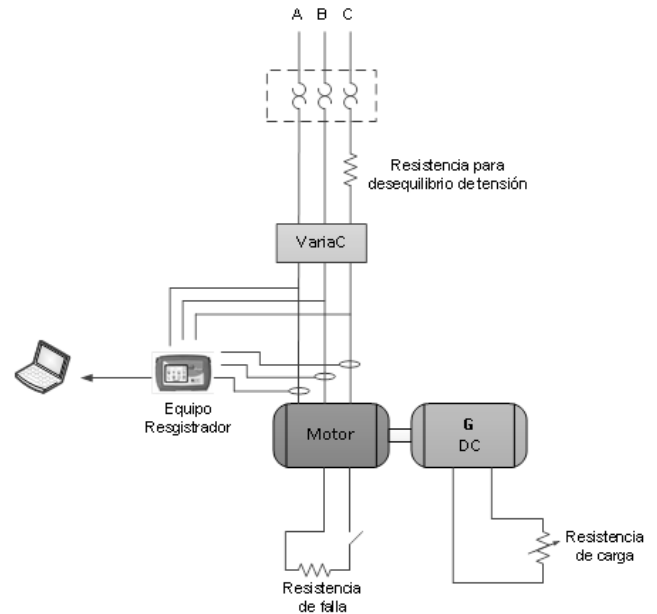

Figura 4. Esquema general del banco de pruebas del laboratorio.
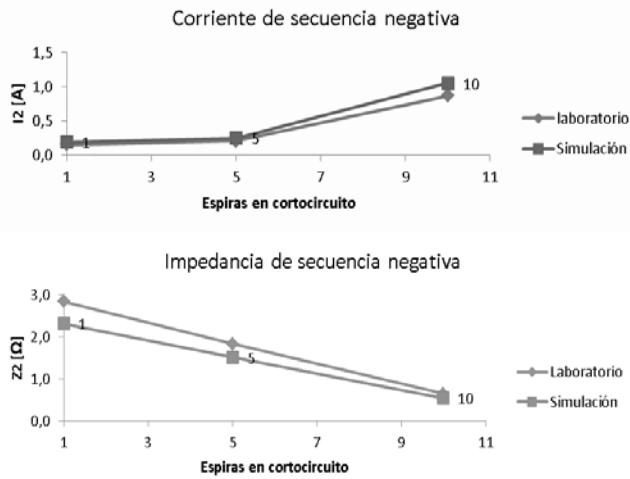Fuente: Los autores.

Figura 5. Corriente e impedancia de secuencia negativa en función del número de espiras en cortocircuito.
Fuente: Los autores.

Tabla 1.
Características tensión aplicada al motor de inducción.

Deslizamiento	Kv	Ángulo θv	V1	V2
VC menor que VA y VB				
0.010	3.97	262.18	122.68	4.87
0.015	4.24	269.72	122.78	5.20
0.019	4.81	272.37	121.58	5.84
VA menor que VB y VC				
0.005	2.85	143.24	123.49	3.52
0.010	3.36	153.39	122.52	4.11
0.015	3.93	160.51	122.02	4.80
0.019	4.07	160.42	121.59	5.07
VB menor que VA y VC				
0.005	2.64	5.82	123.92	3.27
0.010	2.92	19.25	122.66	3.58
0.015	3.51	30.00	121.68	4.27
0.019	3.77	35.13	121.85	4.59

Fuente: Los autores.

era la siguiente: para magnitudes de voltajes en la fase A menores a las demás fases se obtuvieron ángulos entre los 143° y 160°, para magnitudes voltajes en la fase B menores a las demás fases se obtuvieron valores de ángulos entre los 5° y 35°, y para magnitudes voltajes en la fase C menores a las demás fases se obtuvieron ángulos entre los 262° y 272°.

Los resultados de la corriente y la impedancia de secuencia negativa bajo las condiciones mencionadas anteriormente, se encuentran graficados en las Figs. 7 y 8, donde se puede apreciar el efecto del ángulo θ_V en los parámetros de secuencia inversa empleados para el diagnóstico de fallas. Estos cambios en I_2 y Z_2 se deben principalmente a que las variaciones de θ_V pueden producir un efecto de compensación en el motor en falla, el cual enmascara los cambios en los parámetros de secuencia negativa debido al cortocircuito y por lo tanto, producir un error en el diagnóstico de la falla en el motor.
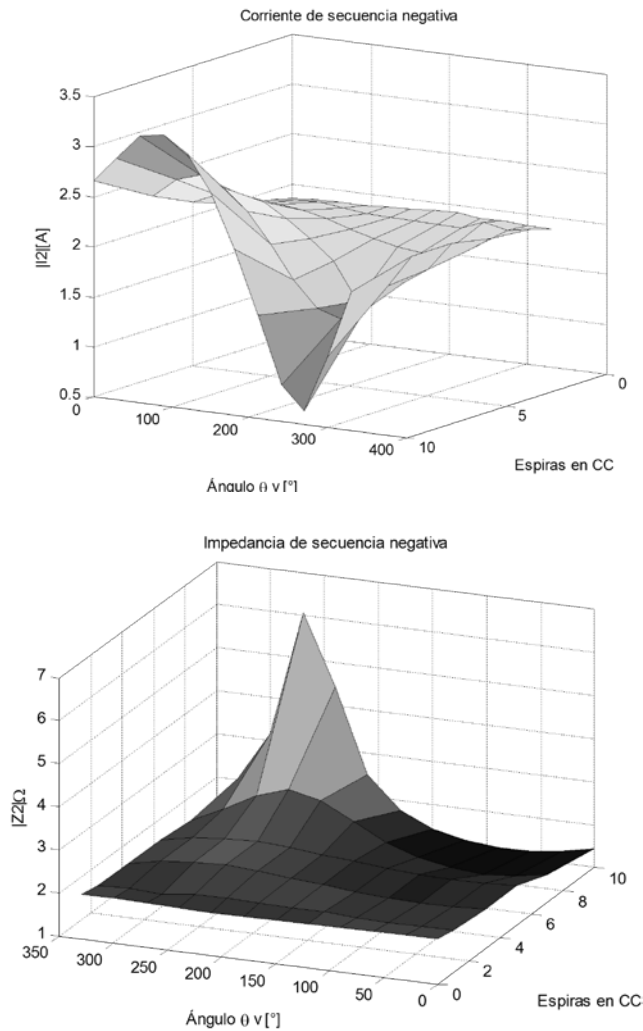

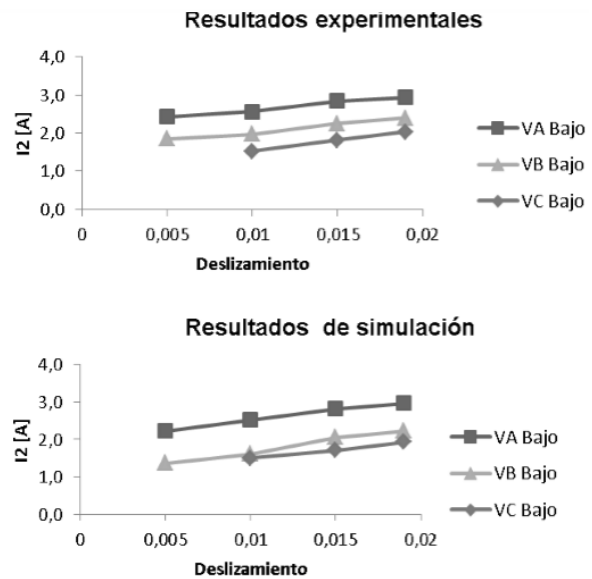Figura 1. Corriente e impedancia de secuencia negativa en función del número de espiras en cortocircuito y del ángulo del factor CVUF.
Fuente: Los autores.

Figura 2. Corriente de secuencia negativa bajo diferentes desequilibrios de tensión.
Fuente: Los autores.

Resultados experimentales
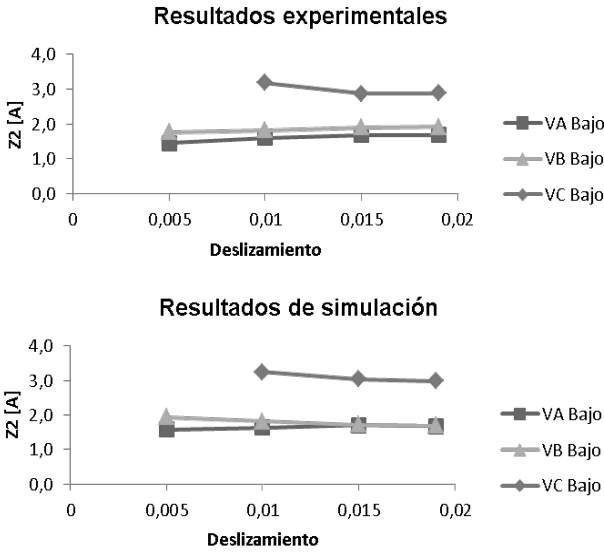

Resultados de simulación

Figura 8. Impedancia de secuencia negativa bajo diferentes desequilibrios de tensión.
Fuente: Los autores.

La Fig. 9, muestra los resultados experimentales y de simulación de la componente de corriente de secuencia negativa debida sólo a fallas en el estator, es decir I_2^f, esta corriente es calculada por medio de las ecuaciones 3 y 4. Z_2^f es obtenida por simulación y de manera experimental, sus valores son a $2.11 \angle 59.2°$ para el motor real y $2.09 \angle 51.9°$ para el motor simulado.

La expresión empleada en el análisis es la siguiente.
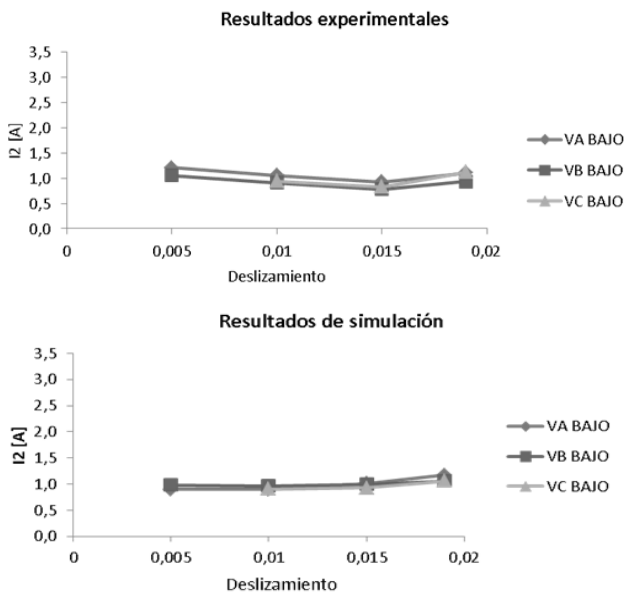
$$I_2^f = I_2 - \frac{V_2}{Z_2^s} \qquad (8)$$

Resultados experimentales

Resultados de simulación

Figura 9. Corriente I_2^f donde I_2^V es calculada por medio de Z_2^s.
Fuente: Los autores.

Resultados de simulación
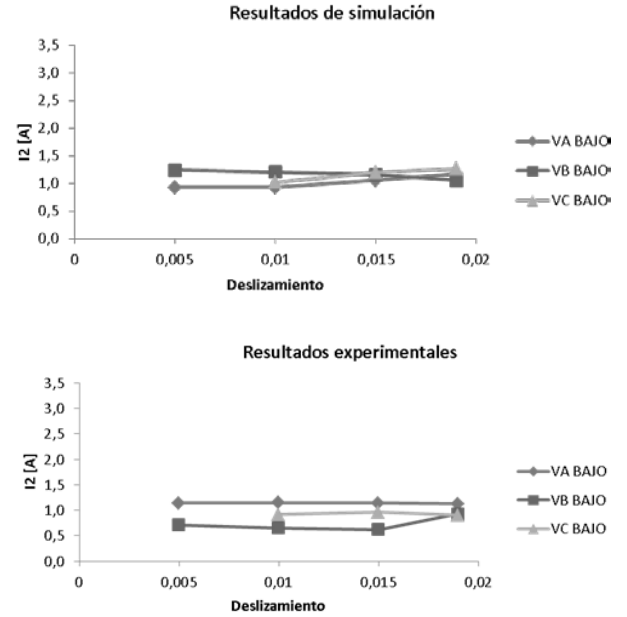

Resultados experimentales

Figura 10. Corriente I_2^f donde I_2^V es calculada por medio del modelo matemático del motor sin falla.
Fuente: Los autores.

La Fig. 10, muestra la componente de la corriente de secuencia negativa debido sólo a fallas en el estator, I_2^f, obtenida por medio de la ecuación 3. A diferencia del método anterior, la componente de la corriente de secuencia negativa debido al desequilibrio de tensión, I_2^V, es obtenida del modelo en elementos finitos (MEF) del motor sin fallas.

Lo anterior indica que I_2^V, es obtenida a partir del modelo matemático del motor (MEF) en estado sano y alimentado con fuentes desequilibradas.

En este caso la expresión empleada en el análisis es la siguiente:

$$I_2^f = I_2 - I_2^V \qquad (9)$$

5. Conclusión

Los resultados experimentales y de simulación muestran que cuando el motor de inducción se encuentra bajo falla, la impedancia y la corriente de secuencia negativa calculada a partir de las corrientes y las tensiones de fase, no solo son afectadas por las variaciones en magnitud del desequilibrio de tensión (VUF), sino también por el ángulo (θ_V) del factor CVUF, lo cual puede producir errores en el diagnóstico de falla. Por lo anterior, se emplearon dos métodos de desagregación de la corriente de secuencia negativa con el fin de obtener la componente debida solo a falla. Los resultados muestran que ésta última componente es menos sensible ante el desequilibrio de tensión.
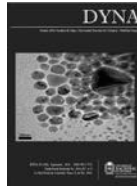
Referencias

[1]   Diaz, D., Amaya, M.C. and Paz, A., Inter-turn short-circuit analysis in an induction machine by finite elements method and field tests,

Electrical Machines (ICEM), 2012 XXth International Conference on, pp.1757-1763, 2012.

[2] Pinjia, Z., Yi, D., Habetler, T.G. and Bin, L., A Survey of condition monitoring and protection methods for medium-voltage induction motors, industry applications, IEEE Transactions on, 47(1), pp. 34-46, 2011.

[3] Paz, A., Reliability on operation of three-phase squirrel-cage induction motors on industry applicationTesis de Dr., s.a.

[4] Report of Large Motor Reliability Survey of Industrial and Commercial Installations, Part I, Industry Applications, IEEE Transactions on, IA-21(4), pp. 853-864, 1985.

[5] Albrecht, P.F., Appiarius, J.C., McCoy, R.M., Owen, E.L. and Sharma, D.K., Assessment of the reliability of motors in utility applications - updated, energy conversion, IEEE Transactions on, EC-1(1), pp. 39-46, 1986.

[6] Albrecht, P.F., Appiarius, J.C., Cornell, E.P., Houghtaling, D.W., McCoy, R.M., Owen, E.L. and Sharma, D.K., Assessment of the reliability of motors in utility applications, Energy Conversion, IEEE Transactions on, EC-2(3), pp. 396-406, 1987.

[7] Thorsen, O.V. and Dalva, M., A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries, Industry Applications, IEEE Transactions on, 31(5), pp. 1186-1196, 1995.

[8] Thorsen, O.V. and Dalva, M., Condition monitoring methods, failure identification and analysis for high voltage motors in petrochemical industry, Electrical Machines and Drives, 1997 Eighth International Conference on (Conf. Publ. No. 444), pp.109-113, 1997.

[9] Thorsen, O.V. and Dalva, M., Failure identification and analysis for high voltage induction motors in petrochemical industry, Industry Applications Conference, Thirty-Third IAS Annual Meeting. The IEEE , 1, pp. 291-298 1998.

[10] Verucchi, C.J. and Acosta, G.G., Fault detection and diagnosis techniques in induction electrical machines, Latin America Transactions, IEEE, 5(1), pp. 41-49, 2007.

[11] Bakhri, S., Ertugrul, N. and Soong, W.L., Practical issues on negative sequence current monitoring for condition monitoring in induction motors, Universities Power Engineering Conference (AUPEC), 2012 22nd Australasian, pp.1-6, 2012.

[12] Bakhri, S., Ertugrul, N. and Soong, W.L., Negative sequence current compensation for stator shorted turn detection in induction motors, IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, pp. 1921-1926, 2012.

[13] Bakhri, S., Ertugrul, N., Soong, W.L. and Arkan, M., Investigation of negative sequence components for stator shorted turn detection in induction motors, Universities Power Engineering Conference (AUPEC), 2010 20th Australasian, pp. 1-6, 2010.

[14] Kohler, J.L., Sottile, J. and Trutt, F.C., Alternatives for assessing the electrical integrity of induction motors, Industry Applications, IEEE Transactions on, 28(5), pp. 1109-1117, 1992.

[15] Kliman, G.B., Premerlani, W.J., Koegl, R.A. and Hoeweler, D., A new approach to on-line turn fault detection in AC motors, Industry Applications Conference, 1996. Thirty-First IAS Annual Meeting, IAS '96., Conference Record of the 1996 IEEE, 1, pp. 687-693,1996.

[16] Arkan, M., Perovic, D.K. and Unsworth, P., Online stator fault diagnosis in induction motors, Electric Power Applications, IEE Proceedings,148(6), pp. 537-547, 2001.

[17] Bouzid, M. and Champenois, G., Accurate stator fault detection insensitive to the unbalanced voltage in induction motor, Electrical Machines (ICEM), 2012 XXth International Conference on, pp. 1545-1551, 2012.

[18] Paz, A., Amaya, M., Palacio, J. y Díaz D., Consideraciones acerca del uso de la corriente de secuencia negativa como herramienta de diagnóstico de fallas estatóricas en motores de inducción de jaula de ardilla, 16va Convención Científica de Ingeniería y Arquitectura, Palacio de convenciones de la Habana, Cuba, 2012.

[19] Wang, Y.-J., An analytical study on steady-state performance of an induction motor connected to unbalanced three-phase voltage, Power Engineering Society Winter Meeting, 2000. IEEE, 1, pp. 159-164, 2000.

[20] Wang, Y.-J., Analysis of effects of three-phase voltage unbalance on induction motors with emphasis on the angle of the complex voltage unbalance factor, Energy Conversion, IEEE Transactions on, 16(3), pp. 270-275, 2001.

[21] Faiz, J. and Ebrahimpour, H., Precise derating of three-phase induction motors with unbalanced voltages, Industry Applications Conference. Fourtieth IAS Annual Meeting. Conference Record of the 2005, 1, pp. 485-491, 2005.

[22] Gnacinski, P., Windings temperature and loss of life of an induction machine under voltage unbalance combined with over- or undervoltages, Energy Conversion, IEEE Transactions on, 23(2), pp. 363-371, 2008.

[23] Quispe, E.C., Efectos del desequilibrio de tensiones sobre la operación el motor de inducción trifásico. Énfasis en la caracterización del desequilibrio de tensiones y el efecto sobre la potencia nominal, Tesis PhD, Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia, 2011.

[24] Filho, A.L.F., Garcia, D.C., Nascimento, F.AO. and Cormane, J.A.A., Study of voltage unbalance conditions based on the behavior of the complex voltage unbalance factor (CVUF), Transmission and Distribution Conference and Exposition: Latin America (T&D-LA), 2010 IEEE/PES, pp. 184-189, 2010.

[25] Williamson, S. and Mirzoian, K., Analysis of cage induction motors with stator winding faults, power apparatus and systems, IEEE Transactions on, PAS-104(7), pp. 1838-1842, 1985.

[26] Kohler, J.L., Sottile, J. and Trutt, F.C., Alternatives for assessing the electrical integrity of induction motors, Industry Applications Society Annual Meeting, Conference Record of the IEEE, 2, pp. 1580-1586, 1989.

J.E. Micolta, es Ing. Electricista de la Universidad del Valle, Colombia, en 2009 y MSc. en Ingeniería Eléctrica de la misma Universidad, en 2015.
ORCID: 0000-0002-0766-3507

J.L. Oslinger, es Ing.Electricista de la Universidad del Valle, Colombia, 1996. Dr. en Ingeniería de la Universidad del Valle en 2007. Profesor de la Escuela de Ingeniería Eléctrica y Electrónica (EIEE) de la Universidad del Valle desde 1996. Su área de trabajo es la teoría, diseño, diagnóstico, mantenimiento, modelamiento de máquinas eléctricas rotativas y el diagnóstico de aislamiento de grandes máquinas eléctricas rotativas. Es miembro del grupo de investigación en Conversión de Energía CONVERGIA de la Universidad del Valle.
ORCID: 0000-0001-9004-3562

E. Franco, es Ing. Electricista de la Universidad del Valle, 1991. MSc. en Automática en 1994. Dr. en Ingeniería de la Universidad del Valle en 2007. Profesor de la Escuela de Ingeniería Eléctrica y Electrónica (EIEE) de la Universidad del Valle desde 1993. Su área de trabajo en investigación son las tecnologías en sistemas de control con aplicaciones a control de sistemas electromecánicos y de procesos. Es miembro del grupo de investigación en Control Industrial GICI de la Universidad del Valle.
ORCID: 0000-0003-4045-3808

A new noncircular gear pair to reduce shaft accelerations: A comparison with sinusoidal and elliptical gears

Libardo V. Vanegas-Useche [a], Magd M. Abdel-Wahab [b] & Graham A. Parker [c]

[a] Facultad de Ingeniería Mecánica, Universidad Tecnológica de Pereira, Colombia. lvanegas@utp.edu.co
[b] Faculty of Engineering and Architecture, Ghent University, Ghent, Belgium. magd.abdelwahab@ugent.be
[c] Faculty of Engineering and Physical Sciences, The University of Surrey, Guildford, UK. G.parker@surrey.ac.uk

Abstract
This article presents a new noncircular gear pair to obtain small shaft accelerations. The centrode contours may be controlled depending on the required maximum acceleration and smoothness of the centrodes. A comparison among elliptical, sinusoidal, and the new gears is provided. Results show that, for two-lobule gears, the maximum and minimum polar radii and gear ratios are the same for the new and sinusoidal profiles but differ for the elliptical ones. Conversely, there are significant differences in the maximum angular acceleration, tangential acceleration, and pressure angle. It is concluded that the novel gears provide not only small shaft accelerations, but also small tangential accelerations and pressure angles, and it is excepted that the elliptical gears may exhibit lower tangential accelerations and pressure angles for large values of the angular speed alternating component. Consequently, shaft and tooth loads and stresses may be lower for the new gears. This may result in more compact systems.

Keywords: noncircular gear centrodes; shaft accelerations; sinusoidal gears; elliptical gears.

Nuevo par de engranajes no circulares para reducir las aceleraciones del eje: Comparación con engranajes sinusoidales y elípticos

Resumen
Se presenta un nuevo par de engranajes no circulares para obtener pequeñas aceleraciones del eje. Las formas de los engranajes pueden ser controladas dependiendo de la máxima aceleración y suavidad de los contornos requeridos. Se hace una comparación con engranajes elípticos y sinusoidales. Los resultados muestran que, para engranajes de dos lóbulos, los valores máximos y mínimos de los radios polares y de la relación de transmisión son iguales para los perfiles desarrollados y los sinusoidales, pero difieren para los elípticos. En contraste, los valores máximos de la aceleración angular, aceleración tangencial y ángulo de presión difieren. Se concluye que los nuevos perfiles proporcionan pequeñas aceleraciones angulares y tangenciales y menores ángulos de presión, excepto que los elípticos pueden exhibir menores aceleraciones tangenciales y ángulos de presión para grandes variaciones de la velocidad. Consecuentemente, las cargas y esfuerzos podrían ser menores. Esto podría resultar en sistemas más compactos.

Palabras clave: perfiles de engranajes no circulares; aceleraciones del eje; engranajes sinusoidales, engranajes elípticos.

1. Introduction

Mechanisms, such as linkages and gears, are used in a huge number of applications: several types of machines [1], devices [2], vehicles, computers, and prosthesis [3], among many others. Noncircular gears (NCGs) are a particular class of gears, the centrodes of which are non-circular; therefore, the speed ratio and the ratio between the angular displacements are not constant. For this reason, they are used for special purposes. For example, they may improve the

efficiency, function, or versatility of a mechanical system; they may be used to obtain vibrations, a certain function for the velocity of the driven shaft, or for controlling the force function in a human powered vehicle. Applications of NCGs include bicycles, agricultural mechanisms, textile machines, mechanical presses, and high torque hydraulic engines for bulkhead drives [4]. The most common types of NCGs are elliptical gears, true or high-order, and logarithmic-spiral gears. Two true elliptical gears in mesh are identical geared ellipses rotating about their foci, and they produce one speed cycle per revolution.

