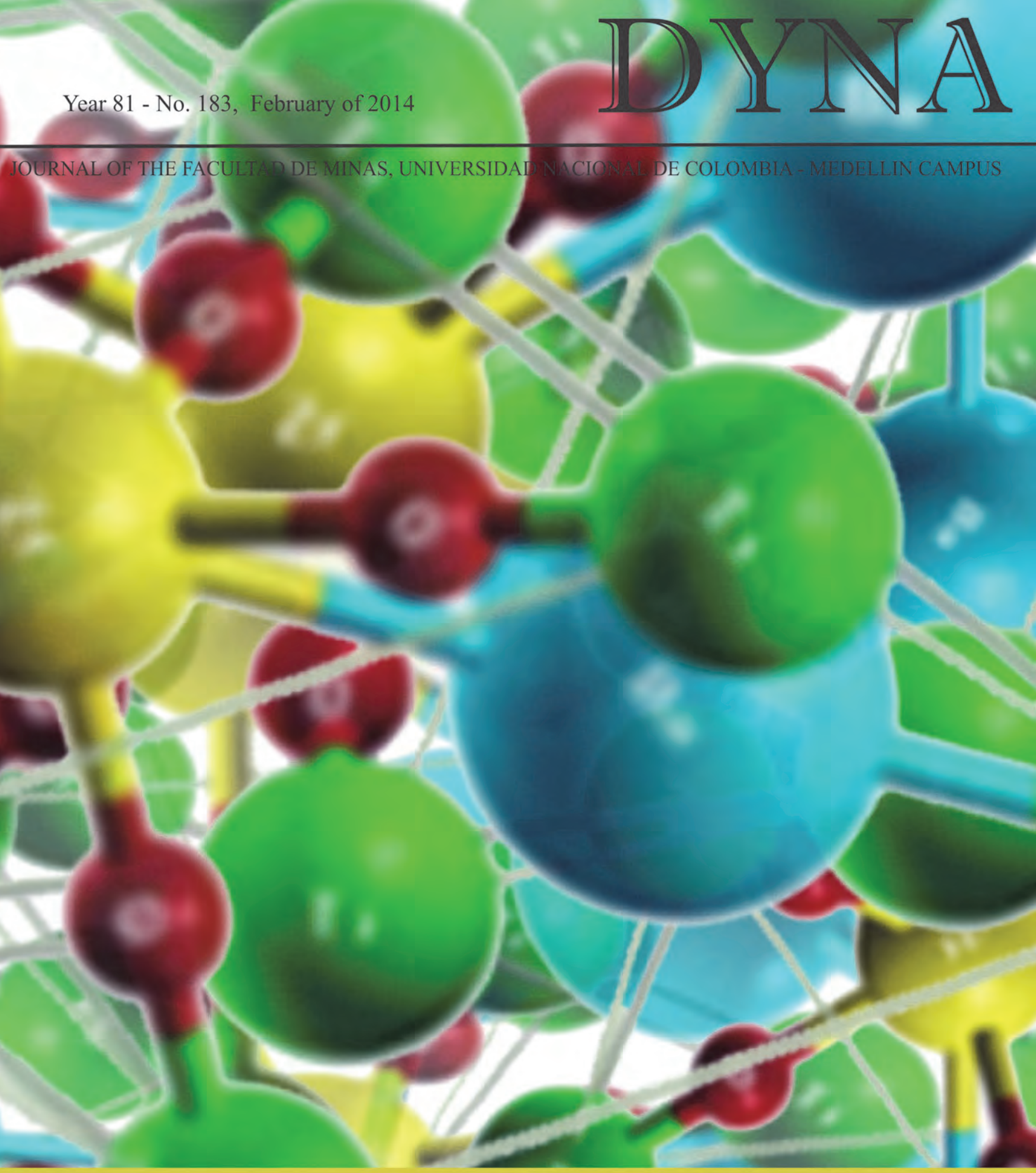
Year 81 - No. 183, February of 2014

DYNA

JOURNAL OF THE FACULTAD DE MINAS, UNIVERSIDAD NACIONAL DE COLOMBIA - MEDELLIN CAMPUS



Year 81 - No. 183, February of 2014 - ISSN 0012-7353
Tarifa Postal Reducida No. 2014-287 4-72
La Red Postal de Colombia, Vence 31 de Dic. 2014.

UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

DYNA
Journal of the Facultad de Minas
Universidad Nacional de Colombia - Medellin Campus

Year 81 - Number 183, February of 2014
ISSN 0012-7353

DYNA is an international journal published by the Facultad de Minas, Universidad Nacional de Colombia, Medellín Campus since 1933. DYNA publishes peer-reviewed scientific articles covering all aspects of engineering. Our objective is the dissemination of original, useful and relevant research presenting new knowledge about theoretical or practical aspects of methodologies and methods used in engineering or leading to improvements in professional practices. All conclusions presented in the articles must be based on the current state-of-the-art and supported by a rigorous analysis and a balanced appraisal. The journal publishes scientific and technological research articles, review articles and case studies. All of the articles published are submitted to a peer or referee review of recognized suitability on the topic.

DYNA publishes articles in the following areas:

Organizational Engineering
Civil Engineering
Materials and Mines Engineering
Geosciences and the Environment
Systems and Informatics
Chemistry and Petroleum
Mechatronics
Bio-engineering
Other areas related to Engineering

DYNA is admitted to:

- The National System of Indexation and Homologation of Specialized Journals CT+I-PUBLINDEX, Category A1
- Science Citation Index Expanded
- Journal Citation Reports - JCR
- Science Direct
- SCOPUS
- Chemical Abstract - CAS
- Scientific Electronic Library on Line - SciELO
- GEOREF
- PERIÓDICA Data Base
- Latindex
- Actualidad Iberoamericana
- RedALyC - Scientific Information System
- Directory of Open Acces Journals - DOAJ
- PASCAL
- CAPES
- UN Digital Library - SINAB
- CAPES

DYNA ISSN 0012-73533
Bimonthly periodicity: February, April, June, August, October, and December
Reduced Postal Fee # 2014-287 4-72. *La Red Postal de Colombia*, expires Dec. 31st, 2014
Circulation License Resolution 000584 de 1976 from the Ministry of the Government
E-mail: dyna@unal.edu.co
Web page: http://dyna.unalmed.edu.co

Facultad de Minas
Universidad Nacional de Colombia - Medellín Campus
Carrera 80 No. 65-223 Bloque M9 - Of.:103
Telephone: (574) 4255068 Fax: (574) 4255343
Medellín - Colombia

INSTITUTIONAL EXCHANGE REQUEST
DYNA may be requested as an institutional exchange through:

The Central Library (Biblioteca Central) "Efe Gómez"
UNIVERSIDAD NACIONAL DE COLOMBIA
Calle 59A No 63-20
Telephone: (57+4) 430 97 86
Medellín - Colombia
E-mail for institutional exchange: canjebib_med@unal.edu.co

Jorunal's Office of Facultad de Minas
Universidad Nacional de Colombia - Medellín Campus
Juan David Velásquez Henao, Director
Catalina Cardona A., Assistant editorial
Amilkar Álvarez C., Diagrammer
Byron Llano V., Assistant editorial
Alberto Gutiérrez, IT Assistant
Mónica del Pilar Rada T., Editorial Choordinator

CONTENT

Our cover
Image alluding to Article:
Electronic, structural and ferroelectric properties of the Ba_2ZrTiO_6 double Perovskite

Authors:
David A. Landínez-Téllez, Críspulo Enrique Deluque-Toro & Jairo Roa-Rojas

CONTENIDO

Nuestra Caratula:
Imagen alusiva al artículo:
Propiedades electrónicas, estructurales y ferroeléctricas de la doble Perovskita Ba_2ZrTiO_6

Autores:
David A. Landínez-Téllez, Críspulo Enrique Deluque-Toro & Jairo Roa-Rojas

DEAN'S NOTE

CHANGES IN THE EDITORIAL BOARD OF THE JOURNAL

Dear authors, reviewers and readers:

Professor Juan D. Velásquez initiated his work as Editor of DYNA in November 2013, succeeding Professor Oscar J. Restrepo, editor for the last ten years. Professor Velásquez has been an active author and reviewer in several journals for the last decade and he was the first editor of the journal Avances en Sistemas e Informática. He received a Bs. Eng in Civil Engineering, an MSc degree in Systems Engineering, and a PhD degree in Energy Systems from the Universidad Nacional de Colombia. Medellin. He worked for electricity utilities and consulting companies within the power sector for seven years. Currently, he is a professor in the Computing and Decision Sciences Department in the Facultad de Minas, Universidad Nacional de Colombia. 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.

On behalf of the faculty council members, I would like to thank, Professor Restrepo for ten years of magnificent work. Many successes have been achieved in the last decade; they include, among others, the creation of the online version and the presence of the journal in several scientific and social networks, the inclusion of DYNA in several bibliographical indexes and databases, but mainly, the inclusion of DYNA in the Journal Citation Reports. Due to the diligent work of professor Restrepo, the authors, reviewers, and readers, DYNA has an impact factor 0.175 for the year 2012.

John W. Branch
Dean
Facultad de Minas
Universidad Nacional de Colombia.

EDITORIAL

REVIEW OF THE AIMS AND SCOPE OF THE JOURNAL

In my first editorial as new editor, I have decided to discuss the revised Aims and Scope of DYNA which is posted in our website and in the paper version of the journal as follows:

> "DYNA is an international journal published by the Facultad de Minas, Universidad Nacional de Colombia, Medellín Campus since 1933. DYNA publishes peer-reviewed scientific articles covering all aspects of engineering. Our objective is the dissemination of original, useful and relevant research presenting new knowledge about theoretical or practical aspects of methodologies and methods used in engineering or leading to improvements in professional practices. All conclusions presented in the articles must be based on the current state-of-the-art and supported by a rigorous analysis and a balanced appraisal. The journal publishes scientific and technological research articles, review articles and case studies."

The new Aims and Scope has a clear and explicit focus on original, useful and relevant research in engineering. By the word "original", I mean that the journal only accepts original research that not has been previously published, except as a thesis in a university or as a short abstract in conference proceedings. Here, the editorial team follows a common rule in many scientific journals. As a consequence, manuscripts based on previously published work with unsubstantial new original contributions will be rejected directly by the Editor; examples of unsubstantial contributions include rewritten manuscripts with the same conclusions of previous published works but varying the data or adding new cases. By the word "original", I also mean that theoretical or practical results and conclusions cannot be derived from current knowledge in a straightforward, direct or trivial way.

The words "useful" and "relevant" are used to mean that conclusions describe important and interesting contributions that will be potentially usable for the readers in their future professional or research activities. As a consequence, the originality is a necessary but not sufficient condition for accepting a manuscript.

By "rigorous analysis" must be understood that authors followed well-accepted and well-applied methodologies for obtained and supporting their conclusions. For example, some manuscripts in engineering education state in their conclusions that the participants are "happier and more motivated" but without a rigorous experimental design demonstrating the advantage of the proposed practice o method in the learning process; as a result, this type of manuscripts are not suitable for publication.

As many other research journals, DYNA is focused on publishing original scientific or technological research articles in engineering. In addition, review articles are considered for publishing, but in this case, it is required the use of at least fifty references, and a in-depth discussion focused in the findings obtained when the literature is analyzed as a whole. Review articles should clearly and explicitly justify the necessity of review in their introduction and explain careful the objectives of the work. Further, it is mandatory that the discussion and conclusions presented cannot be derived from the isolated reading of the manuscript's references.

DYNA accepts case studies for publishing; however, the case considered must be non trivial, possibly unique, and interesting for the community; the manuscript must present new knowledge to improve the understanding of the case or novel methods for tackling it. Authors should clearly and explicitly justify how the work meets previous conditions for publishing. In this sense, technical reports or case studies describing the use of well-known methods or techniques applied in typical cases are not suitable for publishing.

With the new Aims and Scope, authors are encouraged to explicitly discuss how their work impacts the current state of the art of engineering or how to improve current practices. The editor and the reviewers will look for this discussion in the manuscript.

Finally, the objective and contributions of the work must be clearly stated in the abstract. This point is of special importance because of it allows to the editor and reviewers to accelerate the publishing decision.

Juan D. Velásquez, MSc, PhD
Professor
Departamento de Ciencias de la Computación
y de la Decisión
Facultad de Minas
Universidad Nacional de Colombia

DEAN'S NOTE

CHANGES IN THE EDITORIAL BOARD OF THE JOURNAL

Dear authors, reviewers and readers:

Professor Juan D. Velásquez initiated his work as Editor of DYNA in November 2013, succeeding Professor Oscar J. Restrepo, editor for the last ten years. Professor Velásquez has been an active author and reviewer in several journals for the last decade and he was the first editor of the journal Avances en Sistemas e Informática. He received a Bs. Eng in Civil Engineering, an MSc degree in Systems Engineering, and a PhD degree in Energy Systems from the Universidad Nacional de Colombia. Medellin. He worked for electricity utilities and consulting companies within the power sector for seven years. Currently, he is a professor in the Computing and Decision Sciences Department in the Facultad de Minas, Universidad Nacional de Colombia. 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.

On behalf of the faculty council members, I would like to thank, Professor Restrepo for ten years of magnificent work. Many successes have been achieved in the last decade; they include, among others, the creation of the online version and the presence of the journal in several scientific and social networks, the inclusion of DYNA in several bibliographical indexes and databases, but mainly, the inclusion of DYNA in the Journal Citation Reports. Due to the diligent work of professor Restrepo, the authors, reviewers, and readers, DYNA has an impact factor 0.175 for the year 2012.

John W. Branch
Dean
Facultad de Minas
Universidad Nacional de Colombia.

EDITORIAL

REVIEW OF THE AIMS AND SCOPE OF THE JOURNAL

In my first editorial as new editor, I have decided to discuss the revised Aims and Scope of DYNA which is posted in our website and in the paper version of the journal as follows:

> "DYNA is an international journal published by the Facultad de Minas, Universidad Nacional de Colombia, Medellín Campus since 1933. DYNA publishes peer-reviewed scientific articles covering all aspects of engineering. Our objective is the dissemination of original, useful and relevant research presenting new knowledge about theoretical or practical aspects of methodologies and methods used in engineering or leading to improvements in professional practices. All conclusions presented in the articles must be based on the current state-of-the-art and supported by a rigorous analysis and a balanced appraisal. The journal publishes scientific and technological research articles, review articles and case studies."

The new Aims and Scope has a clear and explicit focus on original, useful and relevant research in engineering. By the word "original", I mean that the journal only accepts original research that not has been previously published, except as a thesis in a university or as a short abstract in conference proceedings. Here, the editorial team follows a common rule in many scientific journals. As a consequence, manuscripts based on previously published work with unsubstantial new original contributions will be rejected directly by the Editor; examples of unsubstantial contributions include rewritten manuscripts with the same conclusions of previous published works but varying the data or adding new cases. By the word "original", I also mean that theoretical or practical results and conclusions cannot be derived from current knowledge in a straightforward, direct or trivial way.

The words "useful" and "relevant" are used to mean that conclusions describe important and interesting contributions that will be potentially usable for the readers in their future professional or research activities. As a consequence, the originality is a necessary but not sufficient condition for accepting a manuscript.

By "rigorous analysis" must be understood that authors followed well-accepted and well-applied methodologies for obtained and supporting their conclusions. For example, some manuscripts in engineering education state in their conclusions that the participants are "happier and more motivated" but without a rigorous experimental design demonstrating the advantage of the proposed practice o method in the learning process; as a result, this type of manuscripts are not suitable for publication.

As many other research journals, DYNA is focused on publishing original scientific or technological research articles in engineering. In addition, review articles are considered for publishing, but in this case, it is required the use of at least fifty references, and a in-depth discussion focused in the findings obtained when the literature is analyzed as a whole. Review articles should clearly and explicitly justify the necessity of review in their introduction and explain careful the objectives of the work. Further, it is mandatory that the discussion and conclusions presented cannot be derived from the isolated reading of the manuscript's references.

DYNA accepts case studies for publishing; however, the case considered must be non trivial, possibly unique, and interesting for the community; the manuscript must present new knowledge to improve the understanding of the case or novel methods for tackling it. Authors should clearly and explicitly justify how the work meets previous conditions for publishing. In this sense, technical reports or case studies describing the use of well-known methods or techniques applied in typical cases are not suitable for publishing.

With the new Aims and Scope, authors are encouraged to explicitly discuss how their work impacts the current state of the art of engineering or how to improve current practices. The editor and the reviewers will look for this discussion in the manuscript.

Finally, the objective and contributions of the work must be clearly stated in the abstract. This point is of special importance because of it allows to the editor and reviewers to accelerate the publishing decision.

Juan D. Velásquez, MSc, PhD
Professor
Departamento de Ciencias de la Computación
y de la Decisión
Facultad de Minas
Universidad Nacional de Colombia

SIMULTANEOUS LOCALIZATION OF A MONOCULAR CAMERA AND MAPPING OF THE ENVIRONMENT IN REAL TIME

LOCALIZACIÓN DE UNA CÁMARA MONOCULAR Y MAPEO SIMULTÁNEO DEL ENTORNO EN TIEMPO REAL

ANDRÉS DÍAZ

Magister , Grupo de Percepción y Sistemas Inteligentes, Universidad del Valle-Cali, andres.a.diaz@correounivalle.edu.co

LINA PAZ

Dra en Sistemas e Informática, Grupo de Robótica, Percepción y Tiempo Real, Universidad de Zaragoza, linapaz@unizar.es

EDUARDO CAICEDO

Dr en Informática, Grupo de Percepción y Sistemas Inteligentes, Universidad del Valle-Cali, eduardo.caicedo@correounivalle.edu.co

PEDRO PINIÉS

Dr en Ingeniería de Sistemas, Grupo de Robótica, Percepción y Tiempo Real, Universidad de Zaragoza, ppinies@unizar.es

Received for review November 03 th, 2012, accepted October 18th, 2013, final version december, 12 th, 2013

ABSTRACT: In this work a Visual SLAM system (Simultaneous Localization and Mapping) that performs in real time, building feature-based maps and estimating the camera trajectory is presented. The camera is carried by a person that moves it smoothly with six degrees of freedom in indoor environments. The features correspond to high quality corners parametrized with inverse depth representation. They are detected inside regions of interest and an occupancy criterion is applied in order to avoid feature agglomeration. The association process is developed using active search. The final representation is made in a three-dimensional environment.

Keywords: Localization, mapping, EKF, monocular camera, inverse depth, real time, 6DOF, active search.

RESUMEN: En este trabajo se presenta el desarrollo de un sistema de SLAM Visual (Simultaneous Localization and Mapping) que se desempeña en tiempo real, construyendo mapas basados en puntos característicos y estimando la trayectoria de la cámara. La cámara es transportada por una persona que la mueve suavemente con seis grados de libertad en entornos interiores. Los puntos característicos corresponden a esquinas de alta calidad parametrizados con el inverso de su profundidad. Estos son detectados dentro de regiones de interés y se aplica un criterio de ocupación con el fin de evitar aglomeración de características. El proceso de asociación se desarrolla usando búsqueda activa. La representación final se realiza en un entorno tridimensional.

Palabras Clave: Localización, mapeo, EKF, cámara monocular, inverso de la profundidad, tiempo real, 6DOF, búsqueda activa.

1. INTRODUCTION

Before carrying out tasks such as navigation, path planning, and object and place recognition, a totally autonomous mobile robot must interpret the information obtained by its sensors and then estimate its position and the position of environmental features. The simultaneous localization and map building algorithms face both problems at the same time [1], and they have been the focus of attention of the research community on mobile robotics during the last two decades.

The system described in this article is able to estimate the camera position, which is carried by a person or by a mobile platform, and to represent the trajectory that it makes. The system creates a three-dimensional map composed of the camera model and spatial points that represent object corners in the environment. Moreover, it can be adapted to different mobile platforms -terrestrial, aquatic, and aerial- because it is portable and has six degrees of freedom that reduce motion restrictions. The system is of great importance when GPS information is not available and in applications where is not practical to carry heavy and bulky sensors such as object tracking and mapping of environments in rescue operations.

Section 2 defines the schema of the Visual SLAM system and the general methodology used in this

work. Section 3 presents outstanding projects about Visual SLAM. Sections 4 and 5 explain how the key points were detected and how the radial distortion was corrected, respectively. Sections 6, 7, 8 and 9 present the parametrization process with the inverse depth of the features, the constant velocity model, the prediction of feature location in the image plane and the data association, respectively. Finally, the results and conclusions obtained in this work are presented.

2. VISUAL SLAM

Recently, the use of visual sensors has generated great interest in the research field of SLAM due to the large amount of texture information provided by these sensors of the objects found in a scene [2-4]. Moreover, cameras are compact, accurate, and much cheaper than laser sensors.

Implementations such as the ones developed by Castellanos [5] and Davison [6] proved the EKF (Extended Kalman Filter) in the building of small maps in SLAM systems with stereo vision, working in real time at 5 Hz. The system was able to build three-dimensional maps and to control a mobile robot. Jung and Lacroix [7] developed an autonomous system for mapping terrains using stereo vision as the only sensor and the standard EKF. Saez [8] presented a SLAM system with stereo vision for six degrees of freedom movements and indoor environments.

Some SLAM systems that use a monocular camera have proved to be viable in small environments; the most outstanding systems are the ones designed by Bailey [9], Kwok [10] and Lemaire [11]. Most of them are essentially EKF-SLAM systems and only change the initialization techniques and the kind of interest points extracted from the images (Harris corners, Shi and Tomasi corners, SIFT features, or any mixture of them). The works of Civera [12], Tully [13], Clemente [14] and Marzorati [15] show a tendency to use monocular cameras, inverse depth parametrization, and to perform in real time. The sub-mapping techniques, such as the ones depevoped by Bosse [16], Leonard [17], Paz [18] and Piniés [19], allow the system to achieve a performance in long trajectories.

3. SCHEMA OF OUR SLAM SYSTEM

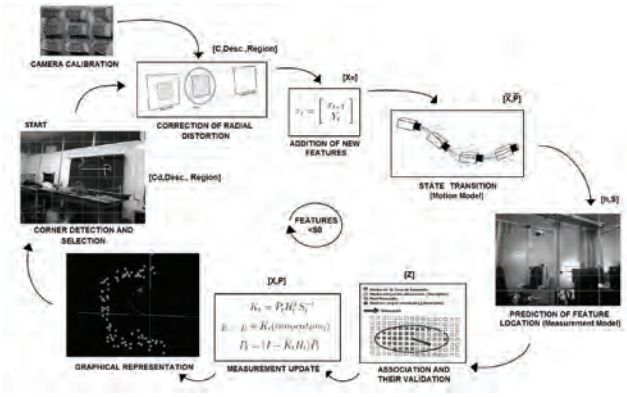


Figure 1. Schema of the SLAM system

The SLAM system involves many processes that work together in sequential way as is shown in Fig. 1. The probabilistic core is the EKF that alternates between a prediction step and an update step. Every process has inputs and outputs that are in a chain that ends up in a state estimate. In the next sections, these variables and their functions in the whole process are explained.

4. CORNER DETECTION

The system begins getting information of the environment through key points; in this work the corners are obtained with the Harris detector, supported by OpenCV. The image is split in 36 region of interest and for each region the Harris detector is applied, returning the best corner. From them, the corners with their minimum eigenvalue over a given threshold are chosen and only five of them are initialized, the best corners. At the beginning all the regions are empty, but after the first iteration, an occupancy algorithm must be used in order to avoid agglomeration of corners and therefore, wrong associations. In this step the coordinates of the five best corners are stored, the regions where they were found and a patch of 15x15 pixels around each corner.

4.1. Occupancy Algorithm

This criterion defines empty and occupied regions of interest. Only empty regions can be used to initialize a new feature. Moreover, when a region becomes empty because both the feature was deleted or the feature moves to another region, 20 time steps must pass in order to consider this region available to be occupied

again. This technique allows the features to be well distributed over the image plane.

5. CORRECTION OF RADIAL DISTORTION

The corner coordinates have radial distortion that affects the location of the pixels and this displacement grows as the pixel nears the image boundary. The model that describes this distortion is shown in (1).

$$\begin{pmatrix} x_{nd} \\ y_{nd} \end{pmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (1)$$

where k_1 and k_2 are the coefficients of radial distortion, r is the radius, x_n and y_n are the normalized coordinates. This model allows the system to include radial distortion. However, the opposite process is needed (remove radial distortion) and there is no analytical function that does this. Therefore, a numerical method is employed, the Newton Raphson method, that use the expression (2) and its derivative in order to calculate an approximation of the radius without distortion.

$$r_d = r + k_1 r^3 + k_2 r^5 \quad (2)$$

Given the radius r, the principal point (C_x, C_y) and the image coordinates with distortion (u_d, v_d), the image coordinates without distortion (u, v) can be computed using the expressions (3) and (4). Hereafter the corners will be called features.

$$u = \frac{u_d - c_x}{1 + k_1 r^2 + k_2 r^4} + c_x \quad (3)$$

$$v = \frac{v_d - c_y}{1 + k_1 r^2 + k_2 r^4} + c_y \quad (4)$$

6. FEATURE INITIALIZATION

This step consists in the corner parametrization and its inclusion to the state vector. The explanation of the corner parametrization using inverse depth representation and the addition of features in the state vector will be presented in this section.

6.1. Inverse Depth Representation

A significant limitation of the initial approaches of Davison [2] and others was that the systems could only use features close to the camera and that had great parallax during the motion. This problem limited the robot navigation (or the camera navigation) to indoors. Montiel [20] proposed a technique to initialize features using the inverse distance between the feature and the camera where it was seen for first time. This technique allows the system to work with both close and distant features from the moment they are detected. The distant features are used to improve the motion estimation, acting initially as an orientation reference. These features are common in outdoor environments.

The coordinates (u, v) are used in the back projection model, obtaining normalized coordinates x_n and y_n:

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \begin{pmatrix} 1/f & 0 & -c_x/f \\ 0 & 1/f & -c_y/f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (5)$$

where f is the focal length and (c_x, c_y) is the principal point. The normalized coordinates give information about the ray h_c that passes through the optical center of the camera and the point in the world whose image coordinates are (u_d, v_d). The ray can be defined by the angles θ and Φ, the azimuth and the elevation angles respectively:

$$\theta = \tan^{-1}(x_n) \qquad \phi = \tan^{-1}(y_n) \quad (6)$$

The camera state is defined with six parameters:

$$y_i = [X_{wca} \quad \theta_i \quad \phi_i \quad \rho_i] \quad (7)$$

The vector $X_{wca} = [x_{wc}\ y_{wc}\ z_{wc}]^T$ corresponds to the camera location, in Cartesian coordinates, from where the features were seen for first time, θ_i is the azimuth angle, ϕ_i is the elevation angle and $\rho_i = 1/d_i$ is the inverse distance between the camera position and the feature.

6.2. Addition of Features to the State Vector

The state vector stores the information of the camera and outstanding features:

$$x(k) = \begin{bmatrix} x_c(k) \\ Y(k) \end{bmatrix} \quad (8)$$

$$x_c(k) = [r^w(k) \quad \psi^w(k) \quad v^w(k) \quad w^c(k)]^T \qquad (9)$$

where r^w corresponds to the three cartesian coordinates of the camera location, ψ^w is the camera orientation in *Roll, Pitch, Yaw* angles $[\psi_x, \psi_y, \psi_z]^T$, v^w is the linear velocity of the camera and ω^c is the angular velocity with respect to the camera frame. The vector $Y(k)$ contains the information of the environment, organized by set of features taken from different camera locations:

$$Y(K) = [y_1(k) \quad ... \quad y_n(k)]^T \qquad (10)$$

where each feature y_i was defined in equation (7). A feature initialized remains in the state vector for the whole execution if this overcomes the following criterion: the feature must be seen at least 17 times in the first 20 iterations, from the time it was detected. If certain feature overcomes this criterion, it will not be deleted from the state vector and will be predicted in every iteration.

7. MOTION MODEL

The camera is connected to a laptop and is carried by a mobile robot or by a person. A program on the laptop determines the trajectory and builds a map with well distributed features in real time. The camera moves freely in three dimensions in an unknown environment. A constant linear and angular velocity model is used. The motion model allows the system to estimate the state transition in order to predict the camera position in the next time step before getting a new observation of the environment. The motion model is a non-linear function that only affects the camera state because the features are assumed to be static. The following transition function is used to pass from the state x_k to the state x_{k+1}:

$$x(k+1) = f(x(k), W(k)) \qquad (11)$$

The vector $W(k)$ represents a zero-mean Gaussian noise with covariance Q that affects the linear and angular velocities of the camera to detect small changes in the model:

$$W(k) = \begin{pmatrix} \Delta v^w(k) \\ \Delta w^c(k) \end{pmatrix} \qquad (12)$$

The camera state x_c evolves according to the following expression:

$$\begin{bmatrix} r^w(k+1) \\ \psi^w(k+1) \\ v^w(k+1) \\ w^c(k+1) \end{bmatrix} = \begin{bmatrix} r^w(k) + v^w(k+1) * \Delta t \\ \psi^w(k) + E_c{}^w * w^c(k+1) * \Delta t \\ v^w(k) + \Delta v^w(k) \\ w^c(k) + \Delta w^c(k) \end{bmatrix} \qquad (13)$$

Where $E_c{}^w$ is a matrix that transforms angular velocities with respect to the camera frame to equivalent angular velocities in the world frame.

8. PREDICTION OF THE FEATURE LOCATION

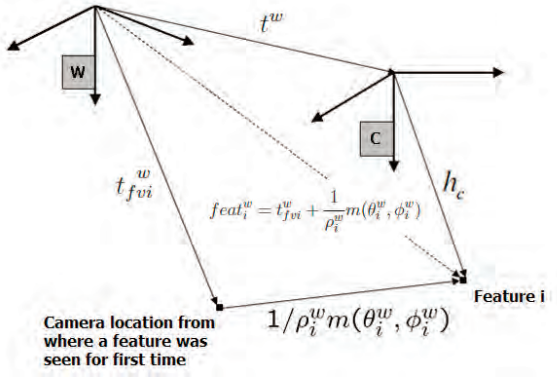


Figure 2. Feature observed from the initial and current camera location

This process consists in predicting the feature location in the next image, without making a new observation. Figure 2 provides a graphical representation of the vectors of the camera and feature location.

The vector t_{fvi} represents the camera location from where a feature i was observed for first time. The vector defined by m, the unitary vector of the bearing of the feature i when this feature was seen for the first time, this represents the feature location with respect to the vector t_{fvi}. The sum of these vectors is equal to vector $feat_i^w$, the feature location with respect to the world frame.

The vector t_w represents the current camera position, estimated with the motion model described in section 7. The difference of t_w and $feat_i^w$ is equal to the vector h_w. This vector has to be transformed to the camera frame, obtaining h_c. The equation used to predict the azimuth and elevation angles of a feature is based on

the components of the vector h_c, [h_{cx}, h_{cy}, h_{cz}]:

$$x_n = \frac{h_y^{\,c}}{h_x^{\,c}} \qquad y_n = \frac{h_z^{\,c}}{\sqrt{\left(h_x^{\,c}\right)^2 + \left(h_y^{\,c}\right)^2}} \tag{14}$$

The coordinates (u, v) are calculated from the normalized coordinates x_n and y_n:

$$u = x_n * f + c_x \qquad v = y_n * f + c_y \tag{15}$$

9. DATA ASSOCIATION

The location in the image plane (u_i, v_i) where the features $feat_i$ will be observed, for i = 1, 2, 3, ..., n, is predicted together with the innovation covariance matrix S_i. This matrix defines an elliptical zone of uncertainty where there is high probability to re-observe the feature. In this zone a correlation algorithm is executed, comparing the distribution of the digital levels of the pixels. The location that shows the strongest similarity will be taken as the equivalent point to the central pixel of a corner patch and will be the *observed position* of the feature from the new camera position.

Figure 3. Prediction of the feature location in the image plane (red points). The ellipses represent the prediction uncertainty.

In Fig. 3 the predictions of feature locations (red points) into the image plane are shown. The blue ellipses indicate failed correlations and therefore, there is no new observation. The green ellipses indicate successful correlations and the new observation is drawn in blue. The yellow point corresponds to a new observation that was parametrized and included into the state vector. This new feature is over an empty and available region and its distance to any other feature is more than 30 pixels.

A joint compatibility test based on the Mahalanobis distance is carried out to deal with spurious associations between observations and predicted features that come from dynamic objects in the mapped environment.

When the uncertainty of a feature increases so much, the search zone is too big and it is not suitable to develop the correlation process. In this case this prediction is not used, but the feature is not deleted, it remains in the state vector.

Finally, the difference between the observed feature (blue point) and predicted feature (red point) is the innovation vector and it is used by the Extended Kalman Filter to update the joint state camera-features. This vector moves the estimated position in the direction in which it is reduced.

10. RESULTS

Figure 4. Hand held camera

The experiments were developed with the Logitech Pro 9000 camera connected to a HP laptop with a 2.2 GHz AMD Dual-core processor. The camera was carried by a person (Fig. 4) that moves it smoothly with six degrees of freedom, in unknown environments.

10.1. Open Trajectory in Indoor environments

Figure 5. Laboratory of PSI group

The first experiment was performed in the Laboratory of the Perception and Smart Systems Group. It is a small room with glass walls, chairs, and desks with monitors, printers, CPUs, among other things (Fig. 5). Some corners over the walls belong to reflections and produce failed correlations (blue ellipses in Fig. 3) so most of them are rejected by the high quality features criterion.

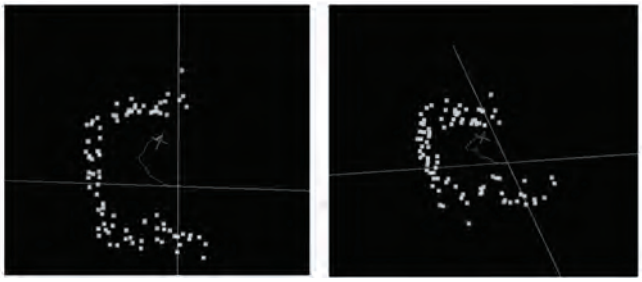


Figure 6. Three-dimensional Graphic

Figure 6 show the corners (points), the camera (triangular prism) and its trajectory (points connected by segments), represented in a three dimensional environment, developed with OpenGL.

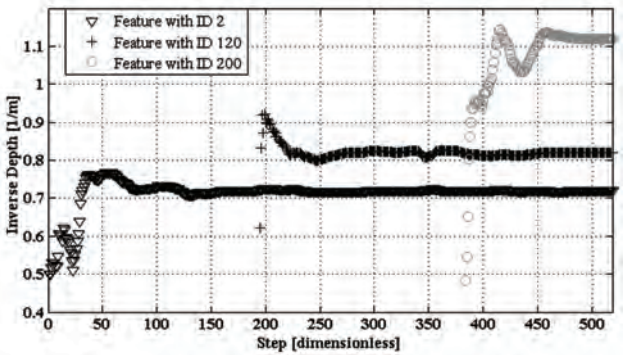


Figure 7. Evolution of Inverse depth estimates of three features

Figure 7 shows how the inverse depth estimates evolve over time. The inverse depth of a feature is initialized with a predefined value with respect to the camera location when the feature is seen for first time. The camera is both rotated and translated and the inverse depth estimate converges to a given value after about 50 iterations. At steady state, the estimates do not vary significantly, which means that the map is consistent. Finally, these estimates are used to compute the feature locations with respect to a global frame.

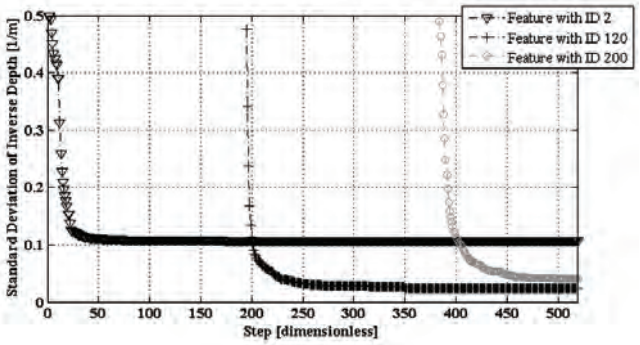


Figure 8. Evolution of Standard Deviation of Inverse Depth Estimates

As time passes, the parallax angles increase, yielding better estimates of the inverse depth, which is evidenced by a reduction in standard deviation, as can be seen in Fig. 8.

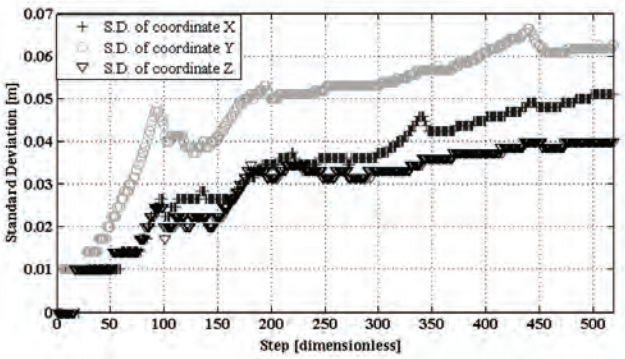


Figure 9. Evolution of Standard Deviation of Camera Location *(X,Y,Z)*

As the camera moves, its own pose uncertainty increases (Figs. 9 and 10). This fact is due to the errors introduced by the motion and observation models and the linear approximations made by the EKF. However, something very interesting happens when a loop is closed. This fact will be seen in the following experiment.

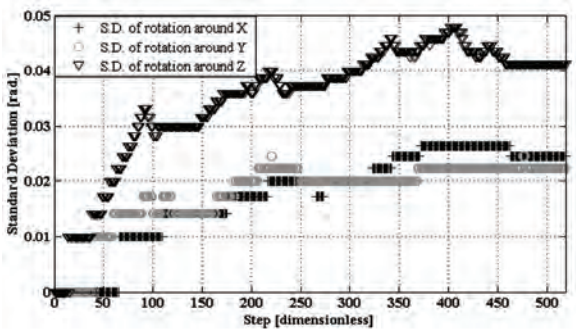


Figure 10. Evolution of Standard Deviation of Camera Bearing around the axis X, Y and Z.

10.2. Closed Loop in Indoor Environments

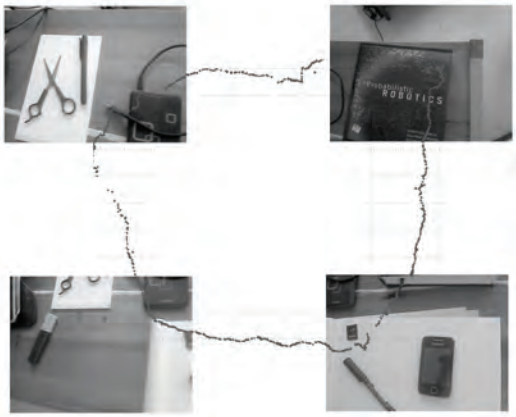


Figure 11. Closed loop with the camera focusing objects over a desk.

This experiment was carried out with the camera focusing objects over a desk (Fig. 11), trying to follow a square trajectory and to keep a constant distance from the camera to the surface of the desk. The scale of the trajectory was fixed by hand because it is not observable with a monocular camera.

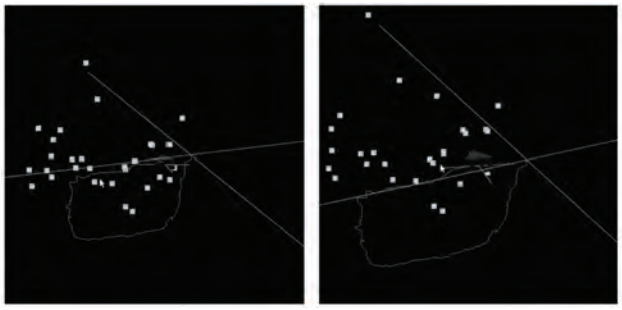


Figure 12. Three-dimensional graphic of a closed loop.

Figure 12 shows the square trajectory and the corners represented with OpenGL.

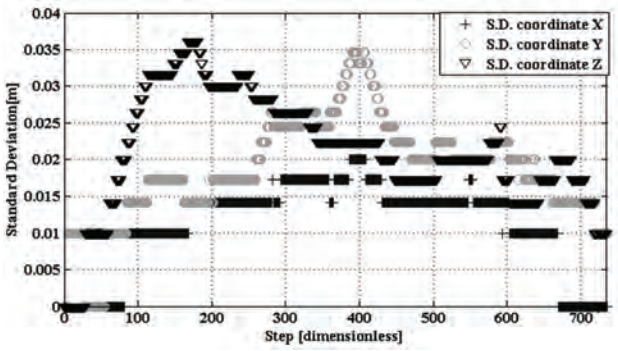


Figure 13. Evolution of Standard Deviation of Camera Location *(X,Y,Z)* in closed loop

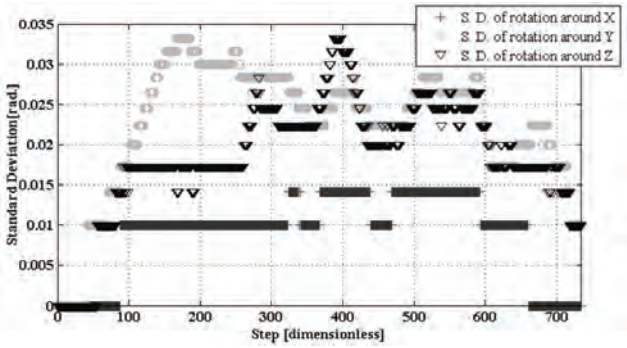


Figure 14. Evolution of Standard Deviation of Camera Bearing around the axis X, Y and Z in closed loop.

The camera observes features that were seen in the beginning of the mapping and whose location is relatively well known. Through these observations the uncertainty in camera position (location and orientation) is reduced as is shown in Figs. 13 and 14. These observations also reduce the uncertainty for other features in the map due to the correlation stated in the covariance matrix.

10.3. Computational Cost

The high computational cost is the main limitation in systems that perform in real time. This problem has been tackled with sub-mapping techniques that allow the system to navigate in large environments and to reduce the errors due to the linear approximations made by the Extended Kalman Filter.

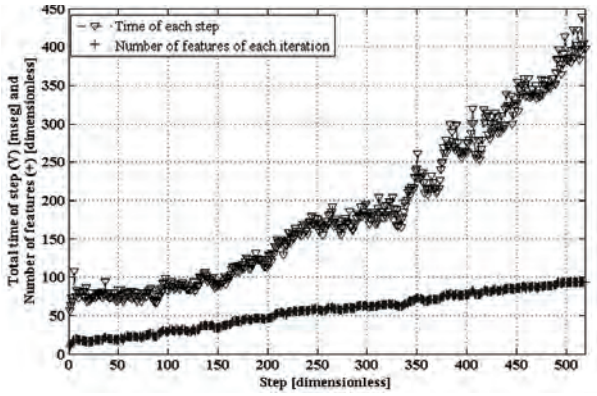


Figure 15. Total Computational Cost (with ▼) and number of features (with +).