This article builds on a previous work [5], which reviews noncircular gears and explains in detail the main variables of NCGs. The reader can, therefore, refer to this or other works to read reviews (e.g., [6]) and descriptions (e.g., [7]) of noncircular gear variables. Despite this, all the relevant information is provided in this paper.

Compared to research on circular gears, the amount of research on NCGs is very small, and it may be the case that articles from the early twentieth century or previous to this are unavailable [8].

Most of the papers on NCGs can be categorized into three research areas: (a) new pitch centrodes that satisfy particular requirements or functions [9-20]. For example, Litvin *et al.* [18] developed NCGs to obtain various functions: when the derivative of the function changes its sign and pitch profiles that have unclosed curves. Ottaviano *et al.* [19] developed and experimentally validated a pair of NCGs to attain a prescribed motion for an application related to external blood circulation in cardiac surgery. The NCGs performance is compared with a cam-follower system, and the paper outlines the advantages and disadvantages. (b) Development of novel profiles of NCG teeth [8,21,22]. (c) Development of mathematical models to describe or manufacture NCGs teeth and their cutters [23-30]. For example, Xia *et al.* [30] developed a method for hobbing noncircular helical gears, which is based on a method with axial shift of the hob. The approach derived from this method includes two schemes: fixed and moving meshing point on hob. The advantages of the methodology are that it can increase machine tool operating range and that gears with large eccentricity and large NCGs may be hobbed by the latter.

Other areas in which research has been conducted include: undercutting analysis [23,24,29,31], pointed teeth [29], computer aided design or engineering [24,32], mechanisms comprising linkages combined with NCGs [33,34]. For example, Mundo *et al.* [33] propose a five-bar linkage combined with NCGs to move a point of the mechanism along a prescribed trajectory and Modler *et al.* [34] present a general methodology to design a desired transmission function.

In this article, a new noncircular gear pair that produces small shaft accelerations is proposed. Their characteristics are compared with their elliptical and sinusoidal gears counterparts, which are two common types of NCGs. In the comparison, gears with two lobules are considered. The following sections describe the equations for these three types of gear centrodes. Then, in Section 6, a comparison is presented. Finally, Section 7 concludes this article.
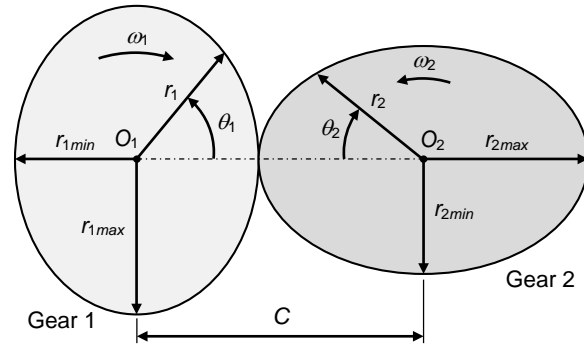

Figure 1. Centrodes of second-order elliptical gears; $n_{lob1} = n_{lob2} = 2$, $\omega_{2a}/\omega_{2m} = 0.3$. $r_{1max} = r_{2max} = 0.577C$ and $r_{1min} = r_{2min} = 0.423C$.
Source: The authors.

2. Geometric and kinematic equations of NCGs

This section presents the relevant equations for the design and analysis of noncircular gears. A more detailed theory on NCGs is presented in other papers [5-7].

The gear ratio in a NCG pair is not constant and may be expressed as

$$g = \frac{\omega_1}{\omega_2} = \frac{d\theta_1}{d\theta_2} = \frac{r_2}{r_1} \tag{1}$$

where ω_1 and ω_2 are the angular speed of gear 1 and gear 2 respectively (hereafter "1" will refer to the driving gear and "2" to the driven gear), and θ and r are the gear polar angle and polar radius (Fig. 1). The gear ratio is normally expressed as a function of θ_1.

The sum of the polar radii of two external gears is constant and is given by

$$r_2 + r_1 = C \tag{2}$$

where C is the fixed center distance (O_1 and O_2 are fixed).

The angles μ_1 and μ_2 (Fig. 2) correspond to the direction of the centrode tangent and are given by
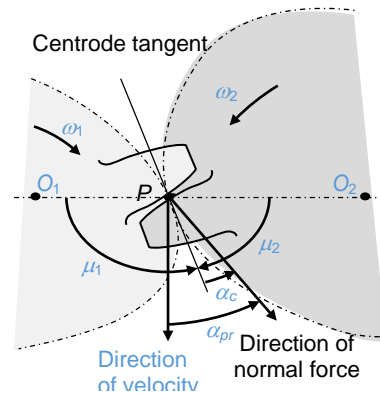

Figure 2. Pressure angle, α_{pr}, profile angle of the rack cutter, α_c, and angles μ_1 and μ_2.
Source: The authors.

$$\tan \mu_1 = \frac{r_1}{\dfrac{dr_1}{d\theta_1}}; \quad \tan \mu_2 = \frac{r_2}{\dfrac{dr_2}{d\theta_2}} = -\frac{r_1}{\dfrac{dr_1}{d\theta_1}} = -\tan \mu_1 \quad (3)$$

The pressure angle may be expressed as

$$\alpha_{pr} = \mu_1 + \alpha_c - \pi / 2 \quad (4)$$

where α_c is the profile angle of the rack cutter.

In a NCG pair, gear 1 usually rotates at a constant speed ω_1, whereas the speed of gear 2, ω_2, varies (Fig. 3); the latter one is characterized by

$$\omega_{2m} = \frac{\omega_{2max} + \omega_{2min}}{2}; \quad \omega_{2a} = \frac{\omega_{2max} - \omega_{2min}}{2} \quad (5)$$

where ω_{2m} and ω_{2a} are the mean and alternating component of ω_2; ω_{2max} and ω_{2min} are the maximum and minimum components, respectively.

The following expression is useful to derive the equations of NCG pair's centrodes. It can be demonstrated that [5]:

$$n_{lob1}\theta_1 = 2\pi f t \quad (6)$$

where n_{lob1} is the number of lobules (protuberances) of gear 1, f is the frequency of ω_2 and t is time.

The shapes of the NCGs can be determined by [5]

$$r_1(\theta_1) = \frac{\omega_2(\theta_1)C}{\omega_1 + \omega_2(\theta_1)} \quad (7)$$

$$r_2(\theta_1) = C - r_1(\theta_1) \quad (8)$$

$$\theta_2(\theta_1) = \frac{1}{\omega_1} \int_0^{\theta_1} \omega_2(\theta_1)d\theta_1 \quad (9)$$

The kinematics of NCGs may be determined by the following expressions

$$v(\theta_1) = \omega_2(\theta_1)r_2(\theta_1) = \omega_1 r_1(\theta_1) \quad (10)$$

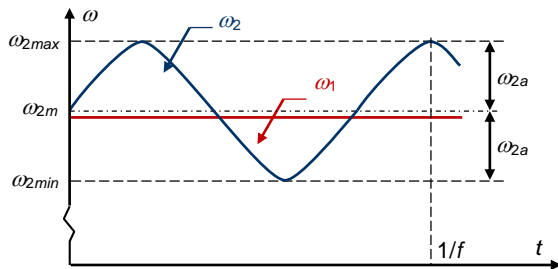where v is the velocity at the contact point P (Fig. 2).

Figure 3. Examples of curves ω-t for gears 1 and 2.
Source: The authors.

The angular acceleration is defined by

$$\alpha_2 = \frac{d\omega_2}{dt} \quad (11)$$

The tangential component of acceleration at P can be calculated by [5]

$$a_t(\theta_1) = \frac{C\omega_1^2}{\left(\omega_1 + \omega_2(\theta_1)\right)^2}\alpha_2(\theta_1) \quad (12)$$

Lastly, it has that [5]

$$\tan \mu_1(\theta_1) = -\tan \mu_2(\theta_1) = \frac{\omega_2(\theta_1)}{\alpha_2(\theta_1)}[\omega_1 + \omega_2(\theta_1)] \quad (13)$$

The following sections present the specific equations for the second order elliptical gears, sinusoidal gears, and the novel gear pair developed.

3. Elliptical gears

Two second-order elliptical gears rotating about their geometric centers are identical. The polar radius of gear 1, r_1, as a function of its polar angle, θ_1, is given by [35]

$$r_1(\theta_1) = \frac{2a_e b_e}{(a_e + b_e) - (a_e - b_e)\cos(2\theta_1)} \quad (14)$$

where a_e and b_e, which satisfy $a_e + b_e = C$ and $b_e < a_e$, are the semi-axes of each centrode. The equations for this pair can be obtained from eq. (14) and some of those in Section 2. Fig. 1 presents an example of a pair of elliptical gears.

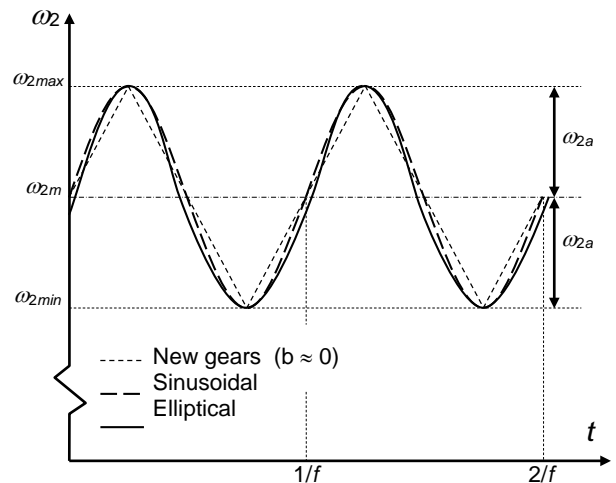Fig. 4, 5 show the shapes of the curves ω-t and α-t for the elliptical gears.

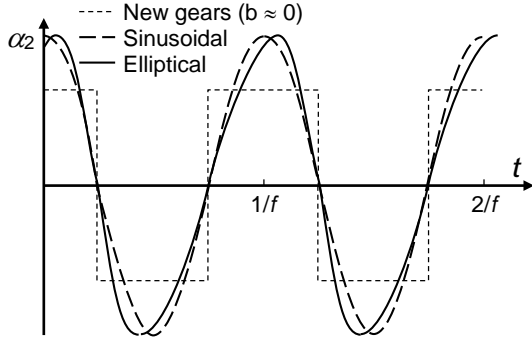Figure 4. Curve ω_2-t for elliptical, sinusoidal, and new gears.
Source: The authors.

Figure 5. Curve α_2-t for elliptical, sinusoidal, and new gears.
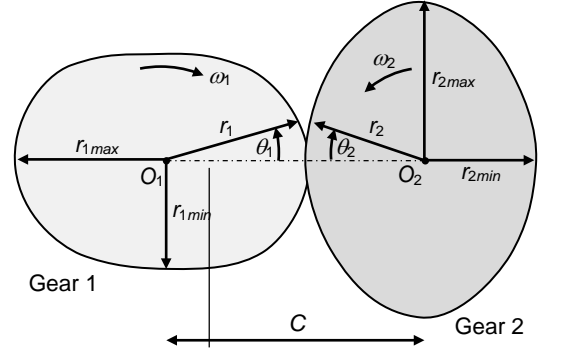Source: The authors.

Figure 6. Centrodes of sinusoidal gears; $n_{lob1} = n_{lob2} = 2$, $\omega_{2a}/\omega_{2m} = 0.3$. $r_{1max} = 0.565C$, $r_{2max} = 0.588C$, $r_{1min} = 0.412C$, and $r_{2min} = 0.435C$.
Source: The authors.

4. Sinusoidal gears

For the sinusoidal gears, the angular speed of gear 2 is given by

$$\omega_2(t) = \omega_{2m} + \omega_{2a} \sin 2\pi ft \tag{15}$$

Also, by using eq. (6), it can be given by

$$\omega_2(\theta_1) = \omega_{2m} + \omega_{2a} \sin n_{lob1}\theta_1 \tag{16}$$

The angular acceleration of gear 2 is calculated by differentiating eq. (15) with respect to time. By using eq. (6), and taking into account that for sinusoidal gears, as well as for the novel gear pair, the following applies [5]

$$f = \frac{n_{lob1}\omega_1}{2\pi} = \frac{n_{lob2}\omega_{2m}}{2\pi} \tag{17}$$

And the following are yielded

$$\alpha_2(t) = 2\pi f \omega_{2a} \cos 2\pi ft \tag{18}$$

$$\alpha_2(\theta_1) = n_{lob2}\omega_{2m}\omega_{2a} \cos n_{lob1}\theta_1 \tag{19}$$

Therefore, the maximum and minimum angular accelerations, α_{2max} and α_{2min}, are respectively

$$\alpha_{2max} = -\alpha_{2min} = 2\pi f \omega_{2a} = n_{lob2}\omega_{2m}\omega_{2a} \tag{20}$$

Finally, substitution of eq. (16) into eq. (9) produces

$$\theta_2(\theta_1) = \frac{1}{\omega_1}\left[\omega_{2m}\theta_1 + \frac{\omega_{2a}}{n_{lob1}}\left(1 - \cos n_{lob1}\theta_1\right)\right] \tag{21}$$

A sinusoidal NCG pair's centrodes can be obtained from eq. (7), (8), (21). Fig. 6 provides an example by showing a sinusoidal NCG pair's centrodes, with two lobules each, when $\omega_{2a}/\omega_{2m} = 0.3$. Also, Fig. 4, 5 show the shapes of the curves ω-t and α-t for the sinusoidal gears.

5. A new NCG pair

For a given set of parameters f, ω_{2a}, and ω_{2m}, the minimum value of the maximum acceleration of the driven shaft is obtained when the velocity increases linearly from its minimum value and then decreases linearly from its maximum value. However, this speed variation will produce pointed centrodes and, subsequently, gears that are not able to produce teeth generation. In view of this, this work presents the mathematical equations for the NCG centrodes of a new pair that has been devised to minimize shaft accelerations, while the centrodes obtained are sufficiently smooth. To the authors' knowledge no other researchers have dealt with NCGs to minimize shaft accelerations.

For the angular speed of gear 2, the following equation has been devised:

$$\omega_2(t) = \omega_{2m} + \frac{4\omega_{2a}ft}{1-b}\left(1 - be^{\frac{1-b}{b}(4ft-1)}\right) \tag{22}$$

which satisfies

$$\omega_2\left(\frac{1}{4f}\right) = \omega_{2m} + \omega_{2a}; \quad \frac{d\omega_2}{dt}\left(\frac{1}{4f}\right) = 0 \tag{23}$$

Substitution of eq. (6) into eq. (22) yields

$$\omega_2(\theta_1) = \omega_{2m} + \frac{2\omega_{2a}n_{lob1}\theta_1}{\pi(1-b)}\left(1 - be^{\frac{1-b}{b}\left(\frac{2n_{lob1}\theta_1}{\pi}-1\right)}\right) \tag{24}$$

The parameter b is called a "smooth parameter" and may vary from a value close to 0 (when the centrodes become pointed) to 1. As b is increased from 0, the centrodes become

222

smoother; however, in order to achieve small shaft accelerations, b has to be much closer to 0 than to 1.

Eq. (22), (24) are valid for a quarter of an ω_2 cycle. These are extended as follows:

$$\omega_2(\theta_1) = \omega_{2m} + \frac{2\omega_{2a}}{1-b} h_1(\theta_1)\left(1 - b\, e^{\frac{1-b}{b}[2h_2(\theta_1)-1]}\right) \tag{25}$$

Where

$$h_1(\theta_1) = \frac{1}{\pi}\arcsin(\sin n_{lob1}\theta_1) \tag{26}$$

And

$$h_2(\theta_1) = \frac{1}{\pi}\arcsin\{\sin[\arccos(\cos n_{lob1}\theta_1)]\} \tag{27}$$

It can be demonstrated that

$$\alpha_2(\theta_1) = \omega_1 \frac{n_{lob1}}{\pi} \frac{2K_1\omega_{2a}}{1-b}$$
$$\times\left(1 - [b + 2(1-b)h_2(\theta_1)]e^{\frac{1-b}{b}[2h_2(\theta_1)-1]}\right) \tag{28}$$

Where

$$K_1 = \begin{cases} 1, & \text{if int}\left(\dfrac{n_{lob1}\theta_1}{\pi}+0.5\right) \text{ is even} \\[2mm] -1, & \text{if int}\left(\dfrac{n_{lob1}\theta_1}{\pi}+0.5\right) \text{ is odd} \end{cases} \tag{29}$$

where "int" is a function that rounds the argument down to the closest integer.

The maximum and minimum values of $\alpha(t)$ are given by

$$\alpha_{2max} = -\alpha_{2min} = \omega_1 \frac{n_{lob1}}{\pi} \frac{2\omega_{2a}}{1-b}\left(1 - b\, e^{\frac{b-1}{b}}\right) \tag{30}$$

It can be demonstrated that the polar angle of gear 2 is given by

$$\theta_2(\theta_1) = 2\frac{\omega_{2m}}{\omega_1}\left(K_2\theta_1 + (K_3 - K_2K_4)\frac{\pi}{n_{lob1}} \right.$$
$$\left. + 2K_5h_3\left(\frac{\pi}{2n_{lob1}}\right) + K_6h_3\left(K_7\left(\theta_1 - \frac{K_8\pi}{n_{lob1}}\right)\right) \right) \tag{31}$$

where h_3 is a function of an angular variable, say θ, given by

$$h_3(\theta) = \frac{\omega_{2m}}{\omega_1}\theta - \frac{\omega_{2a}}{\omega_1}\frac{n_{lob1}}{\pi(b-1)^3}\left[e^{K_6}\frac{b^2(b-1)\pi\theta}{n_{lob1}} \right.$$
$$\left. + \theta^2(b-1)^2 + \frac{b^3\pi^2}{2n_{lob1}^2}\left(e^{K_6} - e^{\frac{b-1}{b}}\right) \right] \tag{32}$$

And

$$K_2 = \begin{cases} 0, & \text{if int}\left(n_{lob1}\theta_1/\pi\right) \text{ is even} \\ 1, & \text{if int}\left(n_{lob1}\theta_1/\pi\right) \text{ is odd} \end{cases} \tag{33}$$

$$K_3 = \text{int}\left(\frac{n_{lob1}\theta_1}{2\pi}\right) \tag{34}$$

$$K_4 = \text{int}\left(\frac{n_{lob1}\theta_1}{\pi}\right) \tag{35}$$

$$K_5 = \begin{cases} 0, & \text{if int}\left(n_{lob1}\theta_1/\pi+1/2\right) \text{ is even} \\ 1, & \text{if int}\left(n_{lob1}\theta_1/\pi+1/2\right) \text{ is odd} \end{cases} \tag{36}$$

$$K_6 = \begin{cases} 1, & \text{if int}\left(n_{lob1}\theta_1/\pi+1/2\right) \text{ is even} \\ -1, & \text{if int}\left(n_{lob1}\theta_1/\pi+1/2\right) \text{ is odd} \end{cases} \tag{37}$$

$$K_7 = \begin{cases} 1, & \text{if int}\left(2n_{lob1}\theta_1/\pi\right) \text{ is even} \\ -1, & \text{if int}\left(2n_{lob1}\theta_1/\pi\right) \text{ is odd} \end{cases} \tag{38}$$

$$K_8 = \text{int}\left(\frac{n_{lob1}\theta_1}{\pi}+\frac{1}{2}\right) \tag{39}$$

The profiles of the NCGs may be obtained from eq. (7), (8), (31). As an example, Fig. 7 shows the centrodes of a NCG pair with $b = 0.1$. Also, Fig. 4, 5 show the shapes of the curves ω-t and α-t for the new gears with $b \approx 0$. It should be noted that these gears produce the smallest angular accelerations.
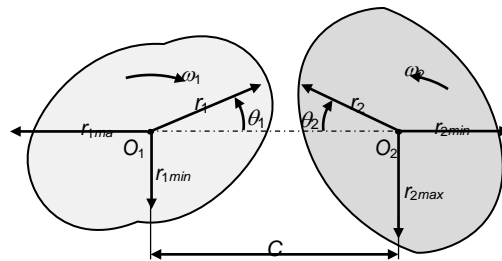

Figure 7. Centrodes of the new gears ($b = 0.1$); $n_{lob1} = n_{lob2} = 2$, $\omega_{2a}/\omega_{2m} = 0.3$. $r_{1max} = 0.565C$, $r_{2max} = 0.588C$, $r_{1min} = 0.412C$, and $r_{2min} = 0.435C$.
Source: The authors.

6. Comparison of geometric and kinematic characteristics of the three profiles

The equations presented in Sections 2 to 5 are applied in order to analyze and compare the three types of gears that this paper addresses. The independent variables for this study are ω_{2a}/ω_{2m}, ω_{2m}, b, α_c, n_{lob1}, and n_{lob2}. However, the NCGs analyzed in this article have two lobules ($n_{lob1} = n_{lob2} = 2$). The dependent variables to be studied are the nondimensional polar radii terms r_1/C and r_2/C, the tangential acceleration term a_t / C, maximum and minimum α_{pr}, and maximum angular acceleration α_2. As the profile angle of the rack cutter only affects the pressure angle as an addend (eq. 4), any value, say the standard value of 20°, may be used for the comparison. It should be noted that ω_{2a}/ω_{2m} has been taken as an independent variable, as r_1/C, r_2/C, α_{pr}, and g are independent of the values of ω_{2a} and ω_{2m} for a certain ratio ω_{2a}/ω_{2m}; the rest of the variables and terms are proportional to ω_{2m}.

Fig. 8, 9 show the variation of the maximum and minimum values of α_{pr} for the elliptical, sinusoidal, and new profiles (indicated by "VAP" in the figures), for $\alpha_c = 20°$ and $n_{lob1} = n_{lob2} = 2$. For the last ones, three values of b are taken: ~0, 0.1, and 0.4. Both figures indicate that the differences in the maximum and minimum pressure angle are not so large for the small ratios ω_{2a}/ω_{2m}. However, they tend to become larger for higher ratios. For small ratios, the new profiles (with $b \leq 0.4$) produce the smallest maximum values of pressure angle, as well as the largest minimum ones, and the sinusoidal ones produce the largest maximum values and the smallest minimum ones. This result indicates that the derived profiles may have an advantage to reduce tooth forces and stresses. In contrast, for example, when $\omega_{2a}/\omega_{2m} > 0.28$ o 0.325 the elliptical gears and the sinusoidal gears, respectively, produce smaller values than the new gears with $b \approx 0$. For $\omega_{2a}/\omega_{2m} = 0.5$, α_{prmax} is 49.0°, 53.9°, 60.3°, and 54.9° for the elliptical, sinusoidal, $b \approx 0$, and $b = 0.1$, respectively.
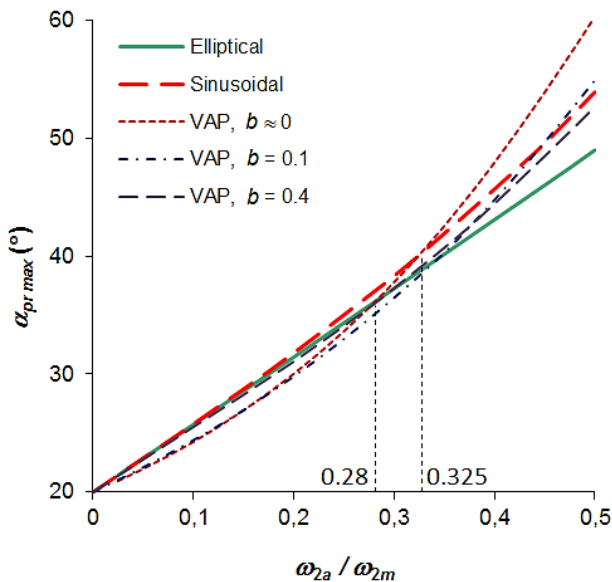

Figure 8. Variation of the maximum pressure angle with the ratio ω_{2a}/ω_{2m} for the three gear pairs studied. $n_{lob1} = n_{lob2} = 2$; $\alpha_c = 20°$.
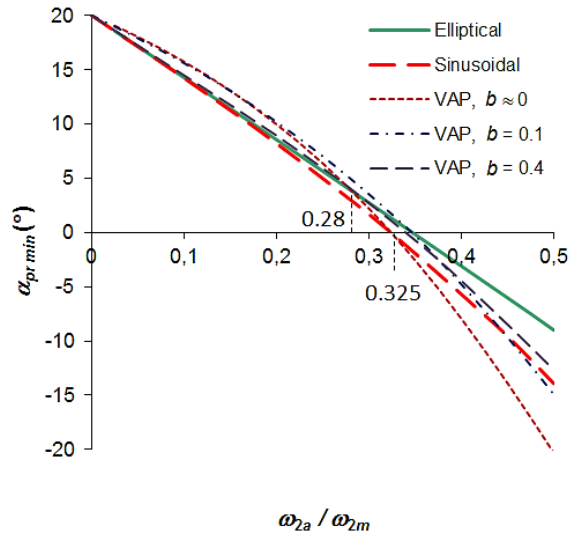Source: The authors.