Figure 15 depicts the quadratic dependence on the number of features in the map. This fact is due to the size of the covariance matrix that is used to update the state. The matrix operations that involve the covariance matrix are computationally expensive and impose a limit of the number of features to 50 in order to perform in real time, managing to process at least 10 images per second (at the critical point).

11. CONCLUSIONS

A Visual SLAM system that works with a monocular camera in real time was developed. The core of the system relies on the well known incremental Extended Kalman Filter such that the positions of camera and a feature-based map can be estimated in real time. The kind of sensor, the 6 DOF and the probabilistic focus used to solve the problem, make it a complex system. The results show that the system performs in indoor environments in real time if the amount of features is under 50, processing from 10 to 20 frames per second.

The estimated state of the camera has low uncertainty: the standard deviation in location is less than 7cm (for each coordinate) and in orientation is less than 3 degrees (for each axis). The inverse depth estimates of landmarks converge to a steady state in about 50 iterations, building consistent maps.

The feature detection is performed using regions of interest and an occupancy algorithm is implemented to avoid feature agglomeration, achieving high quality corners that are well distributed. The elliptical zones defined by the innovation covariance matrix allow the system to carry out an active search of corner patches, optimizing the correlation process. However, the matrix operations increase the computational cost and set a limit to real time performance.

An interesting fact was analyzed, with closed loops the uncertainty decreases when the camera visits a place where it has been before, and recognizes features that were seen before.

REFERENCES

[1] Smith, R. C., and Cheeseman P. On the representation and estimation of spatial uncertainty. Int. J. Robotics Research, 5(4): pp.56–68, 1986.

[2] Davison, A. J., and Murray, D. W. Simultaneous localization and map-building using active vision. IEEE Trans. Pattern Analysis and Machine Intelligence, 24(7): pp.865–880, 2002.

[3] Davison, A. J., Reid, I. D., Molton, N. D., and Stasse, O. Monoslam: Real-time single camera slam. IEEE Trans. Pattern Analysis and Machine Intelligence, 29(6):pp.1052–1067, June 2007.

[4] Quintián, H., Calvo, J. L., and Fontenla, O. Aplicación de un robot comercial de bajo coste en tareas de seguimiento de objetos. Revista Dyna, 175, pp.24-33, 2012.

[5] Castellanos, J. A. Mobile Robot Localization and Map Building: A Multisensor Fusion Approach. PhD thesis, Dpto. de Informática e Ingeniería de Sistemas, University of Zaragoza, Spain, May 1998.

[6] Davison, A.J. Mobile Robot Navigation using Active Vision. PhD thesis, University of Oxford, 1998.

[7] Lacroix, S., and Jung, I. K. High resolution terrain mapping using low altitude aerial stereo imagery. In Proc. of the 9th Int. Conf. on Computer Vision, pages pp.946–951, 2003.

[8] Saez, J.M., Escolano, F., and Penalver, A. First Steps towards Stereobased 6DOF SLAM for the Visually Impaired. In Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05)-

Workshops-Volume 03. IEEE Computer Society Washington, DC, USA, 2005.

[9] Bailey, T. Constrained initialization for bearing-only SLAM. In Proc. IEEE Int. Conf. on Robotics and Automation, (ICRA'03), volume 2, 2003.

[10] Kwok, N. M., Dissanayake, G. An efficient multiple hypothesis filter for bearing-only SLAM. In Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, (IROS'04)., volume 1, 2004.

[11] Lemaire, T., Lacroix, S., and Sola, J. A practical 3D bearing-only SLAM algorithm. In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS'05), pp. 2449–2454, 2005.

[12] Civera, J., Davison, A. J., and Montiel, J. M. Inverse depth parametrization for monocular SLAM. IEEE Transactions on Robotics, 24(5): pp932–945, October 2008.

[13] Tully, S., Moon, H., Kantor, G., and Choset, H. Iterated filters for bearing-only slam. In IEEE International Conference on Robotics and Automation, 2008.

[14] Clemente, L., Davison, A. J., Reid, I. D., Neira, J., and Tardos, J. D. Mapping large loops with a single hand-held camera. In Proc. Robotics: Science and Systems, Atlanta, GA, USA, June 2007.

[15] Marzorati, D., Matteucci, M., Migliore, D., and Sorrenti, D. On the use of inverse scaling in monocular slam. In IEEE International Conference on Robotics and Automation, 2009.

[16] Bosse, M., Newman, P. M., Leonard, J. J., Soika, M., Feiten, W., and Teller, S. An atlas framework for scalable mapping. In Proc. IEEE Int. Conf. Robotics and Automation, pp. 1899–1906, Taipei, Taiwan, 2003.

[17] Leonard, J. J., and Newman, P. M. Consistent, convergent and constant time SLAM. In Int. Joint Conf. Artificial Intelligence, Acapulco, Mexico, August 2003.

[18] Paz, L. M., Tardós, J. D., Neira, J. Divide and Conquer: EKF SLAM in O(n). Accepted in Transactions on Robotics (in Print), 24(5), October 2008.

[19] Piniés, P., and Tardós, J. D. Large Scale SLAM Building Conditionally Independent Local Maps: Application to Monocular Vision. Accepted in Transactions on Robotics (in print)., 24(5), October 2008.

[20] Montiel, J. M., and Civera, J., Davison, A. J. Unified inverse depth parametrization for monocular SLAM. In Proc. Robotics: Science and Systems, Philadelphia, USA, August 2006.

MICROWAVE HEATING AND SEPARATION OF WATER-IN-OIL EMULSION FROM MEXICAN CRUDE OIL

CALENTAMIENTO POR MICROONDAS Y LA SEPARACIÓN DE EMULSIÓN DE AGUA EN ACEITE DEL PETRÓLEO CRUDO MEXICANO

ADRIAN VAZQUEZ V.
Dr(Est.)Inst Politécnico Nacional, México, Centro de Investigación Cienc. avazquezv0802@alumno.ipn.mx,
ARTURO LOPEZ M.
Profesor Titular Inst Politécnico Nacional, México, Centro de Inve Cienc. Avanzada, arlopezm@ipn.mx
LUIS J. ANDRADE C.
MSc(Est) Inst Politécnico Nacional, México, Centro de Inv. en Ciec. landradec1100@alumno.ipn.mx
ARIANA M. VAZQUEZ A.
MSc(Est) Inst Politécnico Nacional, México, Centro de Inv. en Ciencia , ariana_vazquez@hotmail.com

Received for review February 04 th, 2013, accepted October 8th, 2013, final version December, 04 th, 2013

ABSTRACT: Microwave heating and gravity sedimentation are alternatives for demulsification and layer separation into oil and water layers, this process was demonstrated in the laboratory and provides an option for reducing and oil recovering from water-in-oil Mexican oil emulsions. The combinatorial process was implemented in a test lab using Mexican crude oil samples. The Laboratory samples were 100% and 50-50%, crude and crude-water respectively, were heated. The results were encouraging show that microwave heating and gravity sedimentation are alternatives for the separation of Mexican Oil emulsions.

Keyword: Microwave heating, Demulsification, Separation, Water-in-oil emulsions.

RESUMEN: El calentamiento por microondas y la sedimentación por gravedad son alternativas para desemulsificar y separar por capas aceite y agua, este proceso se demostró en el laboratorio y proporciona una opción para la reducción y la recuperación de emulsiones de agua en aceite del petróleo Mexicano.
El proceso se implementó en un laboratorio de pruebas con muestras de petróleo crudo mexicano. Las muestras de laboratorio fueron del 100% y 50-50%, crudo y crudo-agua, respectivamente, se calentaron. Los resultados fueron alentadores muestran que el calentamiento por microondas y la sedimentación por gravedad son alternativas para la separación de emulsiones de petróleo mexicano.

Palabras Clave: Calentamiento por Microonda, Desemulsificar, Emulsión de agua en aceite.

1. INTRODUCTION

A significant portion of world crude oil is produced in the form of a water-in-oil emulsion stabilized by natural surfactants that must be treated (demulsified) before they can be processed [1].

Water and oil sometimes combine (emulsify) during industrial processes. Resulting emulsions are either oil-in-water (o/w) or water-in-oil (w/o), depending on which material is dispersed in the other. An oil-in-water emulsion has water as the continuous phase, while the water-in-oil emulsion has oil in the continuous phase. Either type of emulsion may contain other contaminating materials (solids, dirt, metal particles, emulsifiers, cleaners, soaps, solvents, etc.). Emulsions can be found in a variety of industries and formed by a variety of processes. Since formation of emulsions is so specific to the industry and process, the emulsion breaking product selection is somewhat difficult and requires bench testing [2]. For water-in-oil emulsions which usually have high viscosity, the required mixing of these chemicals with the emulsion is difficult. Also when a high dosage of chemicals is used to overcome the difficulty, it leads to a secondary pollutant, since the separated water may contain too high a level of chemicals to be discharged to public water.

The concept of microwave heating of emulsion was first introduced by Klaila (1978) [3] and Wolf (1986) [4] in their patent applications. Recently, research continues

to develop technology for microwave demulsification chemical plants, as described by Coutinho (2010) [5].

Today the Mexican oil fields have a high content of water and oil, and the water usually contains dissolved salts forming brines.

The use of microwaves is an alternative, effective, clean and chemicals free method for oil desalting and dehydration. This publication shows that microwaves are an efficient alternative to heating oil emulsions.

2. EXPERIMENTAL

The study area is located in the oil drilling fields in Altamira, Tamaulipas, Mexico. Two types of sample of crude, 100% crude and 50-50% crude-water, were heated.

The variables studied are the temperature and the irradiation time during microwave irradiation. Equipment was built to irradiate emulsions and measure the temperature at a predetermined time. Modules were designed for measuring the temperature and controlling the magnetron.

The experimental process is: 1) Collection of Oil Field samples (Figure 4,7), 2) Preparation and Stabilization of emulsions of 100% and 50% crude and crude-water respectively, 3) 50 ml of both samples are irradiated.

Temperature Measurement Module

An infrared sensor (Raytek RAYTXSLTCF1 model) (Figure 1) for measuring the temperature of oil samples, together with a graphical interface. Which can measure temperature in real time while the samples are being irradiated.

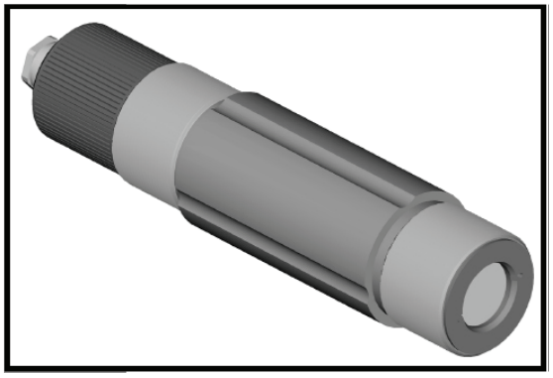


Figure 1. Infrared Sensor

A graphic user interface (GUI) was designed to visualize the temperature inside the sample to be irradiated with microwaves in real time (see Figure 2).

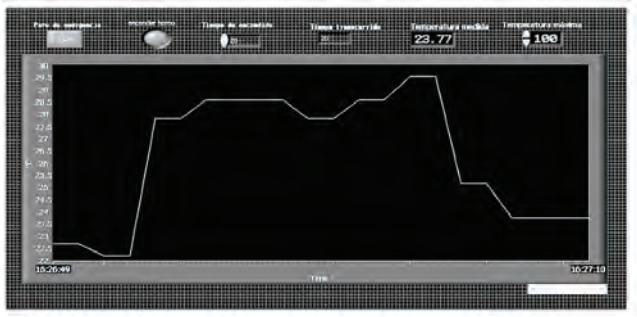


Figure 2. Visual Interface (GUI)

Magnetron Control Module

The magnetron operates at a working frequency of 2.45 GHz and a rated power output of 700 W. The control module magnetron has a graphical interface which controls the magnetron, the duration of microwave irradiation and emergency protections (see Figure 2,3).

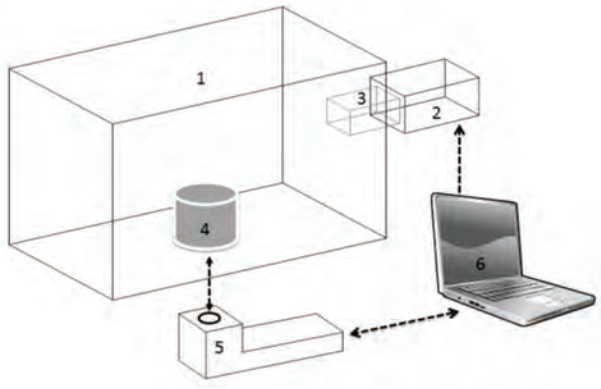


Figure 3. Modules magnetron and temperature control. (1)Resonant Cavity, (2)Magnetron, (3)Magnetron Control, (4) Irradiated sample, (5) Infrared Sensor, (6) Visual Interface (GUI).

3. RESULTS AND DISCUSSION

1) Collection of Oil Field samples

Samples were collected from a cistern in the Tamaulipas-Constituciones Oil Field (Figures 4 and 5).

Figure 4. *Injection plant, Tamaulipas-Constituciones field.*

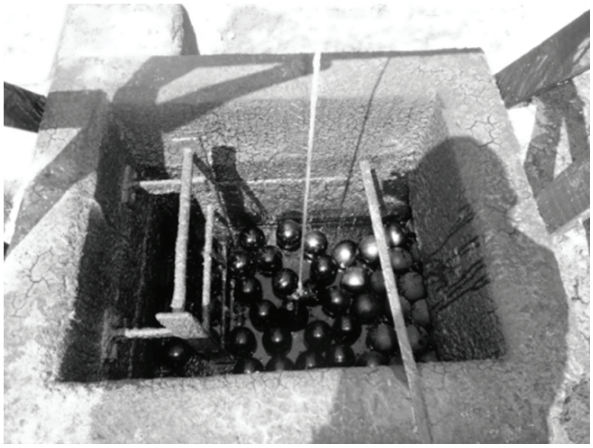


Figure 5. *Cistern Tamaulipas-Constituciones Oil Field.*

2) Preparation and Stabilization of emulsions.

Two types of samples were prepared and stabilized, 50-50% crude-water (Figure 6) and 100% crude (Figure 7). Each sample contained 50 ml.

3) Irradiation of samples.

Oil-Water samples irradiated for 30 to 60 seconds

The sample of oil-water 50-50% was irradiated for 30 seconds (Figure 8), with an initial temperature of 27 °C and a final temperature of 44 °C, with a temperature increase of 17 °C, shown in the Table 1, there is a rapid increase in temperature.

Figure 6. *50-50 % crude-water sample prepared and stabilized.*

Figure 7. *100% crude sample prepared and stabilized.*

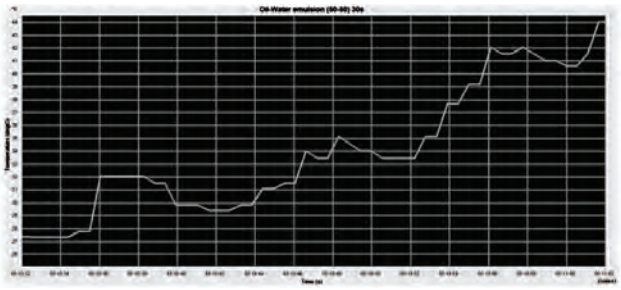


Figure 8. *Oil-Water sample irradiated for 30 seconds*

Then, the other sample of oil-water 50-50% was irradiated for 60 seconds (Figure 9), with an initial

temperature of 26 °C and a final temperature of 140 °C, with a temperature increase of 114 °C, as seen in the Table 1, a rapid increase in temperature. Although at 25 seconds there is a sudden temperature increase associated with interaction between microwaves and chemical properties of oil, as shown in Figure 12 FTIR.

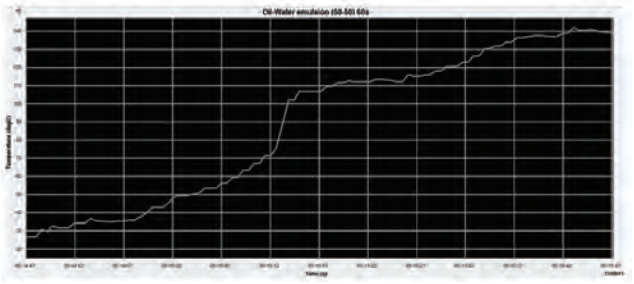


Figure 9. *Oil-Water irradiated for 60 seconds*

Oil samples irradiated for 30 to 60 seconds

The sample of 100% oil was irradiated for 30 seconds, with an initial temperature of 25 ° C and a final temperature of 53 °C, with a temperature increase of 28 °C, shown in the Table 1. Shown in Figure 10 shows that 100% oil also has a substantial absorption of the microwave.

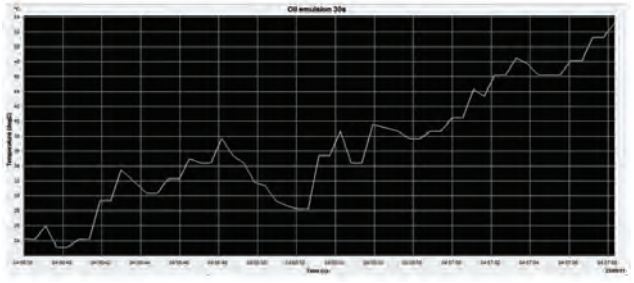


Figure 10. *100 % Oil sample irradiated for 30 seconds*

The second sample of 100% oil was irradiated for 60 seconds, with an initial temperature of 26 °C and a final temperature of 105 °C, with a temperature increase of 79 shown in Table 1. It can be observed in Figure 11 that shows the 100% sample also has considerable energy absorption, but not as much as the sample of 50-50% oil-water because the water accelerates energy absorption thereby causing greater polarization and thus further warming.

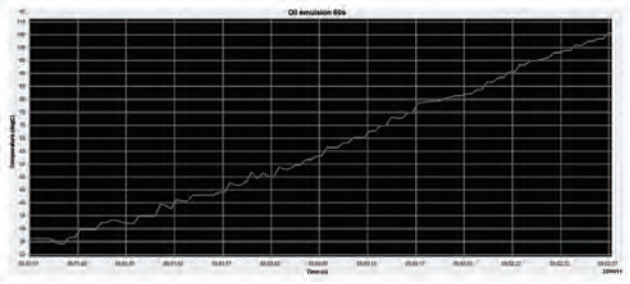


Figure 11. *100% Oil sample irradiated for 60 seconds*

Table 1. Comparison Chart irradiated samples.

	Time of Irradiated (50 ml)	Initial Temperature (°C)	Final Temperature (°C)	Temperature increase
Sample of Oil-Water (50-50)	30s	27	44	17
Sample of Oil-Water (50-50)	60s	26	140	114
Sample of Crude Oil	30s	25	53	28
Sample of Crude Oil	60s	26	105	79

Spectra de FTIR (Fourier Transform Infrared Spectroscopy)

The oil was characterized by FTIR before being irradiated, and subsequently after 30 and 60 seconds of it being irradiated:

- Crude: Samples of crude from Constitutions Tamaulipas field were characterized before being irradiated.
- Oil-water: two characterizations were performed on samples of 50ml of crude oil which were irradiated for 30 and 60 seconds respectively.
- Crude: two characterizations were performed on 50ml samples of Oil-Water (50-50), which were irradiated for 30 and 60 seconds respectively.

The main functional groups of absorption resins and asphaltenes are qualitatively identical. To facilitate the study of the results of this analysis, the spectrum is divided into four regions that can be interpreted as follows:

Table 2. - Organic functional groups infrared spectra of Crude by FTIR

	N° of waves	Group	Description
1	3400 cm-1	OH	Alcohols
2	3059 cm-1	=CH, =CH2	aromatic
3	2953 cm-1	-CH3	methyl
4	2920 cm-1	=CH2	methylene
5	2850 cm-1	CH3	methyl
6	2732 cm-1	C-H	aldehydic hydrogen
7	1774 cm-1	C=O	acids esters and acid anhydrous
8	1642cm-1	NH2	aromatic
9	1600 cm-1	C=C	Aromatic compound
10	1455 cm-1	CH3 + CH2	methyl and methylene
11	1375 cm-1	CH3	methyl
12	1308 cm-1	S=O	Sulphides
13	1250cm-1	C-O	carboxylic acids
14	1175 cm-1	C=S	sulphides
15	1100 cm-1	C-N	Amine in ring
16	1032 cm-1	C-N	Amine
17	949 cm-1	=CH	substituted aromatic
18	871cm-1	=CH	substituted aromatic
19	815 cm-1	=CH	aromatic alkenes
20	750 cm-1	C-H	ring
21	720 cm-1	4 (CH2)	methylene

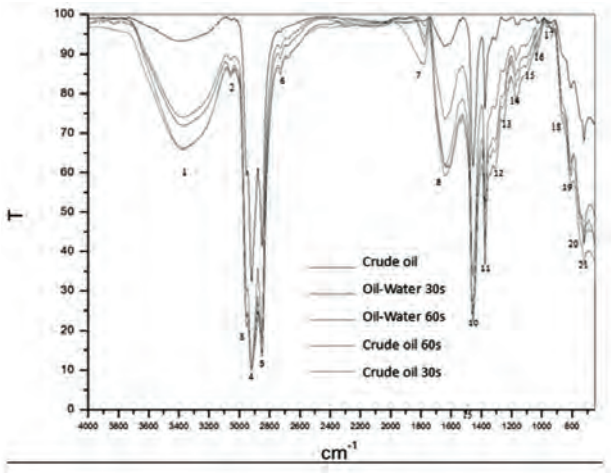


Figure 12. - FTIR spectra of the samples of oil water

Group I: hydrogen valence vibrations (small) in the region of 3100-2700 cm-1. Group II: Vibrations valence double or partially double bonds, this region of 1900-1350 cm-1. Group III: dubbing vibrations out of the plane. Valence vibrations low energy region between 1000-710 cm-1. Group IV: Region valence vibrations of single and double bonds very strongly coupled, a region is observed in the range of 1350-1000 cm-1 band as a matter of oxygen valency vibrations. In the spectra obtained the following functional groups can be identified:

In accordance with the intensity changes of the absorption bands after interaction with microwave radiation shown in Figure 12, the following can be summarized:

- An increase in the content of aromatic structures, manifested in higher band intensity in regions 1600, 3040, 870, 815 and 750 cm-1. This may also relate to an increase in the degree of hydrogen substitution in aromatic structures.

- An increase in the content of aliphatic and alicyclic (naphthenic) groups mainly at the expense of the CH2 group, as indicated by the increased intensity of the bands in the 2930, 2860, and 1470 cm-1 regions and of the doublet at 730 -720 cm-1. A slight decrease in the number of CH3 groups is indicated by weakening of the band in the 1380 cm-1 region and a decrease in the "shoulder" in the 2950 cm-1 region

- A significant increase pf the OH functional group band at 3400 cm-1 and in the region of 1774 cm-1 corresponding to C = O. This is attributed to the recovery of polar compounds solubilized in the added water, such as alcohols, esters, acids and acid anhydride.

- An increase in the bands belonging to the region covered by 1350-1000 cm-1, corresponding to an increase of asphaltenes recovered during microwave irradiation, because of the added water.

4. CONCLUSIONS

Using the temperature module with the infrared sensor the temperature could be measured in real time, without the need to enter any sensor within the sample, as this would contaminate the oil sample or damage the sensor. Using this module the signal representative of

the temperature could be visualized and recorded to be studied later.

The magnetron control module could be programed for the time or time intervals that the samples are to be irradiated using the graphical interface of , and to realize an emergency stop if the temperature rises above the limits established and so control the microwave heating process.

The crude oil, after being irradiated with microwaves, qualitatively retains the functional groups that it is compose of, there are only quantitative changes that are assumed to be due to the recovery of soluble compounds in injection water.

The increased intensity of the OH group (3400 cm-1) and C = O (1774 cm-1) confirms the recovery of soluble polar compounds in the injection water.

Asphaltenes, the main emulsion stabilizers present in the injection water are recovered. Microwave radiation, breaks the film of surfactant present in the water.

The spectra show that the time corresponding to 30 seconds irradiation, gave better results recovery of dissolved hydrocarbons (Figure 12). Establishing the lower irradiation time corresponds to low amount of energy required in the destabilization of the emulsion and therefore energy savings in the process.

ACKNOWLEDGEMENTS

The authors thank the company Pemex Exploration and Production for allowing sample collection in the Tamaulipas-Constituciones field.

REFERENCES

[1] Nour A. Emulsion Stability and Microwave Demulsification of Crude Oil Emulsion: Emulsion characterization, mechanisms of microwave heating technology, chemical demulsification, Lambert Academic Publishing, 2011, p. 25.

[2] Flynn J. The Nalco Water Handbook, 2009, p. 813.

[3] Klaila W. Method and Apparatus for Controlling Fluency of High Viscosity Hydrocarbon Fluids, 1978, U.S. Patent 4,067,683.

[4] Wolf N. Use of Microwave Radiation in Separating Emulsions and Dispersions of Hydrocarbons and Water, 1986, U.S. Patent 4,582,629.

[5] Countinho R. Method for the Microwave Treatment of Water-In-Oil Emulsions, 2010, U.S. Patent 7,705,058 B2.

A STABILITY AND SENSITIVITY ANALYSIS OF PARAMETRIC FUNCTIONS IN A SEDIMENTATION MODEL

UN ANALISIS DE ESTABILIDAD Y SENSIBILIDAD DE LAS FUNCIONES DEFINIDAS POR PARAMETROS EN UN MODELO DE SEDIMENTACION

CARLOS D. ACOSTA

PhD, Departamento de Matemáticas y Estadística, Universidad Nacional de Colombia, Manizales, cdacostam@unal.edu.co

RAIMUND BÜRGER

Dr.rer.nat., CI²MA y Departamento de Ingeniería Matemática, U. de Concepción, Chile, rburger@ing-mat.udec.cl

CARLOS E. MEJIA

PhD, Escuela de Matemáticas, Universidad Nacional de Colombia, Medellín, cemejia@unal.edu.co

Received for review December 13 th, 2012, accepted October 18th, 2013, final version december, 08 th, 2013

ABSTRACT: This paper deals with the reliable and efficient numerical identification of parameters defining the flux function and the diffusion coefficient of a strongly degenerate parabolic partial differential equation (PDE), which is the basis of a mathematical model for sedimentation-consolidation processes. A zero-flux initial-boundary value problem (IBVP) posed for this PDE describes the settling of a suspension in a column. The parameters for a given material are estimated by the repeated numerical solutions of the IBVP (direct problem) under systematic variation of the model parameters, with the aim of successively minimizing a cost functional that measures the distance between a space-dependent observation and the corresponding numerical solution. Two important features of this paper are the following. In the first place, the method proposed for the efficient and accurate numerical solution of the direct problem. We implement a well-known explicit, monotone three-point finite difference scheme enhanced by discrete mollification. The mollified scheme occupies a larger stencil but converges under a less restrictive CFL condition, which allows the use of a larger time step. The second feature is the thorough sensitivity and stability analysis of the parametric model functions that play the roles of initial guess and observation data, respectively.

Keywords: Sedimentation of suspensions, sensitivity analysis, degenerate parabolic equation, parameter estimation, discrete mollification.

RESUMEN. Este artículo se dedica a la identificación numérica confiable y eficiente de los parámetros que definen la función de flujo y el coeficiente de difusión en una ecuación diferencial parcial de tipo parabólico fuertemente degenerada que es la base de un modelo matemático para procesos de sedimentación-consolidación. Para esta ecuación, el problema de valor inicial con valores en la frontera (IBVP) en el que el flujo es nulo, describe el asentamiento de una suspensión en una columna. Los parámetros para un material dado se estiman con base en repetidas soluciones numéricas del problema directo (IBVP) con una variación sistemática de los parámetros del modelo, con el objeto de minimizar sucesivamente un funcional de costo que mide la distancia entre una observación dependiente de tiempo y la correspondiente solución numérica. En este artículo se destacan dos aspectos. El primer aspecto es que en el método propuesto para la solución numérica eficiente y acertada del problema directo, se implementa un esquema explícito monótono bien conocido basado en diferencias finitas que usan tres puntos mejorado por molificación discreta. El esquema molificado utiliza una malla de más puntos pero converge con una condición CFL menos restrictiva, lo cual permite usar pasos temporales más grandes. El segundo aspecto es el exhaustivo análisis de sensibilidad y estabilidad de las funciones definidas por parámetros en el modelo y que juegan los papeles de aproximación inicial y dato observado, respectivamente.

Palabras claves: Sedimentación de suspensiones, análisis de sensibilidad, ecuación parabólica degenerada, estimación de parámetros, molificación discreta.

1. INTRODUCTION

1.1. Scope

Our goal is the numerical identification of unknown parameters in the flux and diffusion terms for the following initial-boundary value problem (IBVP) for a strongly degenerate parabolic equation in one space dimension:

$$u_t + f(u)_x = A(u)_{xx},$$

$$(x,t) \in \Omega_T := (0,L) \times (0,T], L > 0, T > 0 \quad (1a)$$

$$u(x,0) = u_0(x), \ x \in [0,L], \quad (1b)$$

$$f(u) - A(u)_x|_{x=0} = \psi_0(t), \ t \in (0,T], \quad (1c)$$

$$f(u) - A(u)_x|_{x=L} = \psi_L(t), \ t \in (0,T], \quad (1d)$$

where A is an integrated diffusion coefficient, i.e.,

$$A(u) = \int_0^u a(s)\, ds, \quad a(u) \geq 0. \qquad (2)$$

The diffusion function a is assumed to be integrable and is allowed to vanish on u-intervals of positive length, on which (1a) turns into a first-order hyperbolic conservation law, so that (1a) is a *strongly degenerate parabolic*. On the other hand, we assume that f is piecewise smooth and Lipschitz continuous. Under suitable choices of u_0, f, a, ψ_0 and ψ_L the IBVP (1) may describe a variety of real-world applications like traffic flow [9]. We focus our attention on Equation (1) as a model of the sedimentation-consolidation process of a solid-liquid suspension [8].

It is well known that solutions of (1a) are, in general, discontinuous even if u_0 is smooth, and need to be defined as weak solutions along with an entropy condition to select the physically relevant solution, the *entropy solution*. For the definition, existence and uniqueness of entropy solutions of (1) we refer to [7, 8, 10].

In the present work we are interested in a stability and sensitivity analysis of the parametric model functions. In order to perform the tests, we first proceed with a numerical estimation procedure based on repeated numerical solutions of the direct problem (1) under successive variation of parameters appearing in the coefficient functions f and a. In this phase the main components are the efficient and stable solver of the direct problem and the optimization procedure based on the Nelder- Mead Simplex Method. Our goal is the stability and sensitivity analysis of the resulting inverse problem. Theoretical aspects related to identifiability are not our concern in this paper (but cf., e.g., [11]). By sensitivity analysis we mean an intensive set of tests for the numerical identification of parameters with or without noisy observation data. Our approach follows the methodology of [4] but we acknowledge the existence of other ways to perform a sensitivity analysis, for instance [19].

1.2. Related work and outline of the paper

The discrete mollification method is a convolution-based filtering procedure suitable for the regularization of ill-posed problems and for the stabilization of explicit schemes for the solution of PDEs. For the numerical identification of diffusion coefficients by discrete mollification, see [16] and its references.

Inverse problems for strongly degenerate parabolic equations are of particular interest in the context of the sedimentation-consolidation model. In fact, in applications such as wastewater treatment and mineral processing, the reliable extraction of material-specific parameters appearing in the model functions f and a from laboratory experiments allows the operation and control of continuous clarifier-thickeners handling the same material to be simulated [10, 21]. For the special case $A \equiv 0$, i.e., when effects of sediment compressibility are absent or negligible, (1a) reduces to a first-order nonlinear conservation law, portions of the function f can be identified by comparing observed space-time trajectories of concentration discontinuities, with trajectories appearing in closed-form solutions for piecewise constant initial concentrations [6, 14]. In the presence of sediment compressibility, closed-form solutions are not available and one has to resort to numerical techniques to solve the parameter identification problem [5, 11].

The paper is organized as follows. Section 2 presents the sedimentation-consolidation model along with details on the schemes for the solution of the direct problem, including a brief description of the mollification method. Section 3 deals with the parameter identification problem, the proposed algorithm, the sensitivity analysis and the effect of noisy observation data. This section ends with some conclusions.

2. THE APPLICATION OF THE MATHEMATICAL MODEL

2.1. Sedimentation model

According to [8, 10] and the references cited in these works, (1) can be understood as a model for the settling of a flocculated suspension of small solid particles dispersed in a viscous fluid, where $u = u(x,t)$ is the local solid concentration as a function of height x and time t. For batch settling in a closed column of height L we set $\psi_0 = 0$ and $\psi_L = 0$; the function u_0 denotes the initial solid concentration. The material specific function f describes the effect of hindered settling.

We employ here the following typical parametric expression:

$$f(u) = \begin{cases} v_\infty u\left(1 - \dfrac{u}{u_{max}}\right)^C & 0 \le u \le u_{max}, \\ 0 & otherwise, \end{cases} \quad (3)$$

where $v_\infty < 0$ is the settling velocity of a single particle in an unbounded fluid, $C>1$ is a dimensionless exponent that quantifies how rapidly the settling velocity decreases (as an absolute value) with increasing solids concentration, and $0 < u_{\max} \le 1$ is a (nominal) maximal solids concentration. The function A is given by (2), where we define

$$a(u) = -\frac{f(u)\sigma_e'(u)}{(\rho_s - \rho_f)gu}, \quad (4)$$

where ρ_s and ρ_f are the solid and fluid densities, respectively, g is the acceleration of gravity and $\sigma_e'(u) = \frac{d\sigma_e}{du}$ is the derivative of the material specific solid stress function σ_e.

Among several proposed semi-empirical approaches for σ_e we chose the power law type function

$$\sigma_e(u) = \begin{cases} 0 & \text{for } 0 \le u \le u_c, \\ \sigma_0\left[\left(u/u_c\right)^\beta - 1\right] & u_c < u, \end{cases}$$

with material-dependent parameters $\sigma_0 > 0$ and $\beta > 1$. The values of β, σ_0 and u_c characterize the compressibility of the sediment formed by a given material.

Values of the primitive $A(u)$ usually have to be determined by numerical quadrature. However, if f and a are given by (3)–(4) and β is an integer, then $A(u)$ can be evaluated in closed form by $A(u) = 0$ for $0 \le u \le u_c$ (equation (1a) is strongly degenerate) and $A(u) = \boldsymbol{A}(u) - \boldsymbol{A}(u_c)$ for $u > u_c$, where the function $\boldsymbol{A}$ is defined by

$$\boldsymbol{A}(u) := \frac{v_\infty \sigma_0}{\Delta_\rho g u_c^\beta u_{\max}^C} \times \sum_{k=1}^{\beta}\left(\prod_{l=1}^{k}\frac{\beta+1-l}{C+l}\right)(u_{\max} - u)^{C+k}u^{\beta-k}.$$

2.2. Discrete mollification

The discrete mollification method [17, 18] consists in replacing a set of data $y = \{y_j\}_{j\in\mathbb{Z}}$ by its mollified version $J_\eta y$, where J_η is the *discrete mollification operator* defined by

$$[J_\eta y]_j := \sum_{i=-\eta}^{\eta} \omega_i y_{j-i}, \; j \in \mathbb{Z}.$$

The support parameter $\eta \in \mathbb{N}$ indicates the width of the mollification stencil, and the weights ω_i satisfy $\omega_i = \omega_{-i}$ and $0 \le \omega_i \le \omega_{i-1}$ for $i = 1, \ldots, \eta$ along with $\omega_{-\eta} + \cdots + \omega_{\eta-1} + \omega_\eta = 1$. The weights ω_i are obtained by numerical integration of a suitable truncated Gaussian kernel. Details can be found in [1, 2, 3, 16, 20].

2.3. Discretization of the direct problem

The domain Ω_T is discretized by a standard Cartesian grid by setting $x_j := j\Delta x, j = 0, \ldots, \mathcal{N}$, where $\mathcal{N}\Delta x = L$ and $t_n := n\Delta t, n = 0, \ldots, \mathcal{M}$, where $\mathcal{M}\Delta t = T$.

We denote by u_j^n an approximate value of the cell average of $u = u(x, t)$ over the cell $[x_j, x_{j+1}]$ at time $t = t_n$ and correspondingly set

$$u_j^0 = \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u_0(x)\,dx, \;\; j = 0, \ldots, \mathcal{N} - 1.$$

We solve (1) numerically using two convergent finite difference methods. The first one [8, 13] has the following form, where $\lambda := \Delta t/\Delta x$ and $\mu := \Delta t/\Delta x^2$:

$$u_j^{n+1} = u_j^n - \lambda\Delta_+ F^{EO}\left(u_{j-1}^n, u_j^n\right) + \mu\left(A\left(u_{j+1}^n\right) - 2A\left(u_j^n\right) + A\left(u_{j-1}^n\right)\right). \quad (5)$$

Here F^{EO} stands for the well-known Engquist-Osher numerical flux [12], and Δ_+ denotes the standard forward difference operator. Scheme (5) is monotone and convergent under the CFL condition

$$\lambda\|f'\|_\infty + 2\mu\|a\|_\infty \le 1. \quad (6)$$

The second finite difference method is the mollified scheme [2], which is also monotone and convergent and takes the form

$$u_j^{n+1} = u_j^n - \lambda\Delta_+ F^{EO}\left(u_{j-1}^n, u_j^n\right) + 2\mu C_\eta\left([J_\eta A(u_j^n)]_j - A\left(u_j^n\right)\right), \quad (7)$$

with

$$C_\eta := \left[\sum_{j=-\eta}^{\eta} j^2 \omega_{-j}\right]^{-1}.$$

This is an explicit method and has the convenient CFL condition

$$\lambda \|f'\|_\infty + 2\mu\varepsilon_\eta \|a\|_\infty \leq 1, \tag{8}$$

where $\varepsilon_\eta < 1$. (For the particular mollification weights considered herein, we obtain $\varepsilon_3 = 0.7130, \varepsilon_5 = 0.3969 \; and \; \varepsilon_8 = 0.1960$.) Clearly, condition (8) is more favorable than (6) since it shows that for a given value of Δx, mollified schemes may proceed by larger time steps. See [2] for more details on this scheme.

3. SENSITIVITY AND STABILITY ANALYSIS

3.1. Parameter identification problem

The inverse problem can be formulated as follows: given observation data $u^{obs}(x)$ at the final time $T > 0$ and functions u_0, ψ_0 and ψ_L, find the flux f and the diffusion function a such that the entropy solution $u(x, T)$ of problem (1) is as close as possible to $u^{\text{obs}}(x)$ in some suitable norm. The inverse problem is solved by minimizing the cost function

$$J(u(.,T)) = \frac{1}{2}\int_0^L \left|u(x,T) - u^{\text{obs}}(x)\right|^2 dx. \tag{9}$$

Since the functions f and a depend on a vector of parameters, the inverse problem corresponds to the following parameter identification problem:

$$Minimize\, J(\boldsymbol{p})\; w.r.t.\, parameter\; vector\; \boldsymbol{p}. \quad (PI)$$

The functions f and a are associated to the current parameter vector $\boldsymbol{p}$.

We define the piecewise constant function u^Δ by $u^\Delta(x,t) = u_j^n$ for $x \in [x_j, x_{j+1})$ and $t \in [t_n, t_{n+1})$ for $j = 0,\ldots, \mathcal{N}-1$ and $n = 0,\ldots, \mathcal{M}-1$ and replace u^{obs} by a piecewise constant function $u^{\text{obs},\Delta}$ formed by cell averages as follows:

$$u^{\text{obs},\Delta}(x) = u_j^{\text{obs}} = \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u^{\text{obs}}(x)\, dx \quad \text{for } x \epsilon [x_j, x_{j+1}),$$

where $j = 0,\ldots, \mathcal{N}-1$. The parameter dependent cost function is

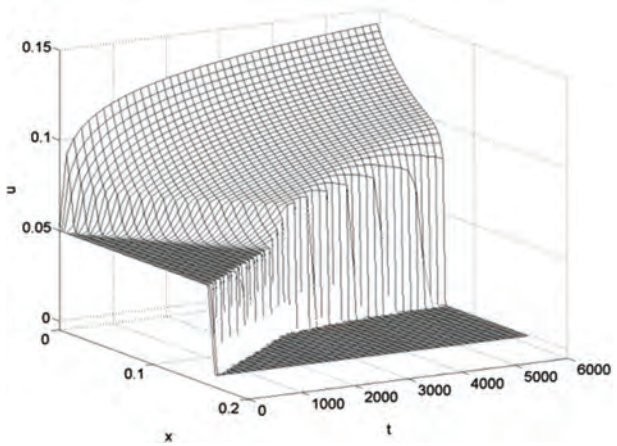


Figure 1. Reference solution

$$J^\Delta(\boldsymbol{p}) = \frac{1}{2}\int_0^L \left|u^\Delta(x,T) - u^{\text{obs},\Delta}(x)\right|^2 dx = \frac{\Delta x}{2}\sum_{j=0}^{\mathcal{N}-1}\left|u_j^{\mathcal{M}} - u_j^{\text{obs}}\right|^2. \tag{10}$$

This yields a discrete version of (PI) given by

$$\begin{cases} minimize\, J^\Delta(\boldsymbol{p})\; w.r.t.\; \boldsymbol{p} \\ u^\Delta, numerical\; solution\; of\; (1) \\ f\; and\; a\; associated\; to\; current\; \boldsymbol{p}. \end{cases} \quad (PI^\Delta)$$

There are many options for the numerical implementation of the optimization procedure. We selected a globalized bounded Nelder-Mead Method with restarts (MATLAB function *fminsearch*, see [15] for details), which is a major improvement over the basic simplex method. The strategy is described by the following algorithm. Suppose $\boldsymbol{p}_j = (p_j^1, \ldots, p_j^K)$, that is, K different parameters are sought.

Step 1 Input $\boldsymbol{p}_0, \epsilon$

Step 2 for $j = 1\; to\; M$

$\boldsymbol{p}_j = \texttt{fminsearch}\,(J^\Delta, \boldsymbol{p}_{j-1})$.

If $\max_{1\leq k\leq K}\left|\frac{p_j^k - p_{j-1}^k}{p_{j-1}^k}\right| \leq \epsilon$ then break, end

Step 3 End.

3.2. Numerical examples

The reference solution is generated by the corresponding numerical scheme (5) or (7) on a very fine grid. For examples 1, 2 and 3 we consider batch settling in a column of height $L = 0.16$ m and parameter values

$u_{\max} = 0.5, g = 9.81\text{m/s}^2, \rho_s - \rho_f = 1660\ \text{kg/m}^3, v_\infty = -2.7 \times 10^{-4}\ \text{m/s}$, $C = 21.5$, $\beta = 5, u_c = 0.07$ and $\sigma_0 = 1.2\text{Pa}$.