Figure 9. Variation of the minimum pressure angle with the ratio ω_{2a}/ω_{2m} for the three gear pairs studied. $n_{lob1} = n_{lob2} = 2$; $\alpha_c = 20°$.
Source: The authors.

Therefore, the elliptical profiles provide the smallest values and the new gears the larger ones. The largest difference in the interval study is 11.3° (23% of 49.0°), which corresponds to the difference in the maximum or minimum pressure angle between the elliptical and the new function with $b \approx 0$.

On the one hand, the developed profiles have been devised to use a value of b greater than 0; therefore, for the new profiles, $\alpha_{pr\,max}$ becomes smaller and $\alpha_{pr\,min}$ becomes larger. On the other hand, it may be more appropriate, as far as the pressure angle is concerned, to use elliptical gears if the ratio ω_{2a}/ω_{2m} is large. Lastly, from Fig. 8, it can be concluded that the derived gears with $b \geq 0.1$ outperforms the other profiles for $\omega_{2a}/\omega_{2m} \leq 0.35$ in terms of the pressure angle.

Fig. 10, 11, 12 present the results regarding the maximum and minimum values of the polar radii and gear ratio. It can be observed that regardless the value of b, the derived gears, as well as the sinusoidal one, have the same maximum and minimum values of r_1, r_2, and g. Only the elliptical gears have different values; these gears tend to have larger polar radii for gear 1 and smaller polar radii for gear 2 than their counterparts with other profiles. However, the variation of the polar radii, i.e., the difference between the maximum and minimum values of each gear (driving and driven) is virtually the same. Additionally, the maximum values increase and the minimum ones decrease as the ratio ω_{2a}/ω_{2m} increases, as expected. In light of this, it may be concluded that the variation of polar radii does not seem to be an important criterion for gear performance.

Regarding gear kinematics, Fig. 13 presents the variation of the maximum angular acceleration of gear 2 for $\omega_{2m} = 100$ rpm. As expected, the angular acceleration increases as the ratio ω_{2a}/ω_{2m} increases. As eq. (20) and (30) show, the relationship between α_{2max} and ω_{2a}/ω_{2m} is linear; this is not the case for the elliptical gears. Also, as expected, the proposed centrodes produce the minimum angular

accelerations for small values of b (when $b \leq \sim 0.5$); the elliptical gears are the ones that produce the highest accelerations. However, the important finding is that the differences in the angular acceleration are large. For example, there are differences between the values for the sinusoidal and new gears with $b = 0.1$ is 29.3%, with respect to α_{2max} of the sinusoidal pair (regardless of the value of ω_{2a}/ω_{2m}). The difference between the values for the elliptical and new gears with $b = 0.1$ has a maximum value of 35.5% with respect to α_{2max} of the elliptical profiles. Therefore, the novel gears outperform the other profiles and, thus, will produce significantly smaller shaft accelerations.

Fig. 14 shows the variation of the ratio maximum tangential acceleration at the contact point - center distance, for $\omega_{2m} = 100$ rpm. It can be observed that the teeth of the new profiles withstand lower tangential accelerations than those of the sinusoidal gears. A similar situation occurs when comparing the elliptical and new gears for $\omega_{2a}/\omega_{2m} < \sim 0.35$.

Figure 12. Variation of the maximum and minimum values of the gear ratio with the ratio ω_{2a}/ω_{2m} for the three gear pairs. $n_{lob1} = n_{lob2} = 2$.
Source: The authors.

Figure 10. Variation of the maximum and minimum values of the nondimensional term r_1/C with the ratio ω_{2a}/ω_{2m} for the three gear pairs studied. $n_{lob1} = n_{lob2} = 2$.
Source: The authors.

Figure 13. Variation of the maximum angular acceleration with the ratio ω_{2a}/ω_{2m} for the three gear pairs studied. $n_{lob1} = n_{lob2} = 2$; $\omega_{2m} = 100$ rpm.
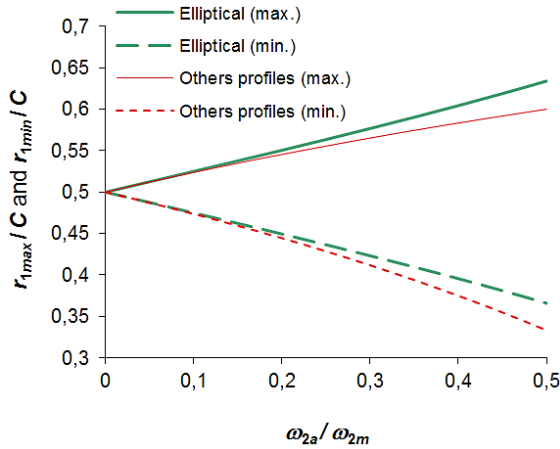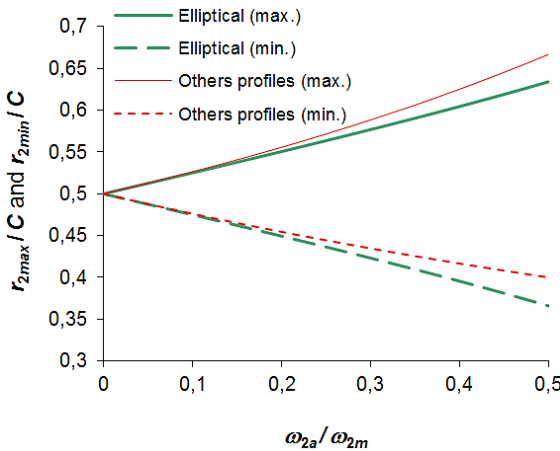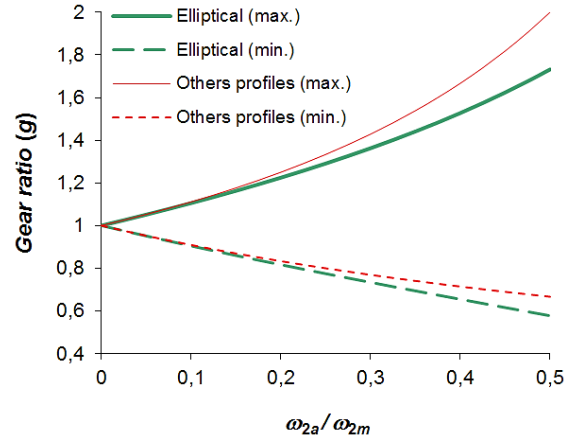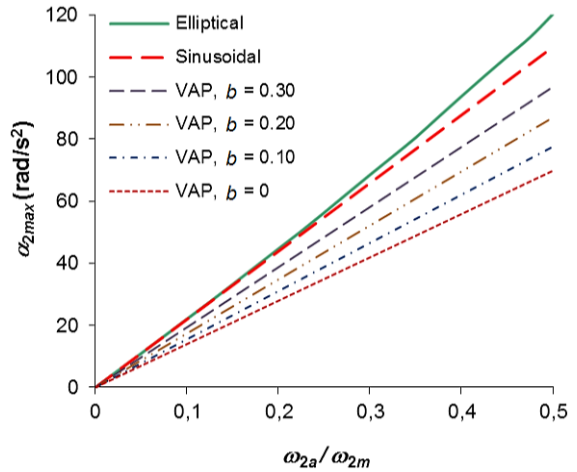Source: The authors.

Figure 11. Variation of the maximum and minimum values of the nondimensional term r_2/C with the ratio ω_{2a}/ω_{2m} for the three gear pairs studied. $N_{lob1} = n_{lob2} = 2$.
Source: The authors.

For larger ratios, the elliptical gears provide smaller tangential accelerations. In view of this, as far as the tangential accelerations at the contact points are concerned, the new gears will outperform the other profiles for $\omega_{2a}/\omega_{2m} < \sim 0.35$, and the elliptical gears will outperform the other gears for $\omega_{2a}/\omega_{2m} > \sim 0.35$.

The discussion in this section indicates that, overall, the derived gear profiles tend to outperform the sinusoidal and elliptical gears, as far as the shaft accelerations, the tangential accelerations, and the pressure angle are concerned. The only exception is that the elliptical gears produce smaller values of the maximum tangential accelerations if $\omega_{2a}/\omega_{2m} > \sim 0.35$ and there are smaller maximum values of the pressure angle for $\omega_{2a}/\omega_{2m} > \sim 0.28$. The new gears would only have large positive or negative pressure angles for very small values of b together with very high ratios ω_{2a}/ω_{2m}, which may increase tooth forces and stresses. However, a suitable value of b might be selected, so that the gears withstand low stresses. The advantage of all this will be reflected in smaller forces,
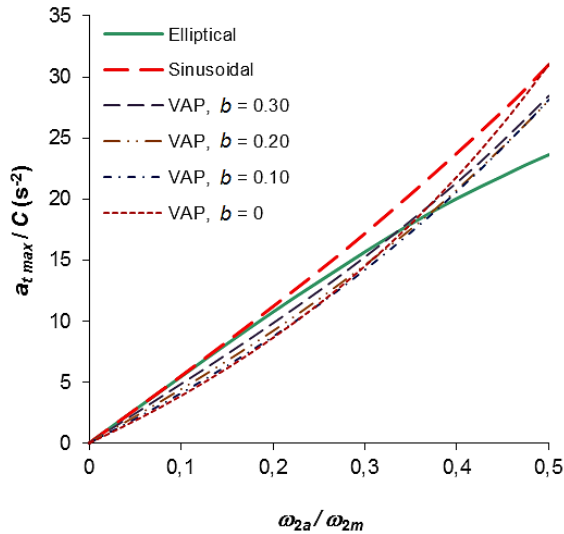
Figure 14. Variation of the maximum tangential acceleration - center distance ratio with ω_{2a}/ω_{2m} for the three gear pairs studied. $n_{lob1} = n_{lob2} = 2$; $\omega_{2m} = 100$ rpm.
Source: The authors.

torques, and stresses in the drive system and in the gear teeth. Thus, the gears may be more compact and provide a smoother operation. However, it is necessary to perform additional research so as to ascertain the effect of the smoothness parameter on tooth forces and stresses.

7. Conclusions

In this article, a new noncircular gear pair has been proposed. This has been devised so that the angular velocity of the driven gear increases and decreases virtually linearly from its minimum and maximum values, respectively, with the aim of producing small accelerations for the driven shaft. The profiles of the new gears are aimed at reducing driven gear accelerations and are based on an exponential function that depends on a smoothness parameter that controls the smoothness of the centrodes in order to avoid pointed lobules.

A comparison of the geometric and kinematic characteristics of three NCG pairs, namely elliptical, sinusoidal, and new gears is presented. The results indicate that for two-lobule centrodes the new gears tend to outperform the other two profiles. This is particularly true for the angular acceleration of the driven shaft; therefore, it is expected that the loads and stresses of the driven shaft are smaller than those produced by the other profiles. Additionally, for certain ranges of the ratio between the alternating and mean angular velocities, the new pair also tends to produce smaller maximum values of the pressure angle and of the tangential accelerations at the contact point. Therefore, the developed gears may exhibit lower tooth forces and stresses.

Acknowledgements

References

[1] Peña-Bustos, F.M. and Álvarez-Vargas, C.A., Design and construction of a torsional fatigue testing machine operated by inertial loads. DYNA, 79(172), pp. 46-55, 2012.

[2] Marulanda, D. and Olaya, J.J., Unbalanced magnetron sputtering system for producing corrosion resistance multilayer coatings. DYNA, 79(171), pp. 74-79, 2012.

[3] Loaiza, J.L. and Arzola, N., Evolución y tendencias en el desarrollo de prótesis de mano. DYNA, 78(169), pp. 191-200, 2011.

[4] Zarebski, I. and Salacinski, T., Designing of non-circular gears. The Archive of Mechanical Engineering, LV(3), pp. 275-292, 2008.

[5] Vanegas-Useche, L.V., Abdel-Wahab, M.M. and Parker, G.A., Design of noncircular gears to minimise shaft accelerations, Proc. of the 8th ASME Conf. Eng. Sys. Des. Analysis ESDA 2006, pp. ESDA2006-95560-1 - ESDA2006-95560-10, 2006. DOI: 10.1115/ESDA2006-95560

[6] Quintero, H.F., Cardona, S. and Jordi, L., Engranajes no circulares: aplicaciones, diseño y manufactura. Scientia et Technica, 24, pp. 133-138, 2004.

[7] Olsson, U., Non-circular cylindrical gears. Acta Polytechnica, 2(10), pp. 1-216, 1953. DOI: 10.1002/zamm.19540340711

[8] Lozzi, A., Non-circular gears - graphic generation of involutes and base outlines. Proc. Inst. Mech. Engrs., 214(C), pp. 411-422, 2000. DOI: 10.1243/0954406001523074

[9] Dooner, D.B., Function generation utilizing an eight-link mechanism and optimized non-circular gear elements with application to automotive steering. Proc. Instn. Mech. Engrs. (ImechE), Part C: J. Mech. Eng. Sc., 215(7), pp. 847-857, 2001. DOI: 10.1243/0954406011524090

[10] Dooner, D.B., Use of noncircular gears to reduce torque and speed fluctuations in rotating shafts. ASME J. Mech. Des., 119(2), pp. 299-306, 1997. DOI: 10.1115/1.2826251

[11] Dooner, D., Yoon, H-D. and Seireg, A., Kinematic considerations for reducing the circulating power effects in gear-type continuously variable transmissions. Proc. Inst. Mech. Engrs., 212(D), pp. 463-478, 1998. DOI: 10.1243/0954407981526118

[12] Tong, S-H. and Yang, D.C.H., Generation of identical noncircular pitch curves. ASME J. Mech. Des., 120(2), pp. 337-341, 1998. DOI: 10.1115/1.2826978

[13] Doege, E., Meinen, J., Neumaier, T. and Schaprian, M., Numerical design of a new forging press drive incorporating non-circular gears. Proc. Inst. Mech. Engrs. (ImechE), Part B: J. Eng. Man., 215(4), pp. 465-471, 2001. DOI: 10.1243/0954405011518430

[14] Guo, L.S. and Zhang, W.J., Kinematics analysis of a rice transplanting mechanism with eccentric planetary gear trains. Mechanism and Machine Theory, 36(11-12), pp. 1175-1188, 2001. DOI: 10.1016/S0094-114X(01)00052-0

[15] Librovich, B.V., Dynamics of rotary vane engine. ASME J. Mech. Des., 125(3), pp. 498-508, 2003. DOI:10.1115/1.1582500

[16] Tucker, R.W., Wang, C. and Librovich, B.V., Mathematical modelling of rotary vane engines. Proc. Inst. Mech. Engrs. (ImechE), Part C: J. Mech. Eng. Sc., 217(6), pp. 687-704, 2003. DOI: 10.1243/095440603321919608

[17] Yao, Y-A. and Yan, S-H., A new method for torque balancing of planar linkages using non-circular gears. Proc. Inst. Mech. Engrs. (ImechE), Part C: J. Mech. Eng. Sc., 217(5), pp. 495-503, 2003. DOI: 10.1243/095440603765226786

[18] Litvin, F.L., Gonzalez-Perez, I., Fuentes, A. and Hayaska, K., Design and investigation of gear drives with non-circular gears applied for speed variation and generation of functions. Comput. Methods Appl. Mech. Engrg., 197, pp. 3783-3802, 2008. DOI: 10.1016/j.cma.2008.03.001

[19] Ottaviano, E., Mundo, D., Danieli, G.A. and Ceccarelli, M., Numerical and experimental analysis of non-circular gears and cam-follower systems as function generators. Mechanism and Machine Theory, 43, pp. 996-1008, 2008. DOI: 10.1016/j.mechmachtheory.2007.07.004

[20] Figliolini, G., Stachel, H. and Angeles, J., Synthesis of the base curves of non-circular gears via the return circle, Int. Gear Conf., pp. 540-550, 2014. DOI:10.1533/9781782421955.540

[21] Yang, D.C.H., Tong, S-H. and Lin, J., Deviation-function based pitch curve modification for conjugate pair design. ASME J. Mech. Des., 121(4), pp. 579-586, 1999. DOI: 10.1115/1.2829502

[22] Danieli, G.A., Analytical description of meshing of constant pressure angle teeth profiles on a variable radius gear and its applications. ASME J. Mech. Des., 122(1), pp. 123-129, 2000. DOI: 10.1115/1.533551

[23] Chang, S-L., Tsay, C-B. and Wu, L-I., Mathematical model and undercutting analysis of elliptical gears generated by rack cutters. Mechanism and Machine Theory, 31(7), pp. 879-890, 1996. DOI:10.1016/0094-114X(95)00121-E

[24] Chang, S-L. and Tsay, C-B., Computerized tooth profile generation and undercut analysis of noncircular gears manufactured with shaper cutters. ASME J. Mech. Des., 120(1), pp. 92-99, 1998. DOI: 10.1115/1.2826682

[25] Bair, B-W., Computer aided design of elliptical gears. ASME J. Mech. Des., 124(4), pp. 787-793, 2002. DOI: 10.1115/1.1485092

[26] Bair, B-W., Computer aided design of non-standard elliptical gear drives. Proc. Inst. Mech. Engrs. (ImechE), Part C: J. Mech. Eng. Sc., 216(4), pp. 473-483, 2000. DOI: 10.1243/0954406021525250

[27] Bair, B-W., Computerized tooth profile generation of elliptical gears manufactured by shaper cutters. J. Mat. Proc. Tech., 122(2-3), pp. 139-147, 2002. DOI: 10.1016/S0924-0136(01)01242-0

[28] Figliolini, G. and Angeles, J., The synthesis of elliptical gears generated by shaper-cutters. ASME J. Mech. Des., 125(4), pp. 793-801, 2003. DOI: 10.1115/1.1631573

[29] Bair, B-W., Computer aided design of elliptical gears with circular-arc teeth. Mechanism and Machine Theory, 39(2), pp. 153-168, 2004. DOI: 10.1016/S0094-114X (03)00111-3

[30] Xia, L., Liu, Y., Li, D. and Han, J., A linkage model and applications of hobbing non-circular helical gears with axial shift of hob. Mechanism and Machine Theory, 70, pp. 32-44, 2013. DOI: 10.1016/ j. mechmachtheory. 2013.07.002

[31] Bair, B-W., Sung, M-H., Wang, J-S. and Chen, C-F., Tooth profile generation and analysis of oval gears with circular-arc teeth. Mechanism and Machine Theory, 44, pp. 1306-1317, 2009. DOI: 10.1016/ j. mechmachtheory. 2008.07.003

[32] Barkah, D., Shafiq, B. and Dooner, D., 3D mesh generation for static stress determination in spiral noncircular gears used for torque balancing. ASME J. Mech. Des., 124(2), pp. 313-319, 2002. DOI:10.1115/1.1470492

[33] Mundo, D., Gatti, G. and Dooner, D.B., Optimized five-bar linkages with non-circular gears for exact path generation. Mechanism and Machine Theory, 44, pp. 751-760, 2009. DOI: 10.1016/j.mechmachtheory. 2008.04.011

[34] Modler, K.-H., Lovasz, E.-C., Bär, G.F., Neumann, R., Perju, D., Perner, M. and Mărgineanu, D., General method for the synthesis of geared linkages with non-circular gears. Mechanism and Machine Theory, 44, pp. 726-738, 2009. DOI: 10.1016/j.mechmachtheory.2008.05.006

[35] Bloomfield, B. Design guide … when you need noncircular gears, in Chironis - N. P. Gear design and application, New York, McGraw-Hill, Inc., 1967. DOI: 10.1016/0022-2569(68)90361-3

L.V. Vanegas-Useche, PhD., is full professor in the Mechanical Engineering Department, Universidad Tecnológica de Pereira, Colombia. He received his BSc. in Mechanical Engineering from the Universidad Tecnológica de Pereira, Pereira, Colombia, in 1994, his MSc. degree in Advanced Manufacturing Technology and Systems Management from the University of Manchester, Manchester (UK), in 1999, and his PhD. degree in Mechanical Engineering from the University of Surrey, Guildford (UK), in 2008. He has published more than 50 scientific papers. His research interests include: Fracture Mechanics, Fatigue, Mechanical Design, and Finite Element Modeling of Machine Elements and Structures. E-mail: lvanegas@utp.edu.co.
ORCID: 0000-0002-5891-8696

M.M. Abdel-Wahab, PhD., is a Professor of Applied Mechanics in the Department of Mechanical Construction and Production at Ghent University, Belgium. He received his BSc., 1988, in Civil Engineering and his MSc., 1991, in Structural Mechanics, both from Cairo University. He completed his PhD. in Fracture Mechanics in 1995 at KU Leuven, Belgium. He was awarded a D.Sc. from the University of Surrey in 2008. He has published more than 200 scientific papers on Solid Mechanics and Dynamics of Structures. His research interests include Finite Element Analysis, Fracture Mechanics, Damage Mechanics, Fatigue of Materials, Durability, and Dynamics and Vibration.
ORCID: orcid.org/0000-0002-3610-865X

G.A. Parker, PhD., is emeritus professor of Mechanical Engineering in the Department of Mechanical and Physical Sciences at The University of Surrey. He holds the following qualifications and titles: BSc., PhD., FIMechE, MEM.ASME, CEng, Eur.Ing. He is also member of the EPSRC Mechanical Engineering College. His research interests include virtual and augmented reality, control and systems integration, machine vision, brushing technology, fluid control systems.
ORCID: 0000-0003-0677-6451

UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería Mecánica

Oferta de Posgrados

Maestría en Ingeniería - Ingeniería Mecánica

Mayor información:

E-mail: acmecanica_med@unal.edu.co
Teléfono: (57-4) 4259262

Roasting of sulphide using carbothermal reduction

Ivana Marcela Orozco [a] & Vanesa Lucia Bazan [b] & Andrea Alejandra Diaz [a] & Rodolfo Francisco Lara [a]

[a] Facultad de Ingeniería, Universidad Nacional de San Juan, San Juan, Argentina iorozco@unsj.edu.ar; aadiaz@unsj.edu.ar, rflara@unsj.edu.ar
[b]CONICET, Facultad de Ingeniería, Universidad Nacional de San Juan, Argentina bazan@unsj.edu.ar

Abstract

The carbothermic reduction process is a direct reduction of sulfides. It allows a roasting in which higher concentrations of metals such as molybdenum are obtained by using both a reducing agent (in this case, carbon C) and a sulfur scavenger, such as CaO so as to prevent air toxic gases pollution such as SO_2. In this paper, we analyze sulfur concentrates rich in copper, iron, and molybdenum that are obtained through rougher flotation and differential floats, which give rise to different laws. The concentrates were subjected to a carbothermic reduction process in which the working parameters were studied. The values of both the capturing agent and the reducing agent were adjusted to retain higher concentrations of sulfur. Analyzing the results obtained, the temperature was set and we reached the optimal working time and optimum molar ratio of CaO and C for the greatest recovery of metals such as molybdenum.

Keywords: carbothermic reduction; sulfides; roasting.

Tostación de sulfuros por reducción carbotérmica

Resumen

El proceso de reducción carbotérmica es una reducción directa de sulfuros. Permite realizar una tostación en donde se obtienen concentraciones mayores de metales como molibdeno, empleando un agente reductor (en este caso, el carbón C) y un captador de azufre como el CaO, evitando la contaminación del aire con gases tóxicos como el SO_2.

En este trabajo se analizan concentrados sulfurados ricos en cobre, hierro y molibdeno, obtenidos por flotación rougher y flotaciones diferenciales, dando origen a diferentes leyes. Los concentrados se sometieron al proceso de reducción carbotérmica en donde se estudiaron los parámetros de trabajo. Se ajustaron los valores del agente captador y reductor, para retener las mayores concentraciones de azufre. Analizando los resultados obtenidos se fijó la temperatura, el tiempo óptimo de trabajo y optimas relación molar del CaO y C para los cuales se alcanzó la mayor recuperación de metales como el molibdeno.

Palabras clave: reducción carbotérmica; sulfuros; tostación.

1. Introduction

1.2. Obtaining Molybdenum

Molybdenum is recovered mainly as molybdenite (MoS_2) from minerals or concentrates that have been obtained by using "rougher" or "differential" flotation. A roasting or leaching is later carried out using nitric acid or nitrate/sulfuric acid, technical grade molybdenum trioxide (MoO_3) is thus obtained [1,2].

However, within the framework of the conventional metallurgical processes, the treatment of concentrates includes one or more oxidation stages to obtain a high-temperature metal, generally by using roasting furnaces. In general, such operations involve SO_2 formation, which generates environmental pollution problems. From an economic standpoint, it is seldom feasible to abate this completely, and as such a considerable part of this gas frequently leaks into the atmosphere as chimney gas or flue gas [3,4].

Oxidizing roasting is a traditional pyrometallurgical method that is not always possible for the whole variety of metallic sulfides, especially if these are low-grade concentrates or when they are forming disseminated mineral

deposits or polymetallic sulfide deposits. This is generally due to the complexity of treatment flow charts, followed by the high cost of eliminating the SO_2 gas that is produced in these pyrometallurgical methods.

An alternative method for the treatment of sulfide concentrates that prevents air from being polluted by SO_2 is the direct reduction of sulfides with appropriate reducing agents in the presence of a desulfurizing agent such as CaO, $CaCO_3$, or Na_2CO_3 [7,8]. Such processes are generally applied to MoS_2 concentrates with grades higher than 21%. Carbon C is the reducing agent that is used worldwide to obtain reduced sulfide [6, 8]. The general equation (1) for carbothermic reduction of metallic sulfides in the presence of CaO and C can be represented by the following global reaction:

$$MeS + C + 2CaO + O_2 (g) \leftrightarrow MeO + CaS + CO_2 (g) \quad (1)$$
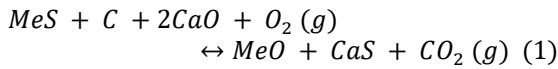
In this work, a method is presented to obtain a greater molybdenum recovery. This method may be applied to concentrates produced directly through various flotation processes, the concentrates containing a Mo percentage that ranges from 0.4-3.0 %.

1.2. The thermodynamics of carbothermic reduction roasting

Carbothermic reduction reactions have been shown to occur mainly through intermediary gaseous compounds, and, to a lot lesser degree, through reactions among solids [6,7,8]. Therefore, it is to be expected that the reduction reactions of metallic sulfides in the presence of CaO may also occur through the CO_2 and CO gaseous species. Based on the above, it can be stated that the chemical equation 2 dominates the carbothermic reduction process for molybdenum:
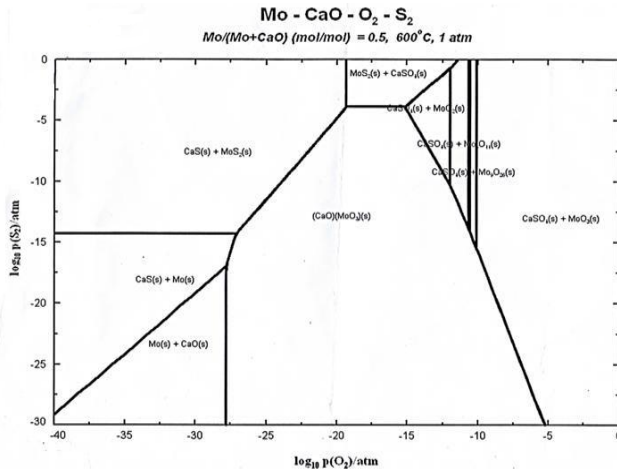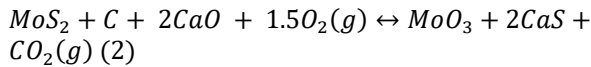
$$MoS_2 + C + 2CaO + 1.5O_2(g) \leftrightarrow MoO_3 + 2CaS + CO_2(g) \quad (2)$$

Figure 1. Stability diagram for the Mo-CaO-S_2-O_2 at 600 °C for the 0<Mo/Mo+Ca<0.5 metallic fraction.
Source: The authors.

Compounds used in the carbothermic reduction such as CaO and C may be analyzed in detail through theMe-CaO-S2-O2 systems phase equilibrium diagrams. For a 600° C condition and a 0<.Mo/Mo+Ca<0.5 metallic fraction, the phase diagram shown in Fig. 1 may be obtained. The area of interest for molybdenum reduction is the superior right area. This is in line with chemical reaction 2, which is carried out at relatively high temperatures.

2. Experimental Procedure / Method

The samples under study come from two different flotation processes, namely, rougher flotation and differential flotation. Each concentrate presents different chemical compositions from those of elements such as Cu, Fe and mainly Mo. The aim is of this research is to study the behavior of each concentrate in the roasting process by using carbothermic reduction for different chemical compositions. The most favorable conditions of the raw material are determined so that these are fed into the roasting furnace.

The mineralogical characterization was obtained using a Philips Diffractometer WP 1011. Chemical composition was determined through volumetric techniques by acid digestion attack and quantification with atomic absorption spectrophotometryby using a Perkin Elmen AA 100 equipment. Total sulfur quantification was performed with the ASTM C-25 technique. For each of the chemical elements, Cu, Fe and Mo, specific techniques were used [9-11].

Based on the results of chemical and mineralogical composition, various experimental tests were designed to simulate the roasting process by using carbothermic reduction. Tests were conducted on a 1g sample with CaO and C for various stoichiometric ratios: these are always referred to as Mo, that is, Mo:CaO:C. The sample was placed in a porcelain crucible that turned out to be inert to the chemical reactions involved in the roasting process through carbothermic reduction. After, the sample was placed in a muffle furnace with an oxidation atmosphere. Finally, the roasting product was analyzed by means of an acid digestion attack and was read with an atomic absorption spectrophotometer.

To summarize, on the laboratory scale, the concentrates are treated with carbothermic reduction for the aim of finding the optimal operational conditions for which the highest percentage of molybdenum recovery can be obtained.

2.1. Tests conducted on concentrates produced with rougher flotation

2.1.1. Temperature variation – Time variation

The samples are introduced into a muffle kiln at a constant temperature with CaO and C to Mo molar ratios of 1:1:1 (Mo:CaO:C) for an established variable time. The sample exposure was evaluated at: 500 ℃, 600 ℃, 700 ℃ and 800 ℃. For each temperature, the same roasting times were run, that is, from 15 min, 30 min, 60 min, 90 min, and 120 min to 180 min.

2.1.2. Variation of molar ratios (Mo:CaO:C)

Once the roasting temperature was set up, for which the maximum molybdenum value was obtained, work was started with molar variations. Initially, CaO varied with respect to Mo, and the C concentration was left fixed. Once the molar ratio for CaO was established, the stoichiometric ratio of C started to vary with respect to Mo.

2.2. Tests conducted on concentrates obtained with differential flotation

2.2.1. Temperature Variation – Time variation – Variation of molar ratios (Mo:CaO:C)

For concentrates produced with differential flotation, the same working conditions that were applied to concentrates obtained with rougher flotation were taken as a basis as the mineralogy of both flotations is quite similar.

The samples exposure was evaluated at different temperatures: 500 ℃, 600 ℃, 700 ℃ and 800 ℃. For each temperature, the same roasting times were run, from 15 min, 30 min, 60 min, 90 min, and 120 min to 180 min. It should be pointed out that the tests were carried out rendering the molar ratios fixed for CaO and C with respect to Mo (Mo:CaO:C).

3. Results

3.1. Mineralogical characterization

Copper and iron ores such as chalcopyrite ($CuFeS_2$) and pyrite (FeS_2) were determined from the mineralogical study and smaller proportions of minerals such as molybdenite (MoS_2) were observed. The diffractograms obtained are shown in Figs. 2 and 3.

3.2. Chemical characterization

Results from the chemical characterizations of each of the concentrates are shown in detail in Tables 1 and 2.
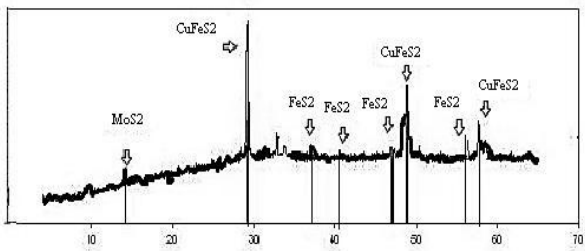

Figure 2. Diffractograms of sulfide products produced with rougher flotation.
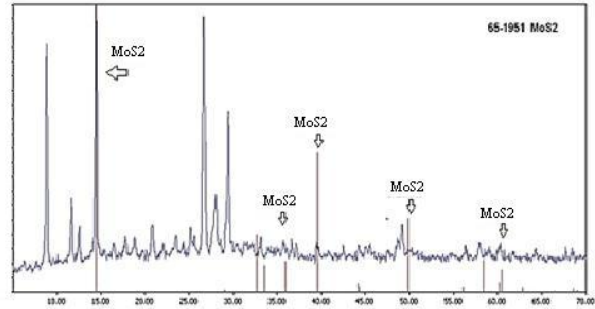Source: The authors.

Figure 3. Diffractograms of sulfide products from differential flotation.
Source: The authors.

Table 1.
Chemical composition of concentrates obtained through rougher flotation – "R" samples

Mtras.	Mo (%)	Cu (%)	S (%)	Fe (%)	CaO (%)	MgO (%)	Other (%)
R1	0,43	27,69	32,51	30,61	0,18	0,17	7,83
R2	0,50	26,81	31,30	28,01	0,18	0,08	12,52
R3	0,64	26,09	31,01	30,45	0,18	0,17	11,35
R4	0,83	25,77	31,01	31,89	0,18	0,10	9,40
R5	0,94	24,71	29,91	31,14	0,22	0,36	12,54

Source: The authors.

Table 2.
Chemical composition of concentrates obtained through differential flotation – "D" samples

Mtras.	Mo (%)	Cu (%)	S (%)	Fe (%)	CaO (%)	MgO (%)	Other (%)
D1	1,22	9,54	33,57	17,00	7,59	5,89	24,34
D2	2,11	9,98	19,53	17,51	6,49	7,69	36,17
D3	2,89	10,5	25,06	11,35	7,75	6,45	35,25

Source: The authors

3.3. Kinetics of roasting at various temperatures

Figs. 4 and 5 show the comparative results of roasting through carbothermic reduction applied both to R1 concentrates obtained with rougher flotation and D1 with differential flotation. R1 and D1´s behavior are taken as a basis.
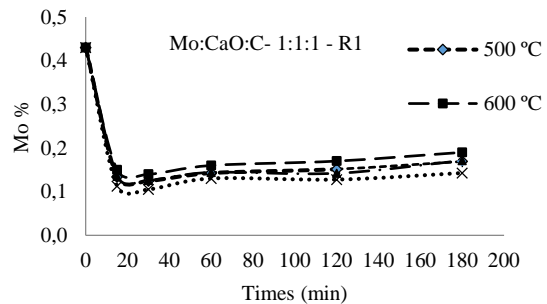

Figure 4. Temperature variation for R1 samples.
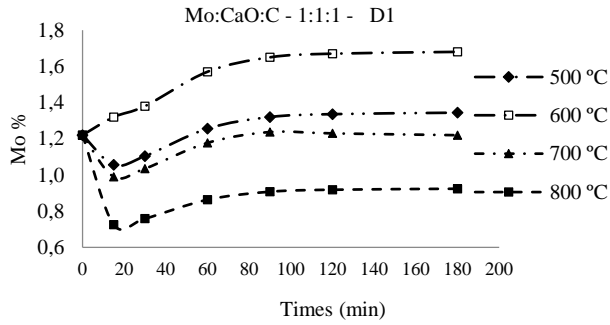Source: The authors.

Figure 5. Temperature variation for D1 samples.
Source: The authors.

3.4. Molar ratio variation for CaO with respect to Mo and C (Mo:CaO:C) for R samples.

Once the working temperature has been established, work with molar variations is started. Figs. 6 to 9 show CaO molar variations with respect to Mo, C molar variation is left constant.

There are a variation of CaO molar ratios with respect to Mo and C (Mo:CaO:C). Variations are: 1:1:1; 1:0.5:1; 1:2:1; 1:3:1 at a constant temperature of 600° C for R1 to R5 samples.
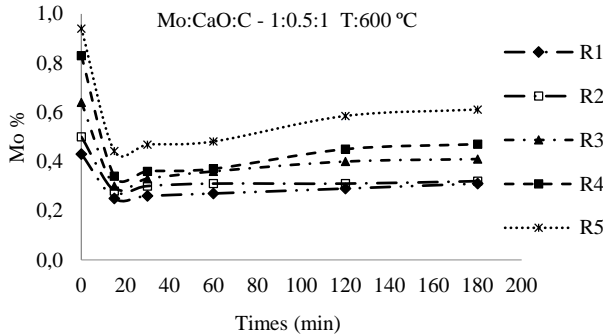
Figure 6. Mo percentage of variation versus time at 600° C for R samples.
Source: The authors.
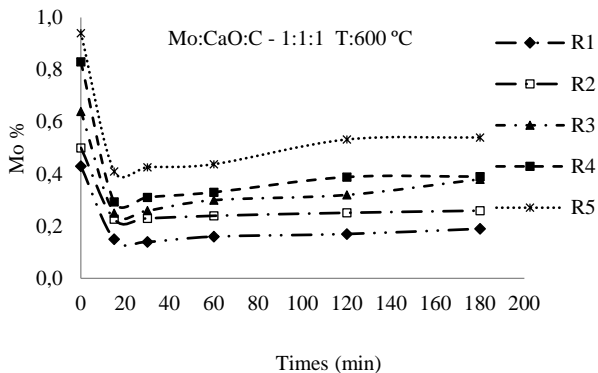
Figure 7. Mo percentage of variation versus time at 600° C for R samples.
Source: The authors.
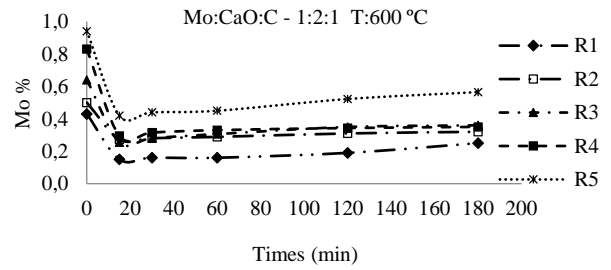
Figure 8. Mo percentage of variation versus time at 600° C for R samples.
Source: The authors.
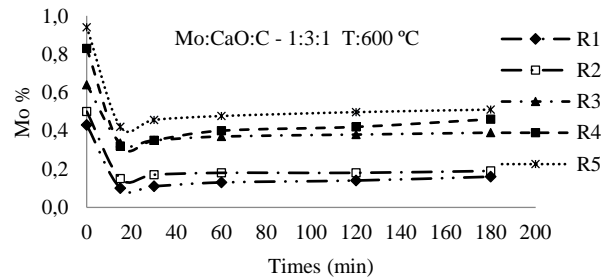
Figure 9. Mo percentage of variation versus time at 600° C for R samples.
Source: The authors.

3.5. Variation of molar ratios for C compared to Mo and CaO (Mo:CaO:C) for R samples

Once the CaO optimal molar ratio has been set up as 0.5, at which point a better molybdenum recovery is obtained, the molar ratio for C starts to altered.

Variation of molar ratios for C with respect to Mo and CaO (Mo:CaO:C) are: 1:0.5:1; 1:0.5:0.5; 1:0.5:2; 1:0.5:3, as shown in detail below in Figs. 10 to 13.

3.6. Variation of molar ratios for C, with respect to Mo and CaO (Mo:CaO:C) for D samples

For concentrates produced with differential flotation, we kept the same optimal conditions as those imposed on concentrates produced with rougher flotation. These conditions are the following: temperature: 600° C, molar ratio: 1:0.5:1, and an optimal time of 90 minutes. The tests performed are shown in Fig. 13.

Figure 10. Variation of the Mo percentage vs time at 600°C for R samples.
Source: The authors.

Figure 11. Variation of the Mo percentage vs time at 600 °C for R samples. Source: The authors.

Figure 12. Variation of the Mo percentage vs time at 600° C for R samples. Source: The authors.
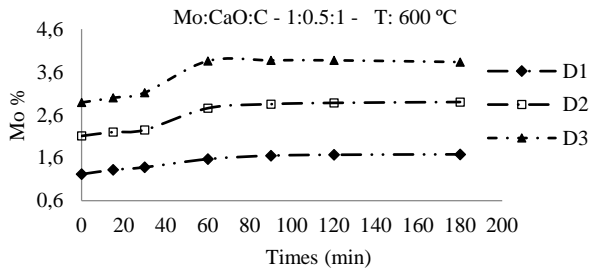
Figure 13. Variation of the Mo percentage vs time at 600° C for D samples.
Source: The authors.

The response from the application of roasting to differential flotation is quite different from that obtained from its application to rougher flotation; rather good recoveries are obtained. This indicates that separation of molybdenum with high concentrations of copper and iron leads to better results when obtaining MoO_3.

3.7. Results of Mo recovery in D samples

The results obtained show that there is no increase in productivity in those concentrates treated with simple flotation productivity (R1 to R5), whereas those treated with differential flotation (D1 to D3) show an increase higher than 35 % with respect to the initial Mo value. These results are shown in detail in Table 3.

Table 3.
Recovery values for "D" samples

Samples	1.1.1	Recovery (%)
D1		35.25
D2		33.64
D3		32.66

Source: The authors.
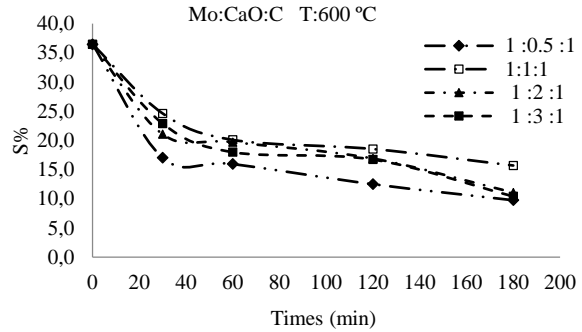
Figure 14. Variation of the S percentage vs time at 600° C for R samples. Source: The authors.

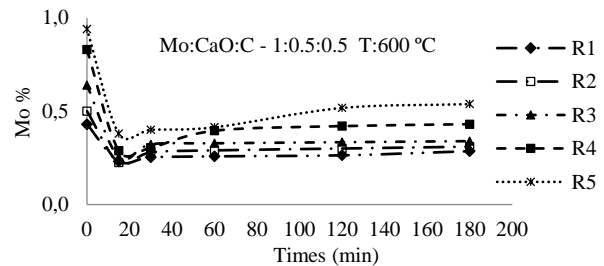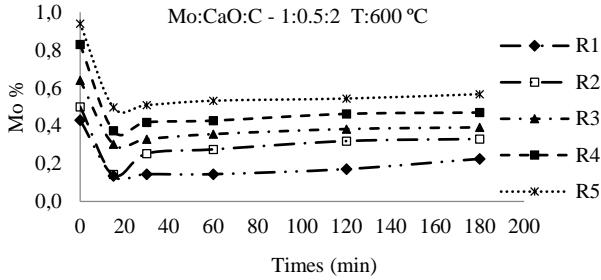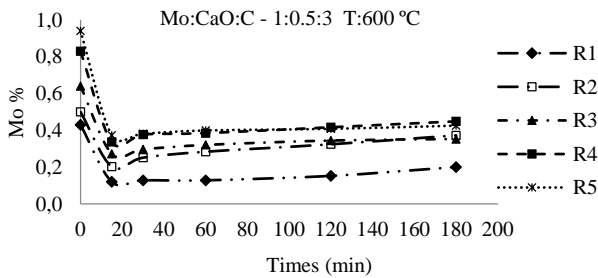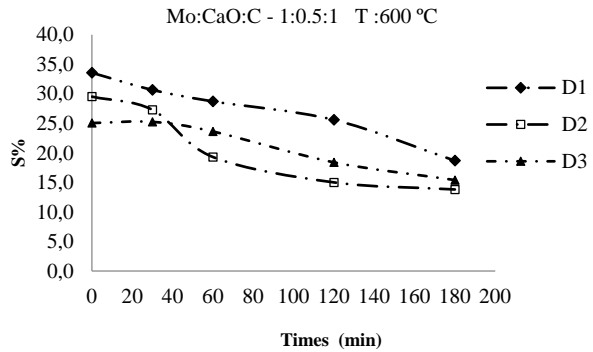Figure 15. Variation of the S percentage vs time at 600° C for D samples. Source: The authors.

3.8. Results for sulfur values of samples R and D

Below are ~~Then~~ the results of chemical analysis for simple flotation productivity, Fig. 14, for the concentrates obtained from the differential flotation, Fig. 15.

When the sulfur capture agent, CaO, range was analyzed for each sample "R", variations were observed and studied.

The results are shown in Fig. 14. The figure shows molar CaO variations under the following stoichiometric conditions Mo:CaO:C: 1: 0.5:1; 1:1:1; 1:2:1 and 1:3:1, all evaluated at 600°C.

The graphs indicate that an increase in CaO stoichiometric conditions does not increase the value of the reactant fraction. Because the function of retaining the CaO is S, once the optimum conditions for which maximum retention is reached is found, increases in molar ratios do not cause increases in the reactant fraction of Mo. It is worthwhile mentioning that the recovery that remains is due to the conditions set by 1:0.5:1, with respect to other variations.

4. Discussion and analysis of results

4.1. Temperature and time variations

Temperature is one of the determining variables in the roasting process to obtaining Mo. On the basis of a thermodynamic analysis we defined the possible and dominant chemical reactions, shown in Fig. 1, as well as in Equation 2. This agrees with the cited references [12, 13].

Fig. 4 shows that, initially, a considerable decrease exists in the Mo %, with values decreasing from 0.87% to 0.17%. Remarkable decreases take place in the first 20 min when roasting with carbothermic reduction was applied. This indicated that the formation of the various compounds quoted by other authors [7] took place.

Later, the Mo percentages are seen to oscillate from 20 to 80 min. These are specific for each of the temperatures studied (500 to 800 °C). The temperature that responded best was 600° C. As time elapses, beginning at approximately 80 minutes the Mo percentage starts to stabilize and its value becomes more constant. The Mo values obtained reach 0.20 %.

The situation is quite different for D samples, which show another behavior, as can be seen in Fig. 5. Higher percentages of Mo are obtained with D samples than they are with R samples.

The variations found for R samples are due to the formation of various compounds such as molybdates, molybdenum carbides, and molybdenum sulfates. Each of the compounds formed produce oscillations in the weights obtained. According to the various references [6, 8], this is accounted for by a complicated chemical reaction mechanism.

From the time and temperatures that are stated, the kinetic curves slowly increase. However, the 1% Mo value in the R samples does not prove to be optimal.

From the results obtained, it may be stated that in the roasting process, Cu and Fe high concentrations interfere with carbothermic reduction. This leads us to consider the roasting results as favorable.

4.2. Discussions on the behavior of variations in molar ratios for CaO with respect to Mo (Mo:CaO:C) for R samples.

When analyzing the sulfur scavenging agent CaO, oscillating variations may be observed for each of the R samples. The results obtained are shown in Figs. 6 to 9.

The function of CaO is to retain S. However, in order to find the optimal operational conditions in the value of its equimolar ratio, the % of Mo is always regarded as a reference. Therefore, in this work, sulfur retention values are not dealt with.

The graphs indicate that an increase in CaO stechiometric conditions does not improve the values of the reactant fraction or the reaction rate.

The recovery shown is for the 1:0.5:1 conditions in comparison with other variations. These conditions are shown in Fig. 6. The highest Mo values obtained ranged from 0.3 to 0.6 %. For the remaining molar ratios (Figs. 7 to 9), the

values found ranged from 0.1 to 0.5 % Mo.

When analyzing the C reducing agent, the work we undertook was similar to that performed with CaO: the results are shown in Figs. 10 to 12. A response quite similar to that of CaO can be observed. The difference noted when the amounts of C are increased is that Mo recovery improves for the1:0.5:1 condition and it is possible to obtain Mo values that range from 0.3 to 0.6%. For the remaining stoichiometric ratios of 1:0.5:0.5; 1:0.5:2 and 1:0.5:3, no improvements exist in Mo percentages higher than those obtained for the 1:0.5:1 conditions. Mo percentages are not higher than those obtained for the 1:0.5:1 conditions.