The objective is to obtain an accurate identification of the parameters u_c σ_0 and C in eight different instances described in Table 1. Our experiments include clean and noisy observation data. Data at the instant $T = 800\ \text{s}$ will play the role of u^{obs}. Figures 1 and 2 (a) show the reference solution over the whole computational domain and the profile at $T = 800\ \text{s}$, respectively. The restarting parameter and the tolerance parameter for the optimization are $M = 10$ and $\epsilon = 10^{-4}$, respectively.

Example 1: Sensitivity to mollification parameters. Clean observation data (no noise added) and $\Delta x = L/256$. The results are summarized in Table 2. Here, j denotes the number of calls to the `fminsearch` algorithm, $\boldsymbol{p_j}$ is the vector of parameter values found, E_J is the required number of computed solutions of the direct problem, e_∞ is the maximum relative error in the result for each parameter (usually due to σ_0), and CPU denotes the total CPU time of each run.

Example 2: Sensitivity to initial guess. We randomly generate 100 initial guesses and carry out the identification task. Each initial guess $\boldsymbol{p_0} = (u_c^0, \sigma_0^0, C^0)^T$ is generated in the form

$$u_c^0 = (1 + 0.3\xi_1)u_c, \sigma_0^0 = (1 + 0.3\xi_2)\sigma_0,$$
$$C^0 = (1 + 0.3\xi_3)C,$$

where $\xi = (\xi_1, \xi_2, \xi_3)^T \in \mathbb{R}^3$ is a uniformly distributed random vectorial variable whose components are between −1 and 1. The results are indicated in Table 3. Here, the average $\bar{e}_\infty$ of e_∞ and its standard deviation σ are included. Additionally, column "# restarts" stands for the number of calls of `fminsearch` and E_J for the number of solutions of the direct problem.

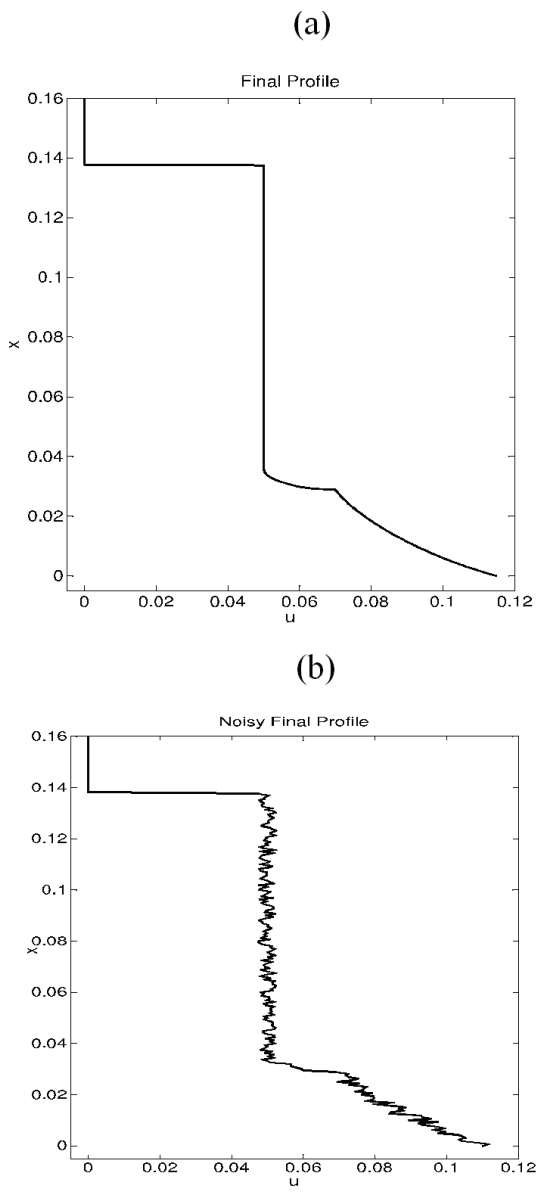


Figure 2. Profiles u^{obs} for Examples 1 and 2 (clean data) and Example 3 (noisy data $\epsilon = 0.05$) respectively

Example 3: Effect of noisy observation data. We randomly generate 100 final profiles and associate them to the previously generated initial guesses. The corrupted profile is generated as follows:

$$u_j^\varepsilon = \left(1 + \varepsilon\varphi_j\right)u^{\text{obs}}\left(x_j\right), \quad j = 0, \ldots, \mathcal{N} - 1,$$

where $\varepsilon = 0.01, 0.03$ and 0.05, and φ_j is a uniformly distributed random variable assuming values between −1 and 1. The results are in Table 4.

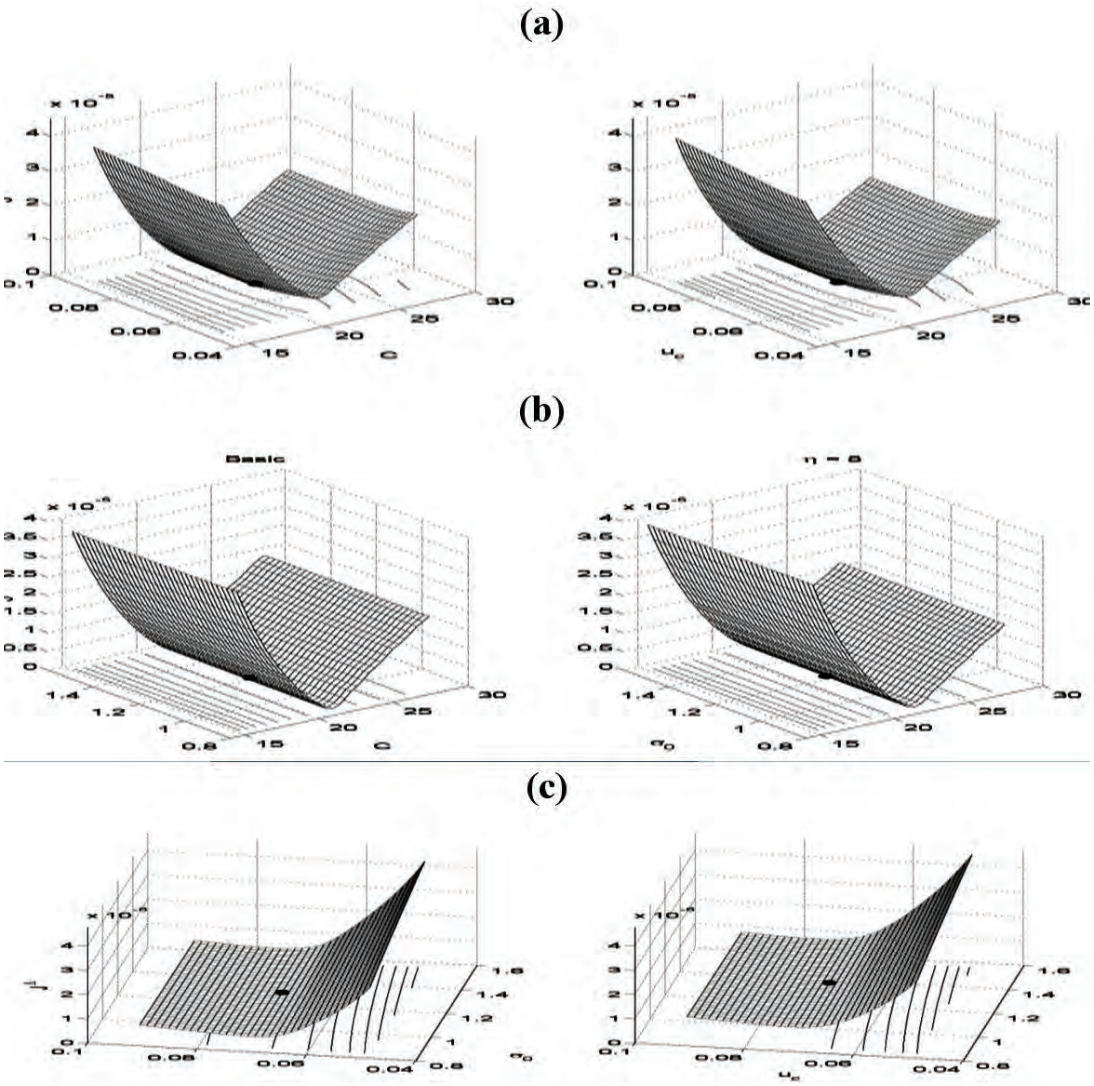


Figure 3. Two parameter cost functionals for parameter sets (a) (C, u_c), (b) (C, σ_0) and (c) (u_c, σ_0).

Table 1. Example 1: initial guesses used for identification experiments.

initial guess	parameter values	initial guess	parameter values
A	$(0.7u_C, 0.7\sigma_0, 0.7C)$	E	$(1.3u_C, 0.7\sigma_0, 0.7C)$
B	$(0.7u_C, 0.7\sigma_0, 1.3C)$	F	$(1.3u_C, 0.7\sigma_0, 1.3C)$
C	$(0.7u_C, 1.3\sigma_0, 0.7C)$	G	$(1.3u_C, 1.3\sigma_0, 0.7C)$
D	$(0.7u_C, 1.3\sigma_0, 1.3C)$	H	$(1.3u_C, 1.3\sigma_0, 1.3C)$

Table 2. Example 1: Results for the basic scheme (5) and the mollified scheme (7) with $\eta = 3$, 5 and 8.

	IG	j	p_j	E_j	e_∞	CPU [s]
Basic scheme (5)	A	2	(0.0697, 1.1219, 21.4706)	290	0.0651	79.863
	B	4	(0.0696, 1.1111, 21.4700)	517	0.0741	90.310
	C	5	(0.0697, 1.1324, 21.4700)	616	0.0564	163.74
	D	3	(0.0696, 1.1252, 21.4699)	669	0.0623	127.96

Continuation Table 2.

	IG	j	p_j	E_j	e_∞	CPU [s]
	E	3	(0.0696, 1.1179, 21.4700)	367	0.0684	59.662
	F	6	(0.0696, 1.1117, 21.4700)	674	0.0736	108.77
	G	3	(0.0696, 1.1114, 21.4700)	421	0.0738	71.391
	H	2	(0.0696, 1.1180, 21.4700)	391	0.0684	64.747
Mollified scheme (7)	A	2	(0.0695, 1.1104, 21.5067)	356	0.0746	89.43
with $\eta = 3$	B	4	(0.0695, 1.1105, 21.5067)	490	0.0746	89.77
	C	2	(0.0695, 1.1104, 21.5067)	296	0.0747	103.1
	D	5	(0.0695, 1.1103, 21.5067)	797	0.0747	155.8
	E	3	(0.0696, 1.1259, 21.5075)	332	0.0617	58.19
	F	3	(0.0696, 1.1294, 21.5066)	358	0.0588	61.25
	G	3	(0.0695, 1.1104, 21.5068)	402	0.0747	73.08
	H	5	(0.0695, 1.1104, 21.5065)	676	0.0747	119.7
Mollified scheme (7)	A	5	(0.0500, 0.1486, 21.5456)	685	0.8762	152.8
with $\eta = 5$	B	5	(0.0697, 1.1651, 21.5465)	664	0.0291	105.8
	C	3	(0.0695, 1.1124, 21.5466)	486	0.0730	103.6
	D	3	(0.0696, 1.1301, 21.5465)	639	0.0582	107.1
	E	3	(0.0696, 1.1301, 21.5466)	445	0.0583	69.26
	F	6	(0.0695, 1.0945, 21.5466)	697	0.0879	107.7
	G	3	(0.0696, 1.1299, 21.5466)	485	0.0584	78.23
	H	4	(0.0696, 1.1472, 21.5465)	807	0.0440	128.4
Mollified scheme (7)	A	2	(0.0696, 1.1174, 21.5777)	293	0.0688	49.008
with $\eta = 8$	B	3	(0.0697, 1.1530, 21.5776)	461	0.0391	64.792
	C	4	(0.0697, 1.1531, 21.5776)	649	0.0391	107.35
	D	2	(0.0697, 1.1530, 21.5777)	418	0.0391	60.759
	E	3	(0.0697, 1.1531, 21.5776)	535	0.0391	74.026
	F	3	(0.0696, 1.1174, 21.5777)	543	0.0688	73.858
	G	4	(0.0697, 1.1531, 21.5776)	576	0.0391	80.604
	H	2	(0.0697, 1.1531, 21.5776)	640	0.0391	89.225

Table 3. Example 2: Results for the basic scheme (5) and the mollified scheme (7) for different values of Δx and η.

$\Delta x/L$	Scheme	# restarts	E_j	$e_{\infty \pm \sigma}$	CPU[s]
1/128	(5)	387	53693	0.1400 ± 0.0329	49.17
	(7), $\eta = 3$	340	49274	0.1295 ± 0.0338	53.92
	(7), $\eta = 5$	333	50659	0.0909 ± 0.0315	50.91
	(7), $\eta = 8$	276	41263	$0.0534 \pm 3.72\text{e-}05$	39.36
1/256	(5)	388	48842	0.0696 ± 0.0096	136.03
	(7), $\eta = 3$	398	53513	0.0739 ± 0.0081	160.74
	(7), $\eta = 5$	355	48280	0.0587 ± 0.0174	125.46
	(7), $\eta = 8$	326	48392	0.0431 ± 0.0190	114.35
1/512	(5)	334	38879	0.0312 ± 0.0038	440.91
	(7), $\eta = 3$	352	40416	0.0332 ± 0.0049	439.55
	(7), $\eta = 5$	346	41128	0.0278 ± 0.0057	352.47
	(7), $\eta = 8$	374	47250	0.0195 ± 0.0072	335.24

Table 4. Example 3, $\Delta x = L/256$: results for the basic scheme (5) and the mollified scheme (7) for different values of ε and η.

ε	Scheme	# restarts	EJ	$\bar{e}_{\infty} \pm \sigma$	CPU[s]
0.01	(5)	385	50310	0.0661 ± 0.0264	140.63
	(7), $\eta = 3$	367	50489	0.0674 ± 0.0238	152.60
	(7), $\eta = 5$	368	52214	0.0533 ± 0.0256	135.85
	(7), $\eta = 8$	331	47153	0.0401 ± 0.0260	110.56
0.03	(5)	376	49025	0.0801 ± 0.0554	137.65
	(7), $\eta = 3$	364	48447	0.0807 ± 0.0533	146.72
	(7), $\eta = 5$	374	50622	0.0760 ± 0.0528	132.10
	(7), $\eta = 8$	353	49136	0.0710 ± 0.0524	115.31
0.05	(5)	378	49716	0.1163 ± 0.0913	140.31
	(7), $\eta = 3$	376	51246	0.1149 ± 0.0897	156.50
	(7), $\eta = 5$	359	50344	0.1129 ± 0.0902	131.10
	(7), $\eta = 8$	332	49139	0.1084 ± 0.0915	115.54

3.3. Conclusions

According to Table 2, most of the identifications are successful and $\eta = 8$ seems to be the best choice. For the initial guess A the method for $\eta = 5$ does not converge, but it does converge when started with initial guesses close to A. The results in Table 3, corresponding to Example 2, illustrate how by improving the spatial resolution (i.e., reducing Δx) the quality of the identification is increased.

Table 4 indicates that the level of noise influences the quality of the recovery but stability is never lost.

Summarizing, this parameter identification procedure yields good results for both the basic scheme and its mollified versions but the mollified approach returned advantages not only in CPU time (in s), but also in the error level, the sensitivity to the initial guess and the effect of noise in the data. This well-posed behavior was already suggested by the convex-shape of the cost functional (Figure 3.)

ACKNOWLEDGEMENTS

CDA and CEM acknowledge support by COLCIENCIAS project 1118-48925120 and by Universidad Nacional de Colombia, Vicerrectoría de Investigaciones, Projects 20201007076, 201010011107 and 20101009545. RB acknowledges support by Conicyt Anillo project ACT 1118 (ANANUM), Fondecyt project 1090456, BASAL project, CMM at Universidad de Chile, Centro de Investigación en Ingeniería Matemática (CI^2MA) at Universidad de Concepción and Red Doctoral REDOC.CTA, project UCO 1202 at U. de Concepción.

REFERENCES

[1] Acosta, C.D., and Bürger, R. Difference schemes stabilized by discrete mollification for degenerate parabolic equations in two space dimensions. *IMA J. Numer. Anal.*, 32: pp.1509–1540, 2012.

[2] Acosta, C.D., Bürger, R. and Mejía, C.E.. Monotone difference schemes stabilized by discrete mollification for strongly degenerate parabolic equations. *Numer. Meth. Partial Diff. Eqns*., 28: pp 38–62, 2012.

[3] Acosta C.D. and Mejía, C.E. Stabilization of explicit methods for convection-diffusion problems by discrete mollification. *Comput. Math. Applic.*, 55:pp 363.–380, 2008.

[4] Berres, S., Bürger, R. and Garcés, R. Centrifugal settling of flocculated suspensions: a sensitivity analysis of parametric model functions. *Drying Technology*, 28: pp.858–870, 2010.

[5] Bürger, R., Coronel, A. and Sepúlveda, M. Numerical solution of an inverse problem for a scalar conservation law modelling sedimentation. In: Tadmor, E., Liu, J-G., Tzavaras, A.E. (eds.): Hyperbolic Problems: Theory, Numerics and Applications. Proceedings of Symposia in Applied

Mathematics, vol. 67, Part 2. American Mathematical Society, Providence, RI, USA, pp.445–454 (2009)

[6] Bürger, R. and Diehl, S. Convexity-preserving flux identification for scalar conservation laws modelling sedimentation. *Inverse Problems* 29 (2013) paper 045008 (30pp).

[7] Bürger, R. Evje, S. and Karlsen. K.H. On strongly degenerate convection-diffusion problems modeling sedimentation-consolidation processes. *J. Math. Anal. Appl.*, 247:pp. 517–556, 2000.

[8] Bürger, R. and Karlsen, K.H. On some upwind schemes for the phenomenological sedimentation-consolidation model. *J. Eng. Math*., 41: pp.145–166, 2001.

[9] Bürger, R. and Karlsen, K.H. On a diffusively corrected kinematic-wave traffic model with changing road surface conditions. *Math. Models Methods Appl. Sci.*, 13:pp. 1767–1799, 2003.

[10] Bürger , R., Karlsen, K.H. and Towers, J.D. A mathematical model of continuous sedimentation of flocculated suspensions in clarifier-thickener units. *SIAM J. Appl. Math*., 65:pp. 882–940, 2005.

[11] Coronel, A., James, F. and Sepúlveda, M. Numerical identification of parameters for a model of sedimentation processes. *Inverse Problems*, 19: pp. 951–972, 2003.

[12] Engquist, B. and Osher, S. One-sided difference approximations for nonlinear conservation laws. *Math. Comp*., 36:pp. 321–351, 1981.

[13] Evje, S. and Karlsen, K.H. Monotone difference approximations of BV solutions to degenerate convection-diffusion equations. *SIAM J. Numer. Anal*., 37: pp. 1838–1860, 2000.

[14] Grassia, P., Usher, S.P. and Scales, P.J. Closed-form solutions for batch settling height from model settling flux functions. *Chem. Eng. Sci*., 66: pp. 964–972, 2011.

[15] Luersen , M.A. and Le Riche, R. Globalized Nelder–Mead method for engineering optimization. *Computers and Structures*, 82: pp. 2251–2260, 2004.

[16] Mejía, C.E., Acosta, C.D., and Saleme, K.I. Numerical identification of a nonlinear diffusion coefficient by discrete mollification. *Comput. Math. Applic*., 62:pp. 2187–2199, 2011.

[17] Murio, D.A. *The Mollification Method and the Numerical Solution of Ill-Posed Problems*. John Wiley, New York, 1993.

[18] Murio, D.A. Mollification and space marching. In: K. Woodbury (ed.), Inverse Engineering Handbook. CRC Press, Boca Raton (2002).

[19] Petzold, L., Li, S., Cao, Y., and Serban, R. Sensitivity analysis of differential-algebraic equations and partial differential equations. *Computers and Chemical Engineering* 30: pp. 1553–1559, 2006.

[20] Pulgarín, J., Acosta, C., and Castellanos,G. Multiscale analysis by means of discrete mollification for ECG noise reduction. *DYNA* 159: pp. 185–191, 2009.

[21] Zhen, Z., Zhichao,W., Guowei, G., and Zhiwei.,W. Parameter estimation protocol for secondary clarifier models based on sludge volume index and operational parameters. *Asia-Pac. J. Chem. Eng.*, 6: pp. 266–273, 2011.

E-MAINTENANCE PLATFORM: A BUSINESS PROCESS MODELLING APPROACH

MODELADO DE UNA PLATAFORMA DE E-MANTENIMIENTO, CMMS Y CBM

MÓNICA LÓPEZ-CAMPOS
Ph.D. School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, mlopezcampos@us.es
SALVATORE CANNELLA
Ph.D. Industrial Engineering, CEG-IST, Technical University of Lisbon, Portugal, salvatore.cannella@ist.utl.pt
MANFREDI BRUCCOLERI
Ph.D. Production Engineering, DICGIM, University of Palermo, Italy, manfredi.bruccoleri@unipa.it

Received for review May 5 th, 2011, accepted December 11th, 2012, final version April, 10 th, 2013

ABSTRACT: The information and communications technologies (ICT) are increasingly used in society and business. ICT have become a key element for business development and the improvement of activities in all departments. Analogously, the rise of Computerized Maintenance Management Systems (CMMS), Condition Based Maintenance (CBM), and the use of remote devices and functions such as e-maintenance, have opened great possibilities for maintenance optimization. Those techniques have resulted in the emergence of technological challenges such as interoperability between different systems in order to exchange and coordinate information. In this context, the application of business process modelling is a crucial methodology to understand and implement those maintenance techniques. In this paper, we use business process modelling approach to design a system architecture that integrates CMMS, CBM and e-maintenance methods, tools, and processes. The data model is based on the MIMOSA standard.

Keywords: CMMS, CBM, E-Maintenance, UML, MIMOSA.

RESUMEN: Las tecnologías de la información y la comunicación (TIC) se utilizan cada vez más en la sociedad y en los negocios. Las TIC se han convertido en un elemento clave para el desarrollo empresarial y la mejora de las actividades en todos los departamentos de una organización. Análogamente, la creciente importancia de los Sistemas de Gestión de Mantenimiento asistido por Ordenador (CMMS), Mantenimiento Basado en la Condición (CBM), y la aplicación de dispositivos remotos y funciones como el e-mantenimiento, han abierto grandes posibilidades para la optimización del mantenimiento. Estas técnicas han dado lugar a la aparición de retos tecnológicos, tales como la interoperabilidad entre diferentes sistemas para intercambiar y coordinar la información. En este contexto, la aplicación del modelado de procesos de negocio es una metodología fundamental para entender y aplicar las mencionadas técnicas de mantenimiento. En este artículo se presenta una arquitectura para mostrar la forma en que el modelado de procesos de negocio permite la implementación de un sistema que integra CMMS, CBM y el e-mantenimiento. El modelo de datos se describe de acuerdo con la estructura de base de datos de MIMOSA.

Palabras clave: CMMS, CBM, e-mantenimiento, UML, MIMOSA.

1. INTRODUCTION

The information and communications technologies (ICT) are a key pillar in an organization, due their capabilities of supporting and facilitating maintenance management [1, 2].

The application of ICT in maintenance emerged in the 1970's, when the first personal computers appeared. At that time computer systems were limited to performing only certain scheduling tasks for preventive actions [3]. Over time, several technologies have been evolving and with them, the integration of maintenance systems has grown. In particular, three main applications have been identified as novel techniques for an effective maintenance operation: (i) Computerized Maintenance Management Systems (CMMS), (ii) Condition Based Maintenance (CBM) and (iii) e-maintenance processes.

A CMMS is a software tool for maintenance management that incorporates open access to information that facilitates the decision making process regarding maintenance activities [4]. CBM is a technique used to monitor machines for the purpose of diagnostics and prognostics [5]. E-maintenance is the set of maintenance processes that uses the e-technologies to

enable proactive decisions in a particular organization (definition partially derived from Levrat et al. [6]).

The implementation of these techniques, tools and processes on the shop floor is not a trivial issue. Overall, integrating them into a single software platform is the most challenging concern. Therefore several academics and professionals underline the need for a structured approach to support their actual implementation, use, and integration [7, 8]. In this paper we use business process modelling to develop such a structured approach. Specifically we use Business Process Modelling Notation (BPMN) to represent the flow of activities that are required for an implementation. Secondly, we use the Unified Modelling Language (UML) to represent the principal objects constituting the e-maintenance platform, their relations and static operation structure. Finally, the variety of technologies related to the proposed platform, implies the existence of multiple communication protocols, data connections, configurations, etc. In this respect, we refer to the Maintenance Information Management Open System Alliance (MIMOSA) standard as a useful reference.

The paper is organized as follows: Section 2 describes the concept of CMMS, CBM and e-Maintenance. Section 3 contains a literature review and problem statement. Section 4 briefly introduces the business process modelling techniques we use in this paper. Section 5 shows the system architecture for the integration of CMMS, CBM, e-Maintenance in the emaintenance integration platform. Finally, Section 6 provides conclusions and future directions.

2. CMMS, CBM and E-MAINTENANCE

The following three subsections detail the most important characteristic of the aforementioned maintenance techniques.

2.1. CMMS

The main objective of a CMMS is to provide a tool to analyze maintenance and equipment information in order to optimize the management and support for strategic, tactical and operational decisions. An important requirement for these systems is the compatibility and integration with other systems that contains organizational information. Fumagalli et al. [9] indicate CMMS as a success factor for the organization of maintenance programs. CMMS allows access to information that can be used to prioritize actions and make better decisions in maintenance activities. The CMMS can also provide adequate control of the system. According to Crain [10] incorporating a complete management system in the maintenance function can reduce 10-30% annual budget of the maintenance department. Ros [11] identifies the top five most implemented CMMS systems around the world: SAP Plant Maintenance, Maximo Asset Management, MP2, Ellipse and PMC.

2.2. CBM

CBM is defined as the preventive maintenance based on monitoring the operation, the parameters of the element, and the subsequently executed actions [12]. Monitoring is the manual or automatic task of observing the current state and behaviour of a given shop floor element (commonly the thermal and vibratory behaviour) [13]. The final goal of the monitoring task is to adopt a condition-based maintenance approach in order to reduce the probability of failure and prevent irreversible damage to the equipment. CBM allows maximizing the service life of equipment and its components, anticipating failures by monitoring significant parameters. For these reasons, CBM is a key component of proactive maintenance. In recent years, both the reduction of the price of sensors and the increase of information processing capacity have enhanced the development of automatic diagnostic tools. In this context, one of the most widespread applications is the e-CBM, as it is technically feasible, economically viable, and offers many benefits. More specifically, it significantly reduces downtime, saves up to 20% in small production losses, improves quality and reduces stocks of spare parts. [14].

2.3. E-Maintenance

Although the term e-Maintenance has been used since 2000 as a component of eManufacturing, at the present time there is not yet a standardized definition of eMaintenance given by an official institution. We may say that e-Maintenance is "the set of maintenance processes that uses the e-technologies to enable proactive decisions in a particular organization" (definition partially derived from Levrat et al. [6]). Those maintenance processes and methods required to operate them (technologies, procedures, standards,

etc) are specific to each organization according to its particular characteristics. This means that, for different organizations, e-Maintenance could involve different spheres of activities. Frequently, these activities include: e-monitoring, e-prognosis, e-diagnosis, e-management, e-service, remote maintenance and, collaborative maintenance. Such activities are supported by means of a variety of hardware and software technologies such as wireless and mobile devices, embedded systems, web based applications, P2P networks, multi-agent applications, specific software architectures, communication protocols, among others.

3. LITERATURE REVIEW AND PROBLEM SCOPE

We performed a literature review to recognise similar previous attempts at integration of CMMS, CBM and e-maintenance on a single system. The search was specially focused to find projects using modelling languages such as BPMN, UML and/or interoperability standards such as MIMOSA. Specifically, we found 10 studies focused on these issues. A list of these papers is reported in Table 1. In this Table we also reported a summary of the specific problem they deal with and the approach and methodology they use to solve that problem.

Although the referred papers focus on CMMS, CBM, e-maintenance, none of those works clearly propose a structured approach to integrate all of these tools and techniques in a unique platform. This literature review underlines that *there is a lack of models* that clearly show "which" specific data should be shared among CMMS, CBM and E-Maintenance tools and techniques and "how" these data should be used in order to create an integrated e-maintenance platform. Motivated by these observations, we wish to unambiguously show how to create a fully coordination mechanism combining CMMS, CBM and E-Maintenance. To this aim, we adopt a business process modelling approach and more specifically we use (*a*) BPMN (a standard for business process modelling) to design the process of integration of the CMMS, CBM and E-Maintenance components; (*b*) the UML class diagram (a standard for software system design) to design the main software components of the e-maintenance platform and their relationships; *(c)* MIMOSA for standardizing the interoperability of the platform. In the next section we give a brief overview of the business process modelling languages and standards.

4. BUSINESS PROCESS MODELLING

Business process modelling, means the use of methods, techniques and software to design, control and analyze operational processes involving humans, organizations, applications, documents and other sources of information [15]. Business process modelling has become an important subject especially since the 1990s, when companies were encouraged to think in terms of "processes" instead of "functions" [16].

From then on, business process modelling has been used in industry to obtain a global vision of processes by means of support, control and monitoring activities [17], to facilitate the comprehension of the business key mechanisms, to be a base for the creation of appropriate information systems [18], to improve the business structure and operations, etc. [19].

Table 1. Literature review

Year	Author	Title	Contribution/approach
1995	*Deb et al. [20]*	Multisignal Modeling for Diagnosis, FMECA, and Reliability	Developing of a complete solution package for Integrated Diagnostics (ID) It includes some reliability equations.
2001	*Finley, and Schneider [21]*	ICAS: the center of diagnostics and prognostics for the US Navy.	The US Navy chose to apply the Integrated Condition Assessment System (ICAS). One of its key requirements to enable CBM is the ability to track machinery performance and diagnose machinery health.
2006	Kothamasu et al. [22]	System health monitoring and prognostics - A review of current paradigms and practices	This paper reviews the philosophies and techniques that focus on improving reliability and reducing unscheduled downtime by monitoring and predicting machine health. It proposes the use of qualitative information, extracted from failure mode and effects analysis (FMEA) or fault tree analysis (FTA) into a quantitative analysis that generates diagnostic recommendations.

Continuation Table 1.

Year	Author	Title	Contribution/approach
2008	Gao et al. [23]	Implementation of OSA-CBM standard based on .NET remoting	An assembly for MIMOSA OSA-CBM standard is developed on the .NET platform. This paper shows that the integration method for CBM is feasible through a demo program.
2008	Nemeth et al. [24]	Complex diagnostic methods for lifetime extension of power transformers	It shows the main elements to perform complex diagnostic analysis in a transformer system. This paper presents a management system for monitoring and diagnosis.
2009	Wessels, and Sautter [8]	Reliability analysis required to determine CBM Condition Indicators	A reliability-centered analysis is presented that defines the method to correctly select the CBM as a technically and economically viable maintenance solution. This paper proposes a diagram to decide whether or not to implement CBM. It does not integrate CMMS.
2010	Li et al. [25]	A fast development framework for condition-based maintenance systems	In this paper, a framework is proposed to reduce the development complexity of CBM systems, enabling CBM modules independently constructed and dynamically integrated.
2011	Trappey et al. [26]	Maintenance chain integration using Petri-net enabled multiagent system modeling and implementation approach	This research proposes a collaborative environment integrated by a service center of diagnosis, prognosis, and asset operations. To realize the automation of communication and negotiation, multi-agent systems, Petri-net modeling and UML are applied.
2012	Galar et al. [27]	Maintenance decision making based on different types of data fusion	This paper proposes a combined data mining-based methodology to collect useful data from a CMMS to be used in CBM processes.
2012	Chang et al. [28]	Development of an e-operation framework for SoPC-based reconfigurable applications	It presents a conceptual framework for e-operations with the capability of e-maintenance, integrating e-diagnostic, among other processes. UML is used as the tool to accomplish the object-oriented analysis and to design the system.

Several benefits deriving from the adoption of business process modelling have been identified in the literature: improvement of the speed of business processes, increase of the clients' satisfaction, optimization and elimination of unnecessary tasks, and incorporation of clients and partners in the business processes [29].

Process modelling is the subject of interest in many different fields, such as the managerial area [30] and software engineering. This is due to the fact that it does not only describe processes, but in addition it represents a preparatory stage for the improvement of business processes, process reengineering, technological transference and processes standardization [31].

There are several standards for business process modelling. To model the e-maintenance platform proposed in this paper, we use the Unified Modelling Language (UML) and Business Process Modelling Notation (BPMN). Both standards are maintained by the OMG (Object Management Group). During the past years, the most used modelling standards were IDEF (Integration DEFinition), RAD (Role-Activity Diagrams), EXPRESS-G and STEP (Standard for the Exchange of Product model data). However, the use of UML and BPMN standards in modelling maintenance models, processes and software systems is largely increasing. Among other advantages of using BPMN and UML, the easy translation of their models into software code is probably the factor most responsible for their diffusion. [32].

5. A PROPOSED SYSTEM ARCHITECTURE

To obtain the best results of ICT in maintenance, we need to apply the

mentioned maintenance techniques in a coordinated manner. Each approach (CMMS, CBM and e-maintenance) is proved to provide benefits to

the maintenance management, improving at least the efficiency of the maintenance processes. This architecture show a maintenance system that combines the registers of CMMS with the information provided by sensors, providing a data model to allow the development of a system that combines the features of CMMS and CBM in an e-maintenance platform. In the next subsections we present the main e-maintenance process and the data model for the integration platform.

5.1. E-maintenance process

The general operation of the proposed system integration is shown in Figure 1. This Figure shows the activity flow of the CMMS—CBM integration expressed using the BPMN standard. Every maintenance element (CMMS and CBM) is represented by a pool division instead of swimlanes, since CBM and CMMS are two different informatics systems, each one using their own protocols. The tangible output documents are indicated by using the artifact notation. As already mentioned, the third element constituting the e-maintenance integrated platform is the e-maintenance process itself. Indeed, although e-maintenance can be viewed as a technique (or set of techniques), the general idea of our approach is to consider e-maintenance as the whole process of integrating all the maintenance components and elements. In accordance with [33] we consider e-maintenance a philosophy supporting the operation of the entire system and making information exchange possible among remote elements. This is why, there is no pool in the BPMN diagram representing this component, but, instead, it is represented by the flow of activities and

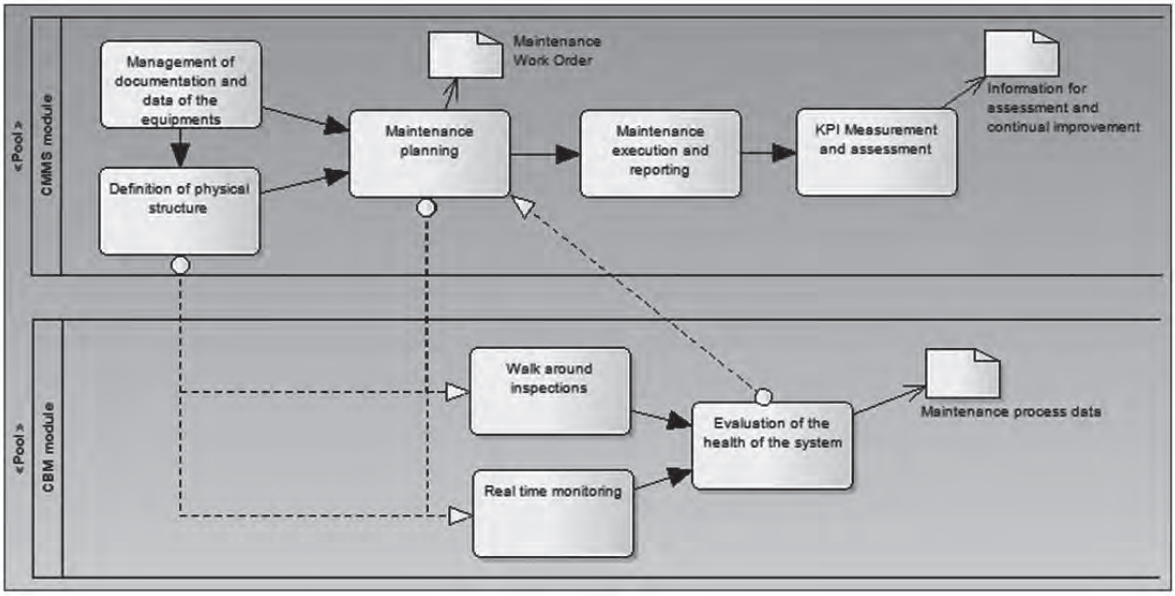


Figure 1. BPMN diagram of the integration system

the exchange of messages (dotted lines in Figure 1) among CMMS and CBM.

5.2. Data model

Figure 2 is a UML class diagram that shows the main software components, their relationships, and the exchanged information. In fact, the UML classes specified for the data model in Figure 2, have been inherited from MIMOSA standard in order to keep the data model compatible with the MIMOSA database structure and to guarantee interoperability.

For the sake of clarity, we distinguish class attributes into NEW and MIMOSA attributes. MIMOSA attributes come from simple elaboration and interpretation of the attributes already defined in the MIMOSA specification. On the contrary, NEW attributes are newly defined and have been added only when necessary. Analogously, we omitted from the model those attributes of the MIMOSA specification that are not useful in the scope of this work. Our experience in implementing the MIMOSA standards in real cases, confirmed that interpreting MIMOSA is not an easy subject. For a company, is quite difficult to match its information and knowledge within the MIMOSA specifications. For this

reason, the model proposed here is based on our elaboration and interpretation of MIMOSA. A short description now follows in order to help the reader navigate the model. The following classes come directly from the MIMOSA specification: Item, Function, MeasLoc, DataEvent, AlertRegion,

ItemRecommendation, ItemRequestForWork and HypotheticalEvent. Two classes have had their names changed with respect to MIMOSA: Alert and EventHistory. Alert represents an alarm. The MIMOSA specification also describes the concept of Alert by a larger number of classes. Here a simplification and synthesis is adopted. EventHistory is a new class introduced to describe the information that is normally stored in a CMMS as records of events happened in the past (like failures, preventive maintenance tasks, etc.); the specification is done by keeping a name conformed to the MIMOSA nomenclature. More in-depth, an item can have more than one MeasLoc (measurement location), i.e. more than one sensor. Hence, the software tool, needed to implement the integration between CMMS-CBM, should be able to automatically identify a MeasLoc, through its IP address and associate it to an Item through a user interface.

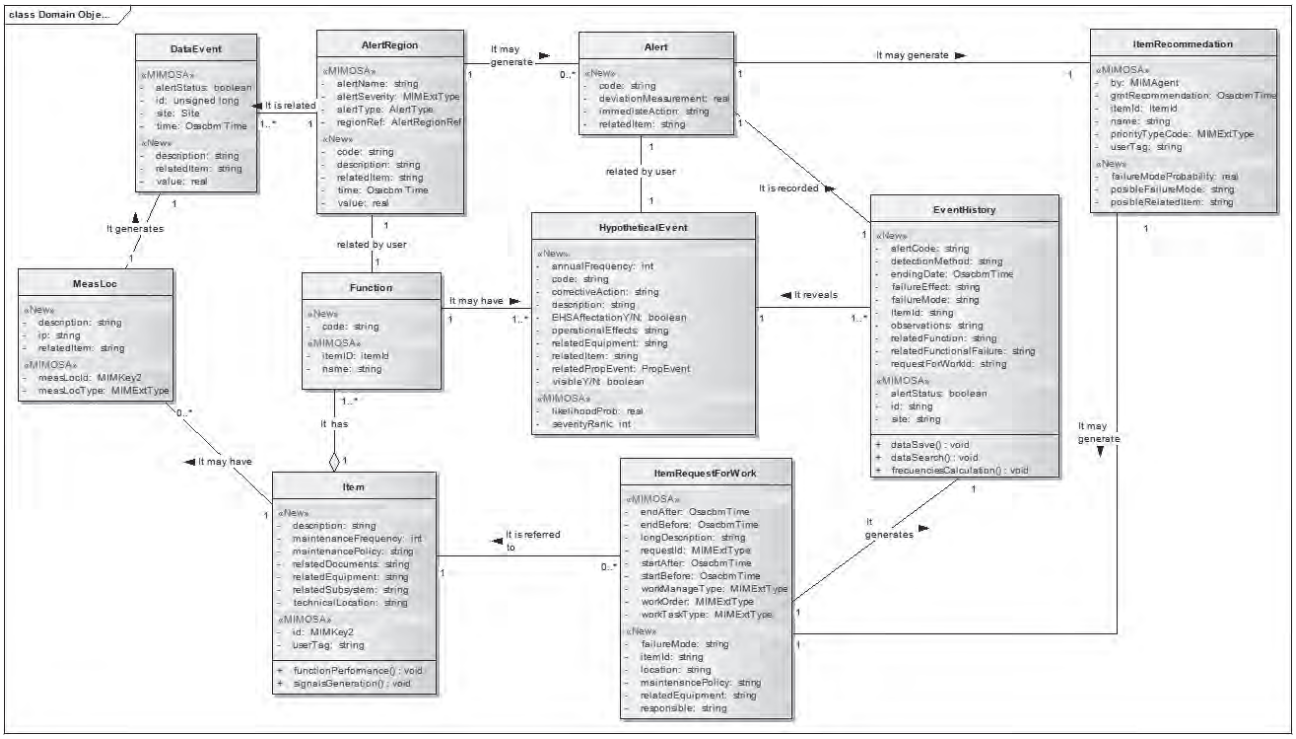


Figure 2. UML Data model of the integrated system

The main idea behind this concept is that common sensors are generally related to an IP. This IP could be the IP of the controller that connects more than one sensor and it is wired to them. There could be also a field-bus to which all the sensors are connected. In these two cases a sub-IP for each sensor can be identified. Moreover, due to new development of smart-sensors [34] it could be envisioned that each sensor is connected to the overall network through its own IP identifier. Following the chain of activities, the MeasLoc is then related with one data event, namely a signal. It is in fact established that a MeasLoc sensor can provide only one DataEvent so the relationship is one to one. If the reader is thinking about sensors providing multiple signals, it is worth remembering that one device may have multiple transducers; this should be considered when utilizing the proposed model . More precisely, in this project the MeasLoc is the soft entity used to identify the transducer in the scope of the integrated system. More than one DataEvent can then be related to an AlertRegion.

This defines the boundaries (upper and lower levels) in which a signal is considered normal; out of these bounds something wrong is occurring. One AlertRegion can then provide one or more Alerts. Following the described chain of relations we can state that one Item

can generate multiple Alerts and this seems correct. However, we should map one Alert to failure modes to be sure that the system can then properly work, in order to provide information about a specific failure coming from an Alert. The Failure mode is represented according to the interpretation of the MIMOSA definitions by means of an HypotheticalEvent class. This describes all the related events stored in the CMMS data base. One or more failure modes are then associated with a Function and an Item can have one or more functions. To be able to identify one single failure mode by an Alert enough information must be collected. It often happens that one alert is associated with multiple failure modes. So, in order to directly connect one alert to a specific failure mode, one solution is that there are enough monitored variables to assign one alert to each failure mode. In this way, the user is able to connect an Alert with each single failure mode, i.e. with a single instance of HypotheticalEvents class.

Besides mapping failure modes to alerts, the user could also be called to connect each failure mode (the HypotheticalEvent in the UML model) with records that represent events that occurred in the real system, related to the failure mode, either as a failure or as a preventive maintenance task planned to avoid the failure (this is the EventHistory in the UML model). When this last relation has been set up, the CMMS data-base can update in real time the occurrence of an event in the HypotheticalEvent and EventHistory classes.

6. CONCLUSIONS

This work proposes a new data model to allow the development of a platform combining the features of the three systems, so fostering the achievement of a synergy among them. More specifically, we show how business process modelling facilitates the process of implementing a system that integrates CMMS, CBM and e-maintenance. A BPMN model was applied to represent the whole process of integration. A UML Class diagram was adopted to show the main elements (and relations) of the e-maintenance integrated platform. MIMOSA was adopted as a reference for designing attributes in the data model.

Both researchers and professionals are expected to benefit from the presented research. Researchers can keep exploring different business process modelling techniques to further design e-maintenance process from different perspectives and also to include different maintenance systems and tools. However, the presented research has a strong practical vein. The proposed framework provides practitioners with a methodological approach that concretely support them in developing, implementing and integrating e-maintenance tools.

Future developments of this work should consider the validation of the model proposed by means of representative case studies.

ACKNOWLEDGEMENTS

This research was funded by a grant from the Portuguese Foundation for Science and Technology, grant SFRH/BPD/68576/2010

REFERENCES

[1] Khatib, A.R., Zuzhu-Dong, Bin-Qui, Yilu-Liu. Thoughts on future Internet based power system information network architecture, Proceedings of the Power Engineering Society Summer Meeting, Seattle, USA, 2000.

[2] Crespo, M.A., Moreu, De L.P., Sanchez, H.A., Ingeniería de Mantenimiento. Técnicas y Métodos de Aplicación a la Fase Operativa de los Equipos, Aenor, España, 2004.

[3] Kelly, A., Maintenance Planning and Control, Butterworths, London, 1984.

[4] López-Campos, M., Gómez-Fernández, J., Holgado-Granados, M., Crespo-Márquez, A., Aplicación de las tecnologías de la información y la comunicación al mantenimiento, Libro de Ponencias del XII Congreso de Confiabilidad, Madrid, Asociación Española para la Calidad, 2010.

[5] Bunks, C., Mccarthy, D. and Al-Ani T., Condition-based maintenance of machines using hidden Markov models. Mechanical Systems and Signal Processing, 14(4), pp. 597-612, 2000.

[6] Levrat, E., Iung, B., and Crespo- Marquez, A., E-maintenance: review and conceptual framework, Production Planning & Control , 19(4), pp. 408-429, 2008.

[7] Wessels, W. and Sautter, F., Reliability Analysis Required to Determine CBM Condition Indicators,

Reliability and Maintainability Symposium, RAMS, 2009.

[8] Niu, G., Yang, B.S. and Pecht, M., Development of an optimized condition-based maintenance system by data Fusion and reliability-centered maintenance, Reliability Engineering and System Safety, 95(7), pp.786-796, 2010.

[9] Fumagalli, L., Macchi, M. and Rapaccini, M., Computerized Maintenance Management Systems in SMEs: a survey in Italy and some remarks for the implementation of Condition Based Maintenance, 13th IFAC Symposium of Information Control Problems in Manufacturing, Moscow, 2009.

[10] Crain, M., The Role of CMMS, Industrial Technologies Northern Digital, Inc, 2003.

[11] Ros Moreno, A., Capítulo 33: Mantenimiento industrial. Software informático (1/2). Disponible en: http://www.mailxmail.com/curso-mantenimiento-industrial-2-3/mantenimiento-industrial-software-informatico-1-2 [consultado 17 de octubre 2010].

[12] UNE-EN 13306 Terminología de Mantenimiento. AENOR, Madrid, 2002.

[13] Aguiar, M., Gómez, J. y Torres, P., Modelamiento térmico y vibratorio de una cápsula para sensores de fibra óptica adaptables a mediciones de sistemas eléctricos de potencia, Dyna, 76(157), pp. 243-250, 2009.

[14] Djurdjanovic, D., Lee, J. and Ni, J., Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction, Advanced Engineering Informatics 17, pp.109-125, 2003.

[15] Van Der Aalst, W., Don´t go with the flow: Web services composition standards exposed. IEEE Intelligent Systems, 18(1), 2003.

[16] López Campos, M., Crespo Márquez, A., Modelling a maintenance management framework based on PAS 55 standard. Quality and Reliability Engineering Int, 27(6), pp. 805-820, 2011.

[17] Russel, N., Vanderalst, W., Hofstede, A. and Wohed, P., On the suitability of UML Activity Diagrams for Business Proces Modelling. Proceedings of the Third Asia-PAcific Conference on Conceptual Modelling (APCCM), vol. 53 of Conferences in Research and Practice Information Technologies, pp.195–104, 2006.

[18] Quiescenti, M., Bruccoleri, M., La Commare, U., Noto La Diega, S. and Perrone, G., Business process-oriented design of Enterprise Resource Planning (ERP) systems for small and medium enterprises. International Journal of Production Research, vol 44, no. 18-19, pp. 3797-3811, 2006.

[19] Beck, K., Joseph, J. and Goldszmidt, G., Learn Business Process Modeling Basics for the Analyst. Disponible en: www-128ibm.com/developersworks/library/wsbpm4analyst [consultado 17 de octubre 2010].

[20] Deb, S., Ghoshal, S., Mathur, R., Shrestha, Pattipati R., Multisignal modeling for diagnosis, FMECA and reliability, Proceedings of the IEEE Int Conference on Systems, Man and Cybernetics 3, pp. 3026-3031, 1998.

[21] Finley, B., Schneider, E., ICAS: the center of diagnostics and prognostics for the United States navy. Proceedings of the SPIE – Int. Soc. Opt. Eng. 4389, pp. 186-193, 2001.

[22] Kothamasu, R., Huang, S. and Derduin, W., System health monitoring and prognostics – a review of current paradigms and practices, Int J Adv Manuf Techol 28, pp.1012-1024, 2006.

[23] Gao, Z., Li, X., Pan, Q. and Zhang, L., Implementation of OSA-CBM standard based on .NET remoting, Chinese Journal of Scientific Instrument 29 (SUPPL. 2), pp. 123-126, 2008.

[24] Németh, B., Benyó, T., Jáger, A., Csépes, G. and Woynárovich, G., Complex diagnostic methods for lifetime extension of power transformers. Conference Record of IEEE Int Symposium on Electrical Insulation, art. no. 4570296, pp. 132-135, 2008.

[25] Li, L., Qian, Y., Du, K. and Yang, Y., A fast development framework for condition-based maintenance systems. ICMEE 2010 - 2010 2nd Int Conference on Mechanical and Electronics Engineering, Proceedings 2, art. no. 5558484, pp. V270-V274, 2010.

[26] Trappey, A.J.C., Hsiao, D.W. and Ma, L., Maintenance chain integration using Petri-net enabled multiagent system modeling and implementation approach, IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, vol. 41 (3), pp. 306-315, 2011.

[27] Galar, D., Gustafson, A., Tormos, B. and Berges, L., Maintenance decision making based on different types of data fusion, Eksploatacja i Niezawodnosc, vol. 14 (2), pp. 135-144, 2012.

[28] Chang, Y., Chen, Y., Hung, M., Chang, A.Y., Development of an e-operation framework for SoPC-based reconfigurable applications, Int Journal of Innovative Computing, Information and Control, vol. 8.(5 B), pp. 3639-3660, 2012.

[29] Pérez, J. et al. Model Driven Engineering Aplicado a Business Process Management. Informe Técnico UCLM-TSI-002, 2007.

[30] Ferreira, L., Del Conte, D. y Poppi, S., Aplicación de la UML como herramienta de auxilio en la gestión empresarial, Revista Avances en Sistemas e Informática, 7 (1), 2010.

[31] Succi, G. Predonzani, P. et al. Business Process Modeling with Objects, Costs and Human Resources. Systems Modeling for Business Process Improvement. Artech House, London, pp. 47-60, 2000.

[32] Muñetón, A. and Zapata, C., Definition of a semantic platform for automated code generation based on UML class diagrams and DSL semantic annotations, Dyna 79(172), pp. 94-100, 2012.

[33] Iung, B., Levrat, E., Crespo Márquez, A. and Erbe, H., Conceptual Framework for e-Maintenance: Illustration by e-Maintenance technologies and platforms. Annual Reviews in Control, 33, pp. 220-229, 2009.

[34] Garetti, M., Macchi, M., Terzi, S. and Fumagalli, L., Investigating the organizational business models of maintenance when adopting self diagnosing and self healing ICT systems in multi site contexts. Paper presented to the IFAC-CEA 07, Monterrey, Mexico, 2007.

LONG TERM PERFORMANCE OF EXISTING PORTLAND CEMENT CONCRETE PAVEMENT SECTIONS – CASE STUDY

DESEMPEÑO A LARGO PLAZO DE SECCIONES DE PAVIMENTOS EXISTENTES DE CONCRETO CON CEMENTO TIPO PORTLAND – ESTUDIO DE CASO

SHAD M. SARGAND
Ph.D., Civil Engineering, Ohio Research Institute for Transportation and Environment, Ohio Univ., Athens, esargand@ohio.edu
CARLOS A. VEGA-POSADA
Ph.D. Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia. e-mail: cvega@udea.edu.co
LUIS G. ARBOLEDA-MONSALVE
Ph.D. Candidate, Dept. of Civil and Environmental Engineering, Northwestern Univ., Evanston, larboleda@u.northwestern.edu

Received for review February 04 th, 2013, accepted October 8th, 2013, final version December, 04 th, 2013

ABSTRACT: This paper presents the performance of seventeen rigid pavements made of Portland Cement Concrete (PCC) located in the State of Ohio, USA. The Falling Weight Deflectometer (FWD) testing method was used in a total of 58 km of PCC pavement sections to assess the pavement structural condition. Each section is classified as excellent, good, fair, or poor based on the normalized deflections obtained from FWD tests, spreadability, load transfer, and joint support ratio parameters. The overall performance of the analyzed sections is from good to excellent. The field testing methodology along with the performance data analyses presented in this publication are a valuable tool to evaluate the actual structural condition of pavements and as a result short or medium term rehabilitation programs can be issued to ensure the serviceability of the pavement.

KEYWORDS: Pavement, Rehabilitation, Rigid pavements, Testing methods, Field tests, Pavement rating, Falling weight deflectometer.

RESUMEN: La presente investigación presenta el desempeño de diecisiete pavimentos rígidos hechos de Concreto con Cemento Portland (PCC) localizados en el Estado de Ohio, USA. La metodología de ensayos con el Deflectómetro de Impacto (FWD) se usó en un total de 58 km de secciones de pavimentos PCC para evaluar la condición estructural del pavimento. Cada sección se clasifica como excelente, buena, regular o deficiente basado en los parámetros de deflexiones normalizadas obtenidas de los ensayos FWD, deformabilidad, transferencia de carga y relación de soporte de las juntas. El desempeño global de las secciones analizadas está entre bueno y excelente. La metodología de ensayo de campo, junto con los análisis de datos del desempeño presentados en esta publicación, son una herramienta valiosa para evaluar la condición estructural actual de pavimentos y como resultado proponer planes de rehabilitación a corto o mediano plazo que garanticen el buen funcionamiento del pavimento.

PALABRAS CLAVES: Pavimentos, Rehabilitación, Pavimentos rígidos, Métodos de ensayo, Ensayos de campo, Evaluación de pavimentos, Deflectómetro de impacto.

1. INTRODUCTION

Rigid pavements are composed of a stiff Portland Cement Concrete (PCC) layer resting on a subgrade (base and subbase layers that need to be evaluated on a per project basis). Due to its stiffness and structure, the traffic load applied to the PCC pavement layer is transmitted under a wider area of subgrade, inducing a moderate stress and strain to the soil [1]. Portland cement concrete pavements are classified into three categories: Joint Plain Concrete Pavement (JPCP), Joint Reinforced Concrete Pavement (JRCP), and Continuously Reinforced Concrete Pavement (CRCP). The most common type of pavement in the State of Ohio, USA, is the JPCP and normally consists of slabs spacing between 12 and 20 ft long, having transversal joints reinforced with dowel bars to improve the performance of the joints. JRCP is not as common as the JPCP and the only difference is that the former consists of slabs with transversal joints at intervals of up to 50 ft. CRCP does not require transversal joints because is reinforced entirely over its length with longitudinal and transverse steel bars to prevent cracking due to variability in environmental conditions and traffic loading. JRCP and CRCP are no longer constructed due to their poor long-term performance [2].

There are several parameters involved in the PCC pavement performance, load transfer being the most

crucial. Load transfer is the ability of the slab to transmit the load to the adjacent slab through the joint to decrease the acting stresses and thus improve the pavement performance. Two of the most common mechanisms used to increase joint efficiency are dowel bars and aggregate interlock. Dowel bars have proved to effectively improve joint performance. Their main advantage is to allow slab movement in the horizontal direction and restrict the movement in the vertical direction while transferring the load. The dowel bars have a diameter between 32 and 38 mm, a length of 450 mm and are normally spaced 305 mm from each other. The dowel bars also limit the most common distress mechanisms in PCC pavement such as faulting, pumping, and corner break. The total load transfer between slab sections is due to the contribution of both the aggregate interlock and the dowel bars. However, the contribution due to the aggregate interlock can be considered negligible in cracks wider than 0.9 mm [2].

A study of the most significant design and construction parameters affecting long-term pavement performance of 2791.6 km (two directions) of CRCP in the State of Illinois, USA, is presented in Gharaibeh et al. [3]. Although some of the sections were exposed to extreme weather and traffic conditions, they had excellent performance during their design life. The study was conducted by analyzing data from field surveys collected since 1977. The database included a variety of information for each one of the pavements such as: section location, slab thickness, steel reinforcement content, base type and thickness, average annual temperature and precipitation. From this study, it was found that among the parameters affecting the CRCP, the reinforcement content and the slab thickness had the most critical influence on the pavement performance.

In a different study [4], the performance of transverse cracking on forty-nine JPCP sections located in the State of Michigan was presented. The purpose of this project was to determine the key parameters influencing the transverse cracking in JPCPs and the conclusions are as follows: a) the average number of cracks per slab increases as the joint spacing increases; b) the type of coarse aggregate of the concrete mix has a significant influence in the number of transverse cracks developed in the slabs; c) joint performance is directly related to temperature (i.e., high temperature increases the load transfer and thereby decreasing the crack width); and d) the load transfer value is an indicator of the crack condition in which values higher than 70% represent satisfactory crack performance (i.e., adequate aggregate interlock load transfer across the cracks).

The causes leading to surface longitudinal cracking and punch-outs on IH-30 (Interstate Highway) in the State of Texas are presented in Chen et al. [5]. Several field and laboratory tests were conducted to determine the causes of distress mechanisms. The authors concluded that longitudinal cracks were developed at early stages due to weather changes and to an increase of 30-50% of the design load. An evaluation of the performance of unbounded JRCP overlays used over existing concrete pavements is presented in Padilla-Llano [6]. Results from this investigation showed that environmental effects are more critical to the pavement than the effect of the traffic load. The strains induced to the pavement by the change in the environment are much greater than the strains induced during the FWD test.

2. FIELD TESTING AND SITE DESCRIPTIONS

A total of seventeen Portland cement concrete sections distributed along the State of Ohio, United States, were investigated. These sections were grouped based on the district location and classified either as excellent or average pavements depending on the structural condition at the time of testing. Figure 1 shows the districts subdivision of the state of Ohio.

Figure 1. District subdivision. State of Ohio

Figure 2 shows the Pavement Condition Rating (PCR) performance data for rigid pavements in the State of

Ohio [7]. The studied sections were a subset of data of those used to define Figure 2. Table 1 presents the location, length, district, year, and initial condition of the pavement sections. The county, roadway and district locations are referred as "Co-Rte" and "Distr." respectively. The directions are referred as upstation (U), downstation (D) or with the dual index (DU) for the cases when the section was tested in both directions. The year refers to construction date. The condition refers to the initial performance condition of the section as giving by the PCR index.

A Falling Weight Deflectometer (FWD) was used to evaluate the structural integrity of these sections. A description of this equipment and testing procedures are presented in Sargand et al. [8]. The FWD device used for this project consisted of seven sensors, aligned radially from the application of the load, to measure and record the deflections induced by the FWD. The separation between the sensors can be adjusted to measure settlements at different points of interest. The settlements measured from sensors 1 – 5 are used to evaluate the pavement structural condition, meanwhile settlements measured from sensors 6 and 7 are used to estimate the stiffness of the subgrade.

2.1. Data interpretation

The modulus of elasticity of the subgrade was calculated by taking an average of the values obtained from sensors 6 and 7, as described below [9], calibrated in U.S. units:

$$M_R(psi) = 9000 \times \frac{0.2892}{24 \times (d_{24}/1000)} \quad (1)$$

$$M_R(psi) = -466 + 9000 \times \frac{0.00762}{(d_{36}/1000)} \quad (2)$$

Eqs. (1) and (2) correspond to the deflections recorded for sensors six and seven, respectively. d_{24} and d_{36} are the settlement readings from the sensors located at a distance of 24 in. and 36 in., respectively from the application of the load. In these equations the values of d_{24} and d_{36} must be given in microinches (μ-in).

The normalized deflections and spreadability for each geophone were computed as follows [10], calibrated in U.S units:

$$Df_{Norm}(mils/kip) = \frac{Df_i(mils)}{Load(kip)} \quad (3)$$

$$Spreadability(\%) = \frac{100 \times \sum_{i=1}^{7} Df_i}{7 \times Df_1} \quad (4)$$

Df_i corresponds to the geophone readings i= 1 to 7 and the load is normalized to 9000 lb (40 kN).

Two additional parameters used to evaluate the structural pavement condition of PCC pavement sections are the load transfer parameter and the joint support ratio. The load transfer parameter is an indicator of the joint performance that depends on the applied load, aggregate interlock, and temperature acting on the pavement. The load transfer is calculated using the following equations for the approaching and leaving joint positions, respectively:

$$LT_A(\%) = (Df_3 / Df_1) \times 100 \quad (5)$$

$$LT_L(\%) = (Df_2 / Df_1) \times 100 \quad (6)$$

Df_1, Df_2, and Df_3 are the geophone readings number 1, 2, and 3, respectively.

The pavement joint condition, as presented in Sargand [10], is classified as good, fair or poor for the corresponding load transfer ranges of 80-100%, 50-80%, and less than 50%, respectively.

The Joint Support Ratio (JSR) parameter is an indicator of the pavement condition under the slabs and is calculated as follows:

$$JSR = Df_{1L} / Df_{1A} \quad (7)$$

Df_{1L} and Df_{1A} are the geophone readings number one at the leaving and approaching positions, respectively. Table 2 shows the values of deflections, spreadability, load transfer, and joint support ratio used to classify the structural pavement condition [11].

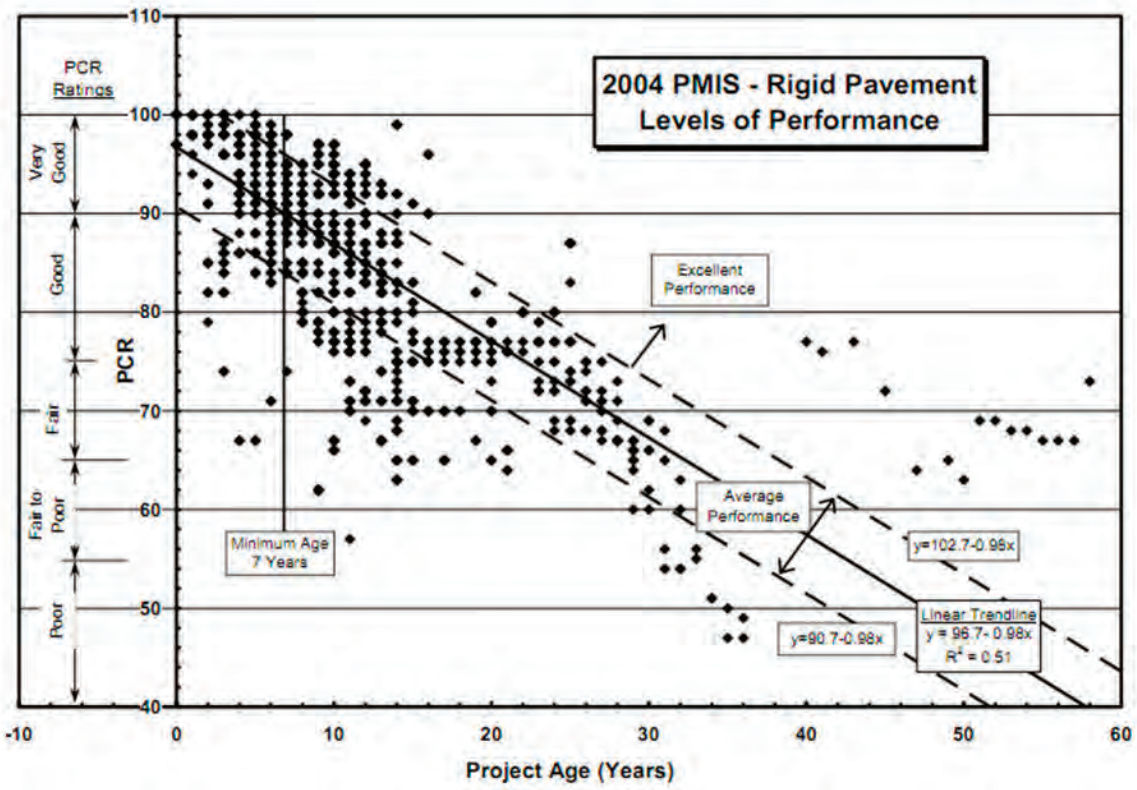


Figure 2. Pavement Condition Rating (PCR) for Rigid Pavements in Ohio. Source Chou et al. [7].

Table 1. Portland cement concrete pavement sections

Proj. No.	Co-Rte	SLM Limits	Dir.	L (mi)	Distr.	Year	Cond.
1	ATH 50	11.46-11.8	U	0.34	10	1986	Avg.
2	ATH 682	0.16-0.64	DU	0.48	10	1976	Avg.
3	CUY 82	3.22-3.66	D	0.44	12	1994	Exc.
		2.05-3.82	U	1.77			Exc.
4	GAL 7	5.71-10.21	U	4.50	10	1946	Exc.
5	HAM 126	11.35-13.31	DU	1.96	8	1990	Exc.
6	JEF 7	18.9-19.21	D	0.31	11	1990	Avg.
7	JEF 22	15.02-16.32	U	1.30	11	1990	Avg.
8	LOG 33	21.79-25.63	D	3.84	7	1994	Avg.
		21.51-25.63	U	4.12			Exc.
9	MOT 35	14.37-15.07	DU	0.70	7	1988	Exc.
10	MOT 202	2-3.25	U	1.25	7	1991	Exc.
11	SUM 76	11.8-13.32	D	1.52	4	1992	Exc.
			U				Avg.
12	SUM 76	13.32-15.32	D	2.00	4	1993	Exc.
			U				Avg.
13	TUS 39	2.84-7.12	U	4.28	11	1990	Avg.

3. PAVEMENT RESPONSES

Figures 3-8 show a typical set of plots for project # 2 (see Vega-Posada [12] for the complete evaluation) including normalized midslab deflection (Df_1 and Df_7), spreadability, normalized maximum joint deflection, joint transfer, joint support ratio, and subgrade modulus of elasticity, respectively. The modulus of elasticity of the subgrade was back-calculated from the FWD collected data.

Figures 3 and 4 show a decrease in pavement stiffness, both concrete and subgrade layers, between SLM 0.3 and 0.4 in the upstation direction. Figures 5-7 show that the overall condition of the pavement joints is from good to excellent.

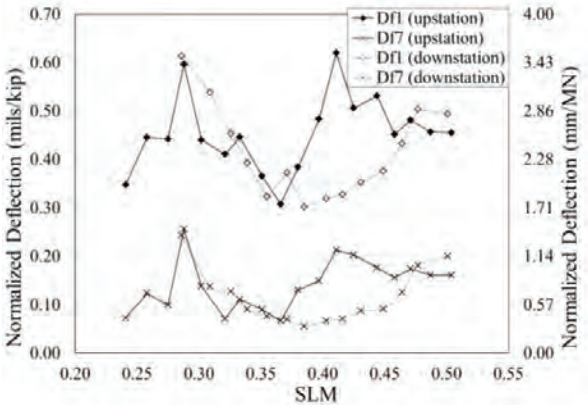


Figure 3. Midslab deflection

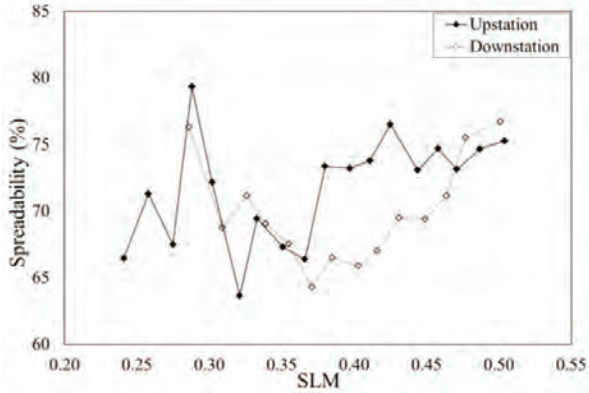


Figure 4. Midslab spreadability

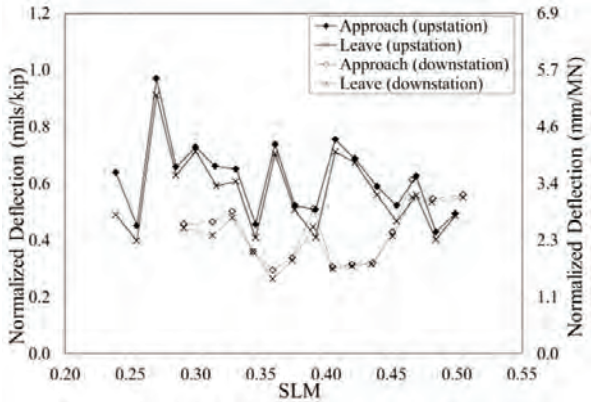


Figure 5. Max. joint deflections

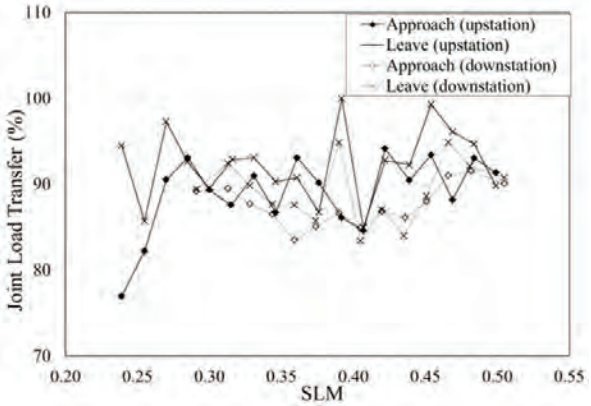


Figure 6. Joint load transfer

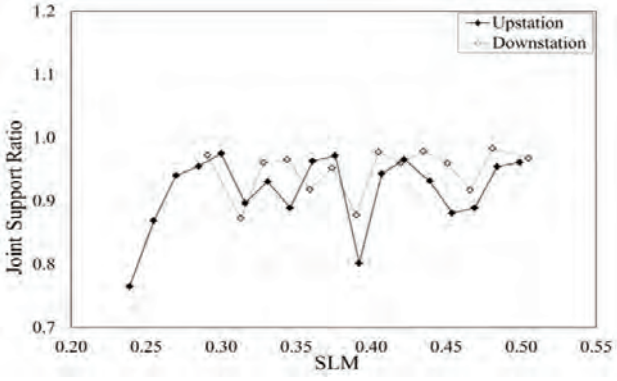


Figure 7. Joint support ratio

Figure 8 presents the back-calculated modulus of elasticity of the subgrade. M_R ranged from 31 ksi (upstation) to 36 ksi (downstation), which in geotechnical terms is considered as a competent layer. From figures 3, 4, and 8, it can be concluded that the stiffness deficiency between SLM 0.3 and 0.4 in the upstation direction is related to a deterioration of the PCC slab and not to a reduction of the subgrade layer capacity. Although the spreadability is classified as good

in the upstation direction, there is a lack in the pavement ability to distribute the load across the section.

The pavement performance in both directions is similar and therefore the conclusions are applicable for both cases. The average modulus of elasticity in the upstation and downstation directions is 31 ksi (214 MPa) and 36 ksi (248 MPa), respectively. The determination of this parameter from actual field observations as presented in this research not only provides an indication of the pavement deterioration in comparison to the initial design modulus but also is a reliable parameter that can be used to estimate the pavement behavior and performance based on computer models.

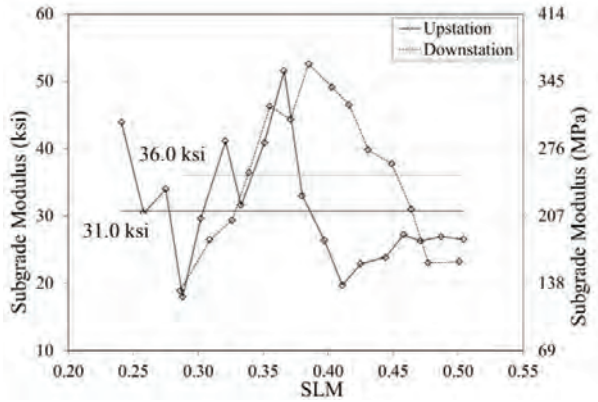


Figure 8. Subgrade modulus

Table 2. Max. values for deflections, spreadability, load transfer, and joint support ratio. Souce: Edward et al. [11]

Cond.	$Df_{1,}$ mm/MN	Df_{7}, mm/MN	SPR (%)	D_{1A}, mm/MN	D_{1L}, mm/MN	LT (%)	JSR
Exc.	< 2.40 (0.42)	< 1.20 (0.21)	> 80	< 3.54 (0.62)	< 3.71 (0.65)	> 85	0.90-1.10
Good	2.45-3.25 (0.43-0.57)	120-1.77 (0.21-0.31)	72-79	3.60-4.45 (0.63-0.78)	3.77-4.74 (0.66-0.83)	70-85	1.11-1.25
Fair	3.31-4.17 (0.58-0.73)	1.77-2.40 (0.31-0.42)	64-71	4.51-5.48 (0.79-0.96)	4.80-5.82 (0.84-1.02)	50-69	1.26-1.49
Poor	> 4.22 (0.74)	> 2.40 (0.42)	<63	> 5.54 (0.97)	> 5.88 (1.03)	< 50	> 1.50

Note: The values in parenthesis are in units *mils/kip*, as obtained directly from Eq. (3)

Table 3 presents the pavement built-up sequence and the back-calculated modulus of elasticity of each pavement layer. The following abbreviations are used for the material specification: a) JRC: Joint Reinforced Concrete Pavement; b) PCC: Portland Cement Concrete Pavement; c) "310": Bituminous Aggregate Base; d) ATB: Asphalt Treated Base; e) DGAB: Dense Graded Aggregate Base; f) NSDB: Non-Stabilized Drainage Base.

A complete definition of these material specifications can be found in ODOT [13]. A Dynatest Model 8000 FWD was used to conduct the field testing program [14]. Tables 4-6 show the structural condition of these pavement sections based on FWD deflections and spreadability, maximum joint deflection, and joint load transfer and joint support ratio, respectively. In general, the structural condition of the studied sections are classified as excellent and good, except for project # 4 that showed a load transfer deficiency in both the approaching and leaving positions.

Table 3. Back-calculated modulus of elasticity

No.	Layer Modulus (M_R in MPa), Thickn. (mm)			
1	Layer	JRC	310	Subgrade
	M_R-Thk.	20,217-229	945-152	148-N/A
2	Layer	JRC	310	Subgrade
	M_R-Thk.	26,151-229	422-152	255-N/A
3	Layer	JRC	DGAB	Subgrade
	M_R-Thk.	26,220-279	1,263-152	274-N/A
4	Layer	JRC	310	Subgrade
	M_R-Thk.	18,975-229	564-152	173-N/A
5	Layer	JRC	310	Subgrade
	M_R-Thk.	31,119-254	5,147-152	324-N/A

Continuation Table 3.

No.	Layer Modulus (M_R in MPa), Thickn. (mm)			
6	Layer	JRC	310	Subgrade
	M_R-Thk.	23,391-229	1,097-152	295-N/A
7	Layer	JRC	310	Subgrade
	M_R-Thk.	26,427-229	642-152	272-N/A
8	Layer	PCC	NSDB	DGAB
	M_R-Thk.	24,357-305	3,222-102	1,035-76
9	Layer	PCC	310	Subgrade
	M_R-Thk.	26,358-229	2,539-152	414-N/A
10	Layer	PCC	310	Subgrade
	M_R-Thk.	24,564-229	547-254	164-N/A
11	Layer	JRC	ATB	Subgrade
	M_R-Thk.	38,502-229	9,798-152	242-N/A
12	Layer	JRC	ATB	Subgrade
	M_R-Thk.	24,219-229	2,187-152	515-N/A
13	Layer	JRC	310	Subgrade
	M_R-Thk.	28,635-229	856-152	167-N/A

Table 4. Pavement condition based on FWD deflections and spreadability

Proj. No.	Dir	L	FWD Defl. (%)				Spreadability (%)			
		(km)	Exc.	Good	Fair	Poor	Exc.	Good	Fair	Poor
1	U	0.55	9	55	27	9	-	73	27	-
2	DU	0.77	57 28	36 61	7 11	- -	- -	21 61	72 33	7 6
3	D	0.71	95	5	-	-	26	63	11	-
	U	2.85	100	-	-	-	94	6	-	-
4	U	7.25	3	55	35	7	7	46	40	7
5	DU	3.16	100 100	- -	- -	- -	- 20	75 60	25 20	- -
6	D	0.5	82	18	-	-	-	36	64	-
7	U	2.09	84	8	8	-	-	23	77	-
8	D	6.18	86	14	-	-	14	79	7	-
	U	6.63	93	7	-	-	22	78	-	-
9	DU	1.13	100 100	- -	- -	- -	- -	69 56	31 44	- -
10	U	2.01	10	90	-	-	-	90	10	-
11	D	2.45	100	-	-	-	29	64	7	-
	U	2.45	91	9	-	-	18	82	-	-
12	D	3.22	100	-	-	-	-	69	31	-
	U	3.22	100	-	-	-	23	65	12	-
13	U	6.89	10	74	16	-	7	79	14	-

Table 5. Maximum joint deflection

Proj. No.	Dir	L	Condition - Approaching (%)				Condition - Leaving (%)			
		(km)	Exc.	Good	Fair	Poor	Exc.	Good	Fair	Poor
1	U	0.55	27	46	27	-	-	64	36	-
2	DU	0.77	100 44	- 50	- -	- 6	100 72	- 22	- 6	- -
3	D	0.71	95	5	-	-	95	-	5	-
	U	2.85	81	19	-	-	100	-	-	-

Continuation Table 5.

Proj. No.	Dir	L	Condition - Approaching (%)				Condition - Leaving (%)			
		(km)	Exc.	Good	Fair	Poor	Exc.	Good	Fair	Poor
4	U	7.25	-	7	25	68	2	23	7	68
5	DU	3.16	19 90	31 10	50 -	- -	25 85	31 15	31 -	13 -
6	D	0.50	100	-	-	-	100	-	-	-
7	U	2.09	92	8	-	-	54	46	-	-
8	D	6.18	50	25	25	-	71	29	-	-
	U	6.63	100	-	-	-	-	100	-	-
9	DU	1.13	100 89	- 11	- -	- -	100 89	- 11	- -	- -
10	U	2.01	80	20	-	-	90	10	-	-
11	D	2.45	100	-	-	-	100	-	-	-
	U	2.45	100	-	-	-	100	-	-	-
12	D	3.22	100	-	-	-	100	-	-	-
	U	3.22	100	-	-	-	100	-	-	-
13	U	6.89	32	38	14	16	41	27	14	19

Table 6. Joint load transfer and joint support ratio

No.	Dir.	Joint Load Transfer									Joint Support Ratio			
		L	Condition - Approaching (%)				Condition - Leaving (%)				Condition - Approaching (%)			
		(km)	Exc.	Good	Fair	Poor	Exc.	Good	Fair	Poor	Exc.	Good	Fair	Poor
1	U	0.55	81	19	-	-	91	9	-	-	64	36	-	-
2	DU	0.77	86 83	14 17	- -	- -	86 94	14 6	- -	- -	86 61	14 39	- -	- -
3	D	0.71	63	32	-	5	79	16	-	5	84	16	-	-
	U	2.85	62	25	13	-	69	19	12	-	63	37	-	-
4	U	7.25	5	2	19	74	10	2	17	71	56	44	-	-
5	DU	3.16	94 100	6 -	- -	- -	81 75	13 25	6 -	- -	63 70	37 30	- -	- -
6	D	0.5	27	-	73	-	9	82	9	-	-	64	27	9
7	U	2.09	100	-	-	-	31	69	-	-	15	54	31	-
8	D	6.18	28	36	36	-	29	50	21	-	64	36	-	-
	U	6.63	100	-	-	-	100	-	-	-	100	-	-	-
9	DU	1.13	62 67	31 33	- -	7 -	54 56	31 44	15 -	- -	77 100	23 -	- -	- -
10	U	2.01	100	-	-	-	100	-	-	-	100	100	-	-
11	D	2.45	14	79	7	-	22	64	14	-	79	21	-	-
	U	2.45	54	46	-	-	44	44	12	-	73	27	-	-
12	D	3.22	50	50	-	-	37	63	-	-	93	7	-	-
	U	3.22	6	94	-	-	24	71	5	-	71	29	-	-
13	U	6.89	83	17	-	-	61	39	-	-	81	14	5	-

4. SUMMARY AND CONCLUSIONS

The methodology presented in this paper is a valuable technique that can be used to determine, with field measurements, the actual structural condition of PCC pavement sections. Based on the structural condition classification, short or long term rehabilitation programs can be implemented to assure satisfactory service and serviceability of the PCC pavement.

The long term performance of seventeen Portland cement concrete sections located in the State of Ohio, USA, was studied. The performance of PCC pavement was influenced by the climate conditions, material properties, construction practices, and traffic loads.

The total length of the PCC pavement sections studied was 36 mi (58 km). The overall structural condition of the analyzed sections was as follows: a) Excellent 67.6% (39 km); b) Good 24.2% (14 km); c) Fair 7.2% (4 km); and d) Poor 1.0% (0.6 km). On the other hand, the overall condition of the pavement stiffness was: a) Excellent 15.8% (10 km); b) Good 63.9% (36 km); c) Fair 19.3% (11 km); and d) Poor 1.0% (1 km).

The overall performance of the pavement system was influenced by the stiffness of the base layer and the thickness of the surface layer. The performance of PCC pavement sections improved as the base stiffness and/or thickness of the surface layer increased. In general, the structural pavement condition of the PCC sections was classified as excellent and good, except for Project 4 in which the load transfer mechanism between the slabs negatively impacted the overall capacity of the rigid pavement.

REFERENCES

[1] Fwa, T. F., The Handbook of Highway Engineering, NY: CRC Press/Lewis Publishers, New York, 2006.

[2] Washington State DOT (n.d., a), Rigid Pavement Types, Available: training.ce.washington.edu/wsdot/modules/02_pavement_types/02-6_body.htm. [April 4, 2008].

[3] Gharaibeh, N. G., Darter, M. I.,and Heckel, L. B., Field performance of continuously reinforced concrete pavement in Illinois. J. of the Transportation Research Board, 1684, Transportation Research Board of the National Academies, Washington, D.C., pp.44-50, 1999.

[4] Frabizzio, M. A., and Buch, N. J., Performance of transverse cracking in jointed concrete pavements, J. of Performance and Constructed Facilities 13 (4), pp. 172-180, 1999.

[5] Chen, D. H., Lin, D. F., Liau, P. H.,and Bilyeu, J., A correlation between dynamic cone penetrometer values and pavement layer moduli, Geotechnical Testing Journal 28 (1), pp. 42-49, 2005.

[6] Padilla-Llano, D., Rehabilitated concrete pavements: performance evaluation of FPCP Overlays on Interstate 86 NY, [Master Thesis], Ohio University, Athens, OH, 2010.

[7] Chou, E., Tack, J., Yu, J., and Wielinski, J., Evaluation of the variation in pavement performance between ODOT Districts, University of Toledo, Ohio, Report FHWA/OH-2004/016, 2004.

[8] Sargand, S. M., Vega-Posada, C. A., and Arboleda-Monsalve, L. G., Long term performance of existing asphalt concrete pavement sections, Rev. Fac. Ing. Univ. Antioquia 66 (3), pp. 49-60, 2013.

[9] Pierce, M. L., Development of a computer program for the determination of the area value and subgrade modulus using the Dynatest FWD, Washington State Department of Transportation, 1999.

[10] Sargand, S. M., Determination of pavement layer stiffness on the OHIO SHRP test road using non-destructive testing techniques, Final Report FHWA/OH-2002/031, Athens, OH, Ohio Department of Transportation and FHWA, U.S. Department of Transportation, 2002.

[11] Edwards, W. F., Green, R. L., and Gilfert, J., Implementation of a dynamic deflection system for rigid and flexible pavements in Ohio, FHWA/OH-89/020, Athens, OH, Ohio, Department of Transportation and FHWA. U.S. Department of Transportation. 1989.

[12] Vega-Posada, C. A., Long term performance of existing AC and PCC pavements in Ohio, [Master Thesis], Ohio University, Athens, OH, 2008.

[13] Ohio Department Of Transportation, Construction and material specifications, Columbus, OH, Department of Transportation, 2008.

[14] Dynatest International, Dynatest FWD/HWD test systems. Available: http://www.dynatest.com/equipment/structural/fwd (Dec. 12, 2013).