The similarity lies in there being a reduction in the Mo percentage during the first 20 minutes of the roasting process. This percentage then starts increasing and finally stabilizes for times longer than 30 minutes. According to the literature [13,14,15], this oscillation occurs because both the CaO and C work synergistically for molybdenum as well as for Cu and Fe.

4.3. Discussion of the molar ratio behavior for CaO and C with respect to Mo for D samples.

The roasting response when applied to concentrates produced with the use of differential flotation (D) is different from that obtained from concentrates produced with rougher flotation (R). These reach higher recoveries with values ranging from 32 to 35% and are shown in Table 3. For rougher (R) flotation, comparison is not possible as the results in the recovery values are not satisfactory.

This indicates that molybdenum separation from copper and iron gives rise to better results in MoO_3 recovery.

The optimal operational conditions for R samples are shown in Fig. 6. Those for D samples are shown in Fig. 12.

The results show that concentrates that used simple flotation or roughers (R1 a R5) do not increase productivity after being subjected to different operational conditions for roasting with carbothermic reduction. However, those concentrates obtained with differential flotation (D1 to D3) show a 35% increase -with respect to Mo- after being subjected to roasting with carbothermic reduction. This can be seen in detail in Table 3. Best recoveries were obtained for differential flotation.

This indicates that the behavior of differential flotation is very different to that of rougher flotation; the best recoveries are achieved in the former. We can conclude that the separation of high concentrations of molybdenum in copper and iron give rise to better results in the production of MoO_3.

4.4. Discussions on the sulfur values of samples R and D

Chemical analysis indicates that there is a decrease in the amount of sulfur in the roasting samples. This decline that can be seen is related to the mass loss of MoS_2 from the total system.

We can say, therefore, that the loss in the R% S in samples where carbothermic reduction was applied is greater, starting with a 34-40% of S and ending at approximately 10% (Fig. 14). This leads us to consider that the R has samples that are easily volatilized compounds throughout the temperature

range for which tests samples were performed. This is of great importance in explaining the differences in reduction.

For sample D we can refer to Fig. 15.It begins with a 30-35% S and ends with a 15-18% S. This indicates that sulfur retention was higher by raising the optimal working conditions and it is shown to contain lower concentrations. Cu% and Fe% weight losses are very different and contain no more volatile R samples.

5. Conclusions

Use of concentrates obtained with differential flotation processes, in comparison with rougher flotation processes, considerably improved molybdenum recovery after the concentrates were subjected to roasting with carbothermic reduction.

The recoveries obtained from the application of carbothermic reduction to concentrates produced with differential flotation were higher than those obtained with rougher flotation. Both samples (R y D) were subjected to identical operational conditions. As such, it can be concluded that a remarkable difference exists that favors the roasting process when concentration of elements such as Cu and Fe are reduced while attempting to recover Mo.

It was found that R samples subjected to toasting with carbothermic reduction had optimal operational conditions, for this condition the recovery the Mo was higher. The conditions were the following: Molar ratio specified for CaO and C was 1:0.5:1, working temperature was 600 ºC, and optimal roasting time was 90 minutes. This indicates that the kinetics is coincident with the optimal values of work, for which higher recoveries are obtained for both samples as R such as D.

The reagent most easily affected by molar ratio variation when applying roasting with carbothermic reduction is CaO and not C. CaO plays a most important role as a sulfur scavenger agent.

On analyzing the results, it can be seen that another variable that it is very important to take into account is the separation of copper and iron from molybdenum prior to the roasting process in order to obtain the desired recovery. This results in the capture of sulfur that is most favored for concentrates from differential flotation that obtained for the roughers flotation, being coincident with behavior of Mo.

References

[1]   Kholmogorov, A.G. and Kononova, O.N., Processing mineral raw materials in Siberia: ores of molybdenum, tungsten, lead and gold. Hydrometallurgy, [Online]. 76, pp. 37-54, 2005. Available at: http://ieg.or.kr/abstractII/I0407601005.html

[2]   Gupta, C.K, Extractive Metallurgy of Molybdenum. CRC Press, Florida, pp. 669-774, 1992. ISBN 0- 8493-4758-0.

[3]   Van den Berg, J.A.M, Yang, Y, Nauta, H.H.K, van Sandwijk, M.A., Comprehensive processing of low grade sulfidic molybdenum ores. Miner.Eng. 15, pp 879-883, 2002. DOI: 10.1016/S0892-6875(02)00135-8.

[4]   Valenzuela F.R., Andrade J.P., Sapag, J, Tapia, C., Basualto, C., The solvent extraction separation of molybdenum and copper from acid leach residual solution of Chilean molybdenite concentrate. Miner.Eng. 8, pp 893-904, 1995. DOI: 10.1016/0892-6875(95)00051-Q.

[5]   Navarro, P. and Alguacil, F., Extracción de Mo(VI) de disoluciones en medio nítrico mediante Alamine 336 ó Aliquat 336. Revista de Metalurgia, 31(6), pp 379-385, 1995. DOI: 10.3989/revmetalm. 1995.v31.i6.937

[6]   Padilla, R., Reducción carbotérmica de sulfuros metálicos, Rev. Inst. Investig. 5(9), ISSN 1561-0888.2002.

[7]   Lovera, D., Núñez, J.P., Zegarra, N.J. et al. Reducción carbotérmica de la calcopirita sin emisión contaminante. Rev. Inst. Investig. Fac. Minas metal cienc. Geogr. 3(5), pp.07-12. ISSN 1561-0888. 2000.

[8]   Padilla, R., Ruiz., M.C., and Sanchez, A., Reduction of molybdenite with carbon in the presence of lime, Chemical Metallurgy, Vol III, IV Meeting of Southern Hemisphere on mineral Technology, pp.383-395. 1994.

[9]   Kolttof I.M. y Sandell, E.B. Tratado de química analítica cuantitativa, 5º ed., Buenos Aires, Editorial Nigar; 1966, pp. 555-630

[10] Skoog, West Holler, Fundamentos de Química Analítica. 8º Ed.; Editorial Reveste, 2005.

[11] Isabel Sierra Alonso. Análisis Instrumental".Vol 1. Editorial Netbiblo SL, 2010.

[12] Orozco, I., Bazán, V., Diaz, A. y Sarquis, P. Reducción carbotérmica de minerales sulfurados. 11º Congreso Binacional de Metalurgia y Materiales SAM / CONAMET 2011.

[13] Kar, B.B., Murphy, B.V.R, Misra, V.N., Extraction of molybdenum from spent catalyst by salt-roasting. Int. J. Miner. Process. 76, pp. 143-147, 2005. DOI: 10.1016/j.minpro.2004.08.017

[14] Bazán, V., Brandaleze, E. y Colque, E., Cinética de tostación de concentrados de baja ley de molibdenita. DYNA, 80(181), pp. 149-152, 2013.

[15] Tamayo, L. Goñi, C. Ruiz, I. Barbés, M.F. Fernández, B., Marinas, E. Barbés, M.A. y Ayala, J., Caracterización fisicoquímica de acreciones presentes en un horno de lecho fluidizado para la tostación de blenda. DYNA, 79(175), pp. 90-96, 2012.

I.M. Orozco, is a BSc. Engineer in Food in 2007 and then in 2013, her MSc. In Extractive Metallurgy, both of these titles were obtained from the National University of San Juan, Argentina. She is currently pursuing a PhD in Mineral Processing Engineering at the UNSJ. She has worked as a teacher and researcher at the Institute of Mining Research, since 2007. She is an Eng in Mines Associate Professor at the UNS. Teachers and she integrate together the Extractive Metallurgy postgraduate program. She also conducts research in the area of instrumental and analytical chemistry at the laboratory of the Institute of Mining Research, Faculty of Eng. UNSJ.
ORCID:0000-0002-9120-3609.

V.L. Bazan, is a BSc. in Engineer Chemistry in 1999, from the National University of San Juan Argentina. In 2005 received degree from PhD in Extractive Metallurgy at the University of Concepción Chile. She is a research assistant CONICET. She is director of the program of Msc in Extractive Metallurgy and Co-Director of the Doctorate in Engineering Program in Mineral Processing UNSJ- in Argentina. She is also codirector of the graduate department of the Faculty of Engineering of the UNSJ. She is an Associate Professor of the School of Eng. UNSJ Mine. It also conducts research and management in the area of Chemical Laboratory of the Institute of Mining Research, Faculty of Eng. UNSJ since 2004.
ORCID:0000-0001-5766-6004.

A.A. Diaz, received a BSc. in Engineer Chemistry in 1991 and in 2001 received MSc in Extractive Metallurgy. Both titles at the National University of San Juan Argentina. In 2011 it received from PhD in Chemistry from the National University of San Luis, Argentina. He is an Associate Professor of the School of Eng. UNSJ Mine integrates the whole program graduate teachers Extractive Metallurgy Magister and Doctorate in Engineering in Mineral Processing in Argentina UNSJ- program. It also conducts research and management at the Laboratory of Physical Chemistry and the Institute of Mining Geomechanics Research School of Eng. UNSJ since 2000..
ORCID:0000-0001-6688-3780.

R.F. Lara, received a BSc. in Chemistry in 1978 and in 2000 received MSc. in Analytical Chemistry. In 2008 he received PhD in Chemistry. Each title was awarded by the National University of San Luis, Argentina. He is Professor of the School of Eng. UNSJ Mine integrates the whole program graduate teachers Extractive Metallurgy Magister and Doctorate in Engineering in Mineral Processing in Argentina UNSJ- program. It also performs tasks of management, research and management in the area of instrumental and analytical chemistry laboratory of the Institute of Mining Research, Faculty of Eng. UNSJ.
ORCID:0000-0002-2763-8763.

Multi-product inventory modeling with demand forecasting and Bayesian optimization

Marisol Valencia-Cárdenas [a], Francisco Javier Díaz-Serna [b] & Juan Carlos Correa-Morales [c]

[a] Institución Universitaria Tecnológico de Antioquia, Medellín, Colombia. mvalencia@unal.edu.co
[b] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. javidiaz@unal.edu.co
[c] Facultad de Ciencias, Universidad Nacional de Colombia, Medellín, Colombia. jccorrea@unal.edu.co

Abstract

The complexity of supply chains requires advanced methods to schedule companies' inventories. This paper presents a comparison of model forecasts of demand for multiple products, choosing the best among the following: autoregressive integrated moving average (ARIMA), exponential smoothing (ES), a Bayesian regression model (BRM), and a Bayesian dynamic linear model (BDLM). To this end, cases in which the time series is normally distributed are first simulated. Second, sales predictions for three products of a gas service station are estimated using the four models, revealing the BRM to be the best model. Subsequently, the multi-product inventory model is optimized. To define the policies for ordering, inventory, costs, and profits, a Bayesian search integrating elements of a Tabu search is used to improve the solution. This inventory model optimization process is then applied to the case of a gas service station in Colombia.

Keywords: Dynamic Linear Models, Inventory Models, Forecasts, Bayesian Statistics.

Modelo de inventario multi-producto, con pronósticos de demanda y optimización Bayesiana

Resumen

La complejidad de las cadenas de suministro exige mejores métodos para programar los inventarios de una empresa. En este trabajo se presenta una comparación entre modelos de pronósticos de demanda de múltiples productos, eligiendo el mejor entre: ARIMA, Suavización exponencial, Regresión Lineal Bayesiana y un Modelo Lineal Dinámico Bayesiano. Para ello, primero se realiza una simulación de casos donde no hay una Distribución Normal en las series de tiempo, segundo, se estiman las predicciones de ventas de tres productos de una estación de servicios de gasolina con los cuatro modelos, encontrando los mejores resultados para la Regresión Lineal Bayesiana. Seguido a esto, se presenta la optimización de un Modelo de Inventarios Multi-Producto. Para definir la política de pedidos, inventarios, costos y ganancias, se utiliza una búsqueda bayesiana, que integra elementos de búsqueda Tabú para mejorar la solución. Dicha Optimización del Modelo de Inventarios se aplica a un caso de una estación de combustibles en Colombia.

Palabras clave: Modelos Dinámicos Lineales, Modelos de Inventarios, Pronósticos, Estadística Bayesiana.

1 Introduction

The increasing complexity of supply chains resulting from the globalization of market economies, changes in customer preferences, and increasing competition among companies is intensifying the search for faster and better methods for decision-making and obtaining optimal solutions for several types of inventory systems.

In inventory models, as noted by Simchi-Levi [1,2], demand represents a very important variable that warrants substantial attention to determine adequate inventory policies, and in some cases, the behavior of this variable is stochastic, generating the

need for accurate forecasting methods. However, problems can arise because on occasion, the forecasting models are inappropriate or mistakes are made, leading to error-laden inventory policies.

Other problems associated with forecasting demand are related to changes in the distribution function, which produce a lack of stability in the time series [3]; indeed, *"... a time series is unstable if there are frequent and significant changes in the distribution"* [3]. This phenomenon has been cited by different authors [4–6]. Other disadvantages are that, sometimes, the models of interest cannot satisfy some theoretical assumptions, such as normality in residuals or constant variance. Alternatively, the researcher may not have sufficient required data for model estimation.

In this sense, Bayesian statistics can be very good alternative for making inferences using different types of models [7–9], as shown for Bayesian forecasting with the Holt winters model [10], the dynamic models proposed in [8,11], and especially, the situation described in [12], in which the authors describe making forecasts in R using a package they developed. Other works report using a combination of Bayesian techniques for forecasting [13–15], and [14] reveals that using such combinations results in increased accuracy and reliability. Bayesian forecasting has many practical applications [8,10,11,16–31]. These methods constitute alternatives to forecast and can be compared with classical methods to identify ways to further increase the accuracy of the required predictions.

In addition to the analytical techniques that can be used to solve inventory models, a practical problem-solving approach exists that is known as heuristics. Heuristics can be programmed according to some rules, but obtaining the optimal solution for a model is not always guaranteed.

Some works relating to inventory models have applied heuristic optimization. The Tabu Search algorithm can be used to search for solutions to inventory problems; for example, according to [32], this algorithm was used to minimize the inventory costs of an organization's final products and gave a better result than the company policy. Indeed, when it was applied to a real situation involving the same products with the same time horizon, it reduced inventory costs by 20% while achieving a 100% service level. In [33], the Tabu Search algorithm is applied to determine the optimum level of orders. Genetic algorithms can be used for efficient supply chain management [34] in multi-product scenarios, but such scenarios are infrequently analyzed using inventory models. A recent work in Colombia [35] proposed a model of multi-product inventory between companies to minimize logistics costs in an urban distribution operational context. In all of these cases, demand was predicted. The measure symmetric mean absolute percentage error (SMAPE), as described in [36], is useful when the response takes values close to zero because using it does not cause the error percentage to increase more when a response is very small.

In their book *Dynamic Linear Models with R*, Petris et al. [12] demonstrate many applications and the use of the package dlm, which they developed. They also list the error indicators used to compare models: mean absolute percentage error (MAPE), mean absolute deviation (MAD) and mean squared error (MSE) or root-mean-square error (RMSE).

In this paper, first, we report a simulation study allowing the recreation and analysis of the behavior of a demand time series with dynamic variation after finding an adequate model to forecast these types of data. The compared models are as follows: autoregressive integrated moving average (ARIMA); exponential smoothing (ES), a novel model developed in a doctoral thesis [37]; a Bayesian regression model (BRM); and a modified Bayesian dynamic linear model (BDLM) presented in that thesis in which a MAPE indicator is used for the comparisons. In this paper, the estimated models are compared using the SMAPE for forecasts once the data have been partitioned. Then, after applying the best model to a real case of combustible demand for a Colombian gas service station, the prediction is saved to do an optimization process.

Finally, we propose a multi-product inventory model that provides not only orders, inventory values, costs, and profits but also transportation durations and costs. The solutions obtained by this optimization process are based on a search that utilizes a Bayesian form to predict orders based on the previous forecasted demands. However, unlike that described in the aforementioned thesis, here, no missing values are considered, and the result for 15 days of planning is shown.

2. Demand forecasting

We consider four models for making predictions: ARIMA, ES, BRM, and BDLM. For this purpose, we program an algorithm using R software to choose the best possible forecast.

The ARIMA(p,d,q) model, which was developed in 1970, by George Box and Gwilym Jenkins [38,39] has been widely studied [40–44]. This model incorporates the characteristics of the same time series according to the autocorrelation results and makes predictions based on historical data.

ES is another oft-used technique [10,45,46] that employs exponential weights of past periods' values in the same series. This model can incorporate the level of the time series, trend and seasonality and it is expressed as follows: $\widehat{Y}_t = \alpha Y_t + (1 - \alpha)\widehat{Y}_{t-1}$, where $\widehat{Y}_t$ is the forecast for the next period, α is the smoothing constant, Y_t is the real value of the series in period t, and $\widehat{Y}_{t-1}$ is the predicted value for the period t–1. The response variable Y_t is adjusted according to a time horizon, and the sum of squared errors of prediction (SSE) value is optimized by searching for the value of α that minimizes it.

Bayesian statistics relies on different assumptions than classical models. For example, the parameters of a probability distribution are random variables, θ, and prior quantitative information is included in a probability distribution [47,48] known as $\xi(\theta)$, with simple information $(y_1, y_2, \ldots, y_n)$ summarized by the likelihood function $L(y_1, y_2, \ldots, y_n | \theta)$. Using the Bayes theorem—the a prior distribution $\xi(\theta)$ multiplied by the likelihood—gives the posterior distribution $\xi(\theta|datos)$. The predictive distribution is the integral of the distribution of the variable to be forecasted and the posterior distribution [49,50].

2.1. Bayesian regression model (BRM)

In [51], Zellner presents a BRM based on a diffuse prior distribution of β parameters. In this work, a different model will be presented: the novel BRM described in thesis above [37]. The

model presented there assumes a Normal prior distribution for the vector parameter β and applies an iterative process to the initial vector β_0 for every time t; however, its accuracy parameter is fixed: $\tau_0 = 1/\sigma$. The predictive distribution used for forecasting is derived in the thesis and is a Student's t-distribution; this derivation is also presented in the appendix of this paper.

The general equation of a linear model is the same in a Bayesian regression, but the estimation process is different. We describe this process here based on the general equation given by eq. (1), where y_t is the demand vector, and $x_1, …, x_k$ are the explanatory variables.

$$y_t = \beta_o + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon \qquad (1)$$

The likelihood function of the data, based in the Normal distribution, is shown in eq. (2), a prior Normal distribution for the β parameter vector, in eq. (3) and the posterior distribution is (4), obtained after the product of the prior times the likelihood, and the algebraic process. Here, $A = \beta'_0 \tau_0 \beta_0 + Y'Y$, and $\tilde{\beta} = (X'X + \tau_0)^{-1}(X'X\hat{\beta} + \beta_0 \tau_0)$

$$L(y_t|y_0, \beta) \propto \tau^T \, exp^{-\frac{\tau}{2}(Y-X\beta)'(Y-X\beta)} \qquad (2)$$

$$\xi(\beta, \tau) \propto \tau \, \tau_0 \, exp^{-\frac{\tau \tau_0}{2}(\beta-\beta_{oi})'(\beta-\beta_{oi})} \qquad (3)$$

$$\xi(\beta, \tau|\beta_0, \tau_0, y_0, y_t) \propto \tau^{\frac{T+1}{2}} exp^{-\frac{\tau}{2}[A(\beta-\tilde{\beta})'A^{-1}(X'X-\tau_0)(\beta-\tilde{\beta})+1]} \quad (4)$$

After taking the integral of the product of the posterior and future data distributions, the resulting predictive distribution is described by eq. (5), which is a Student's t-distribution, with mean $Y_n = X_+\tilde{\beta}$, ν degrees of freedom, and variance according to eq. (6). The forecasting process uses the eq. (5), and its mean considers the designed matrix X based on the adjustment of the regression eq. (1). The analytical formulation of this predictive distribution is also explained in the appendix.

$$f(Y_+|y_0, Y) = [((Y_+ - Y_n)'A^{-1}(Y_+ - Y_n) + 1)A]^{-\frac{T+4}{2}} \quad (5)$$

$$V = \frac{\nu}{\nu-2}A = \frac{\nu}{\nu-2}(Y_+ - Y_n)'(Y_+ - Y_n) \qquad (6)$$

2.2. Bayesian dynamic linear model (BDLM)

In [21], Meinhold and Singpurwalla explain the process used to update θ_t in a recursive form using the posterior Normal distribution to update the observed equation. The model proposed in this work is based on the procedure described in the aforementioned thesis [37], except that the mean value of the posterior Normal distribution are modified based on an average of past values before estimating the observation equation and without the M model used by Meinhold and Singpurwalla. The process is included in the algorithm designed in the R program.

2.3. Simulation of the forecasting process

We simulate different time series under control conditions. For the simulated series, an estimation process is executed using classical and Bayesian techniques.

Subsequently, a comparison is conducted using the SMAPE indicator. The objective is to determine the best forecasting model to use if the time series exhibits high variability and apply it to a real case to obtain predicted values useful for developing and proposing an inventory policy.

The demands analyzed are mainly of a continuous nature, but they do not always follow a Normal distribution. The following cases are used to create the simulated time series:
S1. Regression model with skew Normal distribution for errors, with seasonal and dynamic variation.
S2. Regression model with skew T-distribution for errors, with seasonal and dynamic variation.
S3. Random variable with Poisson distribution.
S4. Random variable with Poisson distribution and seasonal and dynamic variation.

The size of each time series is N = 63 data points and a seasonal variation of seven periods. Each time series consists of 49 data points for the adjustment and 14 data points for forecasting.

The SMAPE of the forecasted values is used as the criterion for identifying the best model, and it is calculated with eq. (7) according to [36], after estimating each of the four models being compared: ARIMA, ES, BRM, and BDLM.

$$SMAPE = \frac{1}{k} \frac{|\hat{z}_{t+1} - Z_{t+1}|}{(Z_{t+1} + \hat{z}_{t+1})/2} \qquad (7)$$

where:
$\hat{z}_{t+1}$: Forecasted value of the demand at period t+1.
Z_{t+1}: Real value of the demand at period t+1.
N : Total number of data points.
k : Total number of forecasted data points.
N–k : Total number of data points used to adjust the model.

The algorithms for the simulation were developed using the R program [52]. For the BRM, a vector of 200 percentile values is calculated, and then, a prediction is made using the Student-t predictive distribution. The error indicator SMAPE is calculated for the adjusted and forecasted data for all models to facilitate finding the minimum SMAPE value.

This process is repeated a thousand times, and the program calculates the percentage of selection for every model and simulated case. The results of the frequency with which each tested model is chosen as the best model are shown in Table 1. For all simulated time series, the BRM is identified as the best model.

Table 2 shows that, for one simulated case, as indicated by the grey color, lower SMAPE values are found for the BRM in all analyzed cases. The results of the two tables are consistent for all scenarios, confirming the reliability of the results.

These results show that when few data points with high variance are used, the BRM model produces the minimum SMAPE values for all simulated series.

Table 1.
Percentage of times each model is chosen as the best based on the simulated data (N = 63).

Model	ARIMA	ES	BRM	BDLM
SN	0	0	100	0
ST	0	0	100	0
Poisson	0	0	100	0

Dynamic Poiss	0	0	100	0

Source: The authors

Table 2.
SMAPE indicators for the simulated data (N = 63).

Model	ARIMA	ES	BRM	BDLM
ST	6.40	6.40	2.05	68.87
SN	5.78	5.78	3.48	69.92
Poisson	25.92	25.92	11.78	48.65
Dynamic Poiss	51.06	51.05	16.88	56.32

Source: The authors

Up to this point, we have shown that in forecasting, when no Normal behavior exists, the BRM is a very good alternative, to the two common used classical models and the new BDLM model, which failed to provide good results in any case.

3. Multi-product inventory modeling

3.1. Previous models

Some classical inventory models were formulated based on Wagner and Whitin's proposed [53] cost minimization. These authors specify that an order could be '0' or the sum of some demands, and they formulated schemes for ordering (i.e., choosing between generating orders or not) in every period t.