ANALYSIS OF THE MOISTURE SUSCEPTIBILITY OF HOT BITUMINOUS MIXES BASED ON THE COMPARISON OF TWO LABORATORY TEST METHODS

ANÁLISIS DE LA SUSCEPTIBILIDAD AL AGUA EN MEZCLAS BITUMINOSAS EN CALIENTE MEDIANTE EL ESTUDIO COMPARATIVO DE DOS MÉTODOS DE ENSAYO DE LABORATORIO

FERNANDO MORENO-NAVARRO
PhD.Civil engineering, Laboratorio Ingeniería de la Construcción. Universidad de Granada. fmoreno@ugr.es
GEMA GARCÍA-TRAVÉ
PhD.Degree in Chemistry, Laboratorio Ingeniería de la Construcción. Universidad de Granada. gemamgt@correo.ugr.es
MARÍA DEL CARMEN RUBIO-GÁMEZ
PhD. Civil engineeringLaboratrorio Ingeniería de la Construcción. Universidad de Granada. mcrubio@ugr.es
MARIA JOSÉ MARTINEZ-ECHEVARRIA
PhD.Civil engineering .Laboratorio Ingeniería de la Construcción. Universidad de Granada. mjmartinez@ugr.es

Received for review July 30 th, 2012, accepted April 18th, 2013, final version May, 3 th, 2013

ABSTRACT: One of the most common road pavement pathologies is caused by water action. The presence of moisture causes a reduction in aggregate-binder adhesion in the mix as well as in the internal cohesion of the bitumen. This leads to problems such as pot holes, aggregate peeling, stripping, etc., which eventually lead to the structural failure of the pavement. Currently, there are numerous laboratory tests that analyze the susceptibility of bituminous mixes to moisture, providing a qualitative or quantitative evaluation.
This study analyzes the performance of bituminous mixes in different experimental conditions. For this purpose, a comparative study of two laboratory tests was carried out. The tests differed in mix compaction method, the conditioning of the test specimens, and the type of load applied. The results obtained showed that in these tests the conditioning temperature had a greater impact on mix performance than the temporal duration of the conditioning process. Furthermore, the application of tensile stress was found to be more suitable for studying moisture susceptibility since mixes were found to be more sensitive to this type of load.

Key Words: Hot bituminous mixes, road, mix design, moisture susceptibility.

RESUMEN: Unas de las patologías más comunes en los firmes flexibles de carretera de todo el mundo son las provocadas por la acción del agua. La presencia de éste elemento provoca una disminución en la adhesividad entre los áridos y el ligante de la mezcla, así como de la cohesión interna del betún, que se traducen en problemas de desprendimiento de áridos, baches, peladuras, etc., que terminan provocando el fallo estructural del firme. Actualmente existen numerosos ensayos de laboratorio que permiten analizar la susceptibilidad al agua de las mezclas bituminosas, proporcionando una medida cualitativa o cuantitativa. Así, desde esta investigación se ha pretendido analizar el comportamiento de las mezclas bituminosas ante diferentes condiciones de ensayo. Para ello se ha llevado un estudio comparativo entre dos de estos ensayos, los cuales difieren en la forma de compactación de la mezcla, el proceso de acondicionamiento de las probetas y el tipo de carga aplicado. Los resultados obtenidos pusieron de manifiesto que, para los ensayos estudiados, la temperatura de acondicionamiento ejerce mayor influencia en el comportamiento de la mezcla que el tiempo de acondicionamiento. Además, la aplicación de cargas de tracción resultan más adecuadas para el estudio de la susceptibilidad al agua de las mezclas bituminosas ya que las mezclas son más sensibles a éstas.

Palabras Clave: Mezcla bituminosa en caliente, carretera, diseño de mezcla, susceptibilidad al agua.

1. INTRODUCTION

Bituminous mixes are some of the most frequently used materials in road construction all over the world [1]. This material is a conglomerate with a mineral skeleton composed of a coarse fraction with a diameter of more than 2 mm, a fine fraction with a diameter of 0.063 - 2 mm, and a mineral powder fraction with a diameter of less than 0.063 mm. It is also composed of a hydrocarbon binder that gives cohesion to the mineral skeleton (bitumen or one of its byproducts from the distillation of crude oil), and sometimes, a

series of additives (fibers, waxes, etc.) that improve the characteristics and mechanical performance of the mixes (greater resistance to plastic deformation, lower susceptibility to fatigue cracking, rutting, raveling, etc.) [2].

In order for the mix to be used in the road surface course, it must pass a series of laboratory tests that certify its suitability and good performance under various types of load or stresses to which it will be subjected during its service life (traffic loads, extreme temperatures, water action, etc.) [3]. Of these stresses, those produced by water action are among the main causes of the pavement deterioration (loss of aggregate, peeling, pot holes, etc.). The presence of moisture in the mix reduces aggregate-binder adhesion (i.e. destruction of the chemical bonds between the binder and the aggregate) [4-10], as well as its cohesion (the emulsification of the binder due to water action separates its particles) [11-14]. This leads to a loss of aggregate, which in the medium and long term, and combined with other negative factors (traffic loads, ice formation, binder ageing, etc.) finally cause most road pathologies and the eventual failure of the road surface course [15-17].

As seen in Lu and Harvey [18], one of the key ways to mitigate the effect of water action on mixes is the design and application of laboratory tests capable of predicting the potential moisture damage during the service life of the pavement. This makes it possible for the mix to have an optimal design. Methods to predict the effect of moisture action on the bituminous mixes were initially developed in the 1930s, and there are now a wide range of such tests available [19-27]. A description of these methods can be found in Solaimanian et al. [28].

The objective of such methods is to reproduce mix performance in laboratory conditions when it is affected by water. They vary in reliability, depending on the characteristics of the test, the mix type, as well as the environmental conditions of the use case (cold or hot temperatures, dry or rainy weather, etc.). Most of these tests are performed with fixed temperatures and loads. This means that it is impossible to simulate the effect of moisture on the pavement when it is subjected to varying traffic loads and weather conditions [22]. Nevertheless, these experiments are often very complicated and time-consuming to perform. Moreover, they are somewhat limited in scope when they try to reproduce the way that moisture attacks mixes. It is thus necessary to analyze and compare the effect of these laboratory tests on the mix in order to evaluate its response. Only in this way will it be possible to predict the performance of a mix when it is actually used in road construction.

In the laboratory, methods for testing moisture susceptibility differ in the mix compaction method, the type of conditioning, and the application of the load. Although there are many tests used for this purpose (e.g. Immersion-Compression, Hamburg Wheel Tracking, Moisture Vapor Susceptibility, Marshall Immersion, etc.), some are more effective than others. Today there is a division of opinion regarding which test is the most suitable for studying the susceptibility of bituminous mixes to moisture action. Consequently, in this study a comparative analysis was made of two of these methods, and the performance of the mix was analyzed in terms of characteristics, such as type of test specimen, compaction method, conditioning process, and load application. In this way it was possible to arrive at conclusions concerning the representativity of the results of each test.

The laboratory tests selected were two that are commonly used in Spain: a) the Immersion-Compression Test (regulated by the Spanish Technical Standard NLT-162); b) the Water Sensitivity Test (regulated by the European standard UNE-EN 12697-12 and implanted as a new reference test to analyze the susceptibility of bituminous mixes to moisture in all European Union countries). Both tests differ in so far as the type of test specimen used (dimensions and compaction methods), conditioning process, and applied load. However the result of both tests is a retained strength value, which makes it possible to compare them.

This article analyzes the results obtained after performing the Immersion-Compression Test and the Water Sensitivity Test on samples of the same mix type with an identical mineral skeleton and the same percentage of bitumen. The number of samples studied had to be sufficient to statistically validate the results obtained and thus extract representative conclusions. A total of 33 samples were used, and the central limit theorem was applied. The results of both laboratory tests were statistically analyzed, as well as the mix response in each.

2. METHODOLOGY

2.1. Materials

The semi-dense mix used in the study was regarded as representative, a generic asphalt concrete that is commonly found in all type of road surfaces (regardless of climate or traffic loads) throughout Europe. It had a continuous grain size with a strong mineral skeleton, and its bitumen content was 4-5% of the total mass. This made it resistant to plastic deformations and water action. More specifically, the mix used was an AC 16-S mix for the surface course [29], manufactured from ophite and limestone aggregate and penetration bitumen 40/60.

The samples were made in different batches though the characteristics of the mix did not vary. The mineral skeleton of each batch was within the grain-size limits specified in the Spanish regulations [29], which establishes a reference particle size for this type of mix. Nevertheless, in our samples, the mineral skeleton was slightly modified to obtain different air void contents so as to better study the effect of the water volume on the mix.

The aggregate used in the mix was ophite for the coarse fraction and limestone for the fine fraction as well as mineral dust. The characteristics of this aggregate are in accord with the PG-3 General Technical Specifications for Roads and Bridge Works in Spain[29].

The filler fraction chosen for the mix was limestone, given its good response to water action. This is due to the fact that the adhesion of limestone with the binder is greater since its alkaline surface reacts better to the acid composition of the bitumen [30]. The limestone had an apparent density in toluene of 0.67 g/cm3 in accordance with the NLT-176.

The manufacturing temperature of the mix was 160-165ºC, whereas the compaction temperature of the test specimens was 150-155ºC. The optimal bitumen content used to make the specimens was 4.5% of the total weight of the mix (based on the Marshall Test results of the job mix formula).

2.2. Experimental design

The experimental design was based on the comparison of the Immersion-Compression Test (NLT-162) and the Water Sensitivity Test (UNE-EN 12697-12).

Both tests offer a similar result (retained strength). Nevertheless, test conditions (compaction method, specimen conditioning, and load applied) are not the same, so they can be used to analyze and compare the performance of the mix to water action

In this study 33 samples of the same mix type (AC 16 – S) were tested. The specimens had identical characteristics but belonged to different batches. They were all subjected to the Immersion-Compression Test and the Water Sensitivity Test. The number of samples was based on the central limit theorem. According to this theorem, when the number of samples of a population is sufficiently large (i.e. more than 30 [31-33]) it is assumed to have a normal distribution. It is thus possible to derive scientifically valid conclusions from the results obtained. The mix response in the two tests was statistically analyzed, and it was possible to observe the representativity of each regarding the mix's susceptibility to moisture action.

The Immersion-Compression Test involves the manufacturing of 10 test cylinders with a diameter of 101.6 mm and a height of 101.6 mm. These cylinders are compacted by means of a static load produced by a double plunger. There is an initial pre-load of 1 MPa (in order to settle the mix) after which the load applied steadily increases for 2.5 minutes until reaching 21 MPa. Then the density of the samples is determined by the saturated surface dry method. As part of the conditioning process before the test, the specimens are divided into two sets of five. One set is submerged in a container with water at 60°C for 24 hours, whereas the other set is left at room temperature. After 24 hours, both sets of specimens are subjected to a simple compression load at a constant deformation velocity of 5.08 mm/min until they break. In this way it is possible to obtain a mean retained strength value for each set. The result of this test, which measures the specimens' sensitivity to water action, this is known as the retained strength index. This index (expressed in %) is calculated by dividing the strength obtained by the group immersed in water at 60°C by the strength obtained by the dry set.

The Water Sensitivity Test involves manufacturing 6 specimens with a diameter of 101.6 mm and a height of 60 mm. The specimens are compacted with an impact compactor by applying 50 blows on each

face. Then the density of the samples is determined by the saturated dry surface method. As part of the conditioning process, the specimens are divided into two sets: a wet set and a dry set. The dry set is stored at room temperature in the laboratory (20±5 °C), whereas a vacuum is applied to the wet set for 30±5 minutes until a pressure of 6.7±0.3 KPa is reached. Then, the specimens in this set are left in immersion conditions of 40°C for 72 hours. Afterwards, the indirect tensile fracture of each specimen (of the dry set as well as the wet set) is performed at a temperature of 15°C after the specimens have previously been acclimated to this temperature for 120 minutes. The division of the tensile strength obtained by wet specimens, by the tensile strength obtained by dry specimens, results in the retained indirect tensile strength (expressed in %).

In this study, the relation between the variables in each test and the mix response was obtained by using a simple linear regression.

Once the straight regression line is fit to the cloud of data points, it is crucial to have a measurement of the goodness of fit, which shows whether the fit is sufficient. The determination coefficient was used for this purpose. When a simple linear regression is performed, this coefficient coincides with the square of the linear correlation coefficient. The determination coefficient is interpreted as the percentage variation of the dependent explained by the model. It is denoted by R^2, and since $0 < R^2 < 1$, this measurement is usually expressed as a percentage. In this type of study, the tests were very heterogeneous, and this meant that the results of the tests were, as well. Consequently, the determination coefficients of the linear fit lines were not very high.

Apart from the simple linear regressions, and given that both tests had different characteristics, the retained strength values were studied according to a normal distribution in order to reinforce the comparative analysis.

3. ANALYSIS OF RESULTS

The study described in this paper was the result of twelve months of laboratory work and shows the results obtained for the 33 samples in the Immersion-Compression Test and the Water Sensitivity Test in relation to the mean values of the wet and dry sets of samples for fracture strength, density, and air void content of the test specimens.

Coefficient of variation (CV) is calculated for the Immersion-Compression Test and the Water Sensitivity Test for fracture strength, density, and air void content for both wet and dry sets obtaining in all cases similar deviations (see tables 1 and 2). Similar results in retained strength values (%) were obtained in both tests as is resumed in table 3.

Table 1.CV Immersion-Compression test.

Compression Strength		Density		Air Void Content	
Dry Set	Wet Set	Dry Set	Wet Set	Dry Set	Wet Set
0.1844	0.1809	0.0222	0.0226	0.3107	0.3190

Table 2. CV Water Sensitivity test.

Indirect Tensile Strength		Density		Air Void Content	
Dry Set	Wet Set	Dry Set	Wet Set	Dry Set	Wet Set
0.1790	0.2065	0.0251	0.0243	0.3508	0.3387

Table 3.CV Retained strength in the tests.

Retained strength	
Immersion-Compression Test	Water Sensitivity Test
0.0279	0.0239

Since the samples in the study were manufactured with the same materials (aggregate and bitumen), there is a close relation between the air void content and the density of the specimens in both tests (despite the fact that they differ in compaction method). Accordingly, the samples with a lower density have a greater content of air voids in the mix.

As can be observed in the results, in the Immersion-Compression Test, the air void content and the density of the specimens do not have a significant effect on the simple compressive strength (Figure 1). Therefore, a dispersed cloud of data points, and the extremely low determination coefficients of the regression lines, reflect the independence of both characteristics in the

mix response to compression forces. This is due to the fact that in bituminous mixes, compression forces are endured mainly by the internal friction of the mineral skeleton. Consequently, the strength of the specimen mainly comes from the aggregate's capacity to withstand the transmitted load, and is independent of the moisture content.

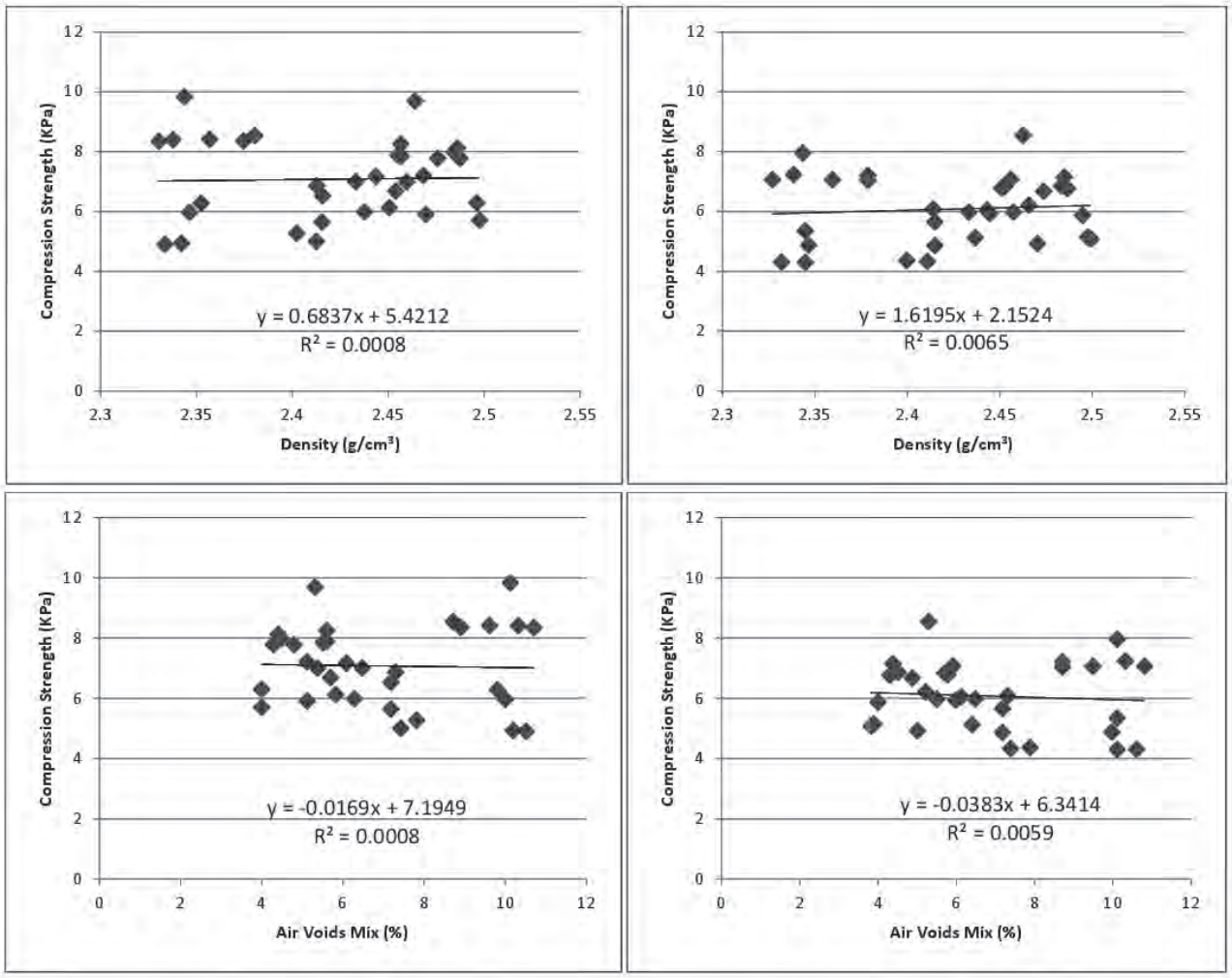


Figure 1. Simple compression strength based on the density and air void content of the mix: dry set (left) and wet set (right)

In contrast, in the Water Sensitivity test, there is a closer relation between the air void content and density of the specimen and its indirect tensile strength. This is reflected by a higher determination coefficient in the linear regression lines (see Figure 2). This tendency implies that when a specimen has a lower density or a higher air void content, its strength is reduced. In this case, the application of the load varied, and the specimen was subjected to tensile stresses such that the internal cohesion of the mix (the sum of binder cohesion and the aggregate-binder adhesion) was the basis of the specimen's fracture strength. This strength decreased as the air void content of the specimen increased. This was because the quantity of air voids (spaces without aggregate or bitumen) on the plane where the break occurred meant that there was less surface to withstand the strain. The test specimen was thus weaker.

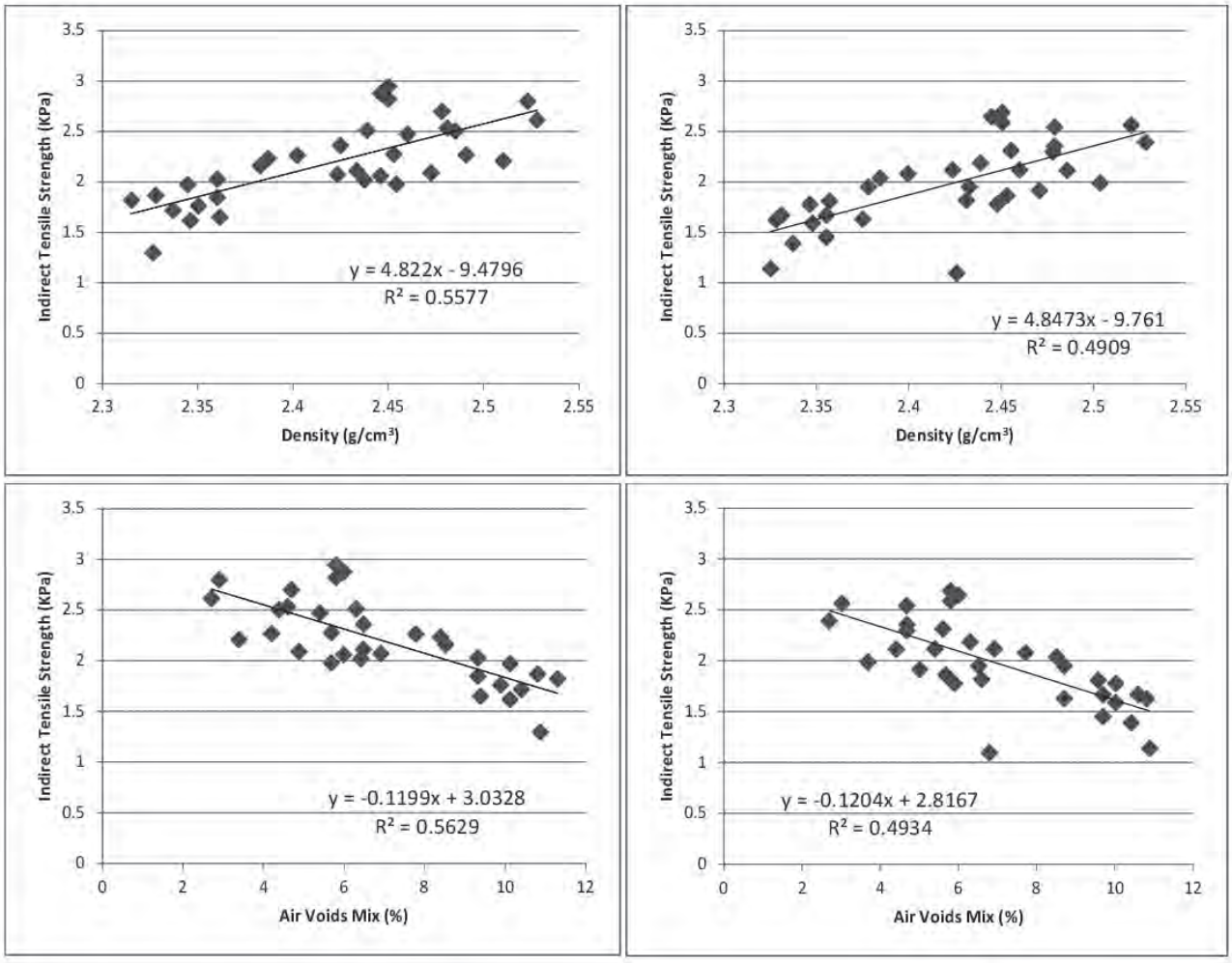


Figure 2. Indirect tensile strength based on the density and air void content of the mix: dry set (left) and wet set (right)

Moreover, it should be noted that the indirect tensile strength of the specimens (in the Water Sensitivity Test) was less than the simple compression strength (in the Immersion-Compression test). This response is logical since concrete-type stone materials, such as bituminous mixes, perform much better under compression loads than under tensile stresses. This is due to the fact that the mineral skeleton makes up about 90% of the weight of the bituminous mix, whereas the binder, which is the source of tensile strength, is only about 5% of the weight of the mix.

Furthermore, the retained strength values (%) obtained in both tests are shown in Figure 3.

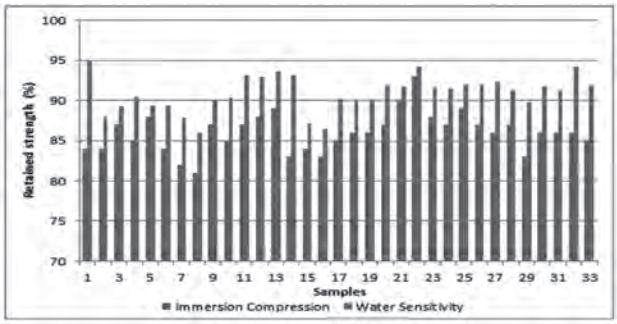


Figure 3. Retained strength of the samples studied in the tests.

As can be observed, the retained strength in the Water Sensitivity Test was greater than in the Immersion-Compression Test. In fact, it was 5% higher, based on the mean value of the 33 samples. When the retained

strength values obtained in each test were compared by using a normal distribution (based on the central limit theorem), this difference in the values can be observed clearly (Figure 4). Thus, in the Water Sensitivity Test, this effect was less than in the Immersion-Compression test. The analysis also showed that the Gaussian bell curve in the Water Sensitivity Test is sharper, which means that the results are more clustered. The samples thus have a lower variability, which signifies that the test is more reproducible (and thus its results are more representative).

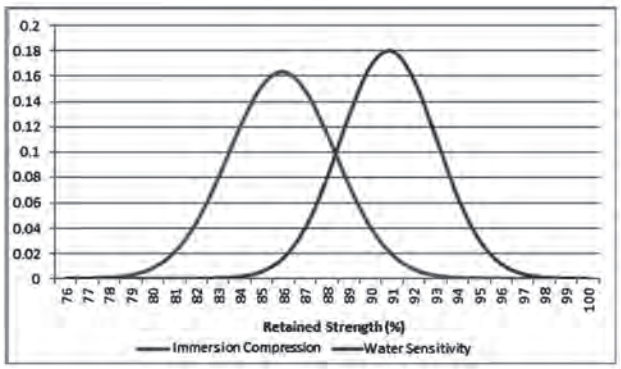


Figure 4. Normal distributions of the Immersion-Compression test and the Water Sensitivity test.

According to these results, the simple compression strength of the bituminous mix was potentially more affected by moisture action than its indirect tensile strength was. However, the simple compression strength mainly depends on the internal friction of the aggregate, and the indirect tensile strength on the internal cohesion of the mix, which is the sum of binder cohesion and aggregate-binder adhesion. Both factors are significantly more sensitive to moisture action than the internal friction of the aggregate (i.e. the presence of water can break the chemical bonds in the bitumen as well as reduce its adhesion to the aggregate) [3-13]. As a result, the retained strength of the Water Sensitivity Test should be lower. This means that the conditioning undergone by the wet set of specimens in the Immersion-Compression Test (24 hours at 60 °C) was more aggressive than that undergone by the wet set in the Water Sensitivity Test (72 hours at 40 °C). The temperature of the conditioning had a greater influence on mix performance than the temporal duration of the conditioning.

The compaction method was the other variable in which the tests differed, apart from the conditioning of the specimens and the type of fracture. However, it did not seem to have a significant influence on the results obtained. In this respect, the densities of the 33 samples in both studies were very similar, though they were slightly higher in the case of the Immersion-Compression Test (Figure 5).

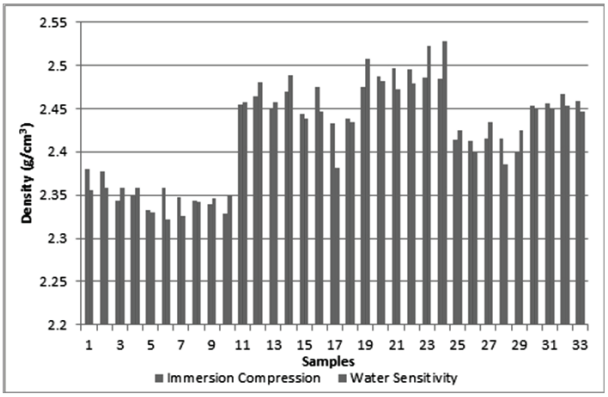


Figure 5. Mean density of the sample specimens studied in each test.

Finally, in the analysis of the influence of the air void content and the density, on the retained strength, it was observed that both tests showed the same tendency. The greater the density (i.e. lower air void content) of the specimens, the greater was the retained strength of the samples (Figure 6). In other words when the air void content of the mix increased, the volume of water that penetrated it was greater. This significantly affected mix performance since it reduced simple compression strength as well as indirect tensile strength.

In the Immersion-Compression Test, the relation between the air void content and the retained strength of the mix was greater than that between the air void content and the simple compression strength (larger determination coefficient). The higher percentage of air voids caused the mix not to respond as well to water action. However, the simple compression strength remained similar to that of a mix with a lower percentage of air voids.

In the Water Sensitivity Test, the relation between the air void content and the indirect tensile strength of the mix was greater than that between the air void content and the retained strength (larger determination coefficient). This shows that an increase in the air void content of the mixes had a greater impact on its capacity to resist indirect tensile stresses than on

its susceptibility to water action. Despite this, the percentage of air voids affected the retained strength of the mix in the Water Sensitivity Test more than in the Immersion-Compression Test. Consequently, it seems that it could be more sensitive to moisture action.

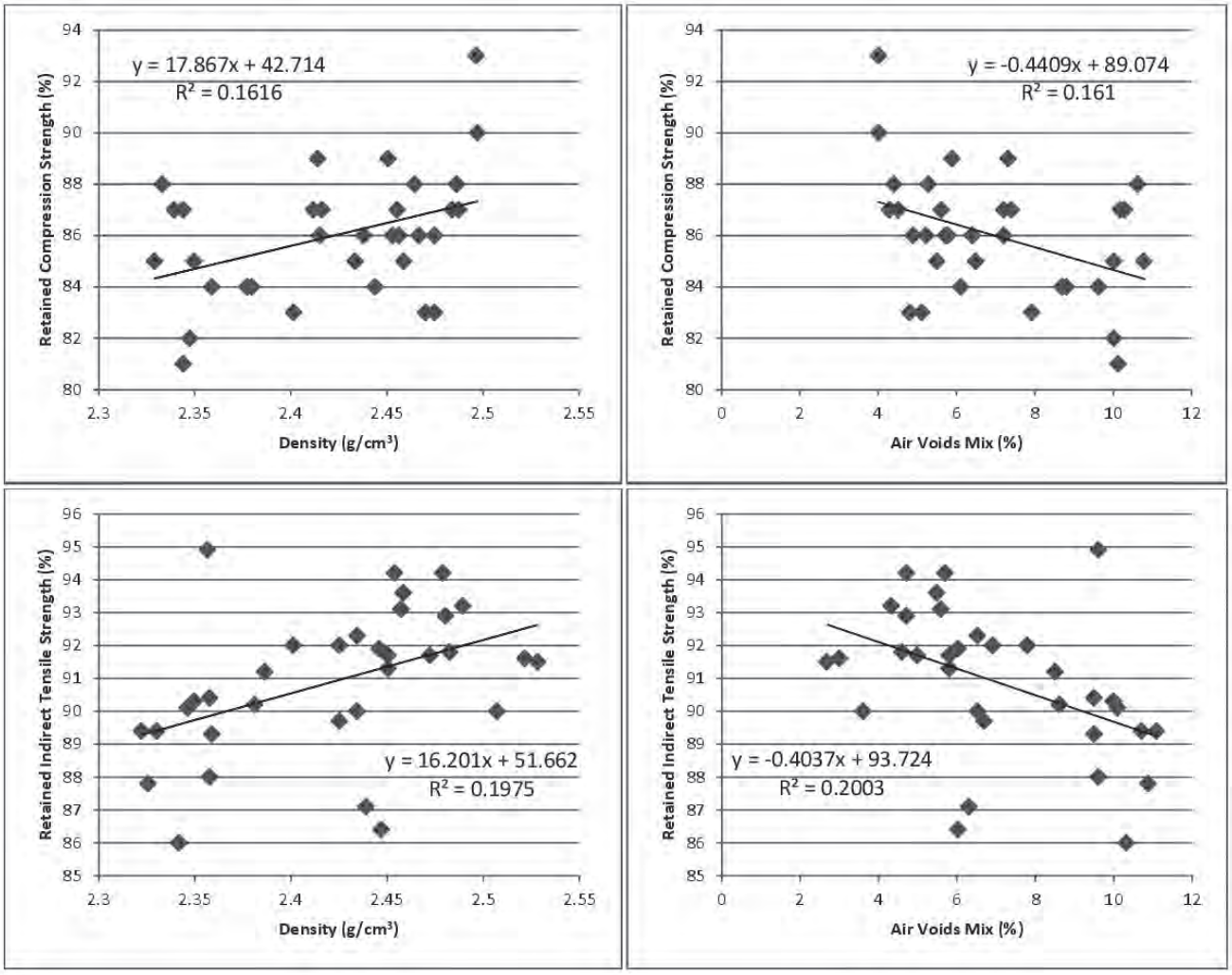


Figure 6.Retained strength based on the air void content and the density of the mix. Immersion-Compression Test (upper graph) and Water Sensitivity Test (lower graphs)

4. CONCLUSIONS

This paper shows the results obtained in a research study that analyzed the moisture susceptibility of hot bituminous mixes based on two different test methods. For this purpose, a comparative study of two laboratory tests was performed, which evaluated the effect of water action on the mechanical performance of the mix. The tests selected for this study were the Immersion-Compression Test and the Water Sensitivity Test. Although these tests differ in the size and compaction method of the test specimens (static load/compaction by blows), their wet conditioning (24 hours at 60ºC/72 hours at 40ºC), and the type of load applied (simple compression/indirect tensile stress).Both evaluated the same parameter, the retained strength value of the wet set and dry set of specimens. Thus, in order to guarantee the reliability of this analysis, we used the central limit theorem (which defines whether the sample population is sufficiently large to assume the normality of the results). Accordingly, the test was carried out on a total of 33 specimens made from bituminous mixes of identical characteristics. The type of mix selected was a standard semi-dense mix that is widely used in many

European countries. Based on the results obtained in our research, the following conclusions were derived:

- For the type of mix analyzed, the simple compression strength of the bituminous mixes was not affected by the density and air void content, but rather by the strength of the mineral skeleton and internal friction. Thus, the fact that the mix has voids where water could be stored was not a determining factor that affected its compression strength. Consequently, the type of load used in the Immersion-Compression Test was not suitable to measure the susceptibility of a bituminous mix to water since the presence of air voids did not affect its response.

- The indirect tensile strength was found to be directly related to the density and air void content of the mix. This means that the lower the density of the mix (or the greater its air void content), and the lower its indirect tensile strength. When the density of the mix became lower, this reduced the cohesion of the mix. The cohesion is what withstands the tensile stresses on the mix. Moreover, since the mix is a stone-like material, its strength against this type of load is not as great as its strength against simple compression loads. As a result, the mechanical performance of the mix is much more sensitive to this type of stress. Thus, indirect tensile stress is better for evaluating the possible moisture damage to the mixes.

- Even though the tests used different compaction methods for the specimens, this did not influence the characteristics of the mix since the samples in both tests produced similar results regarding their density and air voids.

- The type of conditioning given to specimens in the Immersion-Compression Test significantly affected their performance since the retained strength values obtained in this test were lower than those obtained in the Water Sensitivity Test. This seems to indicate that conditioning at high temperatures is more aggressive than conditioning for a longer duration. It was thus found that the effect of the temperature on the mix was greater than that of the duration of the conditioning process.

- The normal distribution obtained from the retained strength results of the Water Sensitivity Test had a sharper peak than that obtained from the Immersion-Compression Test. This indicated that the results had a lower dispersion, and thus, the reproducibility of the test was greater (which points to the fact that the results are more reliable).

- Both tests showed that when the mix had a greater air void content (and thus was less dense), its retained strength was also lower.

- The Immersion-Compression Test showed that these characteristics of mix did not affect its simple compression strength, but did have an impact on its retained strength. This was due to the fact that the increased air void content in the mix caused a larger volume of water to enter. This affected the mechanical performance of the mix and reduced its strength.

- The results of the Water Sensitivity Test showed that the air void content had a greater effect on the indirect tensile strength of the mixture than on its retained strength. Despite this, the retained strength was more closely related to the air void content than the retained strength of the Immersion-Compression Test. This was because the tensile stresses in bituminous mixes are withstood by their cohesion, which is the sum of the binder cohesion and of aggregate-binder adhesion. Both factors are directly affected by the presence of moisture (which breaks the chemical bonds in the bitumen, thus reducing its cohesion as well as its adhesion with the aggregate). For this reason, a larger number of air voids in the mix permits the presence of moisture that reduces the strength of the mix.

- The results obtained in this study point to the fact that the Water Sensitivity Test is more suitable for analyzing the susceptibility of bituminous mixes to water than the Immersion-Compression test. For this reason, the results obtained with the Water Sensitivity Test were more representative of the performance of the mix during its service life.

As a continuation of this research, it would be interesting to make a more in-depth study of the design of more effective laboratory tests for determining the susceptibility of bituminous mixes to moisture. For this purpose, it would be useful to perform further

comparative analyses with other laboratory tests and other types of bituminous mixes.

ACKNOWLEGDMENTS

This research has been funded by the following organisms and companies: Consejería de ObrasPúblicas de la Junta de Andalucía, Aldesa, Eiffage Infraestructuras, Martín Casillas, Ploder, Sacyr and Vera.

REFERENCES

[1] European Asphalt Pavement Association (EAPA). Asphalt in figures 2009. www.eapa.org

[2] Kraemer, C., Pardillo, J.M., Rocci, S., Romana, M., Sánchez, V. y del Val, M.A., Ingeniería de Carreteras. Vol. II. Editorial McGraw Hill. 2004.

[3] Reyes-Ortiz, O.J., Alvarez-Lugo, A.E. and Botella-Nieto, R., Study of a hot asphalt mixture response based on energy concepts. Dyna, Vol.78, N.168, pp. 45-52, 2011.

[4] Majidzadeh, K. and Brovild, F. N., Effect of water on bitumen- Aggregate Mixtures –State of the art. Special HRB Report, No. 98, Highway Research Board, 1968.

[5] Kennedy, T. W., Roberts, F. L. and Lee, K. W., Evaluation of moisture susceptibility of asphalt mixtures using the Texas Freeze-Thaw Pedestal Test. Proceedings of the Association of Asphalt Paving Technologists, Vol. 53, 1982.

[6] Tunnicliff, D. G. and Root, R. E., Antistripping Additives in Asphalt Concrete – State of the Art 1981. Proceedings of the Association of Asphalt Paving Technologists, Vol. 53, 1982.

[7] Hicks, R. G., Moisture damage in asphalt concrete. NCHRP Synthesis of Highway practice, Vol. 175, Transportation Research Board, October 1991.

[8] Rice, J. M., Relationship of aggregate characteristics to the effect of water on bituminous paving mixtures. ASTM STP 240, American Society for Testing and Materials, 1958.

[9] Barksdale, R. D., The Aggregate Handbook. National Stone Association, Washington, D. C., 1991.

[10] Sanderson, F. C., Methylchlorosilanes as Antistripping Agent. Proceedings of the Highway Research Board, Vol. 31, 1952.

[11] Bikerman, J. J., The rheology of adhesion. Rheology, Theory and Application, Vol. 3, 1960.

[12] Kanitpong, K. and Bahia, H. U., Role of adhesion and thin film tackiness of asphalt binders in moisture damage of HMA. Proceedings of the Association of Asphalt Paving Technologists, Vol. 72, 2002.

[13] Terrel, R. L. and Al-Swailmi, S., Water sensitivity of asphalt – Aggregate mixes: Test selection. Report SHRP-A-403. Strategic Highway Research Program, National Research Council, Washington, D. C., 1994.

[14] Fromm, J. H., The mechanisms of asphalt stripping from aggregate surfaces. Proceedings of the Association of Asphalt Paving Technologists, Vol. 43, 1974.

[15] Parr, W. K., Field observation of the behavior of bituminous pavements as influenced by moisture. Symposium on effect of water on bituminous paving mixtures, ASTM Special Technical Publication No. 240, pp. 3-16. 1958.

[16] Sha, Q. L., Observation and prevention of premature failures of asphalt pavements on freeways. People's communication publication house, China. 2001.

[17] Shatnawi, S. R., Premature AC Pavement Distress – District 2 Investigation (Final Report). Report Number FHWA/CA/TL-92-07, Office of Materials Engineering and Testing Services, California Department of Transportation, Sacramento, California, 1995.

[18] Lu, Q. and Harvey, J. T., Investigation of Conditions for Moisture Damage in Asphalt Concrete and Appropriate Laboratory Test Methods. Draft Research Report UCPRC-RR-2005-15, University of California, November 2005.

[19] Terrel, R.L. and Shute, J.W., Summary Report on Water Sensitivity. SHRP-A/IR-89-003, Strategic Highway Research Program, National Research Council, 1989.

[20] ASTM-Road and Paving Material; Vehicle-Pavement Systems. Annual Book of ASTM Standards, V4.03, 1998.

[21] AASHTO- Standard Specification for Transportation Materials and Method of Sampling and Testing, Part 3 edition, 1997.

[22] Epps, J. A., Sebaaly, P. E., Penaranda, J., Maher, M. R., McCann, M. B. and Hand, A. J., Compactability of a test for a moisture-induced damage with superpave volumetric mix design. NCHRP Report 444, Transportation Research Board

- National Research Council, Washington, D. C., 2000.

[23] Richard, P. I. and Maghsound, T., Use of Hamburg Wheel-Tracking Device for evaluating moisture susceptibility of hot-mix asphalt. Transportation Research Record, No. 1681, 1999.

[24] Federal Highway Administration: Superpave Asphalt Mixture Design, Version 8, 2002.

[25] Brown, R. E., Kandhal, S. P. and Zhang, J., Performance testing for hot mix asphalt. NCAT Report 01-05, National Center for Asphalt Technology, November 2001.

[26] Kim, S. H., Evaluation of hot mix asphalt moisture sensitivity using the Nottingham Asphalt Test equipment. MS thesis, Iowa State University, 2004.

[27] Stuart, K. D., Moisture Damage in Asphalt Mixtures-A State of Art Report. Report No. FHWA-RD-90-019, U.S. Department of Transportation, Federal Highway Administration. 1990.

[28] Solamanian, M., Harvey, J., Tahmoressi, M. and Tandon, V., Test method to predict moisture sensitivity of hot mix asphalt pavements. Moisture sensitivity of asphalt pavements: a national seminar, Topic 3. February, pp. 4-6, 2003.

[29] Dirección General de Carreteras: Pliego de Prescripciones Técnicas Generales para Obras de Carreteras y Puentes PG-3, Ministerio de Fomento, Ediciones Liteam, Madrid, 2010.

[30] Heffer, A. W., Adhesion in bitumen-aggregate systems and quantification of the effects of water on the adhesive bond. PhD Dissertation, Texas A&M University, December 2004.

[31] Walpole, R., Myers, R. y Myers, S., Probabilidad y estadística para ingenieros. 6th Edition.Prentice Hall, Pearson. Mexico, 1999.

[32] Montgomery, D. y Runger, G., Probabilidad y estadística aplicadas a la ingeniería. 3rd Edition. McGraw Hill. Mexico, 1996.

[33] Devore, J., Probabilidad y estadística para ingeniería y ciencias. 5th Edition. Thompson. Mexico, 2001.

PATELLOFEMORAL MODEL OF THE KNEE JOINT UNDER NON-STANDARD SQUATTING

MODELO PATELOFEMORAL DE LA ARTICULACIÓN DE LA RODILLA BAJO SENTADILLA NO ESTÁNDAR

GUSZTÁV FEKETE
Ph.D., Laboratory Soete, Ghent University, Zwijnaarde, Belgium, Gusztav.Fekete@ugent.be
BÉLA MÁLNÁSI CSIZMADIA
C.Sc., Institute of Mechanics and Machinery, Szent István University, Gödöllő, Hungary, Csizmadia.Bela@gek.szie.hu
MAGD ABDEL WAHAB
Ph.D., Laboratory Soete, Ghent University, Zwijnaarde, Belgium, Magd.AbdelWahab@ugent.be
PATRICK DE BAETS
Ph.D., Laboratory Soete, Ghent University, Zwijnaarde, Belgium, Patrick.DeBaets@ugent.be
LIBARDO V. VANEGAS-USECHE
Ph.D., Facultad de Ingeniería Mecánica, Universidad Tecnológica de Pereira, Pereira, Colombia, lvanegas@utp.edu.co
ISTVÁN BÍRÓ
Ph.D., Institute of Technology, Szeged University, Szeged, Hungary, biro-i@mk.u-szeged.hu

Received for review February 4 th, 2013, accepted August 12th, 2013, final version September, 16 th, 2013

ABSTRACT: The available analytical models for calculating knee patellofemoral forces are limited to the standard squat motion when the center of gravity is fixed horizontally. In this paper, an analytical model is presented to calculate accurately patellofemoral forces by taking into account the change in position of the trunk's center of gravity under deep squat (non-standard squatting). The accuracy of the derived model is validated through comparisons with results of the inverse dynamics technique.

Keywords: Patellofemoral forces, analytical model, modified squat, knee.

RESUMEN: Los modelos analíticos disponibles para calcular las fuerzas patelofemorales de la rodilla están limitados al movimiento estándar de cuclillas cuando el centro de gravedad es fijo horizontalmente. En este artículo se presenta un modelo para calcular con exactitud las fuerzas patelofemorales, teniendo en cuenta el cambio de posición del centro de gravedad del tronco durante una sentadilla profunda (posición de sentadilla modificada). Se valida la exactitud del modelo desarrollado mediante comparaciones con resultados de la técnica de dinámica inversa.

Palabras Clave: Fuerzas patelofemorales, modelo analítico, sentadilla modificada, rodilla.

1. INTRODUCTION

The widespread occurrence of various types of arthritis results in significant loss of manpower and immeasurable pain and suffering to many patients. Due to the limited understanding of the knee joint movement, clinicians have generally opted for a surgical solution that involves prosthetic replacement arthroplasty.

Analytical models related to the kinetics of the patellofemoral joint are still in short supply. Among the earliest models, Crowninshield et al. [1] and Andriacchi et al. [2] created quasi-static, analytical studies in order to determine the overall stiffness of the joint as a function of flexion angle. However, no investigation of the patellofemoral forces was actually carried out. The ratio between patellofemoral forces such as F_{pf}/F_q and F_{pt}/F_q (ratio of the patellofemoral compression force to quadriceps force and the patellar ligament force to quadriceps force, respectively) has been determined both by experimental methods [3] and with the help of various mathematical models [4, 5]. However, only the standard squat movement was investigated, where the forward movement of the trunk is not considered.

This work is part of a research that has been carried out at Szent István University and Ghent University, where analytical and experimental studies on the knee is being performed. It presents equations for the patellofemoral forces in the knee under non-standard squat, in which the centre of gravity moves both vertically and horizontally, thus the trunk movement during squatting is also considered.

2. EQUATIONS FOR PATELLOFEMORAL FORCES

The assumptions considered in deriving the equations of patellofemoral forces under standard squat are summarized as follows [6]: a) the model is quasi-static, b) it is two-dimensional, i.e., in the sagittal plane, c) contact forces are neglected, d) the femur and tibia are connected through a hinge with one degree of freedom, and e) the load is derived from the total weight of the person.

In this work, the movement of the trunk (i.e., the horizontal movement of the centre of gravity) is taken into account; therefore, the body weight vector (***BW***) can move vertically and horizontally. The movement of the centre of gravity is a known phenomenon; however, how much this movement alters the forces in the knee joint has not been investigated.

Furthermore, the movement of the femur and tibia are independent of each other (neither of them are fixed but can freely rotate during the squat). In this paper, the derived equations describe the forces connected to the femur, tibia, and patella under the non-standard squat as shown in Figure 1.

The arbitrary knee position, in Figure 1, is located at angle α. The ***BW*** force is derived from the body weight. The patella and the tibia are assumed to rotate about point *B*. The line of action ***BW*** intersects the theoretical line of the femur and tibia at points *D* and *E*. Rigid linkages represent the femur (3), the patella (2) and the tibia with foot (1).

The tibia is connected to the foot by a hinge of one degree of freedom (point *N*). The line of action of the centre of gravity intersects the femur at point *D* and the foot at point *A*. These points are not fixed, since the centre of gravity moves horizontally during a deep squat. At point *D*, a roller is assumed, which can move along the axis of the femur, while another roller is applied at point *A*, which can move along the axis of the foot. The rigid bodies are attached to each other by strings of negligible elastic elongation. The free-body diagrams of the three elements are shown in Figures 2,3 and 4.

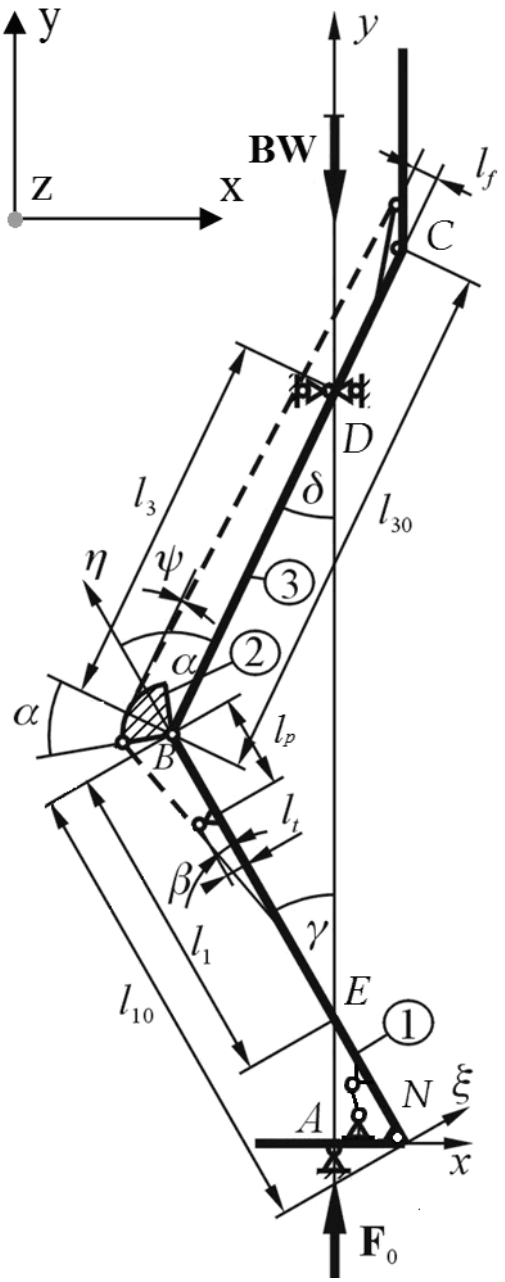


Figure 1. Analytical squat model

The model includes several constants and variables: the notations of the geometric lengths are listed in Table 1.

Table 1. Constants and variables of the model

DESCRIPTION	NOTATION
Length of tibia	l_{10}
Length of femur	l_{30}
Length of patellar tendon	l_p
Moment arm between the axis of tibia and the tibial tuberosity	l_t
Moment arm between the axis of femur and the line of action of the quadriceps force	l_f
Angle between the axis of femur and the quadriceps force	ψ
Intersected length of the axis of tibia and the instantaneous line of action of the ***BW***	l_1
Intersected length of the axis of femur and the instantaneous line of action of ***BW***	l_3
Angle between the axis of tibia and the patellar tendon	β
Angle between the axis of tibia and the line of action of ***BW***	γ
Angle between the axis of femur and the line of action of ***BW***	δ
Angle between the axis of tibia and the tibiofemoral force	φ

The moment equation applied about the z-axis through point *B* on the tibia (Figure 2) is:

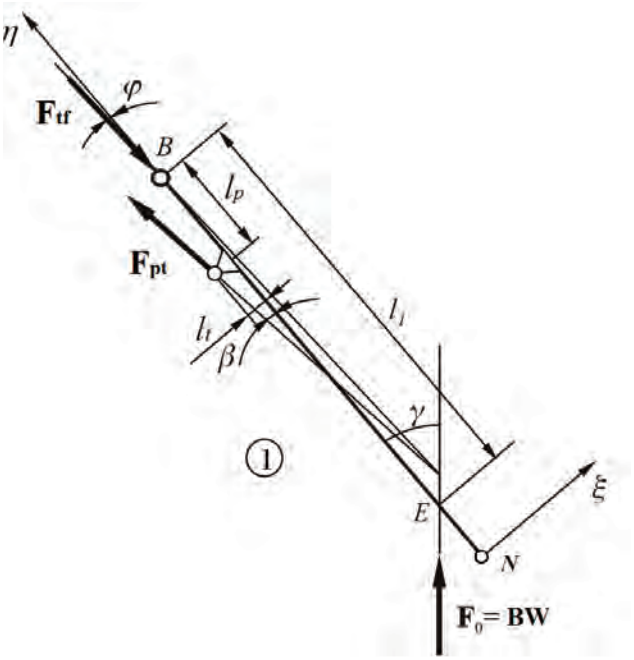


Figure 2. Free-body diagram of the tibia

$$\sum M_{B1z} = 0 = -l_p \cdot F_{pt} \cdot \sin\beta(\alpha) - l_t \cdot F_{pt} \cdot \cos\beta(\alpha) + l_1(\alpha) \cdot BW \cdot \sin\gamma(\alpha) \quad (1)$$

From Eq. (1), the patellar tendon force (F_{pt}) can be derived as:

$$F_{pt}(\alpha) = BW \cdot \frac{l_1(\alpha) \cdot \sin\gamma(\alpha)}{l_p \cdot \sin\beta(\alpha) + l_t \cdot \cos\beta(\alpha)} \quad (2)$$

In order to simplify the results, dimensionless parameters are introduced (Table 2).

Table 2. Dimensionless parameters of the model

DESCRIPTION	FORMULA
Dimensionless, intersected tibia length function	$\lambda_1(\alpha) = l_1(\alpha) / l_{10}$
Dimensionless, intersected femur length function	$\lambda_3(\alpha) = l_3(\alpha) / l_{30}$
Dimensionless length of patellar tendon	$\lambda_p = l_p / l_{10}$
Dimensionless thickness of shin	$\lambda_t = l_t / l_{10}$
Dimensionless thickness of thigh	$\lambda_f = l_f / l_{30}$

By the introduction of these quantities, the normalized force in the patellar tendon is:

$$\frac{F_{pt}(\alpha)}{BW} = \frac{\lambda_1(\alpha) \cdot \sin\gamma(\alpha)}{\lambda_p \cdot \sin\beta(\alpha) + \lambda_t \cdot \cos\beta(\alpha)} \quad (3)$$

The scalar equilibrium equations related to the ξ - η coordinate system (fixed to the tibia) are in the following form (Figure 2):

$$\sum F_{i\eta} = 0 = -F_{tf} \cdot \cos\phi(\alpha) + F_{pt} \cdot \cos\beta(\alpha) + BW \cdot \cos\gamma(\alpha) \quad (4)$$

$$\sum F_{i\xi} = 0 = F_{tf} \cdot \sin\phi(\alpha) - F_{pt} \cdot \sin\beta(\alpha) + BW \cdot \sin\gamma(\alpha) \quad (5)$$

Where F_{tf} is the tibiofemoral compression force and φ (α) is the angle between the axis of the tibia and the tibiofemoral force vector. By substituting Eq. (3) into Eqs. (4) and (5), F_{pt} is eliminated. After this, Eq. (4) is set to F_{tf}, and then it is substituted into Eq. (5). The substitution is followed by some additional simplifications, and finally the angle φ (α) can be

derived as:

$$\phi(\alpha) = arctg\left[\frac{\left(\lambda_1(\alpha)-\lambda_p\right)\cdot tg\beta(\alpha)-\lambda_t}{\lambda_1(\alpha)\cdot tg\gamma(\alpha)+\lambda_p\cdot tg\beta(\alpha)+\lambda_t}\cdot tg\gamma(\alpha)\right] \quad (6)$$

By the use of $\varphi(\alpha)$, the tibiofemoral force can be derived from Eq. (4) or Eq. (5) as:

$$\frac{F_{tf}(\alpha)}{BW} = \frac{F_{pt}}{BW}\cdot\frac{\cos\beta(\alpha)}{\cos\phi(\alpha)} + \frac{\cos\gamma(\alpha)}{\cos\phi(\alpha)} \quad (7)$$

After deriving the F_{pt} and F_{tf} forces, Figure 3 is considered.

Applying the moment equilibrium equation about the *z* axis through point *B* on the femur (Figure 3) gives:

$$\sum M_{B3z} = 0 = l_f\cdot F_q\cdot\cos\psi(\alpha) + l_{30}\cdot F_q\cdot\sin\psi(\alpha) - l_3(\alpha)\cdot BW\cdot\sin\delta(\alpha) \quad (8)$$

Taking into account that $\delta = \alpha - \gamma$, and assuming $\psi \approx 0$, from Eq. (8), the quadriceps force in the tendon becomes:

$$\frac{F_q(\alpha)}{BW} = \frac{\lambda_3(\alpha)\cdot\sin\left(\alpha-\gamma(\alpha)\right)}{\lambda_f} \quad (9)$$

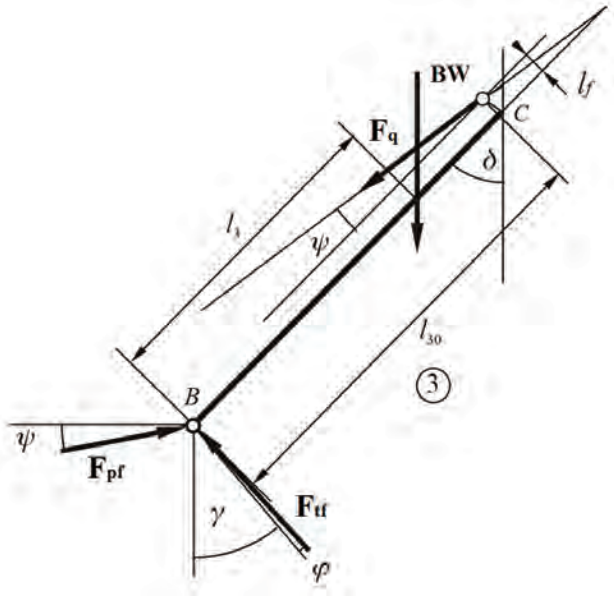


Figure 3. Free-body diagram of the femur

The scalar equilibrium equations related to the patella in the *x* - *y* coordinate system (Figure 4) are:

$$\sum F_{ix} = 0 = F_q(\alpha)\cdot\sin\delta(\alpha) + F_{pt}(\alpha)\cdot\sin\left(\gamma(\alpha)+\beta(\alpha)\right) + F_{pf_x} \quad (10)$$

$$\sum F_{iy} = 0 = F_q(\alpha)\cdot\cos\delta(\alpha) - F_{pt}(\alpha)\cdot\cos\left(\gamma(\alpha)+\beta(\alpha)\right) + F_{pf_y} \quad (11)$$

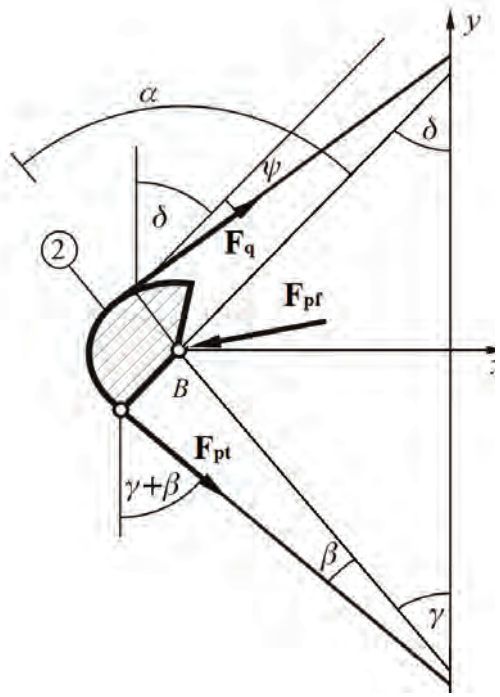


Figure 4. Free-body diagram of the patellofemoral joint

Where F_{pfx} and F_{pfy} are the components of the patellofemoral compression force in the *x* and *y* directions, respectively.

From Eq. (9) and Eq. (10), the magnitude of the patellofemoral compression force is:

$$\frac{F_{pf}(\alpha)}{BW} = \frac{\sqrt{F_{pf_x}^2 + F_{pf_y}^2}}{BW} = \frac{\sqrt{\left(F_q(\alpha)\right)^2 + \left(F_{pt}(\alpha)\right)^2 - 2F_q(\alpha)\cdot F_{pt}(\alpha)\cdot\cos(\Omega)}}{BW} \quad (12)$$

Where .

$$\Omega = \beta(\alpha) + \delta(\alpha) + \gamma(\alpha)$$

For a given ***BW*** force and angle α, Eqs. (3), (7), (9), and (12) can be used to determine $F_{pt}(\alpha)$, $F_{tf}(\alpha)$, $F_q(\alpha)$, and $F_{pf}(\alpha)$, respectively.

The parameters appearing in the above mentioned equations have been experimentally determined [7, 8] and summarized in Table 3.

Table 3. Dimensionless parameters of the model

	C1	C2	SD	r^2
$\lambda_1(\alpha)$ [-]	0.492	0.0024	0.15	0.65
$\lambda_3(\alpha)$ [-]	0.86	-0.0022	0.22	0.63
$\beta(\alpha)$ [°]	26.56	-0.3861	14	0.95
γ/α (α) [-]	0.567	-0.0026	0.081	0.735
λ_t [-]	0.11	0	0.018	-
λ_p [-]	0.1475	0	0.043	-
λ_f [-]	0.164	0	0.028	-

3. RESULTS AND DISCUSSION

As it has been proven by other authors [5, 6], the patellofemoral forces directly depend on the net knee moment; this is a moment about the instantaneous centre of rotation of the knee joint generated by the body weight, in the case of the standard (or non-standard) squat.

Therefore, it is interesting to see how this moment depends on the position of the centre of gravity. By the use of inverse dynamics approach [9, 11, 14] all the movements of the human body can be taken into consideration, thus the effect of the centre of gravity may also be considered. By knowing (measuring) the kinematics of a person during non-standard squat, the measured forces will involve the effect of the moving centre of gravity as well. For this reason, the results are best compared to the results of inverse dynamics method.

In Figure 5, the quadriceps tendon force of the non-standard squat model corresponds well with the inverse dynamics result of Sharma et al. [9] found in the literature.

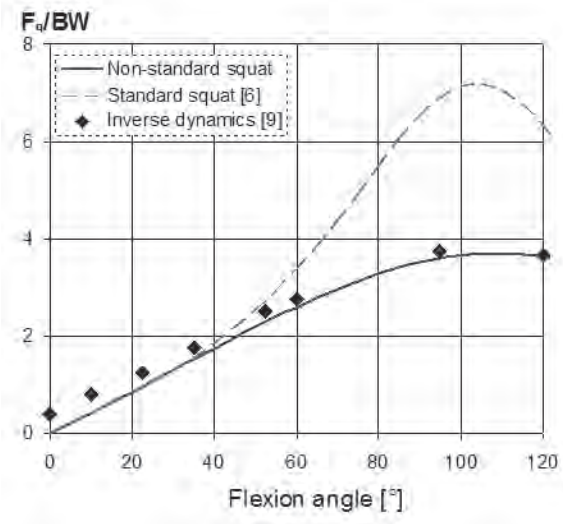


Figure 5. Quadriceps force

The peak force of the non-standard squat model is estimated to be 3.63 ***BW*** at 120˚, while the standard squat model predicts a peak magnitude of 7.2 ***BW*** and peak location between 90˚ and 100˚ of flexion angle.

In practice, human subjects lean forward during squatting, which, apart from helping them to keep their balance, also reduces the patellofemoral forces. This is the reason why every experimental, analytical, or numerical model, which does not incorporate the moving of the centre of gravity tends to overestimate the net knee moment, and results in higher forces in the quadriceps (and in the other muscles or tendons).

All the same, this parameter has only been investigated by Kulas et al. [10], therefore, up to now there has been a very limited amount of information about how the horizontal movement of the centre of gravity influences the patellofemoral forces.

In Figure 6, the patellar tendon force is plotted. The correlation between the standard [6] and non-standard models regarding this force is strong. Their characteristics, magnitudes, and peak locations are in good agreement with each other. According to these corresponding results, the estimated peak force is 6.8 ***BW*** and the peak location is at 120˚ of flexion angle. It has to be mentioned that this is the only force whereupon the movement of the centre of gravity had a relatively small effect.

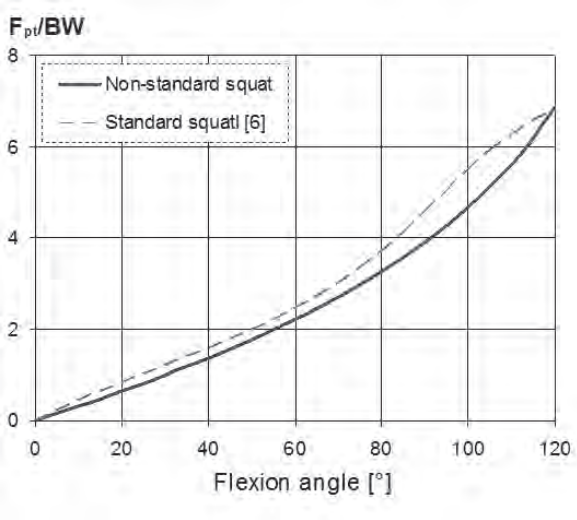


Figure 6. Patellar tendon force

In Figure 7, the patellofemoral compression force is plotted. By considering the plotted results, the non-standard squat model correlates with the results of Escamilla et al. [11] and Churchill et al. [12]. The estimated peak angle of the non-standard squat model, in this case, is located around 110˚ of flexion angle and the peak force is approximately 3.6 ***BW***. However, it should be noted that the result of Escamilla et al. [11] was carried out only up to a 90˚ flexion angle. If we compare the standard squat results with the results provided by the inverse dynamics method and the non-standard squat model, the significant difference related to this force becomes quite apparent.

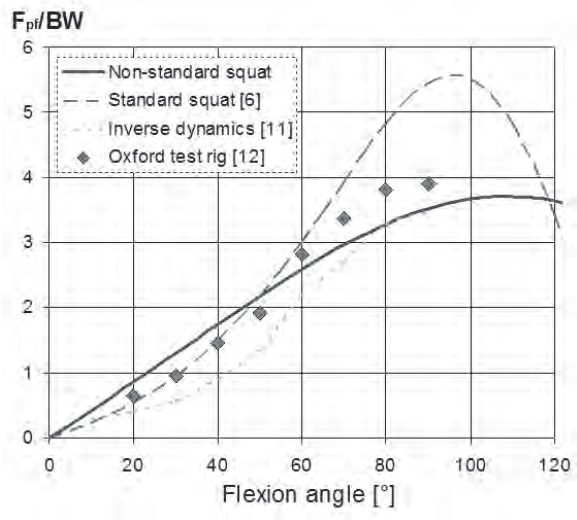


Figure 7. Patellofemoral compression force

The standard squat model is created with similar boundary conditions as an Oxford-type test rig [6], therefore, the apparent difference between the standard and non-standard squat is likely due to this type of modelling technique: according to this method the quadriceps muscles are detached from the femur, and the knee capsule is removed, thus the entire ***BW*** is applied solely through the quadriceps tendon and patellar tendon. Therefore the entire knee moment, due to the ***BW*** is supported by the patella, and the resulting loads may be higher than physiological loads [6]. Further, during a squat motion, the human subject can lean forward to move the centre of mass forward and tilt the hip reducing the joint moment and can therefore lower the patellofemoral force significantly. The standard squat approach assumes that the subject stays perfectly vertical maximizing the knee moment and maximizing the quadriceps force. Finally, the knee moments in Oxford rig studies are often lower than physiological knee moments. The values reported using inverse dynamics method do not suffer from these factors.

In Figure 8, the tibiofemoral force is presented. The standard squat model [6] is not able to predict this force, thus no comparison could be carried out between the two analytical models.

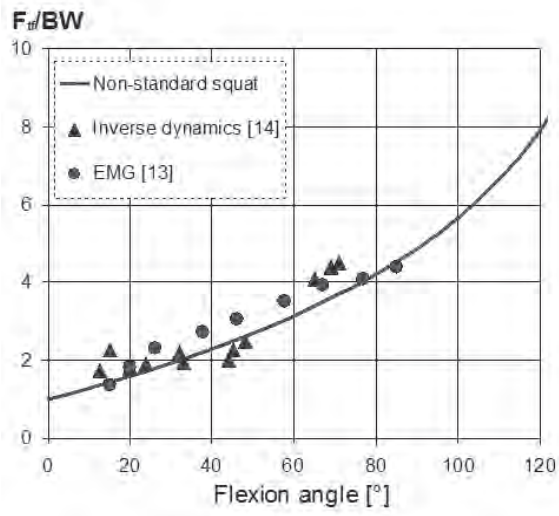


Figure 8. Tibiofemoral compression force

The new non-standard squat model was compared to the results of Zheng et al. [13] and Steele et al. [14]. As it may be observed, the three results are in agreement,

although the experimental result of Zheng et al. [13] and Steele et al. [14] provide predictions only until 90˚ and 70˚ of flexion angle. Here, the peak force is estimated as 7.8 ***BW***.

In order to determine the influence of moving the centre of gravity in numbers, the patellofemoral forces have been calculated and compared as a percentage difference between the standard and non-standard squat:

$$\Delta K = \left(1 - \frac{K_{\text{non-standard}}}{K_{\text{standard}}}\right) \cdot 100 \qquad (13)$$

Where, K can be any quantity (force, moment or displacement). ΔK can provide a percentage difference of a standard quantity compared to a non-standard quantity (here standard and non-standard relates to the squat motion). The obtained results were summarized in Table 4.

The incorporation of the trunk movement significantly lowers the patellofemoral forces (up to 46%) along the calculated domain. This lowering effect on the patellofemoral forces (average 20%) corresponds very well with the result of Kulas et al. [10] who also investigated the effect of moderate forward trunk lean condition and observed 24% lower peak forces in the anterior cruciate ligaments (ACL).

Table 4. Percentage difference between Standard and Non-standard Squat

FLEXIONANGLE	ΔF_q	ΔF_{pf}	ΔF_{pt}
30°	>1%	-27%	20%
60°	23%	13%	10%
90°	46%	37%	15%
120°	42%	6%	>1%

Although, no direct measurement was performed to validate the obtained results, the comparison between the current predictions and the inverse dynamics data found in the literature show appropriate agreement.

The most important feature, which has to be underlined, is the possibility to include the movement of the trunk into the model.

4. CONCLUSIONS

In summary, a new analytical model is presented, which draws the attention to the effect of the movement of the centre of gravity on the knee joint kinetics. This model is capable of accurately predicting the patellofemoral, tibiofemoral, patellar ligament, and quadriceps forces in the knee during standard or non-standard squatting motion.

The new analytical model is derived from theoretical assumptions and experimentally determined parameters based on multiple human participants. The results of the analytical model correlated well with inverse dynamic results taken from the literature.

This new model is limited to the description of the kinetics of squatting motion. More specifically, the model shows how the horizontal movement of the centre of gravity influences the patellofemoral forces, since this phenomenon has not yet been investigated thoroughly by any other author. In addition, the model may be used to easily investigate other types of squat (depending on the λ functions), while the inverse dynamics method requires a measuring system and programs to determine the forces.

Through the modelling approach, together with the creation of the necessary equations, similar modelling issues become more understandable and solvable.

Among the patello- and tibiofemoral forces, the obtained $F_q(\alpha)$ force function can be extended for further use as an input function for isometric motion, since most descriptive relationships [3, 4] found in the literature provide only the ratio of the patellofemoral forces divided by the quadriceps force.

As a further step, a numerical model will be built in MSC.ADAMS using a moving centre of gravity, since the necessity of this parameter is proven by this analytical model. In addition to this feature, the MSC. ADAMS model will be used to investigate the contact forces (friction and normal) between the connecting surfaces and the sliding-roll ratio. These quantities will be given as a function of flexion angle.

ACKNOWLEDGEMENTS

The study was supported by the FWO (project number: G022506), the Universiteit Gent – Labo Soete, and the Szent István University – Faculty of Mechanical Engineering and Mechanical Engineering Ph.D. School.

REFERENCES

[1] Crowninshield, R., Pope, M.H. and Johnson, R.J., An analytical model of the knee. Journal of Biomechanics, 9 (6), pp. 397-405, 1976.

[2] Andriacchi, T.P., Mikosz, R.P., Hampton, S.J. and Galante, J.O., Model studies of the stiffness characteristics of the human knee joint. Journal of Biomechanics, 16 (1), pp. 23-29, 1983.

[3] Miller, R.K., Murray, D.W., Gill, H.S., O'Connor, J.J. and Goodfellow, J.W., In vitro patellofemoral joint force determined by a non-invasive technique. Clinical Biomechanics, 12 (1), pp. 1-7, 1997.

[4] Van Eijden, T.M.G.J., Kouwenhoven, E., Verburg, J. and Weijs, W.A., A mathematical model of the patellofemoral joint. Journal of Biomechanics, 19 (3), pp. 219-229, 1986.

[5] Cohen, Z.A., Roglic, H., Grelsamer, R.P., Henry, J.H., Levine, W.N., Mow, V.C. and Ateshian, G.A., Patellofemoral stresses during open and closed kinetic chain exercises – An analysis using computer simulation, The American Journal of Sports Medicine, 29 (4), pp. 480-487, 2001.

[6] Mason, J.J., Leszko, F., Johnson, T. and Komistek, R.D., Patellofemoral joint forces. Journal of Biomechanics, 41 (11), pp. 2337-2348, 2008.

[7] Fekete, G., Csizmadia, B.M., Wahab, M.A. and De Baets, P., Experimental determination of horizontal motion of human center of gravity during squatting, Experimental Techniques, doi:10.1111/ j.1747-1567.2011.00768.x.

[8] Fekete, G., Kinetics and kinematics of the human knee joint under standard and non-standard squat movement. PhD Thesis, Ghent University, pp. 101-133, 2013.

[9] Sharma, A., Leszko, F., Komistek, R.D., Scuderi, G.R., Cates, H.E. and Liu, F., In vivo patellofemoral forces in high flexion total knee arthroplasty. Journal of Biomechanics, 41 (3), pp. 642-648, 2008.

[10] Kulas, A.S., Hortobágyi, T. and DeVita, P., Trunk position modulates anterior cruciate ligament forces and strains during a single-leg squat. Clinical Biomechanics, 27 (1), pp. 16-21, 2012.

[11] Escamilla, R.F., Zheng, N., MacLeod, T.D., Edwards, W.B., Hreljak, A., Fleisig, G.S., Wilk, K.E., Moorman III, C.T. and Imamura, R., Patellofemoral compressive force and stress during the forward and side lunges with and without stride. Clinical Biomechanics, 23 (8), pp. 1026-1037, 2008.

[12] Churchill, D.L., Incavo, S.J., Johnson, C.C. and Beynnon, B.D., The influence of femoral rollback on patellofemoral contact loads in total knee arthroplasty. Journal of Arthroplasty, 16 (7), pp. 909-918, 2001.

[13] Zheng, N., Fleisig, G.S., Escamilla, R.F. and Barrentine, S.W., An analytical model of the knee for estimation of internal forces during exercise. Journal of Biomechanics, 31 (10), pp. 963-967, 1998.

[14] Steele, K.M., DeMers, M.S., Schwartz, M.H. and Delp, S.L. Compressive tibiofemoral force during crouch gait. Gait & Posture, 35 (4), pp. 556-560, 2012.

IDENTIFYING DEAD FEATURES AND THEIR CAUSES IN PRODUCT LINE MODELS: AN ONTOLOGICAL APPROACH

IDENTIFICANDO CARACTERÍSTICAS MUERTAS Y SUS CAUSAS EN MODELOS DE LÍNEAS DE PRODUCTOS: UN ENFOQUE ONTOLÓGICO

GLORIA-LUCIA GIRALDO
PhD, Associate Professor, Universidad Nacional de Colombia Sede Medellín, Colombia, glgiraldog@unal.edu.co

LUISA RINCÓN-PEREZ
Master of Engineering- Engineering Systems, Universidad Nacional de Colombia Sede Medellín, Colombia, lufrinconpe@unal.edu.co

RAUL MAZO
PhD, Associate Professor, CRI, Université Paris 1 Panthéon Sorbonne, raul.mazo@univ-paris1.fr

Received for review December 30 th, 2012, accepted August 22th, 2013, final version September, 23 th, 2013

ABSTRACT: Feature Models (FMs) are a notation to represent differences and commonalities between products derived from a product line. However, product line modelers could unintentionally incorporate dead features in FMs. A dead feature is a type of defect, which implies that one or more features are not present in any product of the product line. Some authors have used ontologies in product lines, but they have not exploited ontology reasoning to identify and explain causes for defects in FMs in natural language. In this paper, we propose an ontology that represents FMs in OWL (Web Ontology Language). Then, we use SQWRL (Semantic Query-enhanced Web Rule Language) to identify dead features in a FM and identify and explain certain causes of this defect in natural language. Our preliminary empirical evaluation confirms the benefits of our approach.

Key words: Product lines, feature models, ontologies, dead features, SQWRL.

RESUMEN: Los modelos de características (en inglés *Feature Models FMs*) son una notación para representar diferencias y similitudes entre productos derivados de una línea de productos. Sin embargo, quienes modelan la línea de productos pueden introducir sin intención en los *FMs* defectos como las características muertas. Una característica es muerta si no puede estar presente en ningún producto derivado de la línea de productos. Algunos autores han identificado características muertas en los FMs, pero ninguno ha aprovechado las capacidades de razonamiento de las ontologías para identificar y explicar las causas de estos defectos en lenguaje natural. En este trabajo, se propone una ontología para identificar las características muertas en un FM y se proponen consultas sobre la ontología, para identificar y explicar en lenguaje natural ciertas causas de las características muertas detectadas. Nuestra evaluación empírica preliminar confirma los beneficios de nuestra propuesta.

Palabras clave: Líneas de productos, modelo de características, ontologías, características muertas, SQWRL.

1. INTRODUCTION

A product line is a family of related products distinguished by different sets of features that each product provides [1]. A particular application of product line is the software product line (SPL). Software Product Line Engineering (SPLE) is thus the software development paradigm geared for the construction of SPLs. Extensive research and industrial experience widely prove the significant benefits of SPLE practices, which among them are reduced time to market, increased asset reuse and increased software quality [2]. In order to do that, SPLE usually uses Product Line Models (PLMs) to represent the correct combination of features that represent valid products.

A common notation to represent PLMs is Feature Models (FMs). FMs describe the features, their relations and the valid feature combinations of a product line [3]. FMs have also proven useful to communicate with customers and other stakeholders, such as marketing representatives, managers, production engineers, system architects, etc. Consequently, having FMs that correctly represent the domain of the product line is of paramount importance for ensuring quality in SPLE. However, creating feature models with features that correctly represent the domain described by the model is not trivial [4]. In fact, when a FM is constructed, defects may be unintentionally introduced, which decreases the quality of the FM and hence also the

expected benefits of product line. Dead features are one such defect. A feature is dead if it cannot appear in any products of the product line [5–8].

Some studies in the literature automatically identify whether a FM present dead features or not [5, 9–11]. However, few studies have focused on identifying causes for such defect [12, 13]. Identifying the cause consist in identifying the dependencies that, combined in a certain manner, produce a dead feature. Such identification helps product line engineers to understand the problem and to determine the best solution to fix dead features [4, 14]. In addition, in an end-users configuration process, it is important to identify defects and explain the cause of these defects to users [15].

FMs and ontologies are comparable because both represent concepts of a particular domain and their dependencies [16]. However, FMs only offer a graphical means to represent a particular domain, whereas ontologies also offer an efficient mechanism to reason on domain models.

In this paper, we discuss an approach based on OWL-DL (*Web Ontology Language–Description Logic*) [17] ontology and SQWRL (*Semantic Query-enhanced Web Rule Language*) [18]. The ontologies are formal domain models, which have powerful inference mechanisms. The ontologies are recommended for sharing terminologies and understanding. Therefore, modeling with ontologies offers interoperability, reusability and extensibility. We represent our ontology in OWL-DL because this formalism provides computational completeness and expressiveness for representing knowledge [19]. Moreover, we use a rule-based language because the rules are a natural and declarative way to represent knowledge [20]. SQWRL is a rule-based language to extract information from OWL ontologies. Therefore, we use this language to identify dead features in FMs, and some of their causes from the proposed ontology.

Our proposal has two main contributions. First, it provides an ontology that represents FM concepts; second, it identifies and explains in natural language certain causes that produce dead features.

The remainder of the paper is as follows. In Section 2, we give a brief overview of the necessary concepts for understanding the study described herein. Section 3 presents our approach to identify the causes for dead features in FMs. In Section 4, we present the preliminary validation of our proposal. In Section 5, we present related research. Finally, Section 6 presents the conclusions and suggests future research directions.

2. GENERAL CONCEPTS

Feature Models

Feature Models (FMs) are a notation for representing product line models. Using this notation, a feature is a distinctive element, which is visible to users. Each feature is a node in a tree structure, and the model dependencies are arcs [3].

The tree structure represents hierarchical organization of the features, wherein only one feature is the model root feature. In addition, except for the root, each feature has a parent feature [3]. Figure 2 shows a FM, which exemplifies the application of our proposal. Features can have different types of dependencies. Table 1 describes and graphically represents each type of dependency.

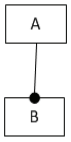
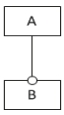
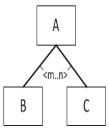
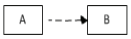
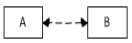
After Kang et al. [3] reported a first notation of FMs, other authors proposed extensions to the original notation [21] (e.g., the group cardinality dependency [22]).

Figure 1 shows an UML-based meta-model for a cardinality base FM. This meta-model relates the concepts presented in Table 1.

of *Utility Functions*, *Settings*, and *Media* features, among others. As shows Figure 2, each child feature in the *Media* feature is optional. Additionally, each child feature in the *Utility Functions* and *Settings* features are mandatory. Each dependency connects two features with a unique nomenclature for easy identification.

In order to illustrate our approach, we intentionally introduced in the original model four dead features (*MSN, Camera, VGA and Megapixels*). For that purpose, we use two additional dependencies (R16 and R17).

Table 1. Types of dependencies in FMs

Notation	Type of Dependency
	**Mandatory** [3] Child feature B should be included in all valid products containing the parent feature A and vice versa. It a feature is mandatory and all its ancestors are also mandatory, then, this feature is a full-mandatory feature [5].
	**Optional** [3] Child feature B may or may not be included in valid products containing parent feature A. However, if feature B is included in a product, its parent A should be included too.
	**Group cardinality** [22] Represents the minimum (m) and the maximum (n) number of child features (B...C) grouped in a cardinality (<m..n>) that a product can have when the parent feature (A) is included in the product. If at least one of the child features is included into a product, the parent feature should be included too.
Transverse Dependencies	
	**Requires** [3] Feature B should be included in valid products with feature A. This dependency is unidirectional.
	**Exclude**s [3] Features A and B cannot be in valid products at the same time. This dependency is bidirectional.

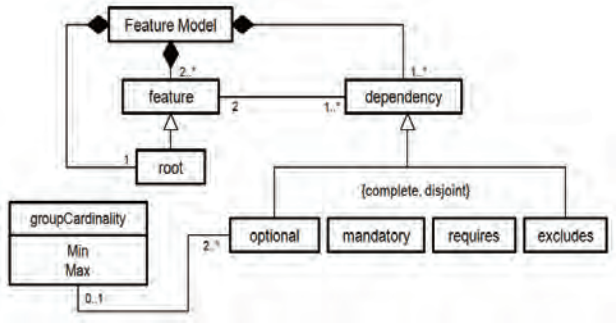


Figure 1. FM meta-model based on the one proposed by Mazo *et al.* [15]

2.2. Application Example

Figure 2 shows a reduced version of a FM based on the one proposed by Segura for mobile phones [23]. In this example, a *Mobile Phone* is composed

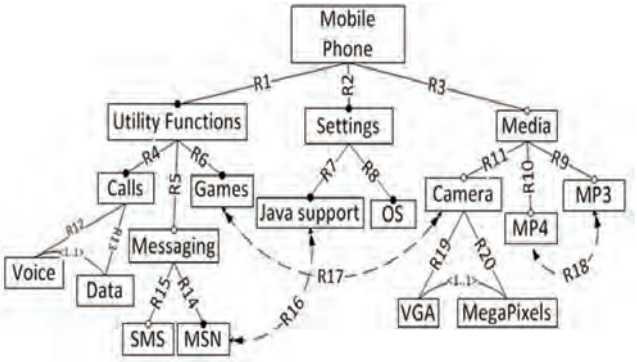


Figure 2. Summary of a mobile phone FM based on Segura's proposal [23]

2.3. Dead features

Features are distinctive elements that are visible to user [3]. A feature is dead when it is not present in any product of the product line [5–8]. When a FM has dead features, the model is not an accurate representation of the domain. Indeed, if a feature belongs to a FM, the feature is important for the domain that we want to represent. Therefore, it should be possible to incorporate that feature in at least one product of the product line [7].

2.4. Ontologies

An ontology is a formal explicit specification for a shared conceptualization [24, 25]. In the same way that FMs, ontologies help to identify and define the domain basic concepts and the dependencies among them.

Representing information with ontologies aids the identification and definition of the basic terms of a domain. In addition, ontologies represent the dependencies and rules for combining such terms, and provide a common vocabulary for the domain model. Ontologies comprises classes, instances, properties and constraints [26].

Classes are the main concepts related to the ontology domain. Instances represent objects in the domain of interest. Properties are object properties or data-type properties: Object properties relate ontology instance among them, whereas data-type properties relate ontology instances with concrete values, for example, an integer value. Finally, constraints describe the restrictions that instances must satisfy in order to belong to a class [26].