The model formulated by Wagner and Whitin, which is cited in [54], assumes that a sequence of orders over a planning horizon with a duration of T exists. The model assumptions are shown in Table 3.

$$Minimize \ Z = \sum_{i=1}^{T}[K\delta(y_t) + hI_t] \qquad (8)$$

Subject to: $\quad I_t = I_{t-1} + y_t - d_t \qquad (9)$
$$I_t = 0, \quad with \ y_t, I_t > 0, \ t = 1, \dots \dots T.$$

where the eq. (9) describes the inventory-balance constraint for every period t.

3.2. Proposed multi-product inventory model

A general mathematical model of inventory management is formulated, and subsequently, an optimization algorithm is developed.

We propose performing profit maximization using a formula similar to the Wagner-Within model with the addition of transportation costs CTR_t. Almost equivalent results can be obtained by considering the problem as one of cost minimization.

Table 3.
Assumptions of Wagner and Whitin's model [54].

Parameter	Assumption
d_t	Demand for period t, $d_t > 0$
C	Cost per unit order
K	Fixed order cost when an order is placed
$\delta(y)$	1 if an order is placed: y>0; 0 otherwise
H	Holding cost per unit per period
I_0	Initial inventory (zero)
Ld	Lead times (zero)
	Decision Variables
y_t	Order placed at the start of period t
I_t	Inventory charged at the end of the period

Source: (54)

Table 4.
Parameters for the proposed inventory model.

Parameters	
T	Total periods of the horizon time
c_i	Cost of the i-th product (i=1, 2, ..., K)
h_i	Cost of the holding inventory for the i-th product
Y_{ij}	Demand for product i in period j (j=1,2, ...,T)
CTR	Transportation costs
$Porc\_$	
ct	Percentage of the holding cost depending on the c_i costs
K	Number of products.
I_{i0}	Initial inventory of the i-th product
Cap_i	Storage capacity for the i-th product
S	Stock, a random variable, which is a minimum quantity added to each order
Decision Variables	
X_{ij}	Ordered quantity to be supplied at the beginning of period j
$Comp_j$	Number of occupied compartments in the transportation vehicle for period j
Output variables	
I_{ij}	Final inventory during period j
$Vehicle_j$	Number of vehicles to be sent in period j

Source: The authors

Objective function for the inventory model.

$$Maximize \quad Z = \sum_{i=1}^{K}\sum_{j=1}^{T}p_{it} * min(y_{ij}, x_{ij} + I_{i(j-1)}) -$$
$$\left[\sum_{i=1}^{K}\sum_{j=1}^{T}c_i x_{ij} + \sum_{i=1}^{K}\sum_{j=1}^{T}h_i(I_{ij}) + \sum_{j=1}^{T}CTR * Vehicle_t\right] \quad (10)$$

Subject to an inventory balance constraint

$$I_{ij} = max\left\{0, I_{i(j-1)} + x_{ij} - y_{ij}\right\} \qquad (11)$$

Capacity constrains for the orders

$$x_{ij} \leq Cap_i \qquad (12)$$

The number of vehicles

$$Vehicles_t = \sum_{i=1}^{k}\left[\!\left[\frac{1}{nc}\left[\!\left[\frac{x_{ij}}{Capc}\right]\!\right]\right]\!\right] \qquad (13)$$

where y_{ij} represents the forecasted demand for product i in period j. The transportation quantities and their respective costs will depend on the number of vehicles to be used ($Vehicles_t$), which depend on the number of compartments, nc, and these, depend on those compartments' capacities, $Capc$, assuming equal values. Dividing the quantity to order x_{ij} into $Capc$, generates a number of compartments, and the maximum integer value of this number is divided into nc, producing the total number of vehicles required, as represented by eq. (13), which can then be substituted into the objective function shown in eq. (10).

The final inventory I_{ij} can be 0 or positive, depending on the orders placed, as described in this work. The inventories are limited by the storage capacity. Therefore, the orders cannot exceed that capacity and can be equal to that maximum when there is zero inventory

3.1.1. Algorithm

We describe the scenarios analyzed to solve the proposed multi-product inventory model. The algorithm was programmed in R software.

The schemes to order are of two kinds, S1 and S2, and the model was explained in the past section. These schemes are replaced in the same model of the equations (10) to (13), and the best solution is kept, in order to be compared and finding the maximum possible profits. The process is repeated, and the random variable S, added to the orders (see Table 5), helps to find a best possible solution, after a period prepared to simulate.

Two types of schemes are used for ordering: S1 and S2; the model is described above. These schemes are substituted into the model described by eqs. (10)–(13), and the best solution is retained for posterior comparison and profit maximization with the results obtained after the process repetition, by using the random variable S, which is added to the orders (Table 5), and it facilitates finding the best possible solution for the simulated period.

We calculate a maximum solution for each S1-model and S2-model combination and save the best of them. After performing a simulation of size "*sim*", a list of solutions is created, and subsequently, the saved maximums are compared, the higher value is selected, and the process is repeated until the best possible solution is obtained according to the following convergence criterion: a difference of zero between *ten* consecutive values of Z.

This process is presented in Fig. 1.

- Ordering schemes

The first ordering scheme, S1, is based on theorem 2 of [53], which states that *"there exists an optimal program such that for all t":*

$$x_t = 0 \ \text{ or } \ x_t = \sum_{j=d}^{k} d_j, \text{ for some k, t<k<N}$$
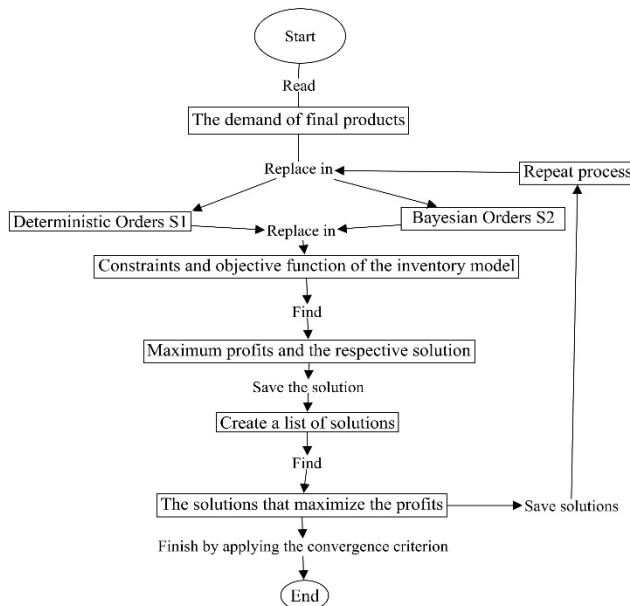

Figure 1. Flowchart of the heuristic procedure.
Source: The authors

Table 5.
Ordering schemes.

Scheme /period	1	2	3	…	t	T
1	$\sum_{i=1}^{T} Y_{it}$	0	0	…	0	0
2	$(Y+S)_1$	$\sum_{i=2}^{T} Y_{it}$	0	…	0	0
3	$(Y+S)_1$	$(Y+S)_2$	$\sum_{i=3}^{T} Y_{it}$	…	0	0
…	…					
5	$\sum_{i=1}^{2} Y_{it}$	0	$\sum_{i=3}^{4} Y_{it}$	0	$\sum_{i=5}^{6} Y_{it}$	0
t	$(Y+S)_1$	$(Y+S)_2$	$(Y+S)_3$	…	$(Y+S)_t$	$\sum_{i=T-1}^{T-1} Y_{it}$
R–1	Q	Q	Q	Q	Q	Q
R	$(Y+S)_1$	$(Y+S)_2$	$(Y+S)_3$	…	$(Y+S)_t$	$(Y+S)_T$

Source: The authors

Thus, the formulation of S1 considers the dynamics shown in Table 5. These R vectors of values are replaced in the model described by eqs. (10)–(13), saving all of the equations and the objective function, and finding the solution that maximizes the profits.

The second scheme, S2, generates orders based on a predictive Bayesian distribution, as explained in section 3.2.2. This allows a random variable based on the Bayesian process that depends on previously forecasted demands to update the posterior distribution. Subsequently, scheme S1 is used with random order values.

Ordering schemes: Explanations corresponding to Table 5:
1. Ordering in the first period to satisfy all the demands estimated for the planning horizon.
2. Ordering in the first period to satisfy only demand 1, and ordering in the second period to satisfy the sum of the remaining demands.
3. Ordering in the first two periods, and ordering in the third period to satisfy the sum of the remaining demands.
4. Same as 2 and 3 for resting periods, until T–1; no ordering in the T-th period.
5. Ordering every two periods.
6. Ordering every three periods.
7. Ordering every four periods and so on.
8. Ordering the economic order quantity (EOQ).
M. Ordering in all periods.

3.1.2. Bayesian optimization

This is the procedure involving the second scheme proposed (S2). The assumptions of this process are as follows:

The data distribution is uniform (Data-unif(a1,b1)). Let 'μ' be the mean value of this distribution, with a prior Normal truncated distribution with the following parameters: μ_o, mean; σ_o, standard deviation (assumed to be constant); a, inferior limit; and b: superior limit (10).

$$\xi(\mu) = \frac{f(\mu|datos)}{\Phi(\mu,\sigma,b,x)-\Phi(\mu,\sigma,a,x)} = \frac{\frac{1}{\sqrt{2\pi}\sigma_o}e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}}{F(b)-F(a)}\alpha e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}} \quad (13)$$

The product of the prior distribution and the likelihood function is the Truncated Normal (TN), that is, the posterior distribution (eq. (14)) of the mean parameters of the final predictive distribution.

$$\xi(\mu,\sigma_o|datos)\alpha\,\xi(\mu,\sigma_o)*\frac{n}{b_1-a_1}$$
$$\xi(\mu,\sigma_o|datos)\alpha\,\frac{n}{b_1-a_1}*$$
$$\frac{f(\mu,\sigma_o|datos)}{F(b)-F(a)}\alpha\frac{1}{\sigma_o}e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}} \quad (14)$$

The predictive distribution is the integral of the Uniform distribution of the data and the posterior TN (14). This distribution is uniform (15) and is used for forecasting, after updating the means, with the posterior distribution for every time t.

$$P(X_{t+1}|x)\alpha\int_{-\infty}^{\infty}\frac{1}{b_1-a_1}*\frac{1}{\sigma_o}e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}d\mu=\frac{\sqrt{2\pi}}{b_1-a_1}\,\alpha\,\frac{1}{b_1-a_1} \quad (15)$$

4. Results

4.1. Comparison of the forecast results obtained using a real case

The dataset of the fuel sales includes 92 daily values for each product: Regular, Extra, and Diesel. In total, 77 data points will be used to adjust the models, and 15 data points will be used for forecasting. The last forecasted values will be used to optimize the models. For the BRM, a vector of 200 percentiles was used to identify the best possible position of the predictive Student T-distribution based on the minimum SMAPE value of the 15 forecasts.
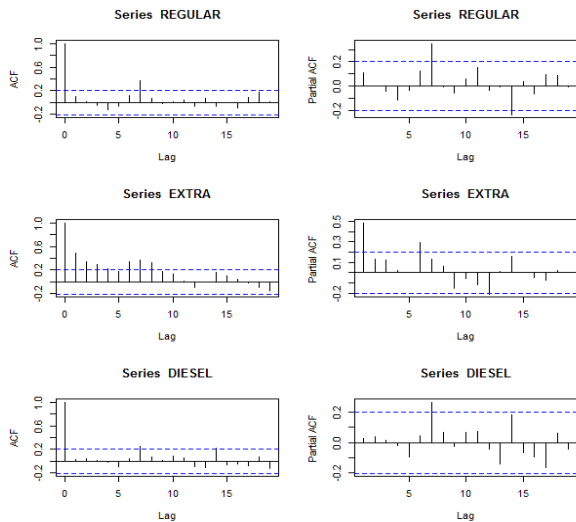

Figure 2. ACF and PACF of the three series: Regular (line 1), Extra (line 2) and Diesel (line 3).
Source: The authors

Table 6.
SMAPE indicators (%) for adjusted-forecasted data.

	ARIMA	ES	BRM	BDLM
ADJUSTED SMAPE				
REGULAR	11.859	12.184	10.625	53.421
EXTRA	39.105	39.419	38.320	99.459
DIESEL	27.873	28.129	22.082	75.749
FORECASTED SMAPE				
REGULAR	12.78	12.29	8.85	62.53
EXTRA	31.53	29.21	32.16	98.00
DIESEL	23.67	25.29	9.73	20.65

Source: The authors

In the description of the Regular and Diesel combustibles, the behavior seems to have a seasonal pattern with seventh-order dependence, whereas for Extra, the behavior of the autocorrelation shows a dependence of first order (Fig. 2), and it also shows seasonality every 7 periods (days). These aspects are considered to estimate the models and are used to define covariables to represent these behaviors.

We also found that the residuals of the ARIMA models estimated do not follow the normality assumption.

- Description of the time series analysis

Fig. 2 presents the autocorrelation and partial autocorrelation functions (ACF and PACF, respectively), they show that the time series exhibits dependence of different orders (i.e., Regular, Extra and Diesel fuels).

Table 6 presents the adjusted and forecasted SMAPE indicators (adjusted-forecasted), when the criterion used to make a choice is based on one of these. Some advantages of the BRM models are clearly evident.

According to the adjustment criteria, the BRM is the best at forecasting the three products. Additionally, among the forecasted values, the BRM produces the best results for Regular and Diesel fuels, whereas for Extra, the ES is better, followed by ARIMA (it should be noted that ARIMA's result is close to that of the BRM). The minimum adjustment criterion is important because a researcher interested in forecasting would not be able to choose the future data. Indeed, when the BRM is selected, the probability of obtaining good performance is high. This approach leads us choose the BRM as the model to forecast to these time series.

The results obtained by applying the optimization process to a 15-day time horizon of fuel distribution and inventory planning are presented below.

4.2. Optimization of the application of the inventory model to a real case

If there were no holding costs and if the inventory capacity were very high, the optimum strategy would be to send all the demand for the T periods on the first day.

However, based on the real cost and capacity data, the results obtained depend on the state of the initial inventories. The designed algorithm compares the four combination models proposed and requires no more than 1 minute to produce results when the initial inventories exceed zero:

- When the initial inventories are set according to real data (4300, 1000, and 2150 gallons of Regular, Extra, and Diesel, respectively), the maximum profit is $69,432.115,

and the orders are not sent every day. Instead, Regular fuel orders are sent on days 4, 7, 8, 10, 11 and 13, and Extra fuel orders are sent on days 7, 10 and 13. The model did not indicate that sending diesel fuel was necessary because of the initial inventory.

- When the initial inventories are all equal to zero, the optimal profit is $12,737.930. However, companies typically prefer to have non-zero initial inventories.

4.3. Validation

- Using real data, the profits were $30,334.447 for a 15-day planning period.
- Using the model proposed here, with current initial inventories for each fuel (4300, 1000, and 2150 gallons of Regular, Extra, and Diesel fuel, respectively) and by substituting the planned orders for the real sales, the profits are $53,619.824, which is higher than the original value. This difference represents a saving of 58.7% relative to the current situation ($30,334.447).

This result shows that the model can provide a better solution for inventory management than what was achieved in the real case.

5. Discussion

The proposed Bayesian forecasting methods establish various accurate alternatives for making predictions compared to some classical methods, such as ARIMA and ES. The modified version of the BRM proposed in the aforementioned thesis [37] (in this case, the SMAPE indicator is used) was capable of generating satisfactory results, as was also shown using MAPE in the thesis.

This alternative Bayesian method can be improved by altering its parameters or probability distributions to obtain more accurate values, which would result in a better optimization solution. Additionally, different research lines could potentially benefit from this optimization technique, which can also be applied to the Dynamic Linear Models described by other authors, including Petris, Harrison and Stevens, and Harrison and West [8,9,11].

6. Conclusions

In addition to maximizing its profits, the proposed inventory model could also consider the fuel transportation costs incurred by the gas service station, and it functions a very good planning tool, especially compared to other models, such as EOQ, which was not found to be the best model in any case. Transportation costs change in every period, and it is not necessary to send vehicles every day. This model also revealed that sending all of the 15-day demands at the beginning of each period was non-optimal, since these quantities involved could exceed the available storage capacities. Furthermore, the model indicated that, if the related costs were zero and the capacity were very high, satisfying all the demands of the T periods during the first period would be optimal.

Appendix

Here, we show the transformation of the exponent in eq. (4) to obtain the predictive Student T-distribution (5).

$$I = \left[A(\beta - \tilde{\beta})' A^{-1} (X'X - \tau_0)(\beta - \tilde{\beta}) + 1 \right]$$

Where $A = \beta'_0 \tau_0 \beta_0 + Y'Y$, and $\tilde{\beta} = (X'X + \tau_0)^{-1}(X'X\hat{\beta} + \beta_0\tau_0)$. The exponent can also be expressed as:

$$I = Y'_+ Y_+ - \beta'_n M \beta_n + Y'Y + \beta'_0 \tau_0 \beta_0$$

Where $M^{-1} = (X'_+ X_+ + X'X + \tau_0)^{-1}$, and $\beta_n = M^{-1}(X'Y + X'_+ Y + \beta_0 \tau_0)$

$$I = Y'_+ Y_+ - M^{-1}(X'Y + X'_+ Y_+ + \beta_0\tau_0) M M^{-1}(X'Y + X'_+ Y_+ + \beta_0\tau_0) + Y'Y + \beta'_0 \tau_0 \beta_0$$

$$I = Y'_+ Y_+ + Y'Y + \beta'_0 \tau_0 \beta_0 - M^{-1}[(X'Y)'(X'Y) + (X'_+Y_+)'(X'_+Y_+) + (\beta_0\tau_0)'\beta_0\tau_0 + 2(X'Y)'(X'_+Y_+) + 2(X'Y)'(\beta_0\tau_0) + 2(X'_+Y_+)'(\beta_0\tau_0)]$$

$$I = Y_+(I - X_+ M^{-1} X_+)Y_+ - 2Y_+ X_+ M^{-1}(X'Y + \beta_0\tau_0) + \beta'_0 \tau_0 \beta_0 - M^{-1}(Y X'X'Y + \tau_0\beta'_0\beta_0\tau_0 + 2X'Y\beta_0\tau_0) + Y'Y$$

$$I = [Y_+Y_+ - 2Y_+X_+(I - M^{-1}X'_+X_+)^{-1}M^{-1}(X'Y + \beta_0\tau_0)] * (1 - M^{-1}X'_+X'_+) + (\beta'_0\tau_0\beta_0 - M^{-1}\tau_0\beta'_0\beta_0\tau_0) + (Y'(I - M^{-1}X'X)Y - 2M^{-1}X'Y\beta_0\tau_0)$$

$$I = [(Y_+ - Yn)'(Y_+ - Y_n) - Yn'Yn](1 - M^{-1}X'_+X'_+) + (\beta'_0\tau_0\beta_0 - M^{-1}\tau_0\beta'_0\beta_0\tau_0) + (Y'Y - 2M^{-1}(I - M^{-1}X'X)^{-1}X'Y\beta_0\tau_0)(I - M^{-1}X'X)$$

$$I = [(Y_+ - Yn)'(Y_+ - Y_n) - Yn'Yn](1 - M^{-1}X'_+X'_+) + (\beta'_0\tau_0\beta_0 - M^{-1}\tau_0\beta'_0\beta_0\tau_0) + [(Y - Y_m)'(Y - Y_m) - Y'_m Y_m](I - M^{-1}X'X)$$

Where $Y_n = (I - X_+ M^{-1} X_+)^{-1} X_+ M^{-1}(X'Y + \beta_0\tau_0)$, according to a definition of inverted difference in matrices [51].

$$Y_n = (I + X_+(X'X)^{-1}X_+)M^{-1}X_+(X'Y + \beta_0\tau_0)$$

Therefore, $Y_n = (I + (X'X)^{-1}X'_+X_+)(X'X + \tau + X'_+X_+)^{-1}X_+(X'Y + \beta_0\tau_0)$

$$= (I + (X'X)^{-1}X'_+X_+)(I + (X'X + \tau_0)^{-1}X'_+X_+)^{-1}X_+(X'X + \tau_0)^{-1}(X'Y + \beta_0\tau_0)$$

τ_0 is a very small quantity. Thus, $Y_n = (X'X + \tau_0)^{-1}X_+(X'Y + \beta_0\tau_0) = X_+\tilde{\beta}$, and

$\tilde{\beta} = (X'X + \tau_0)^{-1}(X'Y + \beta_0\tau_0)$. Let $Y_m = M^{-1}(I - M^{-1}X'X)^{-1}X'\beta_0\tau_0$

Some terms disappear because of the proportionality:

$$I = [(Y_+ - Y_n)'(Y_+ - Yn) - Yn'Yn](1 - M^{-1}X'_+X_+) + [(Y - Y_m)'(Y - Y_m) - Y_mY_m](I - M^{-1}X'X)$$

$$I = [(Y_+ - Yn)'(Y_+ - Yn) + (Y - Y_m)'(Y - Y_m)$$

$$f(Y_+|y_0, Y) = [((Y_+ - Y_n)'(Y_+ - Y_n) + A]^{-\frac{T+4}{2}}$$

$$f(Y_+|y_0, Y) = [((Y_+ - Y_n)'A^{-1}(Y_+ - Y_n) + 1)A]^{-\frac{T+4}{2}}$$

which is the Student t-distribution described by eq. (5).

Acknowledgments

Colciencias is acknowledged for providing a scholarship (567) in support of obtaining a Doctorate in Engineering-Industry and Organizations at Universidad Nacional de Colombia, Sede Medellín.

References

[1] Simchi-Levi, D., Kaminski, P. and Simchi-Levi, E., Designing and managing the supply chain. 3rd ed. New York: McGraw-Hill; 2008.

[2] Chen, X. and Simchi-Levi, D., Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. Operations Research, 2004, 52(6), pp. 887-896. DOI: 10.1287/opre.1040.0127

[3] Hillier, F. y Hillier, M., Métodos cuantitativos para administración. Third Ed. City: México. McGraw-Hill; 2007.

[4] Garcia, C.A., Ibeas, A., Vilanova, R. and Herrera, J., Inventory control of supply chains: Mitigating the bullwhip effect by centralized and decentralized internal model control approaches. European Journal of Operational Research, 224(2), pp. 261-272, 2013. DOI: 10.1016/J.EJOR.2012.07.029

[5] Sarimveis, H., Patrinos, P., Tarantilis, C.D. and Kiranoudis, C.T., Dynamic modeling and control of supply chain systems: A review. Computers and Operations Research. 35(11), pp. 3530-3561, 2008. DOI: 10.1016/J.COR.2007.01.017

[6] Braun, M.W., Rivera, D.E., Flores, M.E., Carlyle, W.M. and Kempf, K.G., A model predictive control framework for robust management of multi-product, multi-echelon demand networks. Annual Reviews in Control, 27(2), pp. 229-245, 2003. DOI: 10.1016/j.arcontrol.2003.09.006

[7] Pole, A., West, M. and Harrison, J., Nonnormal and nonlinear dynamic Bayesian modeling. In Bayesian analysis of time series and dynamic linear models. New York: Marcel Dekker; 1988, pp. 167-198.

[8] West, M. and Harrison, J., Bayesian forecasting and dynamic models. Second ed. New York: Springer Series in Statistics; 1997.

[9] Petris, G., An R package for dynamic linear models. Journal of Statistical Software [online], 36(12), pp. 1-16, 2010. Available at: http://www.jstatsoft.org/

[10] Bermúdez, J.D., Segura, J.V. and Vercher, E., Bayesian forecasting with the Holt–Winters model. Journal of the Operational Research Society, 61(1), pp. 164-171, 2009. DOI: 10.1057/jors.2008.152.

[11] Harrison, J. and Stevens, C., Bayesian forecasting. Journal of the Royan Statistical Society, 38(3), pp. 205-247, 1976.

[12] Petris, G., Petrone, S. and Campagnoli, P., Dynamic linear models with R [online]. 2009. Available at: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-0-387-77237-0

[13] Kociecki, A., Kolasa, M. and Rubaszek, M., A Bayesian method of combining judgmental and model-based density forecasts. Economic Modelling, 29, pp. 1349-1355, 2012. DOI: 10.1016/j.econmod.2012.03.004

[14] Coelho, C., Pezzulli, S, Balmaseda, M., Doblas-Reyes, F. and Stephenson, D., Forecast calibration and combination: A simple Bayesian approach for ENSO. Journal of Climate. 17(7), pp. 1504-1516, 2004. DOI: DOI: 10.1175/1520-0442(2004)017<1504:FCACAS>2.0.CO;2

[15] Andersson, M. and Karlson, S., Bayesian forecast combination for VAR models. Sveriges Riskbanc-working Papers [online]. pp. 1-17, 2007. Available at: http://www.riksbank.se/Upload/Dokument_riksbank/Kat_publicerat/WorkingPapers/2007/wp216.pdf

[16] Bijak, J., Bayesian methods in international migration forecasting. CEFMR Working Papers. Warsaw: Central European Forum for Migration Research, 2005.