2.5. SQWRL Queries

The Semantic Query-enhanced Web Rule Language (SQWRL) is a language that provides query operations for ontologies represented in OWL [18]. A SQWRL query comprises an antecedent and a consequent expressed in terms of OWL classes and properties. The antecedent defines the criteria that instances must satisfy to be selected, and the consequent specifies the instances to select in the query results [18]. each SQWRL uses classes and properties defined in the proposed ontology to query for information of the FM represented as ontology instances. A semantic reasoner, such as JESS (Java Expert System Shell) [27], executes SQWRL queries.

3. PROPOSED SOLUTION

In the previous section, we described the basic concepts underlying our work. Following sub-sections present our approach, which uses ontologies and SQWRL to identify certain dead features in FMs, and to explain their causes in natural language.

3.1. Ontology-based representation of product line models

Figure 3 shows the proposed OWL ontology to represent the FMs concepts as an ontology. This representation allows us to exploit the semantic relationships among the concepts involved in FMs. For instance, we can ask for features that have the same parent, or features related by mandatory and exclude dependencies at the same time. We develop this ontology using the methodology proposed by Noy Noy & McGuinness (2001)and McGuinness [28] with a top-down approach, and we take the FM concepts from the meta-model presented in Figure 1.

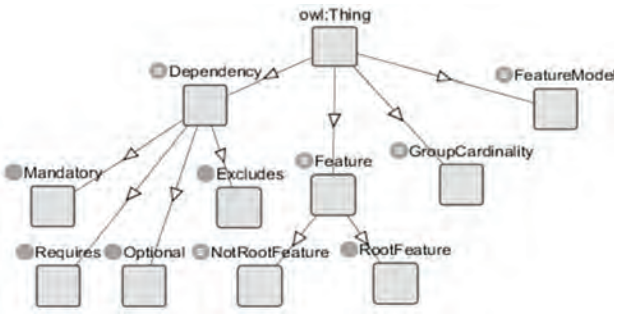


Figure 3. Proposed ontology to represent FMs

In our ontology, meta-model classes correspond to classes of the ontology. In addition, we separate the class *Feature* in classes *RootFeature* and *NotRootFeature* to represent that a FM only has one root feature. We represent the attributes of the *groupCardinality* meta-model class with ontology datatype properties, and we relate ontology classes with ontology object properties. For example, every dependency within the FM comprises an origin and a destination feature. Then, we create the object properties *hasDependencySource* and *hasDependencyDestination* to relate the *Dependency* ontology class with the *Feature* ontology class. Thus, we can relate each dependency with its feature source and its feature destination. Furthermore, in the optional and mandatory dependencies, the property *hasDependencySource* identifies the parent feature, and the property *hasDependencyDestination* identifies the child feature. In the example shown in Figure 2, the *Mobile Phone* feature is the origin feature of the optional dependency *R1*, and the *Utility Functions* feature is the destination feature. Moreover, *Mobile Phone* is the parent feature of the *Utility Functions* feature.

3.2. Rules for identifying dead features

According to the literature [4, 7, 13], misuse of dependency in FMs causes dead features. Our proposal considers that a feature can become a dead feature if a full-mandatory feature excludes an optional feature (see Rule 1), or if the parent feature is a dead feature (see Rule 2). Other cases that cause dead features are outside the scope of this initial proposal.

***Rule 1: Full-mandatory feature excludes an optional feature*:** An optional feature becomes dead when a full-mandatory feature excludes it.

In the example, the *Camera* feature is optional due to its dependency (R11) with the *Media* feature. Furthermore, product cannot have the *Games* and *Camera* features simultaneously due to the exclude dependency (R17) between both features. Because *Games* is a full-mandatory feature [5] (i.e., it is present in all products), the *Camera* feature is a dead feature.

***Rule 2: The parent feature is a dead feature*:** If a child feature is included during product configuration, the parent feature should be included too [3]. If the parent

feature is already a dead feature, its children features cannot be included in any product. In the example, features *VGA* and *MegaPixels* are children of dead feature *Camera*. Thus, these children features are dead features too.

We use SQWRL to implement the rules proposed in this section. For the sake of space, we only present and explain in the Table 2 the source code of the Rule 1, in which full- mandatory features exclude an optional feature. Nevertheless, both rules have a similar structure.

It is important to highlight that in queries, we use `WILDCARD` word as an argument that depends on each rule (e.g. in rule 1 `WILDCARD` belongs to full-mandatory features, while in rule 2 it belongs to dead features). Each statement SQWRL requires a constant value in the argument `WILDCARD.` Therefore, we create dynamically a SQWRL for each possible value of `WILDCARD.` For instance, `WILDCARD` can take seven different values in our application example (see Table 2); hence, we create seven different SQWRL queries.

Table 2. SQWRL query for dead features that satisfy Rule 1.

<table>
<tr><td colspan="3">(1)Excludes(?y) ^
(2)Optional(?w) ^
(3) NotRootFeature(?x) ^
(4) NotRootFeature(WILDCARD) ^
(5)hasDependencySource(?y, WILDCARD)^
(6)hasDependencyDestination(?y,?x) ^ (7)hasDependencyDestination(?w,?x)-> (8)sqwrl:selectDistinct(?x)</td></tr>
<tr><td colspan="3">Rule 1: Consequent result</td></tr>
<tr><td colspan="3">selectDistinct(?x):Optional feature, which is excluded by a full-mandatory feature.
Example Value:Camera ,MSM</td></tr>
<tr><td colspan="3">Rule 1: Antecedent construction</td></tr>
<tr><td>SQWRL instruction</td><td>Definition</td><td>Example Value</td></tr>
<tr><td>Excludes(?y)</td><td>Excludes dependencies</td><td>R16,R17,R18</td></tr>
<tr><td>Optional(?w)</td><td>Optional dependencies</td><td>R3,R9,R15,R19,R20</td></tr>
<tr><td>NotRoot Feature (?x)</td><td>Features non-root of the FM</td><td>Utility Functions, Settings, Calls, Messaging, Games, Java support, OS, Media, MP3,MP4,
Camera,Voice, Data,SMS,MSM,VGA,
Megapixels</td></tr>
<tr><td>NotRoot Feature (WILDCARD)</td><td>In this rule, WILDCARD correspond to full-mandatory features</td><td>Utility Functions, Settings, Calls, Messaging,Games, Java support, OS</td></tr>
<tr><td>has Dependency Destination (?w,?x)</td><td>Data are restricted, so x corresponds to features destination of optional dependencies</td><td>Value of x
Media, MP3, Camera,Voice,Data,SMS, MSM,VGA, Megapixels</td></tr>
<tr><td>has Dependency Source (?y, WILDCARD)</td><td>Data are restricted, so y corresponds to excludes dependencies whose source feature is full-mandatory</td><td>Values of y
R16,R17</td></tr>
<tr><td>has Dependency Destination (?y,?x)</td><td>Data are restricted so x now corresponds to features excluded by the dependency y.</td><td>Values of x
Camera and MSM Both are dead features</td></tr>
</table>

3.3. Natural Language Explanations

We have a predefined explanation text for each proposed rule to identify dead features. Then, after identifying the dead features that satisfy rules 1 or 2, we explain the defect in natural language, as follows:

a) We determine if the dead feature satisfies rule 1 or rule 2.

b) We generate the predefined text that explains rule 1 or rule 2 in natural language.

Text to explain rule 1 is "*Optional feature **featureName** is dead because the full-mandatory feature **fullMandatoryFeatureName** excludes it through the dependency **exclusionDependencyName**".*

Text to explain rule 2 is "*Feature **featureName** is dead because **parentFeatureName**, its ancestor feature, is a dead feature too*"

c) We execute a new SQWRL query to get dependencies and features names related to the predefined text, which explains the dead feature.

d) We replace information from the FM at hand as needed in the predefined text. Figure 4 shows and example of each explanation applied to our application example.

3.4. Implementation Details

We implemented our approach in two stages. In the first

stage, we created the proposed ontology with Protégé 3.4.8 to represent concepts of the FMs meta-model. In the second stage, we developed a tool to integrate our proposed OWL ontology with Java. This integration allows us to manage and query information of each analyzed FM.

The implementation process of our second stage was as follows:

a) We read the proposed ontology in Java.

b) We use Jena [29] to populate and manage the ontology with information of the analyzed FM.

c) We use JESS library, as reasoner engine, to execute from Java SQWRL queries to identify dead features (i.e., features that satisfy the rule 1 or rule 2).

d) We produce a natural language text, which explains the cause of each dead feature. The explanatory text depends on whether the property satisfies rule 1 or rule 2. We complete the explanations from information gather from SQWRL queries.

Figure 4 presents one snapshot of the developed tool with a feedback obtained when we analyzed dead features in our application example. This case comprises features with mandatory, optional, excludes and group cardinality type dependencies. It also comprises an exclusive dependency that does not generate dead features (R18) and additional dependencies that generate dead features.

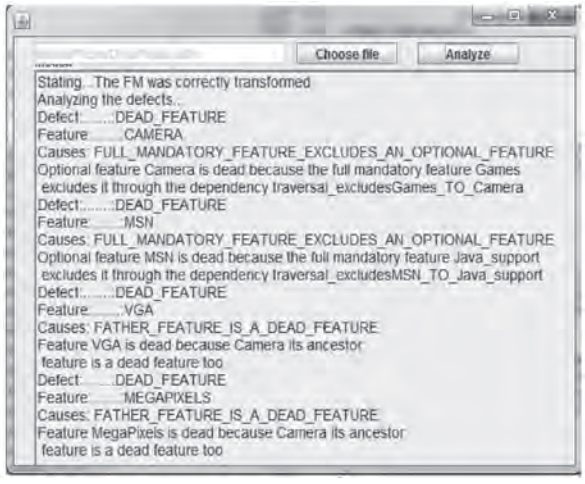


Figure 4. Snapshot corresponding to results generated from analyzing our FM running example

4. PRELIMINARY EVALUATION

4.1. Correctness

We assessed the correctness of our approach with 5 FMs with 25 features and 5 FM with 50 features. We generated these FMs with the BEnchmarking and TesTing on the analYsis (BeTTy) tool [30].

We manually tested our approach in three steps. First, we verified that it did not generate false positives. Second, we verified that the proposed solution identified 100% of dead features considered in our two rules. Finally, if the FMs had dead features, we validated that the cause corresponds to the case that produced the defect, and that the filled spaces in the explanation text corresponded to the correct situation for each one of the models.

In the first stage, we manually compared the dead features with the results obtained using FaMa [12]. We found that our proposal identified the 100% of the dead features that satisfied our rules, with 0% false positive. For the second and third stage, we made a manual inspection of correctness over 2 models (randomly selected) of 25 features and 2 models of 50 features.

We found that our proposal constructed correct explanations; i.e., they corresponded to the cause(s) that originated each defect. Results are available online in https://sites.google.com/site/raulmazo/.

4.2. Comparison of results

We compared results obtained in our proposal with the proposals of Trinidad *et al.* [4] and Rincón *et al*. [31] for the example application.

Table 3 presents the comparison of the results. The first column shows dead features identified by all approaches. The second column shows causes, in natural language, found by our proposal. Finally, the third column shows corrections proposed by Trinidad *et al.* [4] and Rincón *et al*. [31] (In this case, both approaches identified the same corrections).

Table 3. Comparison our proposal vs other approaches

Dead feature	Our proposal	Trinidad *et al.* [4] Rincón *et al.* [31]
	Causes in natural language	**Corrections**
Camera	*Optional feature* ***Camera*** *is dead because the full-mandatory feature* ***Games*** *excludes it through the dependency* ***traversal_Games_TO_Camera***	R1
		R6
		R17
MSN	*Optional feature* ***MSN*** *is dead because the full-mandatory feature* ***Java support*** *excludes it through the dependency* ***traversal_ MSN _TO_ Java support***	R2
		R7
		R16
Mega pixels	***Megapixels*** *is dead feature because* ***Camara*** *its ancestor feature is dead feature too*	R1
		R6
		R17
		R19
VGA	***VGA*** *is dead feature because* ***Camara*** *its ancestor feature is dead feature too*	R1
		R6
		R17
		R19

The results obtained shows that in the application example all approaches identified the same dead features. However, in other FMs, Trinidad *et al.* [4] and Rincón *et al.* [31] could identify other dead features that our approach will not identify. This is because we have not implemented all the cases to identify all dead features. Trinidad *et al.* [4] and Rincón *et al.* [31] identify all cases because they use a constraint satisfaction approach. However, our rule-based approach is extensible: we can create new rules for identifying and explaining in natural language other cases of dead features. Regarding explanations, Trinidad *et al.* [4] and Rincón *et al.* [31] identify the list of dependencies that must be deleted to remove dead features (Corrections). Our work instead focuses on explaining the cause of each dead feature in natural language. This information helps feature modelers to understand why dead features appear. Therefore, our approach is complementary to proposal of Trinidad *et al.* [4] and Rincón *et al.* [31] because the feature modeler could find dead features, their causes in natural language and possible corrections combining those proposals.

5. RELATED RESEARCH

We divide the research studies on identifying causes for dead features into two types: studies related to using ontologies in product line models, and those related to identifying causes of dead features. For the first type, Wang et al. [32] Wang et al., (2007)propose representing FMs and their constraints in OWL ontology language. In their proposal, the authors represent each feature as an ontology class, and each dependency as an ontology property. Their study identifies inconsistencies in particular FMs configurations and provide explanations for inconsistencies. However, their approach does not analyze the FM itself to identify the shortcomings, but each particular configuration.

In [15] Abo et al. propose two SWRL rules to validate model consistency. The first one detects features that excludes and requires the same feature, and the second one detects cycles in the FM, i.e, feature x requires feature y, and feature y requires feature x. Authors define inside ontology, as an antecedent, each situation that creates an inconsistency, and define as the consequence, the elements involved. Our work as the proposal of Abo et al. [15] uses ontology to represent FM in a formal way. However, additionally to use the ontology for formal representation, we use ontologies for two different purposes: (a) we exploit the ontological representation to perform dynamic SQWRL in FMs (i.e. Table 2); and (b) we explain defects in natural language. This is possible due to integration of our approach with Java. Abo et al. [15] implemented their proposal only in Protégé-OWL, therefore they do not have those advantages.

Lee *et al*. [33] Lee, Kim, Song, & Baik (2007) use ontologies to represent FMs and to analyze their variability and commonality. However, they use ontologies to analyze the semantic similarity of the FM, whereas that our approach uses ontologies to identify dead features and explain their causes.

Regarding the second category, several works were carried out to automatically identify dead features (and other defects) on [5, 9–11]. However, none of these works deals with identification of causes that explain in natural language why each dead feature occurs.

Trinidad *et al.* [4] present an automated method for identifying and explaining defects, such as dead features in FMs. For Trinidad *et al.* [4], an explanation is the minimal subset of dependencies that should be modified to remove the defect. They implemented their

approach transforming FMs into a diagnostic problem and then into a constraint satisfaction problem. This Implementation is available in FaMa [12], an Eclipse Plug-in for automatic analysis of FMs.

Rincón *et al.* [31] propose a method to identify corrections in FMs. In this approach, authors transform FMs into a constraint problem, then they identify all minimal corrections subsets (MCSes) [34] of dependencies that could be modified to correct each dead feature of the FM. This approach like FaMa[12], identify the list of dependencies that entail the fewest changes to fix the defect, but also identify others set of dependencies that imply more changes and fix the defect. This information provides more complete information about how to correct each dead feature.

Trinidad and Ruiz-Cortés [13] Trinidad & Ruiz-Cortes, (2009)use abductive reasoning to identify dead features and their causes. Unfortunately, the authors did not describe a method or algorithm to support their proposal.

Constraint satisfaction techniques are not enough to explain causes of a dead feature in natural language because, for instance, the structure needed to provide these explanations, is lost when authors transform the models into constraint programs. In fact, explanations generated by Trinidad *et al.*[4] and Rincón *et al*. [31] are not the causes that explain why a feature is dead, but corrections to apply in order to remove the defect.

6. CONCLUSIONS AND DISCUSSION

In this paper, we present an approach, which takes advantage of the inherent characteristics of ontologies, in order to identify dead features in FMs and explain certain causes of that defect in natural language. We use OWL to create an ontology to represent the FMs and their dependencies, and SQWRL as an ontology query language. We validate our proposal through the implementation and application of two SQWRL rules on a well-known case study and ten other FMs.

Our approach, in contrast to the black box-like approaches found in literature, can be easily extended with other rules to identify and explain other causes that create dead features.

Although ontologies were initially proposed for the semantic web, given their expressive power and formal semantics, they are useful in product lines to support identification of defects in feature models and to obtain information to produce explanations in a human understandable form.

We are currently extending this approach to identify other causes to explain dead features and other defects on FMs (e.g., false optional features or void FM). Other future directions include validating performance, accuracy, and scalability of the proposed approach for application to industrial cases.

ACKNOWLEDGMENTS

We perform this research under the project “Formation, evolution, and consistency of solutions based on the concept of product lines in organizations”. This project is part of the scientific cooperation program between France and Latin America.

REFERENCES

[1] Clements, P. and Northrop, L., Software Product Lines: Practices and Patterns 1st ed., Addison-Wesley Professional, 2001.

[2] Pohl, K., Böckle, G. and Linden, F.J., van der: Software Product Line Engineering: Foundations, Principles and Techniques Springer-Verlag New York, Inc., 2005.

[3] Kang, K.C., Cohen, S.G., Hess, J.A., Novak, W.E. and Peterson, S.P., Feasibility Study Feature-Oriented Domain Analysis (FODA). Technical Report 1990.

[4] Trinidad, P., Benavides, D., Duran, A., Ruiz-Cortés, A. and Toro, M., Automated Error Analysis for the Agilization of Feature Modeling. Journal of Systems and Software., 81, pp. 883–896, 2008.

[5] Von Der Massen, T. and Lichter, H., Deficiencies in Feature Models , En: Workshop on Software Variability Management for Product Derivation - Towards Tool Support. (Eds.Mannisto T, Bosch J). 2004.

[6] Trinidad, P., Benavides, D. and Ruiz-Cortés, A., Isolated Features Detection in Feature Models , En: Proceedings of Conference on Advanced Information Systems Engineering (CAiSE 2006). pp. 1–4, 2006.

[7] Benavides, D., Segura, S. and Ruiz-Cortés, A., Automated analysis of feature models 20 years later: A literature review. Information Systems., 35, pp. 615–636, 2010.

[8] Sun, J., Zhang, H. and Wang, H., Formal Semantics and Verification for Feature Modeling , En: Proceedings of the 10th IEEE International Conference on Engineering of Complex Computer Systems. IEEE Computer Society, Washington, DC, USA, pp. 303–312, 2005.

[9] Salinesi, C. and Mazo, R., Defects in Product Line Models and how to identify them , En: Software Product Line - Advanced Topic.(Eds.Elfaki A). pp. 97–122, 2012.

[10] Thüm, T., Kastner, C., Benduhn, F. and Meinicke, J., SAAK: FeatureIDE: An extensible framework for feature-oriented software development. Science of Computer Programming., 2012.

[11] Mazo, R., Salinesi, C. and Diaz, D., VariaMos : a Tool for Product Line Driven Systems , En: Proceedings of the 24th International Conference on Advanced Information Systems Engineering (CAiSE Forum'12). Springer Press, Gdansk-Poland, pp. 25–29, 2012.

[12] Trinidad, P., Benavides, D., Ruiz-Cortés, A., Segura, S. and Jimenez, A., FAMA Framework , En: Proceedings of the 12th International Software Product Line Conference. IEEE Computer Society, Washington, DC, USA, 359, 2008.

[13] Trinidad, P. and Ruiz-Cortés, A., Abductive Reasoning and Automated Analysis of Feature models: How are they connected , En: Proceedings of the Third International Workshop on Variability Modelling of Software-Intensive Systems. pp. 145–153, 2009.

[14] Batory, D., Benavides, D. and Ruiz-Cortés, A., Automated analysis of feature models: challenges ahead. Communications of the ACM., 49, pp. 45–47, 2006.

[15] Abo, L., Kleinermann, F. and De Troyer, O., Applying semantic web technology to feature modeling , En: Proceedings of the 2009 ACM symposium on Applied Computing (SAC '09). ACM, New York, pp. 1252–1256, 2009.

[16] Czarnecki, K., Peter Kim, C.H. and Kalleberg, K.T., Feature Models are Views on Ontologies , En: Proceedings of the 10th International on Software Product Line Conference. IEEE Computer Society, Washington, DC, USA, pp. 41–51, 2006.

[17] Mcguinness, D.L. and Van Harmelen, F., OTHERS: OWL web ontology language overview. W3C recommendation., 10, 10, 2004.

[18] O'connor, M., and Das, A., SQWRL: a Query Language for OWL , En: Proceedings of the 6th International Workshop OWL: Experiences and Directions. Chantilly, 2009.

[19] Tsetsos, V., Papataxiarhis, V. and Hadjiefthymiades, S., Personalization based on Semantic Web Technologies. Semantic Web Engineering in the Knowledge Society, Information Science Reference., pp. 52–75, 2008.

[20] Breitman, K.K., Casanova, M.A. and Truszkowski, W., Semantic web: concepts, technologies and applications Springer-Verlag, pp. 105–127, 2007.

[21] Djebbi, O. and Salinesi, C., Criteria for Comparing Requirements Variability Modeling Notations for Product Lines , En: Proceedings of the Fourth International Workshop on Comparative Evaluation in Requirements Engineering. IEEE, pp. 20–35, 2006.

[22] Czarnecki, K., Helsen, S. and Eisenecker, U., Formalizing Cardinality-based Feature Models and their Specialization. Software Process: Improvement and Practice., 10, pp. 7–29, 2005.

[23] Segura, S., Automated Analysis of Feature Models Using Atomic Sets , En: Proceedings of the First Workshop on Analyses of Software Product Lines (ASPL08). 2008.

[24] Borst, W., Construction of Engineering Ontologies for Knowledge Sharing and Reuse: Ph.D. Dissertation [Ph Thesis].Enschede, Netherlands: University of Twente,1998.

[25] Gruber, T., Toward Principles for the Design of Ontologies Used for Knowledge Sharing , En: International Workshop on Formal Ontology.(Eds.Guarino N, Poli R). Padova,Italy, 1993.

[26] Horridge, M., Knublauch, H., Rector, A., Stevens, R. and Wroe, C., A Practical Guide To Building OWL Ontologies Using The Protege-OWL Plugin and CO-ODE Tools Edition 1.02004.

[27] Friedman-Hill, E., Java expert system shell, Available:http://herzberg.ca.sandia.gov/jess.

[28] Noy, N. and Mcguinness, D., Ontology Development 101 : A Guide to Creating Your First Ontology, pp. 1–25, 2001.

[29] APACHE SOFTWARE FOUNDATION: Apache Jena, Available:http://jena.apache.org/documentation/ontology/.

[30] Segura, S., Galindo, J.A., Benavides, D., Parejo, J.A.

and Ruiz-Cortés, A., BeTTy: benchmarking and testing on the automated analysis of feature models , En: Proceedings of the Sixth International Workshop on Variability Modeling of Software-Intensive Systems. ACM, New York, NY, USA, pp. 63–71, 2012.

[31] Rincón, L.F., Giraldo, G.L., Mazo, R., Salinesi, C. and Diaz, D., Subconjuntos Minimos de Correccion para explicar caracteristicas muertas en Modelos de Lineas de Productos. El caso de los Modelos de Caracteristicas , En: Proceedings of the 8th Colombian Computer Conference (CCC). Armenia-Colombia, 2013.

[32] Wang, H., Li, Y., Sun, J., Zhang, H. and Pan, J., Verifying feature models using OWL. Web Semantics: Science, Services and Agents on the World Wide Web., 5, pp. 117–129, 2007.

[33] Lee, S., Kim, J., Song, C. and Baik, D., An Approach to Analyzing Commonality and Variability of Features using Ontology in a Software Product Line Engineering , En: Proceedings of the Fifth International Conference on Software Engineering Research, Management and Applications. IEEE, pp. 727–734, 2007.

[34] Liffiton, M. and Sakallah, K., Algorithms for Computing Minimal Unsatisfiable Subsets of Constraints. Journal of Automated Reasoning., 40, pp. 1–33, 2008.

ANAEROBIC DIGESTION OF FIQUE BAGASSE: AN ENERGY ALTERNATIVE

DIGESTION ANAEROBIA DEL BAGAZO DE FIQUE: UNA ALTERNATIVA ENERGÉTICA

HUMBERTO ESCALANTE H.
Ingeniero Químico PhD, Profesor Escuela de Ingeniería Química, Universidad Industrial de Santander, escala@uis.edu.co
CAROLINA GUZMÁN L.
Bacterióloga y Laboratorista Clínico, Profesora Escuela de Bacteriología y Laboratorio Clínico, Universidad Industrial de Santander, caguzlun@uis.edu.co
LILIANA CASTRO M.
Ingeniera Química PhD, Escuela de Ingeniería Química, UniversidadIndustrial de Santander, lcastrousta@gmail.com

Received for review October 26 th, 2012, accepted September 9th, 2013, final version October 28 th, 2013

ABSTRACT: In Colombia the agro-industrial process of fique generates approximately 20 800 kg of waste / ha planted, consisting of bagasse and juice. These wastes are discarded over soil and water generating a problem of environmental pollution. The fique bagasse (FB) has a calorific value of 3 297.91 kcal / kg, high concentrations of cellulose, hemicellulose and C / N ratio that make it appropriate for biogas production. However, the presence of lignin in the FB requires specific microbial consortia for its degradation. Therefore, in this research the biogas production from FB on a laboratory scale was studied through the Anaerobic Digestion (AD) process using a consortium of ruminal fluid (RF) and pig manure (PM). A methane production of 0.35 $m^3 CH_4$/kg volatile solids (VS) added during two weeks, equivalent to 1.38 kWh/kg VS added, indicated that FB is an attractive residual to be used as a source of renewable energy.

Key words: Anaerobic digestion, fique bagasse, inoculums, lignocellulosic waste, ruminal fluid.

RESUMEN: En Colombia, el procesamiento agroindustrial de fique genera aproximadamente 20 800 kg de residuos/ha sembrada que corresponden a. jugo y bagazo. Estos residuos son descartados al ambiente generando problemas de contaminación. El bagazo de fique tiene un valor calorífico de 3 297.91 kcal/kg, altas concentraciones de celulosa, hemicelulosa y una relación C/N favorable para tratar este residuo mediante conversión anaerobia. Sin embargo, la presencia de lignina en el bagazo hace que se requiera un consorcio microbiano específico para llevar a cabo la degradación. En este trabajo se estudio la producción de biogás a partir del bagazo de fique, empleando como inóculo una mezcla de líquido ruminal y lodo estiércol de cerdo. Se alcanzó una producción de metano de 0.35 $m^3 CH_4$/kg Sólidos Volátiles (SV) adicionados durante quince días de digestión, equivalente a 1.38 kWh/kg SV adicionado, indicando que el bagazo de fique es un residuo atractivo para ser usado como fuente de energía renovable.

Palabras clave: Digestión anaerobia, bagazo de fique, inóculos, residuo lignocelulósico, líquido ruminal.

1. INTRODUCTION

Agricultural residues are an alternative source of energy that reduces the depletion of fossil fuels [1]. The molecular structure of this residual biomass is responsible for their energy content which varies between 3000 and 3500 kcal / kg for lignocellulosic wastes, and from 2000 to 2500 kcal / kg for urban waste [2]. The agro industrial process of fique (Agave family) generates 15000 tons of waste (bagasse) per hectare. Bagasse is being left on soil and thrown in rivers causing environmental pollution problems. Waste from the fique process have been evaluated for the production of pharmaceutical active compounds (hecogenin and tigogenin), surfactants, bioinsecticides, paper and fique fiber reinforced materials [3].

Due to the physicochemical composition of fique bagasse (FB), this residue is considered as a lignocellulosic waste biomass. [4].

Actually, one the most important impacts of renewable energy is the anaerobic digestion process from different substrates. The wastes´s physico chemical characteristics condition the biomethane potential.

The Table 1 shows the yields values of 0.03 and 0.48 m^3 CH_4/kg VS added for urban municipal solid and cooked meat wastes, respectively.

Wastes with composition similar to Fique Bagasse, such as sisal, maize silage and grass silage, reached

high biogas production, with yields of 0.24 m^3 CH_4/ kg VS added, 0.36 m^3 CH_4/kg VS added and 0.6 m^3 CH_4/kg VS, respectively [1, 5]

In terms of methane yields, the Cattail aquatic plant reached a conversion of 66% from volatile solids, using ruminal fluid as inoculums [6]. Sisal and corn digestion attain methane values of 60% v/v [5, 7]. Whey, barley and rice residues showed high biomethane potential with values of 501, 229 and 195 L CH_4/kg VS, respectively [8].

The high carbon content in FB makes it a potential substrate for methane production by anaerobic bioconversion systems [9]. Therefore, the aim of this research was to evaluate the production of biogas through anaerobic digestion on a laboratory scale, using a consortium of bacteria from ruminal fluid and pig manure.

Table 1. Potential for biogas production from different wastes

RESIDUE	METHANE PRODUCCTION (m^3CH_4/Kg VS)	REFERENCE
Cooked meat	0.48	Cho *et al.,* (1995)
Office Paper	0.37	
Municipal Solid Waste	0.03	Forster-Carneiro *et al.,* (2007)
Oilseeds	0.42	Petersson *et al.,* (2007)
Straw	0.44	
Corn (silage)	0.36	Amon *et al*., (2007)
Sunflower	0.30	
Sisal	0.32	
Plantain peel	0.27	Nallathambi Gunaseelan (2004)
Rotten tomatoes	0.30	
Grass silage	0.60	Liu et al., (2009)

Source: Ward *et al.*, 2008 modified by the authors.

Anaerobic digestion takes place in four stages: a) hydrolysis b) acidogenesis c) acetogenesis d) methanogenesis. These stages are carried out by microbial consortia formed from different populations of microorganisms. The products generated in each stage are the nourishment of another [10].

In hydrolysis, the organic matter composed of complex molecules must be broken to simpler compounds. The microorganisms involved in this stage produce acetic acid-carbon compounds, fatty acids and other organic polycarbonate compounds. In this way, carbohydrates are converted into simple sugars, fats into glycerol and fatty acids and proteins are hydrolysed to peptides and amino acids, releasing carbon dioxide and hydrogen [11]. At the acidogenesis stage, the monosaccharides produced are converted into organic acids of acetate, propionate, butyrate, valerate type, CO_2 and H_2. In acetogenesis, acetate, H_2 and CO_2 are generated, mostly. In methanogenesis, the methanogenic consortiums convert acetate into methane and CO_2, mainly [12, 13].

The methane production depends on its hydrolytic activity (HA) and specific methanogenic activity (SMA). The HA indicates the inherent ability of a microbial population to degrade carbon and it is quantified as the specific rate of substrate consumption [14]. The SMA refers to the ability of the microbial biomass to convert organic matter into methane and it is expressed as the mass of substrate in terms of chemical oxygen demand (COD) that is converted into methane per biomass unit per unit of time (gCOD-CH_4/g volatile suspended solids- VSS / day) [15].

Physico-chemical composition of FB indicates that these residues are composed of complex polymers such as cellulose, hemicelluloses and lignin. Therefore, FB digestion requires a specialized hydrolytic consortium. Different microbial consortiums have been used in biogas production from lignocellulosic materials, such as anaerobic sludge from primary wastewater treatment plants, ruminal fluid, pig manure or cattle manure, compost, and pure cultures of microorganisms [16, 17].

Previous studies showed that during anaerobic digestion from fique bagasse, the mixture Ruminal Liquid (RL) and Pig Waste Sludge (PWS), as consortia, showed high hydrolytic and methanogenic activities and the best biomethane potential.

During the anaerobic digestion process, the organic matter is converted into soluble fractions, which can be

expressed as total reducing sugar (TRS), total volatile fatty acids (TVA) and cumulative methane volume.

The aim of this research, was to describe anaerobic digestion from fique bagasse, used as inoculum the mixture ruminal liquid and pig manure sludge, through evolution of total reducing sugar, total volatile fatty acids and cumulative volume methane.

2. MATERIALS AND METHODS

2.1. Substrate

FB, as substrate, was collected in a fique processing plant located in Santander –Colombia. FB chemical composition was evaluated by: total alkalinity (TA), concentration of total solids (TS), volatile solids (VS), volatile fatty acids (VFA), carbon / nitrogen ratio (C / N), cellulose, hemicellulose and lignin according to procedures established by Van Soest and the Standard Methods for Examination of Water and Wastewater [18, 19].

2.2. Inoculum

A mixture of 1:1 ruminal fluid (RF) and pig manure sludge (PMS) was used in the bioproduction process. RF was obtained from bovine stomachs collected in a livestock processing plant. PMS was collected from pig septic tanks. The inoculums physicochemical composition was evaluated according to protocols established by the Standard Methods for Examination of Water and Wastewater [18].

The microbiological characterization quantify the major microbial groups present in the inoculum and it was carried out using the technique of Most Probable Number (MPN) according to protocols established [15]. Serial dilutions were made from the mixture RF-PMS. Each dilution were inoculated in five hungate tubes, additionally, five un-inoculated tubes were considered as control. The tests were performed in a CO_2 atmosphere to ensure anaerobic conditions. A positive result was identified according to the characteristics of each trophic group (Table 2). MPN values were calculated using the Mac Grady statistical table.

Table 2. Trophic group quantification determined by the MPN technique

Trophic group	Abbreviation	Substrate	Incubation Time at 35°C ± 2°C (days)	Detection of positive tubes
Glucose-fermenting bacteria	(GFB)	Glucose	5 to 8	Color change from green to yellow
Lactate fermenting bacteria	(LFB)	Lactate	5 to 8	Color change from green to yellow
Acetate sulfate-reducing bacteria	(ASRB)	Acetate	7 to 15	FeS Production
Hydrogenophilic methanogenic bacteria	(HMB)	H_2/CO_2	15 to 45	Methane Detection
Acetoclastic Methanogenic bacteria	(AMB)	Acetate	30 to 60	Methane Detection
Methanogenic bacteria for methanol	(MBM)	Methanol	30 to 60	Methane Detection

2.3. AD process for FB

Methane production from FB using RF-PMS was carried out in 0.5 L reactors, with an operating volume of 0.35 L. substrate / inoculum ratio of 1 g VS / g VS was used. The operation time was 15 days, at 39 ± 2°C. The concentration of total reducing sugars (TRS), volatile fatty acids (VFA), biogas volume and the percentage of methane produced were considered as variables response.

TRS concentration was determined according to the colorimetric method of dinitrosalicylic acid, using a GENESYS 20 Thermo Spectronics spectrophotometer at a wavelength of 540 nm [20]. VFA concentration was quantified according to the titration procedure [21]. Methane volume was measured by the alkaline shift method [22] at standard conditions and the quality of biogas produced was determined by a PGD3-IR Status Scientific Controls infrared gas detector. All fermentations were performed in duplicate. Experimental result was analyzed with standard deviation.

3. RESULTS AND DISCUSSION

3.1. Characterization of FB

FB has similar physicochemical characteristics to sisal waste, cattail and sunflower oil residues (Table 3), all of them with high potential for biogas production [6, 23, 24].

FB has an acidic pH that could inhibit the start of AD process. However, the inoculum's buffering capacity regulates this inhibitory effect. According to the concentration of VS, cellulose and hemicelluloses content, FB has the capacity to produce methane and it is coherent with the biodegradability test in sisal [25].

The C/N content varies with the type of waste and causes inhibition in an inappropriate range. A C/N optimal range of 15 to 25 has been recommended for microbial growth. As examples, the co-digestion of onion and digested sludge has a value of 15; mixtures of corn crops with the sludge reach ratios between 15 and 18, and 21 in adverse operational conditions [13]. In this study FB has a C/N of 14.

In energy terms, olive waste has a calorific value of 4240 kcal / kg and reached the maximum biogas production of 54.26 l / l olive residue containing 83% methane [26, 27]. Sunflower oil residues has a calorific value of 3700 with productions of 0.20 L CH_4/kg VS [24]. FB has a high calorific value, 3000 kcal / kg, which can be exploited for energy production, this value corresponds to agricultural biomass with the low content of sulfur and ash.

Table 3. Physicochemical characteristics of FB

Parameter	Units	Fique Bagasse	Sisal Pulp	Cattail	Sunflower oil residues
pH		4	5.60	nd	Nd
TS	% (p/p)	24	14.2	90.2	11.6
VS	% (TS)	87.1	82.3	91.2	87
C/N		14	65	nd	18
Celulose	%	41.81	47.1	20.8	40.7
Hemicellulose	%	22.17	23.1	22.6	8.5
Lignin	%	15.56	8.60	10.5	11.5
Sulfur	%p/p	0.006	Nd	Nd	Nd
Ash	%p/p	10	Nd	Nd	Nd
Heating Power	kcal/kg	3300	Nd	Nd	3700

Nd: not determined

3.2. Physicochemical and microbiological characterization of the inoculums

RF-PMS characterization is shown in Table 4. Values obtain from this inoculums confirm its application in AD process in terms of pH: 8.0 TVFA: 3100 mg/L and VSS: 21880 mg/L, among others parameters [28].

Microbial distribution of populations in RF-PMS is shown in Table 5. High levels of GFB, LFB and ASRB confirm the enzyme activity required for AD starting up (hydrolytic and acidogenic stages). The last group guarantees acetate metabolism since its ability to grow using this substrate of the incomplete oxidation of ethanol [29]. HBM, AMB and MBM concentrations between 10^4 and 10^5 NMP/g VSS maintain a partial pressure of hydrogen at a level that allows syntrophic degradation of ethanol and propionate [30, 31]. The percentage distribution of hydrogenophilic methanogenic archaea group (40%) is responsible for methane production and shown a symbiotic balance between the trophics groups.

Table 4. Physicochemical characteristics, HA and SMA of inoculums RF-PMS

Parameter	Units	Value
pH	--	8
TVFA	mg/L	7 200
TA	mg $CaCO_3$/L	3 100
TS	mg/L	43 770
VSS	mg/L	21 880
TVS	mg/L	23 640
HA	g COD/g VSS day	0.051
SMA	g COD/g VSS day	0.144

Table 5. Inoculums group trophic quantification by NMP method

Trophic group	Abbreviation	RF-PMS (NMP/g VSS)
Glucose-fermenting bacteria	(GFB)	9.8×10^{11}
Lactate fermenting bacteria	(LFB)	1.5×10^{10}
Acetate sulfate-reducing bacteria	(ASRB)	1.3×10^{10}
Hydrogenophilic methanogenic bacteria	(HMB)	1.8×10^{5}
Acetoclastic Methanogenic bacteria	(AMB)	4.0×10^{4}
Methanogenic bacteria for metanol	(MBM)	2.0×10^{4}

3.3. TRS and VFA variation in AD from FB

TRS are soluble compounds, easily metabolized by microorganisms, which allow the AD first stage to take place [24]. The high concentration of TRS at the beginning encourages the process to start (Figure 1). The rapid consumption of sugars, until the fourth day, is consistent with microorganisms metabolism in the hydrolysis and acidogenesis stages and shows its enzymatic capacity [26]. The TRS concentration was kept constant during the fermentation.

During VFA production, a simultaneous TRS consumption was observed. From the eighth day, VFA concentration remained constant (from 4000 to 4320 mg VFA / L) and avoids inhibition by acidification in the reactor. These results are consistent with VFA and pH values from other studies. For example, the variation of pH for biomethane potential of maize in a batch test, ranged from 7.2 to 8.0, similar results were obtained with fruit/vegetable with maximum values before inhibition of a pH of 7.8 and 7800 mg/l of VFA [7,33].

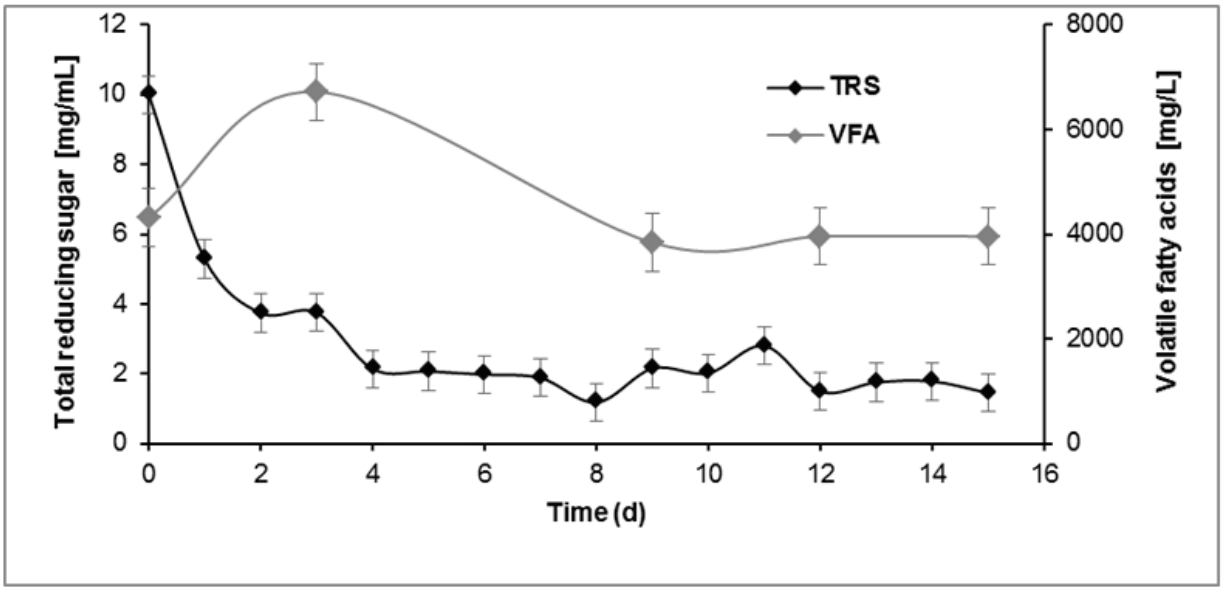


Figure 1. TRS and VFA concentration during the digestion time

3.4. Methane production from FB

Methane production, for the first two weeks, was 3.6 L (Figure 2) and confirms the AD success using FB as an organic substrate. Yields values obtained were $0.45 m^3$ CH_4/kg VS and are comparable with the AD of grass, corn and agro-industrial wastes (0.40, 0.32 and 0.32; respectively) [8, 23, 34]. The percentage composition of the biogas produced from FB (Table 6), is corroborated with research about anaerobic digestion from sisal experiments, the methane production reached was above 50% [5]. These results indicate that FB can be considered as a viable alternative for recovering energy in the form of biogas with 60–65% methane content.