[17] Clements, M.P. and Hendry, D.F.H., Forecasting non-stationary economic time series. Cambridge: MIT Press; 2000, pp. 1-6.

[18] Craig, P., Goldstein, M., Rougier, J. and Seheult, A.H., Bayesian forecasting for complex systems using computer simulators. Journal of the American Statistical Association, 96(454), pp. 717-729, 2001.

[19] Duncan, G., Gorr, W. and Szczypula, J., Bayesian unrelated time forecasting series: For seemingly to local forecasting application government revenue. Management Science, 39(3), pp. 275-293, 1993.

[20] Li, G., Shi, J. and Zhou, J., Bayesian adaptive combination of short-term wind speed forecasts from neural network models. Renewable Energy, 36(1), pp. 352-359, 2011. DOI: 10.1016/j.renene.2010.06.049

[21] Meinhold, R.J. and Singpurwalla, N.D., Understanding the Kalman Filter. The American Statistician, 37(2), pp. 123-127, 1983. DOI: 10.2307/2685871

[22] Neelamegham, R. and Chintagunta, P., A Bayesian model to forecast new product performance in domestic and international markets. Marketing Science [online]. 18(2), pp. 115-136, 1999. Available at: http://bear.warrington.ufl.edu/centers/mks/articles/684541.pdf

[23] Oracle, Inc., The Bayesian Approach to Forecasting [online], 2006. Available at: http://www.oracle.com/us/products/applications/057028.pdf

[24] Pedroza, C., A Bayesian forecasting model: Predicting U.S. male mortality. Biostatistics, 7(4), pp. 530-550, 2006.

[25] Pezzulli, S., Frederic, P., Majithia, S., Sabbagh, S, Black, E, Sutton, R, et al., The seasonal forecast of electricity demand: A simple Bayesian model with climatological weather generator. Applied Stochastic Models in Business and Industry, 22(2), pp. 1-16, 2006. Available at: http://empslocal.ex.ac.uk/people/staff/dbs202/publications/2005/pezzullib.pdf

[26] Popova, I., Popova, E. and George, E., Bayesian forecasting of prepayment rates for individual pools of mortgages. Bayesian Analysis, 3(2), pp. 393-426, 2008.

[27] Putnam, B., Practical experiences in financial markets using Bayesian forecasting systems [online], 2007. Available at: http://www.math.uchicago.edu/~cfm/BP-papers/Lessons_from_Bayesian_Experiences.pdf

[28] Sloughter, J.M., Raftery, A.E. and Gneiting, T., Probabilistic quantitative precipitation forecasting using Bayesian model averaging. Monthly Weather Review, 135(9), pp. 3208-3220, 2006.

[29] Valencia, M., and Correa, J. Un modelo dinámico bayesiano para pronóstico de energía diaria. Revista Ingenieria Industrial. 2013;12(2), pp.7–17.

[30] Valencia, M., Correa J.C., Díaz F. y Ramírez, S., Aplicación de modelación bayesiana y optimización para pronósticos de demanda. Ingenería y Desarrollo, 32(2), pp. 179-199, 2014.

[31] Yelland, P.M., Bayesian forecasting of parts demand. International Journal of Forecasting, 26(2), pp. 374-396, 2010. DOI: 10.1016/j.ijforecast.2009.11.001

[32] Valencia, M., González, D. y Cardona, J., Metodología de un modelo de optimización para el pronóstico y manejo de inventarios usando el metaheurístico Tabu. Revista de Ingeniería. 24(1), pp. 13-27, 2014. DOI: 10.15517/ring.v24i1.13771

[33] Urrea, A. y Torres, F., Optimización de una política de inventarios por medio de búsqueda tabú. In: III Congreso colombiano y I Conferencia Andina internacional [online], 8 P. 2006, Available at: http://dspace.uniandes.edu.co:9090/xmlui/handle/1992/822

[34] Jeyanthi, N. and Radhakrishnan, P., Optimizing multi product inventory using genetic algorithm for efficient supply chain management involving lead time. International Journal of Computer Science and Network Security [online], 10(5), pp. 231-239, 2010. Available at: http://paper.ijcsns.org/07_book/201005/20100534.pdf

[35] Palacio, O. y Adarme, W., Coordinación de inventarios: Un caso de estudio para la logística de ciudad. DYNA, 81(186), pp. 295-303, 2014.

[36] Wallström, P. and Segerstedt, A., Evaluation of forecasting error measurements and techniques for intermittent demand. International Journal of Production Economics, 128(2), pp. 625-636, 2010.

[37] Valencia, M., Dynamic model for the multiproduct inventory optimization with multivariate. PhD. Thesis. Department of Engineering, Facultad de Minas, Universidad Nacional de Colombia, Sede Medellín, Colombia, 2016.

[38] Bowerman, B.L. y Oconnell, R.T., Pronósticos, series de tiempo y regresión: Un enfoque aplicado. Editores. CL, editor. México, 2007.

[39] Valencia, M., Ramírez, S., Tabares, J. y Velásquez, C., Métodos de pronósticos clásicos y bayesianos con aplicaciones. Report. Universidad Nacional de Colombia, 2014.

[40] Wei, W.W.S., Time series analysis. Reading: Addison-Wesley, 1994.

[41] Makridakis, S., Hibon, M. and Moser, C., Accuracy of forecasting : An empirical Investigation. Journal of the Royal Statistical Society. Series A, 142(2), pp. 97-145, 1979. DOI: 10.2307/2345077.

[42] Makridakis, S., Wheelwright, S. and McGee, V., City: NewYork. Second Ed. John Wiley and Sons, 1983.

[43] Diebold, F., Elementos de pronósticos. City: México. International Thomson editors, 1999.

[44] Wilson, J.H., Keating, B. and Galt, J., Pronósticos en los negocios. Fifth Ed. McGraw-Hill; 2007. 461 P.

[45] Wang, S., Exponential smoothing for forecasting and bayesian validation of computer models [online]. Thesis, Georgia Institute of Technology, [Online]. 2006. Available at: https://smartech.gatech.edu/bitstream/handle/1853/19753/wang_shuchun_200612_phd.pdf

[46] Yelland, P.M. and Lee, E., Forecasting product sales with dynamic linear mixture models. Sun Microsystems, 2003.

[47] Gill, J., Bayesian methods: A social and behavioral sciences approach. Second Ed. Boca Raton: Chapman and Hall, 2007.

[48] Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., Bayesian data analysis. Second Ed. Boca Raton: Chapman and Hall, 2004.

[49] Barrera, C. y Correa, J., Distribución predictiva bayesiana para modelos de pruebas de vida vía MCMC. Revista Colombiana de Estadística [online], 31(2), pp. 145-155, 2008. Available at: http://www.emis.ams.org/journals/RCE/ingles/V31/bodyv31n2/v31n2a01BarreraCorrea.pdf

[50] Congdon, P., Bayesian statistical modelling. London: Wiley Series in Probability and Statistics, 2002.

[51] Zellner, A., An introduction to bayesian inference in econometrics. Second Ed. New York: Wiley, 1996.

[52] R Core Team., A Language and environment for statistical computing [online]. Vienna: R Foundation for Statistical Computing, 2014. Available at: http://www.r-project.org/

[53] Wagner, H.M. and Whitin, T.M., Dynamic version of the economic lot size model. Management Science, 5(1), pp. 89-96, 1958.

[54] Simchi-Levi, D., Chen, X. and Bramel, J., The logic of logistics: Theory, algorithms, and applications for logistics and supply chain management. Second Ed. New York: Springer-Verlag, 2005.

UNIVERSIDAD **NACIONAL** DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería Administrativa e Ingeniería Industrial

Oferta de Posgrados

Especialización en Gestión Empresarial
Especialización en Ingeniería Financiera
Maestría en Ingeniería Administrativa
Maestría en Ingeniería Industrial
Doctorado en Ingeniería - Industria y Organizaciones

Mayor información:

E-mail: acia_med@unal.edu.co
Teléfono: (57-4) 425 52 02

M. Valencia-Cárdenas, es Ing. Industrial en 2000, Esp. en Estadística, en 2002, MSc. en Ciencias-Estadística, en 2010. PhD en Ingeniería, Industria y Organizaciones, de la Universidad Nacional de Colombia, Sede Medellín. Ha trabajado en aplicaciones estadísticas, estadística industrial, optimización. Actualmente docente de la Institución Universitaria Tecnológico de Antioquia, Medellín, Colombia. Áreas de interés: métodos estadísticos, optimización con aplicaciones a la industria.
ORCID:orcid.org/0000-0003-3135-3012

F.J. Díaz-Serna, es Ing. Industrial en 1982, Esp. en Gestión para el Desarrollo Empresarial en 2001, MSc. en Ingeniería de Sistemas en 1993, PhD en Ingeniería en 2011. Áreas de trabajo: ingeniería industrial, administrativa y de sistemas. Actualmente, profesor asociado del Departamento de Ciencias de la Computación y la Decisión, Facultad de Minas, Universidad Nacional de Colombia, Sede Medellín, Colombia. Áreas de interés: investigación de operaciones, optimización, sistemas energéticos, producción, dinámica de sistemas.
ORCID:orcid.org/0000-0003-1057-1862

J.C. Correa-Morales, es Estadístico en 1980, MSc. en Estadística en 1989, PhD en Estadística en 1993. Áreas de trabajo: estadística, bio estadística, estadística industrial. Profesor asociado de la Escuela de Estadística, Universidad Nacional de Colombia, Sede Medellín, Colombia. Áreas de interés: análisis multivariado de datos, bioestadística, estadística bayesiana.
ORCID: orcid.org/0000-0002-9368-4725

Thermal dielectric and Raman studies on the KNO$_3$ compound high-temperature region

Fabian Fernando Jurado-Lasso, Natahly Jurado-Lasso, Jaime Alonso Ortiz-Gómez & Jesús Fabian Jurado

Lab. de Propiedades Térmicas Dieléctricas de Compositos, Departamento de Física y Química, Universidad Nacional de Colombia, Sede Manizales, Colombia. fernandoj_27@hotmail.com, natha_sw@hotmail.com, jaortizg@unal.edu.co, jfjurado@unal.edu.co

Abstract

Calorimetric measurements for a heating-cooling cycle determine the transition temperature and enthalpy of the phases present in the KNO$_3$ compound. The effects correlated within the ionic conduction of the KNO$_3$ compound were studied by impedance spectroscopy measurements in a frequency range from 0.1 to 10 MHz for a cooling cycle. The imaginary part of the impedance with a frequency between 200 and 100°C showed a displacement of the Debye-like peak in the lower frequency direction. This displacement indicates an increase in the relaxation times of ionic conductivity by jump. In the dielectric formalism module, the imaginary part showed an asymmetric peak as a correlation consequence in the cationic diffusion. Also the registers demonstrated that the process is thermally activated, with activation energy that is very close to the one obtained for dc conduction. From these results, it can be inferred that both, diffusion and conductivity mechanisms have the same origin. The Raman spectroscopy measurements, based on temperature (when cooling), allowed for correlation on each of the adopted phases and for changes in normal active modes of the isolated groups D$_{2h}$ through the evolution of the active modes v_3 and v_2.

Keywords: ionic conduction; impedance spectroscopy, Raman spectroscopy.

Estudios térmicos dieléctricos y Raman del compuesto KNO$_3$ en la región de alta temperatura

Resumen

Medidas DSC para un ciclo de enfriamiento-calentamiento permiten determinar las temperaturas y entalpías de transición de fases del compuesto KNO$_3$. Efectos correlacionados en la conducción iónica del compuesto KNO$_3$, fueron estudiados por medidas de espectroscopia de impedancia en el rango de frecuencia de 0.1 a 10 MHz, para un ciclo de enfriamiento desde 200 a 100°C. La tendencia de la impedancia parte imaginaria con la frecuencia y la temperatura, mostró un desplazamiento del pico "Debye-like" hacia la región de menor frecuencia, este corrimiento indica un aumento en los tiempos de relajación en la conducción iónica por salto. El módulo dieléctrico parte imaginaria, mostró un pico asimétrico como consecuencia de correlaciones en la difusión catiónica, así también, se evidencia que el proceso es térmicamente activado con energía de muy próxima a la obtenida para la conducción dc. Los resultados apuntan a inferir que los dos mecanismos de difusión y conductividad tienen el mismo origen. Medidas de espectroscopia Raman en función de la temperatura (enfriando), permitió correlacionar en cada una de las fases adoptadas los cambios en los modos activos normales de los grupos aislados D$_{2h}$, a través de la evolución de los modos activos v_3 y v_2.

Palabras claves: conducción iónica; espectroscopia de impedancia; espectroscopia Raman.

1. Introduction

KNO$_3$ is a compound that strongly depends on temperature and pressure. For this reason enormous efforts have been undertaken in order to establish the dynamics and stability of the phases' diversity that the compound has. For example, in normal conditions, KNO$_3$ has a non-polar rhombic structure, known as aragonite phase II (α), with a

Pmcn space group [1-6]. In a heating process, at around 130°C, the compound has a rhombohedral formation with a non-polar space group R-3m, called phase I (β). If the compound is stabilized in phase I and then cooled, the material transitions to Phase III at a temperature close to 124 °C, with R-3m symmetry. The most important particularity of this phase is the polarization of the charge along the direction <001> (-axis). This phase is considered slightly metastable. The technological interest of the compound is focused on phase III (ferroelectric), in which there is a possibility of being used as an active element in electro-devices (capacitors and volatile memories, among others) due to its excellent piezoelectric and ferroelectric properties. Great experimental efforts have been made in order to establish the dynamic of the transition order-disorder, which is thermally activated and mainly related to the rotation of NO_3^- group around the c axis (trigonal axis) [7,8]. Within the phase transitions of the compound, there have also been some theoretical models applied in the interest of understanding and clarify the physics, [9-11]. Calorimetric measurements are reported for heating-cooling cycles, and they precisely determine the transition temperature and its nature [12-14]. In many cases, a dependence on the thermal history has been detected. By using Raman spectroscopy measurements, we demonstrated the correlation between the origin of the crystal structure and the active Raman modes in the compound for all of the existent phases [10,15-18]. In this work, we report the results obtained for differential scanning calorimetry (DSC), spectroscopy impedance (EI), and Raman spectroscopy as a function of temperature for KNO_3, in order to help establish the existent correlations among the structural organization, vibrational order, and dielectric transportation in the high temperature region.

2. Experimental Set-up

The potassium nitrate powder with a high purity (≥99.9 %) used in this study was supplied by Panreac Química SAU.

The structural order of the compound at room temperature was established by powder X-Ray diffraction by using the A8 Advance Bruker AXS equipment (2θ Bragg –Brentano geometry), which worked at 40 kV and 40 mA. The X-ray diffraction patterns were collected in step-scanning mode with 0.02 steps using a $CuK_{\alpha 1}$ (λ=1.54056 Å) wave. Raman scattering spectra were measured using a Raman apparatus (LabRam HR800 Yvon Horiba) with laser (λ=473 nm). The incent and scattering beams were focused with a microscope that had a x20 objective, which allowed for the sample surface to be probed with a ~20μm-spot- sized laser. The spectral range from 150 to 400 cm^{-1} and temperature range from 20 to 200°C. The heating system was performed by using a micro-furnace built in our laboratory, with a precision of ±1.0ºC. It was heated in situ for each register and was kept in the isotherm for 10 min. The thermal behavior was studied with TA Instruments DSC-Q2000 apparatus. The sample was closed in hermetically-sealed aluminum pans in the range from -90 to 350°C at a heating rate of 10°C/min for both heating and cooling cycles under a nitrogen flow of 50 ml/min.

The impedance measurements were carried out by using a LCR meter Solatron 1260 controlled by a computer over a frequency range from 100 to 10 MHz. There was a potential of 500 mV. The possible nonlinear effects due to the amplitude of the voltage were checked 1.0 V in all of the frequency range for a typical configuration of a two-electrode iron cell with a steel/material/steel configuration. The cell was placed into a sealed temperature-controlled (homemade) chamber. The temperature of the sample was registered by using a K-type thermocouple and then located close to the sample with a precision of ±0.5°C. Each sample was equilibrated at the experimental temperature for at least 10 min before the measurements were made. The impedance of the KNO_3 sample was recorded at various isotherms over a temperature range from 23 to 200°C in air.

3. Results and discussion

3.1. X-Ray diffraction

Fig.1 shows the diffractogram at room temperature for KNO_3. If we consider the initial crystals given by data-base code ICSD -260601 [19] and the Rietveld analysis with Topas software V.4.2 as parameters, the crystallographic parameters could be established as a=5.213(59) Å and c=8.162(10) Å; this is in agreement with some published data [9,19].

3.2. DSC Calorimetry

Fig. 2 shows the DSC thermogram of KNO_3 for a cooling-heating cycle from -90 to 350°C.

The DSC measurements for an endothermic reaction also show that in the high temperature region, the compound goes from liquid state to phase I(β) at around 332.1°C and enthalpy of 97.3 J/g. This is followed by the transition from phase III(γ) at around 121.7°C and enthalpy of 27.1 J/g. When heating the compound that is departing from the minor temperature area, it presents a transition from the exothermic phase at around 62.4°C with a formation enthalpy of 11.8 J/g in the recovering phase II(α). When the temperature is increased, the system shows a transition from phase I(β) at
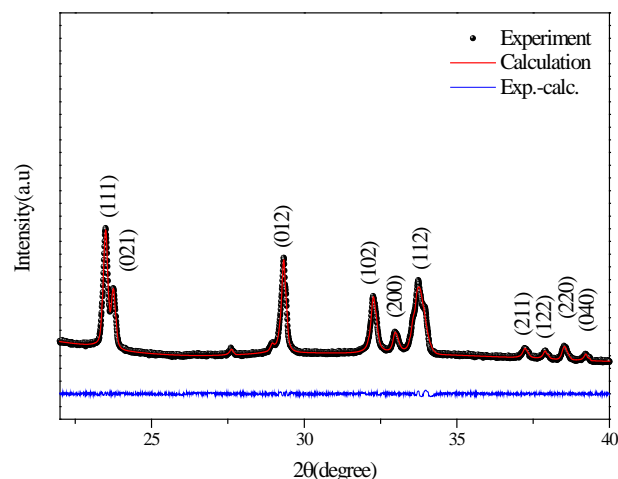

Figure 1. X-Ray powder pattern of KNO_3 at room temperature. The spectral lines were indexed considering the R-3m group.
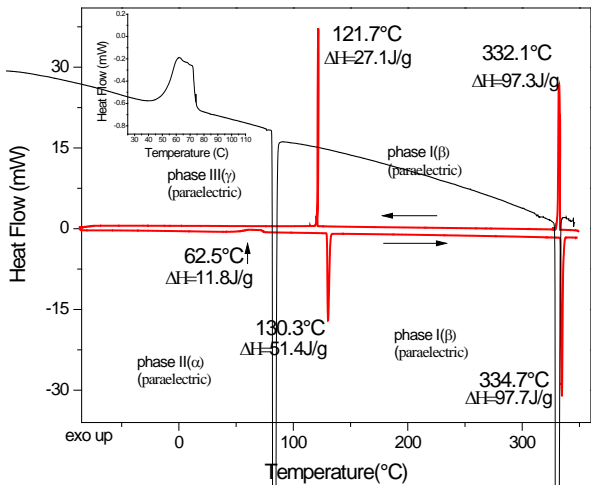Source: The authors.

Figure 2. DSC thermogram of KNO₃, for a cooling-heating cycle from -90 to 350°C. The inset shows enlargement of the exothermic transition near 62.5°C. The arrows indicate the direction of the heating-cooling cycle.
Source: The authors.

Table 1.
The transformation enthalpies during heating and cooling modes.

Modes (Cooling/Heating)	Transition temperature °C	Phase	Enthalpy J/g
Cooling	332.1	I(β)	97.3
Cooling	121.7 124[28]	III(γ)	27.1
Heating	62.5 110[28]	II(α)	11.8
Heating	130.3 130 [28]	I(α)	51.4
Heating	334.7	I(β)	97.7

Source: The authors.

about 130.3°C and an enthalpy of 51.5 J/g. When the temperature is increased, the system melts at about 334.7°C with a transformation enthalpy of 96.7 J/g. The results obtained in this study agree with those reported in the literature [4,12,28]. The transformation enthalpies during heating and cooling modes are show in Table 1.

Comparing the enthalpy transformation values during phase III(β), phase II(α), and phase I(β) respectively, we can concluded that the three transitions come from a different origin. Phase III's transition is an order-disorder transition, which occurs in a very narrow temperature region.

3.3. Impedance spectroscopy

Typical Nyquist (–Z´´versus Z´) plots of impedance data at different isotherms are shown in Fig. 3 between 190 to 130°C for a KNO₃ sample. The dielectric response is represented by an *RC* circuit that is connected in parallel, which is reduced to a pure resistance for high and low frequencies. For all measurements in every temperature range, the circle we described goes through the origin in the high frequency region. For the low frequency region, the measurement does not show any effects related to the electrode/material interface. The dc conductivity was calculated from the Nyquist plots, as is usual in these cases.
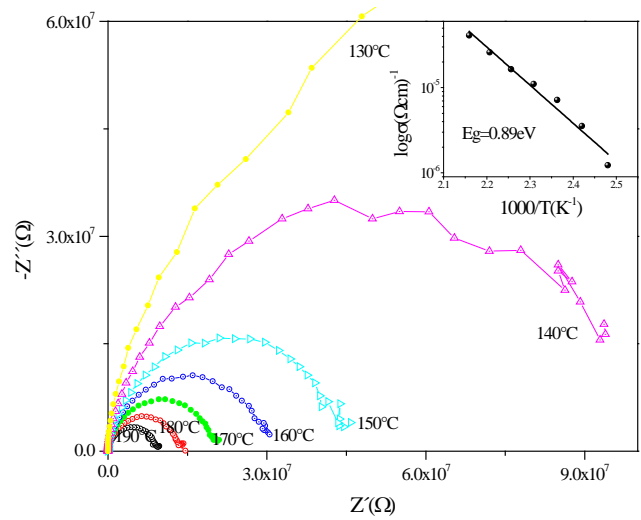

Figure 3. Cole-Cole (Nyquist) plot of KNO₃ impedance measured at different temperatures. The inset shows plots of dc conductivity of Vs 1000/T(K). The line represents the better lineal adjustment.
Source: The authors.

Additionally, the extrapolation arc to the Re*Z* axis (which is in the interception) was the bulk resistance *R* of the KNO₃. The most appropriate ones were selected and the formula $\sigma=d/RA$ was used, where *d* is the thickness and *A* is the contact area of the material and the electrodes.

Fig. 4 shows the imaginary part of impedance variation with the frequency for different temperatures. The movement of the Debye-like peak, which goes in the direction the lower frequency when temperature decreases, supposes an increase in the relaxation times for ionic conductivity by jump. Moreover, the linear dependence of the impedance on the frequency with slope -1, when the temperature is 130°C, shows that conductivity is independent of the frequency for temperatures below this temperature value. The dc conductivity measurements, according to the temperature of the KNO₃ compound, are plotted as: lnσ vs. 1000/T, and they are displayed in the inset of Fig.3. We found an Arrhenius model by using the least square analysis for corresponding data. From the Arrhenius type, we determined the activation energy Eg, calculated using the formula: $\sigma(T)=\sigma_0\exp(-E_g/kT)$, where *k* is the Boltzmann constant, T is the absolute temperature and σ_o the exponential pre-factor. The activation energy calculated in this case was of $E_g=0.89$ eV, which is close to what the references [20,21] report.

An alternative method in the time domain, to distinguish the correlation effects in the conductivity relaxation, can be obtained from the complex modulus formalism [22,23] $M^*(\omega)=1/\varepsilon^*=j\omega C_0 Z^*(\omega)$, where $j=\sqrt{-1}$, $\omega=2\pi f$ is the angular frequency and C_o is the vacuum capacitance of the cell. The most relevant part of this formalism is that the electrode effects can be suppressed at a low frequency.

Fig. 5 shows a M´´/M´´$_{max}$ normalized data versus log(*f*) relative to the temperature given. It shows a displacement in the maximum (M´´$_{max}$) towards the minor frequency region when the temperature decreases. In this formalism, the peak marked the transition regime for the ion mobility that is present in the compound: from a long range (dc) to a short
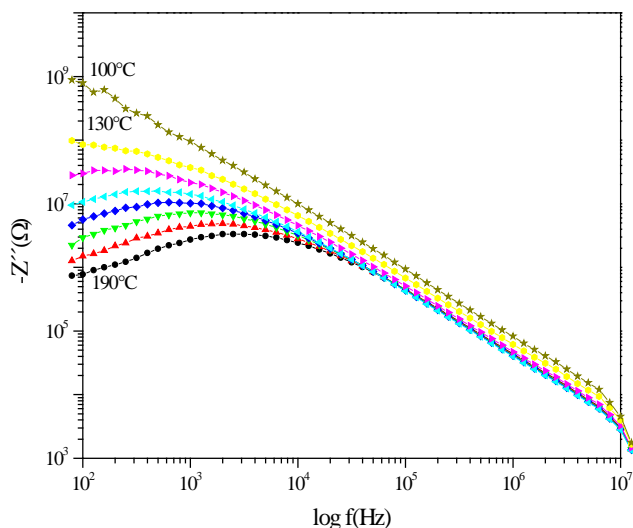
Figure 4. Variation in the imaginary part of impedance with frequency for KNO₃. When the temperature is 100°C, the dependence of impedance with frequency has slope -1.
Source: The authors.

range. Furthermore, the asymmetry of the width of the peak, moving away from the ideal value (one), is an indicator that the ionic conduction is a distribution of relaxation times. This moves away even further from the Debye ideal model while the temperature decreases. From the maximum M''_{max}, the relaxation time was determined by using the expression $\omega_{max}\tau = 2\pi f_{max}RC = 1$, where f_{max} is the frequency relative to M''_{max}, which is the peak corresponding to bulk relaxation.

An Arrhenius-type law is shown to be within the temperature range as it is also exposed that the process is thermally activated (inset Fig.5). The activation energy value related to the relaxation process is close to the energy activation value obtained from lnσ vs. 1000/T. The shape and the width of the peak do not change considerably with temperature decreases. This allows it to be affirmed that, in ionic diffusion, the relaxation process has a correlated effect [22]. Additionally, in the conduction by ion jump, we can find that there is an intrinsic mechanism and that relaxation times are affected by the availability of places. This mechanism is common in many ion conductors [24]

From these results, it can be inferred that conduction is not affected by blocking grain boundaries. It can also be affirmed that the conductivity relaxation peak is not a consequence of dipolar relaxation in bulk materials, as normally happens (bulk materials) [24-26]. However, it is very common that in the M''/M''_{max} vs. log(f) formalism there could be one or more peaks. These represent more relaxation processes of which allocation and interpretation are still very controversial [22,23].

3.4. Raman spectroscopy

The structural change in KNO₃ is mostly related to the change in NO⁻₃ groups f as well as the natural vibration modes. For example, in normal conditions, KNO₃ is in phase II, and NO⁻₃ contains a symmetry of D_{2h} [6,27]. According to the selection rules for Raman and infrared absorption, which

are characterized eleven normal modes that group into: $v_1(A_1)$, $v_3(E')$, $v_4(E'')$ and $2v$. Experiments have shown that they locate around: 1054, 1383, 715 and 1664 cm⁻¹, respectively [6, 27].

Fig. 6 shows the presence of the three phases adopted in KNO₃, due to the change in the compound temperature, for a cooling process from 173 to 22°C. Each phase is characterized by the presence and location of the peaks that are related to the isolated D_{2h} group's natural mode of vibration. During the temperature lowering process the change dynamic of v_3 and $2v$ modes is displayed, respectively. When the temperature reaches approximately 173°C the material is in phase I. According to the selection rules, it has been determined that, there are 9 existent Raman modes in this phase [10,15–18], from which it is possible to determine the presence of the ones located in: 719, 1052, 1356, 1420 and 1667 cm⁻¹. In this phase, the most relevant factor of these measurements is the overlapping v_3 mode, which is located approximately in 1356 and 1420°C, respectively (Fig. 7).

When the temperature of the compound is around 43°C, the material adopts phase III. The most important factor in these conditions is the splitting of the v_3 mode into four main peaks that are located around: 1351, 1381, 1406 and 1441 cm⁻¹, respectively. Progressive splitting of the vibrational mode occurs while the temperature decreases. This dynamic is considered to be a strong indicator of the nature of the phase change, which gradually assimilates when its temperature lowers. When the temperature is at 24°C, the material is in phase II (Fig.7).

The most important characteristic of this phase change dynamic is centered on v_3 mode changing, while temperature varies, to finally define at this temperature, a double located around 1347 and 1362 cm⁻¹, respectively. It is also important to note that there is a splitting of the $2v_2$ mode in the peaks located around 1654 and 1687 cm⁻¹, respectively. However, there is a persistence of the peaks close to 1387 and 1441 cm⁻¹, respectively. This indicates the presence of a residual
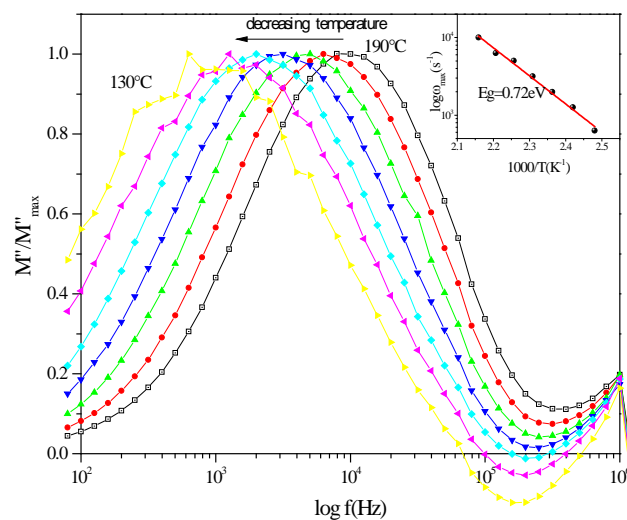

Figure 5. Plots of modulus M''/M''_{max} versus log(f) for KNO₃ at various temperatures. The inset shows an Arrhenius plot of logω_{max}, which is plotted against 1000/T. The line represents the better lineal adjust.
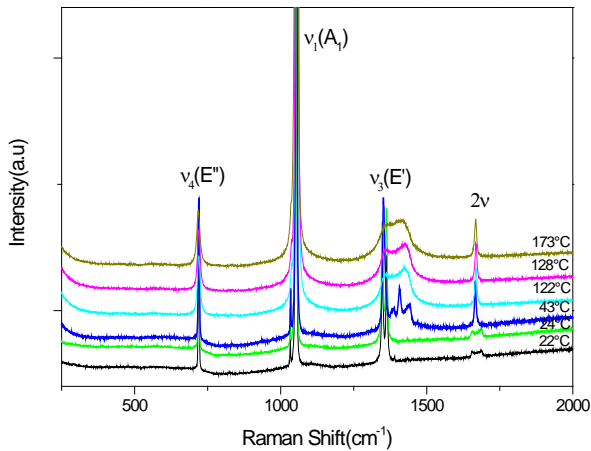Source: The authors.

Figure 6. The evolution of the spectra profiles for the $v_4(E'')$, $v_1(A_1)$, $v_3(E')$, and $2v$ of NO^-_3 vibration between 173°C to 22°C during phase I, II, and III transitions.
Source: The authors.

phase III and concurs with the affirmation of the coexistence of phases. From the three phases presented in Fig. 7 it can be inferred that phase III has a better definition of the main basic vibration modes. This could coincide with the concept that states this kind of structural order is now a charge order (as in the case of the ferroelectric phase).

4. Conclusions

The calorimetric measurements allowed us to determine the phase transition temperatures and compare the respective enthalpy values of the transitions. This comparison allows us to distinguish the transition order-disorder, which is characterized by its very narrow transformation enthalpy value. KNO_3 in the high temperature region showed that dc ionic conductivity is thermally activated by the Arrhenius-type. The impedance (imaginary part) revealed that the dc ionic conduction is caused by the charge carriers jumping. With
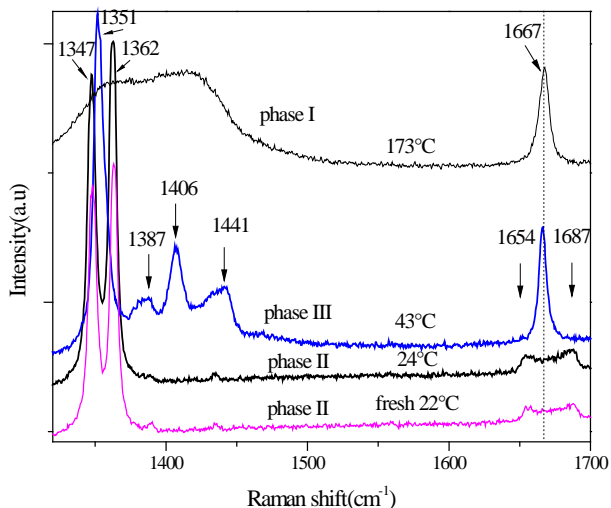

Figure 7. Comparison of Raman data between KNO_3 phases I, II and III for a cooling cycle. The arrow marked their symmetry labels under are phase mode.
Source: The authors.

the temperature increase, the distribution of the relaxation times displaces towards the lower frequency region. Dielectric formalism modulus showed how the relaxation processes were modified, causing a change of structural order and a short-range temperature effect. Additionally, this was an Arrhenius type process and thermally activated the process. The respective activation energies' values lead us to the conclusion that they have the same dc conductivity origin, as well as the ion jump probability, among available sites.

The Raman spectroscopy allowed us to establish the change of phases dynamic as well as determine that phase III presents the best definition of the basic vibrational modes. These results concur with structural organization and charge organization (as in the ferroelectric phase case).

References

[1] Megahid, N.M., Field dependence of electrical conductivity of powder compacts of (KNO3)1-xCux mixtures, Egyto J.Sol., [Online]. 25, pp. 125-136, 2002. Available at: http://egmrs.powweb.com/EJS/PDF/vo251/125.pdf.

[2] Bartholomew, R.F., A study of the equilibrium KNO3--KNO2(l)+1/202(g) over the temperature range 550-750°C, J.Phys. Chem., [Online]. 72(1), pp. 3442-3446, 1966. Available at: http://pubs.acs.org/toc/jpchax/70/11

[3] Dabra, N., Hundalw, J.S., Sekhar, K.C., Nautiyal, A. and Nath, R., Ferroelectric phase stability studies in spray deposited KNO3: PVA composite films, J. Am. Ceram. Soc., 92(4), pp.834-838, 2009. DOI: 10.1111/jace.2009.92.issue-4/issuetoc

[4] Aquino-Olivos, M.A., Jean-Pierre, E. Grolier, S.L. Randzio, A., Aguirre-Gutiérrez, J. and Garcia-Sanchez, F., Transitiometric determination of the phase diagram of KNO3 between (350 and 650) K and at pressures up to 100 MPa, J. Chem. Eng. [Online]. 55, pp.5497-5503, 2010. Available at: http://pubs.acs.org/toc/jceaax/55/12

[5] Baryshnikov, S.V., Charnaya, E.V., Yu, A., Milinskii, E., Stukova, V., Cheng, T. and Michel, D., Dielectric properties of crystalline binary KNO3–AgNO3 mixtures embedded in nanoporous silicate matrices, Physics of the Solid State, 52(2), pp. 392-396, 2010.

[6] Xu, K., Application of Raman in phase equilibrium studies: The structures of substitutional solid solutions of KNO3 by RbNO3, J. Mat. Science, 34, pp. 3447-453, 1999.

[7] Kumar, N. and Natha, R., Ferroelectric properties of potassium nitrate-polymer composite films, J. Pure Appl. & Ind. Phys. [Online]. 1(1), pp. 21-35, 2010. Available at: http://physics-journal.org/archive-1-1.html

[8] Erdinc, B. and Akkus, H., Ab-initio study of the electronic structure and optical properties of KNO3 in the ferroelectric phase, Phys. Scr. [Online]. 79, pp. 025601-025006, 2009. Available at: http://iopscience.iop.org/issue/1402-4896/79/2

[9] Lu, H.M. and Hardy, J.R., Principles study of phase transitions in KNO3, Phys. Rev. B, [Online]. 44, pp. 7215-7224, 1991. Available at: http://journals.aps.org/prb/issues/44/14

[10] Liu, D., Ullman Behlen, F.G. and Hardy, J.R., Raman scattering and lattice-dynamical calculations of crystalline KNO3, Phys. Rev. B., [Online]. 45(5), pp. 2142-2147, 1992. Available at: http://journals.aps.org/prb/issues/45/5

[11] Porto, M., Maass, P., Meyer, M., Bunde, A. and Dieterich, W., Hopping transport in the presence of site-energy disorder: Temperature and concentration scaling of conductivity spectra, Phy. Rev.B., [Online]. 61(9), pp. 6057-062, 2000. Available at: http://journals.aps.org/prb/issues/61/9

[12] Abdulagatova, I.M., Dvoryanchikov, V.I. and Kamalov, A.N., Measurements of the heat capacities at constant volume of H2O and (H2O + KNO3), J. Chem. Thermodynamics, [Online]. 29, pp. 1387-1407, 1997. Available at: http://www.sciencedirect.com/science/journal/00219614/29/12

[13] El-Kabbany, F., Abdel-Kader, M.M., Tosson, M. and El-Khwass, E., Kinetics of the ferroelectricalbehaviour of a thin layer of phase III KNO3, Thermochimica Acta, [Online]. 256, pp. 281-289, 1995. Available at: http://www.sciencedirect.com/science/journal/00406031/256/2

[14] Font, J. and Muntasell, J., Thermobarometric study of KNO3 phase transitions, Thermochimica Acta, [Online]. 293, pp. 167-170, 1997. Available at: http://www.sciencedirect.com/science/journal/00406031/293/1-2

[15] Balbaski, M., Teng, M.K. and Nusimovici, M., Raman scattering in KNO3 phases I, II, and III, Phy. Rev. [Online]. 176(3), pp. 1098-1106, 1968. Available at: https://journals.aps.org/pr/issues/176/3

[16] Scott, J.F. and Zhang, M.-S., Raman spectroscopy of submicron KNO3 films, Phys. Rev.B, [Online]. 35, pp. 4044-4051, 1987. Available at: http://journals.aps.org/prb/issues/35/8

[17] Murugan, R., Huang, P.J., Ghule, A. and Chang, H., Studies on thermal hysteresis of KNO3 by thermo-Raman spectroscopy, Thermochimica Acta, [Online]. 346, pp. 83-90, 2000. Available at: http://www.sciencedirect.com/science/journal/00406031/346/1-2

[18] Loudon, R., The Raman effect in crystal, Advan. Phys., 13, pp. 423-482, 1964. Available at: http://www.tandfonline.com/toc/tadp20/13/52?nav=tocList

[19] Freney, E.J., Garvie, L.A.J., Groy, T.L. and Buseck, P.R., Growth and single-crystal refinement of phase –III potassium nitrate, KNO3, Acta Cryst., [Online]. B65, pp. 659-663, 2009. Available at: http://scripts.iucr.org/cgi-bin/paper?S0108768109041019

[20] Singh, A. and Smith, A.M., Dielectric and electrical conductivity studies in potassium nitrite. J. phys. D: Appl. Phys. [Online]. 4, pp. 560-566, 1971. Available at: http://iopscience.iop.org/issue/0022-3727/4/4

[21] Mansingh, A. and Smith, A.M., Dielectric dispersion in the paraelectric phase of potassium nitrate, J. Phys. D: Appl. Phys., [Online]. 4, pp. 1792-1796, 1971. Available at: http://iopscience.iop.org/issue/0022-3727/4/11

[22] Elliott, S.R., Use of the modulus formalism in the analysis of ac conductivity data for ionic glasses, Journal of Non-Crystalline Solids, [Online]. 170, pp. 97-100, 1994. Available at: http://www.sciencedirect.com/science/journal/00223093/170/1

[23] Svare, I., Borsa, F., Torgeson, D.R., Martin, S.W. and Patel, H. Use of the modulus formalism in the analysis of ac conductivity data for fast ion conductors, Journal of Non-Crystalline Solids, [Online]. 185, pp. 297-300, 1995. Available at: http://www.sciencedirect.com/science/journal/00223093/185/3

[24] Jurado, J.F., Trujillo, J.A., Mellander, B-E. and Vargas, R.A., Correlated ion diffusion in gamma-Ag0.7Cu0.3I, S. State. Ionic, [Online]. 176, pp. 985-990, 2005. Available at: http://www.sciencedirect.com/science/journal/01672738/176/9-10

[25] Mellander, B.-E. and Albinsson, I., in Solid State Ionics: New Developments edited. Chowdari, B.V.R et al., World Scientific, Singapore, [Online]. 1996, pp. 83-87. Jurado, J.F., Játiva, J.A., Metal-insulator transition and hopping conduction mechanisms in the La0.7Ba0.3MnO3 compound, J. Mag. Magn. Mate., 335, pp. 6-10, 2013. Available at: http://www.sciencedirect.com/science/journal/03048853/335

[26] Nakamoto, K., Infrared and Raman spectra of Inorganic and Coordination Compounds, 4th ed., Wiley, New York, 1986, pp. 6-20.

[27] Kumar, E. and Natha, R., Ferroelectric properties of potassium nitrate-polymer composite films, J. Pure Appl. & Ind. Phys., [Online]. 1, pp. 21-35, 2010. Available at: http://physics-journal.org/archive-1-1.html

F.F. Jurado-Lasso, received a First Class MSc. degree in Telecommunications Engineering after attending The University of Melbourne, Australia in 2015. Also, he obtained a bachelor degree in Electronics Engineering after studying in The Universidad del Valle, Colombia in 2012. He has also worked with projects related to new materials with an emphasis in Telecommunications.
ORCID: orcid.org/0000-0002-5005-781X.

N. Jurado-Lasso, Bacteriologist for The Universidad del Valle, Colombia. She has also worked with projects related to new materials.
ORCID: orcid.org/0000-0001-6091-4160.

J.A. Oriz-Gomez, is Engineer -Physicist of the Universidad Nacional de Colombia. He has also worked with projects related to new materials.
ORCID: orcid.or/000-0001-8972-1805

F.J. Jurado, is Dr. Ciencias Física (titular professor) worked on Condensed Matter, total experience –20 years with about fifteen original article research publications in international journal as a co-author. He is the leader of Laboratorio de Propiedades Térmicas Dieléctricas de Compositos, Universidad Nacional de Colombia Sede Manizales.
ORCID:orcid.org/0000-0001-5193-8566.

Entregando
lo mejor de los
colombianos

4
72

Línea de atención al Cliente Nacional: **01 8000 111 210**
Línea de atención al Cliente Bogotá: **(57-1) 472 2000**
> **www.4-72.com.co**

4°N   72°O

- Occurrence of chloritoid-bearing metapelitic rocks and their significance in the metamorphism of the Silgará Formation at the Central Santander Massif
- Geology, litostratigraphy, determination of the quality and classification of carbons of Espino-Boyacá, Colombia
- Stratigraphic analysis and calculation of reserves in the Colorado field, MMV, Colombia
- Drilling grid blasting upgrading based on Geological Strength Index (GSI), case "La Maruja" mine, Colombia
- Failure potential index for slopes conditioned by landforms
- Modeling and behavior of the simulation of electric propagation during deep brain stimulation
- A hybrid partitioning method for multimedia databases
- Characterization of supply chain problems
- Algorithm for wideband spectrum sensing based on sparse Fourier transform
- Transient analysis of mixed wind parks with different turbine types
- New robust capability ratios approaches for quality control
- Effect of elastoplastic behavior on the impact response of expanded metal tubes
- Phenomenological modeling and parametric identification applied to the monitoring of a belt conveyor system in a copper mine
- State observer design for biomass and ethanol estimation in bioreactors using cybernetic models
- Use of white rot fungi in the degradation of an azo dye from the textile industry
- The mechanical properties of Portland cement mortars blended with carbon nanotubes and nanosilica: A study by experimental design
- Influence of solid and liquid antioxidants on the formation of space charge in the XLPE insulation of medium voltage cables
- Method for determining total losses in distribution transformers from the percentage of carbon silicon steel sheet
- Methodology for hydrodynamic model selection. Case study: spatial variability of the thermal structure in the Riogrande II tropical reservoir, Colombia
- Facile one-pot synthesis of uniform silver nanoparticles and growth mechanism
- Adsorption and catalytic oxidation of asphaltenes in fumed silica nanoparticles: Effect of the surface acidity
- Development of a simulation model as a decision support system for sugarcane supply
- Algorithm for detection of overlapped red blood cells in microscopic images of blood smears
- An online algorithm for the container stacking problem
- Marine safety: Asbestos in the maritime industry
- Contributions to the online fault diagnosis of interturn short circuit in three-phase induction motor by of means negative sequence components
- A new noncircular gear pair to reduce shaft accelerations: A comparison with sinusoidal and elliptical gears
- Roasting of sulphide using carbothermal reduction
- Multi-product inventory modeling with demand forecasting and Bayesian optimization
- Thermal dielectric and Raman studies on the KNO3 compound high-temperature region

- Ocurrencia de rocas metapelíticas con presencia de cloritoide y su significado en el metamorfismo de la Formación Silgará en la región central del Macizo de Santander
- Geología, litoestratigrafia, determinación de la calidad y clasificación de carbones de El Espino-Boyacá, Colombia
- Análisis estratigráfico y cálculo de reservas del Campo Escuela Colorado, VMM, Colombia
- Modificación de las mallas de perforación de voladuras a partir del índice de esfuerzo geológico (GSI), caso mina "La Maruja", Colombia
- Indicador potencial de falla para taludes condicionados por las geoformas
- Modelado y comportamiento de la simulación de propagación eléctrica durante la estimulación cerebral profunda
- Un método de fragmentación híbrida para bases de datos multimedia
- Caracterización de las problemáticas de la cadena de abastecimiento
- Algoritmo para sensado de espectro de banda ancha basado en transformada dispersa de Fourier
- Análisis de transitorios de parques eólicos mixtos con diferentes tipos de turbinas
- Nueva propuesta de índices de capacidad robustos para el control de la calidad
- Efecto del comportamiento elastoplástico sobre la respuesta al impacto de tubos de metal expandido
- Modelado fenomenológico e identificación paramétrica aplicada al monitoreo de una correa transportadora en una mina de cobre
- Diseño de un observador de estado para estimación de biomasa y etanol en biorreactores a partir de modelos cibernéticos
- Utilización de hongos de la podredumbre blanca en la degradación de un colorante tipo azo de la industria textil
- Propiedades mecánicas de morteros de cemento Portland adicionados con nanotubos de carbono y nanosilica: Estudio por diseño de experimentos
- Influencia del antioxidante sólido y líquido sobre la formación de carga de espacio en el aislamiento de XLPE de cables de media tensión
- Método para determinar las pérdidas totales en transformadores de distribución a partir del porcentaje de carbono de la chapa de acero al silicio
- Metodología para la selección de modelos hidrodinámicos – Caso de aplicación: variabilidad espacial de la estructura térmica en el embalse tropical Riogrande II, Colombia
- Fácil síntesis en un paso y mecanismo de formación de nanopartículas de plata
- Adsorción y oxidación catalítica de asfaltenos en nanopartículas de sílice fumárica: Efecto de la acidez superficial
- Desarrollo de un modelo de simulación como un sistema de soporte de decisiones para el abastecimiento de caña de azúcar
- Algoritmo para la detección de glóbulos rojos superpuestos en imágenes microscópicas de extendidos de sangre periférica
- Un algoritmo en línea para el problema de apilamiento de contenedores
- Seguridad marítima: Asbestos en la industria marítima
- Aportes al diagnóstico en línea de fallas de cortocircuito entre espiras de un motor de inducción mediante parámetros eléctricos de secuencia inversa
- Nuevo par de engranajes no circulares para reducir las aceleraciones del eje: Comparación con engranajes sinusoidales y elípticos
- Tostación de sulfuros por reducción carbotérmica
- Modelo de inventario multi-producto, con pronósticos de demanda y optimización Bayesiana
- Estudios térmicos dieléctricos y Raman del compuesto KNO3 en la región de alta temperatura