In comparison with other lignocelullosic wastes, FB is one of the most efficient biomasses in terms of electrical energy (Figure 3) [5, 8, 35].

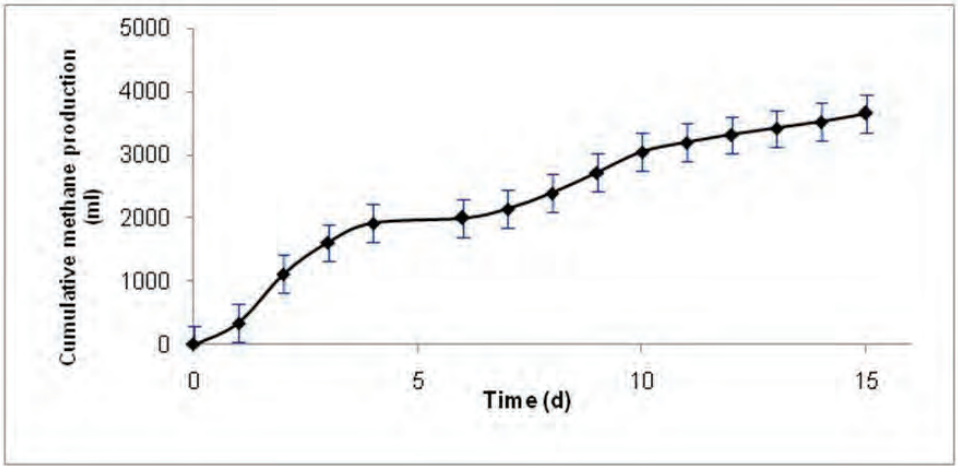


Figure 2. Accumulated methane production during digestion time (STP conditions)

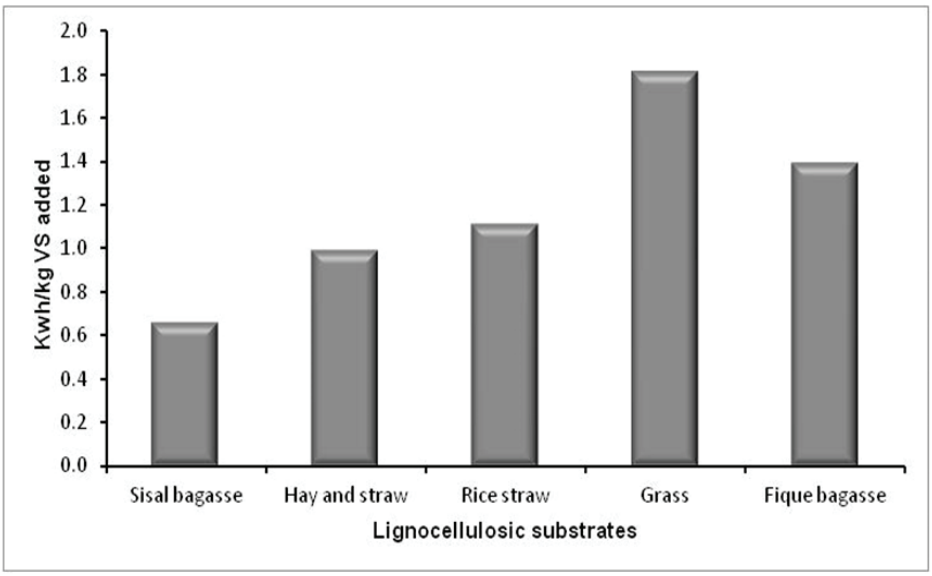


Figure 3. Methane yields, during anaerobic digestion, expressed in kWh / kg VS added

Table 6. Percentage composition of biogas obtained from FB

Composition	Units	Value
CH_4	%	65
CO_2	%	30
Other gases	%	5

4. CONCLUSIONS

Anaerobic Digestion of FB, as the lignocelullosic substrate, produces 1.38 kwh/kgVS using a mixture of ruminal fluid-pig sludge manure with high hydrolytic activity and specific methanogenic activity potential. Anaerobic digestion of fique bagasse is an alternative, not only as a real source of energy but also it contributes to reduce the environmental contamination.

ACKNOWLEDGEMENTS

The authors thank the Universidad Industrial de Santander, Ministerio de Agricultura y Desarrollo Rural y al Departamento de Ciencia y Tecnología de Colombia.

REFERENCES

[1] Ward A J., Hobbs P J., Holliman, PJ., Jones, D. L., "Optimisation of the anaerobic digestion of agricultural resources". Bioresource Technol, 99(17), pp.7928–7940, 2008.

[2] Crotti, C. and Milera, S., "Implementación del modelo cropsyst para la simulación del rendimiento del cultivo de

maíz en una región argentina". Centro de Investigación observación y monitoreo territorial y ambiental, pp. 2-4, 2006.

[3] Ministerio de Agricultura y Desarrollo Rural; CADEFIQUE; CORPOICA. Guía Ambiental del Subsector Fiquero, Colombia, pp. 43-53, 2006.

[4] Mata-Alvarez, J., Macé, S. and Llabrés, P., "Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives". Bioresource Technol., 74(1), pp. 3-16, 2000.

[5] Mshandete, A., Kivaisi, A., Mugassa, R. and Bo, M.,"Anaerobic batch co-digestion of sisal pulp and fish wastes". Bioresource Technol, 95(1), pp. 19-24, 2004.

[6] Zhen-Hu, H. and Han-Qing, Y.,"Anaerobic digestion of cattail by rumen cultures". Waste manage, 26(11), pp. 1222-1228, 2006.

[7] Raposo, F., Banks C.J., Siegert, I., Heaven, S. and Borja, R., "Influence of inoculum to substrate ratio on the biochemical methane potential of maize in batch tests". Process biochem, 41(6), 2006, pp. 444–1450.

[8] Dinuccio, E., Balsari, P. Gioelli, F. and Menardo, S., "Evaluation of the biogas productivity potential of some Italian agro-industrial biomasses". Bioresource Technol, 101(10), pp. 3780–3783, 2010.

[9] Barrera, P., Salas, X., Castro, L., Ortiz, C. and Escalante, H., "Estudio preliminar de la bioproducción de metano a partir de los residuos del proceso del beneficio del bagazo de fique". Revista Ion, 22(1), pp. 21-25, 2009.

[10] Kumar Khanal Samir. Anaerobic Biotechnology for Bioenergy Production: Principles and Applications. Editorial W. Blackwell, USA, pp. 43-53, 2008.

[11] Appels Lise; Baeyens Jan; Degrève Jan; Dewil Raf. "Principles and potential of the anaerobic digestion of waste-activated sludge". Prog energ combust, 34(6), pp. 755–781, 2008.

[12] Mata Alvarez, J., Biomethanization of the Organic Fraction of Municipal Solid Wastes, Editorial IWA, España, pp.5-20, 2002.

[13] Yebo, Li., Stephen, Y. P. and Jiying, Z., "Solid-state anaerobic digestion for methane production from organic waste". Renew sust energ rev, 15(1), pp. 821–826, 2011.

[14] Hu, Z-H., Wang, G. and Yu, H-Q., "Anaerobic degradation of cellulose by rumen microorganisms at various pH values". Biochem Eng J, 21(1), pp. 59–62, 2004.

[15] Díaz, M. C., Espitia, S. E. y Molina, F., Digestión Anaerobia: Una aproximación a la tecnología. Editorial Universidad Nacional de Colombia UNIBIBLOS. Colombia, pp. 43-45, 2002.

[16] O´Sullivan Cathryn, A., Burrell C. Paul; Clarke William, P., Blackall, Linda L., "Comparison of cellulose solubilisation rates in rumen and landfill leachate inoculated reactors". Bioresource Technol, 97(18), pp. 2356–2363, 2006.

[17] Forster-Carneiro, T., Pérez, M., Romero, L.I. and Sales, D., "Dry-thermophilic anaerobic digestion of organic fraction of the municipal solid waste: Focusing on the inoculum sources". Bioresource Technol, 98(17), pp. 3195–3203, 2007.

[18] American Public Health Association, APHA. Standard Methods for the Examination of Water and Wastewater. Edition 20th, USA, 1998.

[19] Goering H.K. and Van Soest, P.J., Forage. Fiber Analyses (Apparatus, Reagents, Procedures and some Applications. Agricultural Handbook 379, ARS-USDA, United States of America, 1970.

[20] Miller, G., "Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar". Anal Chem, 31(3), pp.426-428, 1959.

[21] Anderson, G.K. and Yang, G., "Determination of bicarbonate and total volatile acid concentration in anaerobic digesters using a simple titration". Water Environ Rese 64, pp. 53–59, 1992.

[22] Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J., Guwy, A. J., Kalyuzhnyi, S., Jenicek, P. and Van Lier, J. B. "Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays". Water Sci Technol., 59 (5), pp. 927-934, 2009.

[23] Mshandete, A., Bjornsson. L., Kivaisi, K.A. and Rubindamayugib, S.T., Mattiassona Bo. "Enhancement of anaerobic batch digestion of sisal pulp waste by mesophilic aerobic pre-treatment". Water Res., 39 (8), pp. 1569–1575,2005.

[24] Galí, A., Benabdallah, T., Astals, S. and Mata-Alvarez, J., "Modified version of ADM1 model for agro-waste application". Bioresource Technol., 100 (1), pp. 2783–2790, 2009.

[25] Kivaisi, K. A. and Rubindamayugi, M.S.T., "The potential of agro-industrial residues for production of biogas and electricity in Tanzania" Renew Energ., 9 (1-4), pp. 917-921,1996.

[26] Fernández, J., "Los residuos de las agroindustrias como biocombustibles sólidos (I)". Vida rural. pp. 14-18, 2006.

[27] Fezzani, B. and Ben, CR., "Two-phase anaerobic co-digestion of olive mill wastes in semi-continuous digesters at mesophilic temperature" Bioresource Technol., 101 (6), pp. 1628-1634, 2010.

[28] Quintero, M., Castro, L., Ortiz, C., Guzmán, C. and Escalante, H., "Enhancement of starting up anaerobic digestion of lignocellulosic substrate: fique´s bagasse as an example". Bioresource Technology. Vol. 108, pp. 8-13, 2012.

[29] García, M.C., Díaz Báez, M.C., "Evaluación de la toxicidad de un efluente cervecero mediante ensayos de inhibición de la actividad metanogénica". Rev. Colomb. Biotecnol., 5(2), pp. 23-31, 2003.

[30] Valdez-Vazquez Idania; Poggi-Varaldo. "Hydrogen production by fermentative consortia". Renew Sust Energ Rev., 13(5),pp.1000–1013, 2009.

[31] Gerardi, M., The Microbiology of Anaerobic Digesters. Editorial Wiley Interscience. Estados Unidos, pp. 81-117, 2003.

[32] De La Rubia, M.A., Raposo, F., Rincón, B. and Rincón, Borja R., "Evaluation of the hydrolytic–acidogenic step of a two-stage mesophilic anaerobic digestion process of sunflower oil cake". Bioresource Technol., 100(18), pp. 4133–4138. 2009,

[33] Dinsdale, R. M., Premier, G.C., Hawkes, F.R. and Hawkes, D., "Two-stage anaerobic co-digestion of waste activated sludge and fruit/vegetable waste using inclined tubular digesters". Bioresource Technol., 72(2), pp. 159-168, 2000.

[34] Mata-Alvarez, J., Dosta, J., Macé, S. and Astals, S., "Codigestion of solid wastes: A review of its uses and perspectives including modeling". Crit Rev Biotechnol., 31(2), pp. 99-111, 2010.

[35] Álvarez, J., Otero, L. and Lema, J., "A Methodology for Optimising Composition for Anaerobic Co-digestion of Agro-industrial Wastes". Bioresource Technology. Vol. 101, No. 1, pp. 1153-1158. 2010

ELECTRIC VEHICLE EMULATION BASED ON INERTIAL FLYWHEEL AND A DC MACHINE

EMULACIÓN DE UN VEHÍCULO ELÉCTRICO BASADA EN VOLANTE DE INERCIA Y MÁQUINA DE CD

IVAN ALCALÁ
M.Sc. Ing. Electrónica, Centro Nacional de Investigación y Desarrollo Tecnológico, México, alcala@cenidet.edu.mx
ABRAHAM CLAUDIO
Ph.D. Ing. Eléctrica, Centro Nacional de Investigación y Desarrollo Tecnológico, México, peabraha@cenidet.edu.mx
GERARDO GUERRERO
Ph.D. Ingeniería, Centro Nacional de Investigación y Desarrollo Tecnológico, México, gerardog@cenidet.edu.mx
JESUS AGUAYO ALQUICIRA
Ph.D. Ing. Electrónica, Centro Nacional de Investigación y Desarrollo Tecnológico, México, aguayo@cenidet.edu.mx
VICTOR HUGO OLIVARES
Ph.D. Ing. Electrónica Centro Nacional de Investigación y Desarrollo Tecnológico, México, olivares@cenidet.edu.mx

Received for review August 6 th, 2012, accepted July 29th, 2013, final version August, 6 th, 2013

ABSTRACT: This paper describes the analysis, modeling, and simulation of an electric vehicle (EV) focused on developing a test bench for reproducing EV dynamics. The EV test bench is composed of an induction motor (IM) drive, which is directly coupled to a dc machine and to an inertial flywheel through a transmission system with bands. The inertial flywheel and the DC machine reproduce the dynamic and forces acting on the EV. A methodology is proposed to design EV test bench to study the EV behavior close to the real test conditions. The EV forces and DC machine analysis can define the different conditions of the EV operation. The modeling and simulation are developed in MATLAB/Simulink, the test bench implemented is controlled by a DSP. Finally, simulation and experimental obtained results can validate the operation of the test bench emulator.

Key words: Electric vehicle, flywheel, model, emulator, DSP, DC machine, induction motor.

RESUMEN: Este artículo describe el análisis, modelado y simulación de un vehículo eléctrico (EV) enfocado al desarrollo de un banco de pruebas para reproducir la dinámica del EV. El banco está formado por un drive de motor de inducción (IM) acoplado directamente a una máquina de DC y a un volante de inercia a través de una transmisión. El volante de inercia y la máquina de DC reproducen la dinámica y las fuerzas que actúan en el vehículo. Se propone una metodología para diseñar un banco de pruebas de EV para estudiar el comportamiento de vehículos eléctricos cercano a las condiciones de operación reales. El análisis de las fuerzas en el EV en conjunto con la máquina DC define las condiciones de operación del EV. El modelado y la simulación son desarrollados en MATLAB/Simulink, el banco de pruebas implementado es controlado por un DSP. Finalmente, los resultados de simulación y experimentales obtenidos validan el funcionamiento del banco de prueba.

Palabras clave: Vehículo Eléctrico (EV), volante, modelo, emulador, DSP, máquina de CD, motor de inducción.

1. INTRODUCTION

Nowadays, worldwide environment contamination due to industrial applications, to transportation and to the strong oil dependence for electrical energy generation, has become a big ecological and health problem [1,2]. Additionally, related to the dependence on oil as a main source of energy, it is important to consider the fact that oil is a non renewable source, which is becoming exhausted worldwide. However, very interesting renewable alternative and clean energy sources are being considered for different applications. A solution for transportation, to solve the oil dependence, is to use electric energy for electric vehicles based on different energy sources such as biodiesel, ethanol, solar, fuel cells, among others, which are replacing the traditional non renewable energies [3-6]. Thus, electric vehicles (EVs) have had a strong development for different required applications, such as cars, scooters, bicycles, hoist, and electrical wheelchairs, among others. These

types of EVs are typically driven by stored electric energy (battery, fuel cell, and ultra-capacitor); however, one of the main EV problems is the short driving range due to the energy storage capacity. So, the way to use storage energy efficiently is an important issue for designing EVs. Recently, researches about the design and implementation of EVs have been developed; some of them have been focused on the efficient way to process energy in order to increase the EV autonomy [7-11]. From the point of view of EV energy management design, a whole analysis of the system, considering the different energy processing elements, it is essential to know how to manage the energy in order to improve the EV performance [8,9,11,12]. Several control techniques are proposed in the literature for adjustable speed drives, trying to improve the dynamic response of speed, torque, and slip ratio output regulation [12-14]. However, studies related to the energy management are not always taken into account. Energy management analysis is an important issue in the development of electric vehicles; it is necessary to understand the relations and interactions of the energy process during acceleration/braking in the different parts of the system. In the above mentioned research some are based on two interesting approaches: simulation by models [13] and experimentation using a test bench of the system [7] in order to correctly describe the operation of the different EV parts. Different approaches have been reported to study the EV behavior; some are based on models of the system and others on test to emulate the behavior of EV by using dynamic loads [15-19]. EV simulation is an interesting method to analyze the behavior, this method is feasible and cheap, but the results obtained depend on the model accuracy to represent the real system; so it is related to the number of parameters considered and consequently the simulation can take a relative long time. The EV test bench can reproduce dynamic conditions close to the real operation behavior of the automobile, but this representation has to be scaled to the real system and the cost is more expensive than the model. These two approaches allow us to propose improvements to the EV system design, from the point of view of energy management by using alternative energy sources associated with efficient energy storage technology, such as batteries, fuel cells, ultracapacitors, flywheel, among others.

The most interesting contributions of this work, with regard to other works from the literature, are: a) the study of EV systems is based on modeling and experimentation approaches; and b) the test bench includes an inertial flywheel that allows the EV dynamic behavior to be emulated. In this sense, this paper presents the analysis, modeling and Matlab/Simulink implementation of an EV. Furthermore, the design and experimental implementation of a test bench to emulate a scaled EV. The simulation and experimental obtained results are analyzed based on the energy balance approach.

2. ELECTRIC VEHICLE ANALYSIS

The EV analyzed considers the EV dynamics, the transmission, the source voltage, the control and the induction motor drive, Figure 1. The EV dynamics takes into account all the forces on the EV, and it is possible to observe the behavior under linear or angular acceleration, this behavior is important to compare the EV performance and the rated characteristics of the motor drive. The transmission is represented by a gain relation in order to simplify the system. And finally, the induction motor drive is modeled taking into account the inverter and the induction motor.

2.1. Vehicle Dynamics

The EV model is based on the mechanical and aerodynamical theoretical bases [13]. Figure 2 shows the different forces acting on the EV [4]. The equation (1) describes the EV dynamic, which is defined by the total traction force minus the road load forces acting on the EV.

$$m\frac{dv}{dt} = F_t - F_g - F_a - F_r \quad (1)$$

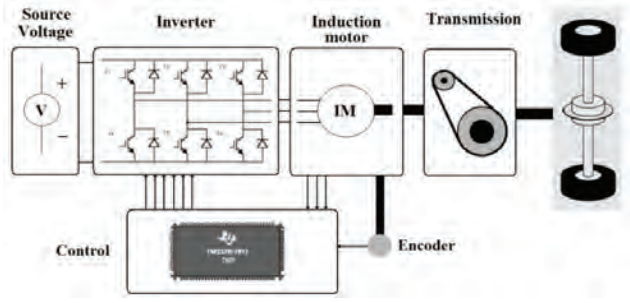


Figure 1. Block diagram of the EV drive system

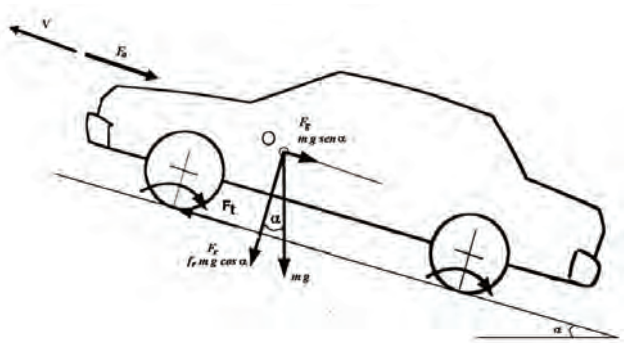


Figure 2. Forces acting on the vehicle.

Equations (2) to (4) describe the opposed forces to the movement and the parameters considered for each one.

$$F_r = f_r mg\cos\alpha \quad (2)$$

$$F_a = \frac{1}{2}\rho ACv^2 \quad (3)$$

$$F_g = mg\,sen\alpha \quad (4)$$

Therefore, the EV dynamic equation is given by:

$$m\frac{dv}{dt} = F_t - mg\,sen\alpha - \frac{1}{2}\rho ACv^2 - f_r mg\cos\alpha \quad (5)$$

2.2. Transmission system

One of the most important issues is the torque imposed to the electric motor. The EV total force (F_t) of equation (5) for linear speed, can be converted to the equivalent motor torque and the angular speed by means of a transmission system (G), Figure 3 shows simplified transmission.

Angular speed (ω) in the electric motor can be related to linear speed (v) of the electric vehicle by (7) through the tire radius (r) and the transmission ratio.

$$\omega = \frac{v}{r}G \quad (6)$$

In the same way, motor torque (τ_t) in the electric motor can be related to the total force (Ft) applied to the electric vehicle by (9) through the tire radius (r) and the transmission ratio (G).

$$\tau_t = \frac{F_t\, r}{G} \quad (7)$$

From (5), the EV dynamic equation (8) can be obtained by introducing the equivalences given in (6) and (7), which is referred to the shaft of the electric motor. Equation (8) is given in terms of torque and angular speed.

$$J\frac{d\omega}{dt} = \tau_t - \tau_g - \tau_a - \tau_r \quad (8)$$

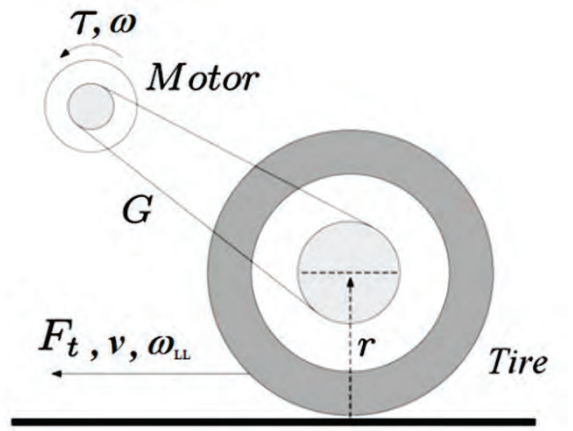


Figure 3. Transmission system of an EV

Hence, the equivalent inertial moment (J) of the total mass of the electric vehicle can be given by (9), also based on the tire radius (r) and the transmission ratio (G).

$$J = \frac{m r^2}{G^2} \quad (9)$$

Also, equivalent expressions of forces acting on the vehicle can be deduced by multiplying by r and divided by G.

2.3. Induction motor drive

Important researches about electric motors have been reported in literature [20,21]; however, the induction motor remains an interesting alternative with a lot of control strategies as proposed in [17]. So, an induction motor drive is developed in this work. The motor drive system consists of an induction motor (IM) and a source voltage inverter with space vector modulation-SVPWM.

The inverter circuit consists of three legs with two current bidirectional switches (IGBT-diode); Figure 4, which is connected to the induction motor by the ABC terminals.

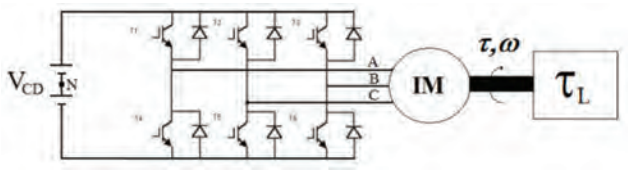


Figure 4. EV drive system.

V_{CD} represents the dc link voltage of the inverter. The three phase voltages needed to supply the induction motor is obtained from ABC output terminals of the inverter and by SVPWM gate signals of the switches (IGBT). A mathematical representation of the inverter is given by (10) [22]. The phase voltages V_{AN}, V_{BN} and V_{CN} are dependent of dc link voltage (V_{CD}) and the modulation signals m_a, m_b and m_c. The modulation signals m_a, m_b and m_c can be only values of 0 or 1, which indicate 0 the switch is opened and for 1 the switch is closed.

$$\begin{bmatrix} V_{AN} \\ V_{BN} \\ V_{CN} \end{bmatrix} = \frac{V_{CD}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} m_a \\ m_b \\ m_c \end{bmatrix} \tag{10}$$

Figure 5 shows the equivalent circuit of a squirrel cage induction motor (IM). The equivalent circuit of the IM is composed of three phases represented by three L-R circuits, for stator and rotor, connected in a Y arrangement.

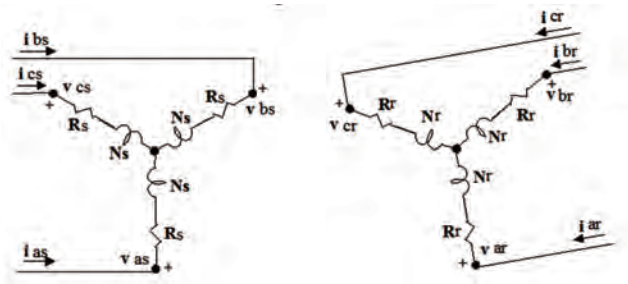


Figure 5. Electric diagram of IM

Each winding produces a magnetic flux, voltages and currents are represented by *as*, *bs*, *cs* (stator), *ar*, *br* and *cr* (rotor). The angle of the magnetic axes of rotor and stator of the same phase represents the angular position of rotor θ and the angular speed by ω. Figure 5 shows electric diagram of the IM, this diagram is used in order to obtain the mathematical representation of the IM.

Based on the analysis of the electric diagram of IM, the stator and rotor behavior can be represented by (11) and (12):

$$v_{abcs} = i_{abcs} R_s + \frac{d\lambda_{abcs}}{dt} \tag{11}$$

$$v_{abcr} = i_{abcr} R_r + \frac{d\lambda_{abcr}}{dt} \tag{12}$$

Where the flux linkage λ for a linear magnetic system is defined by (13)

$$\lambda = \boldsymbol{L}\boldsymbol{i}\,, \tag{13}$$

Hence, the relation between the position and angular speed of the rotor is given by expression (14).

$$\frac{d\theta_r}{dt} = n_P \omega_r \tag{14}$$

And the equation (15) represents the electromagnetic torque, which is dependent in the number of poles (P), the stator currents (i_{abcs}).

$$\tau_e = \left(\frac{P}{2}\right)(i_{abcs})^T \frac{\partial}{\partial \theta_r}[\mathrm{L}'_{sr}]i'_{abcr} \tag{15}$$

Where:

$$\mathrm{L}'_{sr} = L_{ms} \cdot M \tag{16a}$$

$$M = \begin{bmatrix} \cos\theta_r & \cos\left(\theta_r + \frac{2\pi}{3}\right) & \cos\left(\theta_r - \frac{2\pi}{3}\right) \\ \cos\left(\theta_r - \frac{2\pi}{3}\right) & \cos\theta_r & \cos\left(\theta_r + \frac{2\pi}{3}\right) \\ \cos\left(\theta_r + \frac{2\pi}{3}\right) & \cos\left(\theta_r - \frac{2\pi}{3}\right) & \cos\theta_r \end{bmatrix} \tag{16b}$$

$$L_{ms} = \frac{N_s^2}{2} \cdot \frac{\pi \mu_0 r l}{g} \tag{16c}$$

The electromagnetic torque is an important parameter to define the power specification of the electric motor in order to achieve the dynamic imposed by the vehicle. If no other torque is considered then $\tau_t = \tau_e$.

The modeling process developed is carried out for each subsystem with the equations (2-4) and (7-16c),

which represent the electric vehicle. The subsystems of the EV are the induction motor, inverter, powertrain of the vehicle, controller strategy, among others, which are implemented in Matlab/simulink in blocks [23-24]. The developed models are implemented with Simulink tools or "S" functions according to the particular requirements of the model. This allows the subsystem models to be used in other applications. The inter connection of the model subsystems is possible by defining the input and output variables for each subsystem; for example the model of the electric vehicle and the induction motor (IM) is connected by the electromagnetic torque ($\tau_t = \tau_e$). The three phase power supply of the IM is generated by a mathematical function given by (10), which consider voltage level V_{CD} that represents the storage system, such as batteries.

3. TEST BENCH OF EV EMULATOR

The design methodology of the test bench has been focused on the reproduction of the EV dynamics close to the real application. Two approaches of analysis have been taken into account in this methodology: a) the simulation of an EV by the implementation of a mathematical model and b) an experimental way using a test bench implementation, which is controlled by a Digital Signal Processor (DSP). The EV model developed is close to the real application with a great flexibility to modify all the parameters of the EV. However, in spite of having a test bench close to the real application, the parameters to be changed are limited by the natural configuration of the test bench. However, flexibility of variation is possible in the moment of inertia (J), variation of the equivalent forces acting on the EV (represented by torques) and also the flexibility to investigate different types of control strategies. Once the analysis of the dynamic movement of the electric vehicle has been defined, the EV test bench can be designed from the mathematical representation obtained, Figure 6.

Based on the equation (8) the dynamic of movement of the EV can be emulated by the proposed test bench, which is composed of an inertial flywheel and a DC machine. The flywheel can emulate the dynamic behavior of the EV through the moment of inertia (J), which represents the equivalence of the mass of the EV. The forces acting on the EV (F_g, F_a and F_r) are represented by the equivalent torques referred to the motor shaft (τ_g, τ_a and τ_r)

$$J\frac{d\omega}{dt} = \tau_t - \tau_g - \tau_a - \tau_r \tag{17}$$

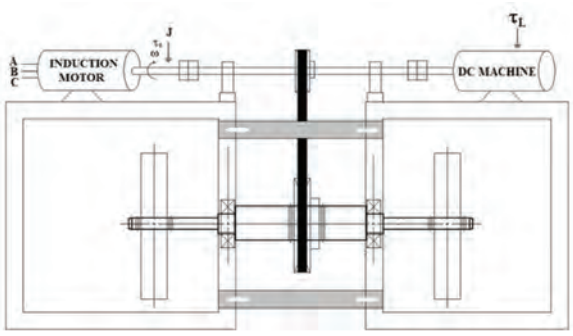


Figure 6. EV test bench

In order to achieve the emulation of the EV by the proposed test bench, some considerations have to be taken into account:

a) A transmission ratio (G) has to be included to have a relation of angular speed-torque in the IM to the EV tires. It is possible with two pulleys in the IM shaft and the tires shaft, which are interconnected by a transmission band.

b) The moment of inertia and the torques have to be affected by a scaling factor (Rt) in order to reproduce the dynamic of a particular EV of interest. This can be defined by the rates of the motor power, the EV of interest and the motor of the test bench.

$$\frac{P_{m2}}{P_m} = Rt \tag{18}$$

The torque emulation is possible using the DC machine. Firstly, it is necessary to identify the different terms from the dynamic equation (8) of the EV. Secondly, the dc machine model is analyzed to define the parameters to be controlled in order to match torques defining the forces acting on the EV. The torques can be emulated with the dc machine operating as a generator with a separate dc power supply to the field windings. A load has to be designed in order to emulate the different torques desired.

1) Torque due to the road resistance force (τ_g)

This torque can be emulated by the dc machine, as it can be seen, this kind of torque is a constant value depending only of constant coefficients.

Vehicule	Test bench
$\frac{JGg\sin\alpha}{r}$	$\frac{JGg\sin\alpha}{r}(Rt)$

2) Torque due to the aerodynamic resistance force (τa)

This kind of torque is dependent on the angular speed (ω) or linear speed and also can be emulated by the dc machine.

Vehicle	Test bench
$\frac{\rho ACr^3\omega^3}{2G^3}$	$\frac{\rho ACr^3\omega^3}{2G^3}(Rt)$

3) Torque due to the rolling resistance force (τr).

This torque is defined by constant parameters and is easily emulated by a dc machine, it depends on the angle (α), if α is constant, this torque remain constant too.

Vehicle	Test bench
$\frac{JGgf_r\cos\alpha}{r}$	$\frac{JGgf_r\cos\alpha}{r}(Rt)$

Analysis and design of torques

In order to correctly emulate the above mentioned torques, it is important to consider the DC machine model. So the torque of the dc machine (τ_L) has to be equal to the torques due to the forces acting on the EV (τ_g, τ_a and τ_r). Therefore:

$$\tau_L = \left[\frac{JGg\sin\alpha}{r} + \frac{\rho ACr^3\omega^2}{2G^3} + \frac{JGgf_r\cos\alpha}{r}\right](Rt) \quad (19)$$

In order to simplify (19) a substitution can be made. Hence,

$$\tau_L = Y\sin\alpha + Z\cos\alpha + X\omega^2 \quad (20)$$

where:

$$\frac{JGg(Rt)}{r} = Y \;,\; \frac{\rho ACr^3}{2G^3}(Rt) = X \;,\; \frac{JGgf_r}{r}(Rt) = Z \;.$$

The DC machine torque (τ_L) is given by (21),

$$\tau_L = L_{af} I_F I_{ar} \quad (21)$$

But considering L_{af} and I_F as constants, then (21) can be simplified replaced by (22),

$$\tau_L = KI_{ar} \quad (22)$$

Equations (20) and (22) are equaled and I_{ar} is given by,

$$I_{ar} = \frac{Y\sin\alpha + Z\cos\alpha + X\omega^2}{K} \quad (23)$$

In equation (23), the current I_{ar} is the parameter that can be adjusted for defining the three terms of the left side of the expression (23). However, (23) has two constant terms that can be reduced to one constant term and the other term depends on the angular speed (ω).

The term I_{ar} can be adjusted using a commutation (converter with current input). Figure 7 shows the proposed circuit configuration. According to the circuit of Figure 7, the following analysis can be done in order to define I_{ar}. Two parameters are introduce R_{load} and duty cycle D

$$I_{load} = I_{ar}(1-D) \quad (24)$$

$$V_a = R_{load}\, I_{ar}(1-D) \quad (25)$$

$$I_{ar} = \frac{V_a}{R_{load}(1-D)} \quad (26)$$

The armature voltage can be given by (27), under constant excitation and considering the armature resistance r_a to be negligible.

$$V_a = K\,\omega \quad (27)$$

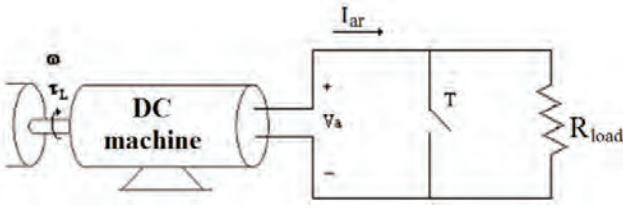


Figure 7. Circuit proposed to emulate the EV forces.

By substituting (27) in (23), the expression (28) can

be obtained. It is interesting to carry out an analysis in order to optimize the R_{load} value and the modulation of duty cycle D by defining a signal reference for D.

$$I_{ar} = \frac{K\ \omega}{R_{load}\left(1-D\right)} = \frac{Ysen\alpha + Z\cos\alpha + X\ \omega^2}{K} \tag{28}$$

$$\therefore R_{load} = \frac{K^2\ \omega}{\left(Ysen\alpha + Z\cos\alpha + X\ \omega^2\right)\left(1-D\right)} \tag{29}$$

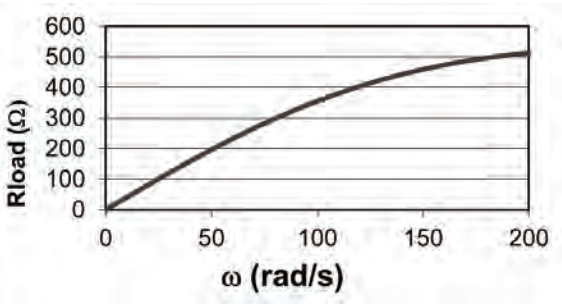


Figure 8. Simulation result for R_{load} behavior versus angular speed (ω) for D=0

Figure 8 shows a simulation varying R_{load} with D=0. This situation is a limiting condition and the current only depends on the R_{load} and the angular speed with constant parameters of the EV. The R_{load} value is selected to be a value within the rated characteristics of the dc machine (voltage and power). In a practical implementation, R_{load} cannot vary continuously. For this reason R_{load} is defined to be 500 Ω. This value allows the angular speed to be varied from zero beyond the nominal speed (Figure 9). This R_{load} only depends on angular speed and the electric vehicle parameters. Equation (30) gives the duty cycle (D) as a function of the forces acting on the EV. Equation (30) is obtained from (28) for a given I_{ar}, it is possible to follow the dynamic behavior of the EV.

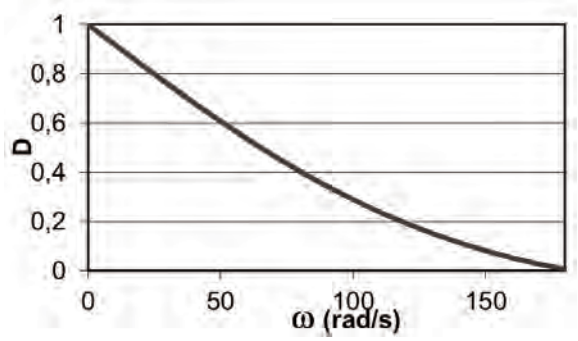


Figure 9. Simulation result for duty cycle behavior versus angular speed (ω) for R_{load} = 500 Ω

Figure 9 shows the evolution of duty cycle D when angular speed (ω) varies from zero to the rated speed.

$$\therefore D = 1 - \frac{K^2\ \omega}{\left(Ysen\alpha + Z\cos\alpha + X\ \omega^2\right)R_{load}} \tag{30}$$

Finally, the EV test bench has been implemented according to the design considerations taken into account; Figure 10 shows the test bench implementation.

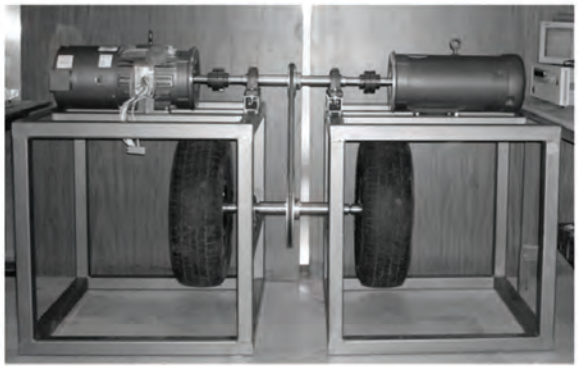


Figure 10. Test bench implementation

4. SIMULATION RESULTS

Now, some simulations are presented for three different drive cycles: 1) City, 2) Road, and 3) under extreme conditions. The results obtained, about instant power and energy, allow the test bench characteristics to be designed and defined.

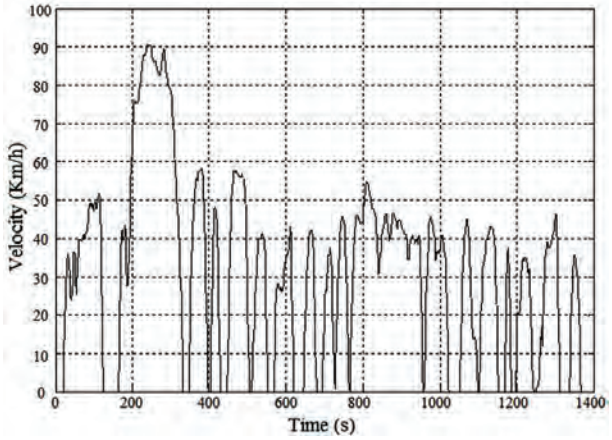


Figure 11. EV city drive cycle

1) City drive cycle

The energy calculated for a city drive cycle is 0.78 MJ at 1400s, this energy allows to design the minimal capacity of battery bank.

2) Road drive cycle

The energy consumption for this type of drive cycle is 1.2 MJ at 800s. Proportionally, this result shows that the energy consumed is greater than in the city drive cycle. This is due to the different driving conditions imposed on the EV, the recovered energy during regenerative braking for road driving conditions is lower than the city drive cycle.

3) Simulation under extreme conditions

The energy consumed for a single cycle defined in Figure 15 is of 34 KJ. In order to compare the energy consumed in this case with the two cases presented before is necessary to consider the same time of driving.

For a case of 800s, the cycle mentioned before has to be repeated 20 times, so the total energy consumed is 0.680MJ. The recovered energy is 15 KJ by cycle, so the total recovered energy during the drive cycle is 0.3MJ.

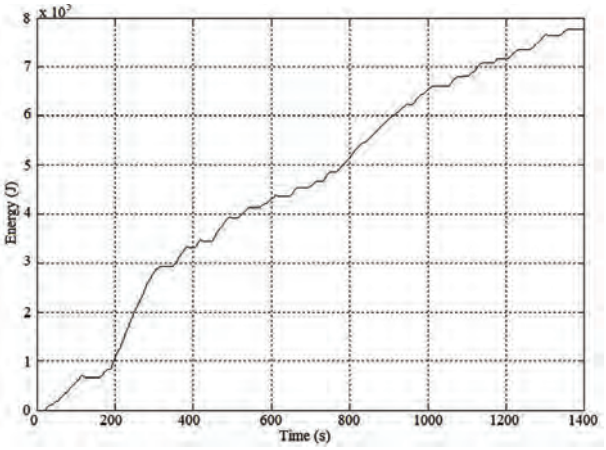


Figure 12. Energy for a city drive cycle

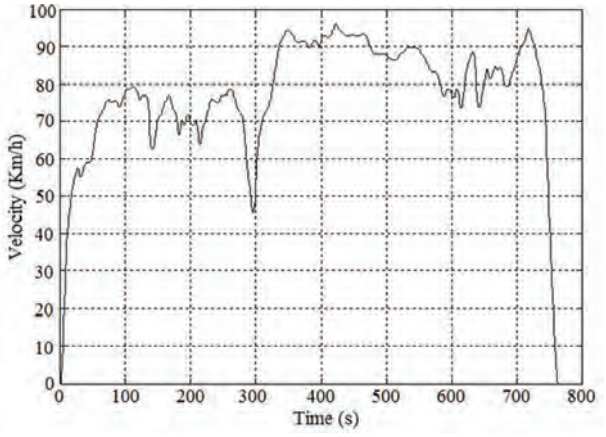


Figure 13. Road drive cycle of EV

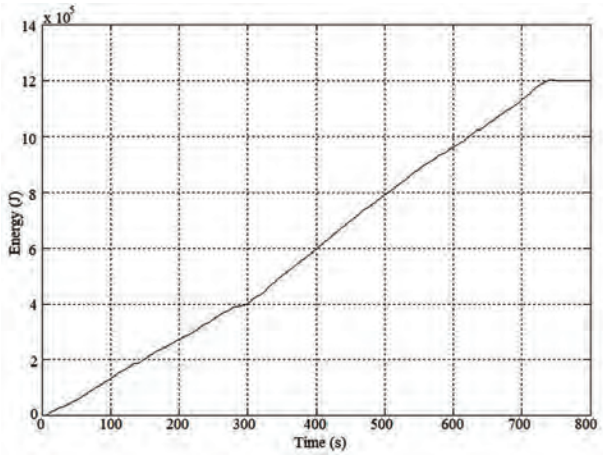


Figure 14. Energy of road drive cycle

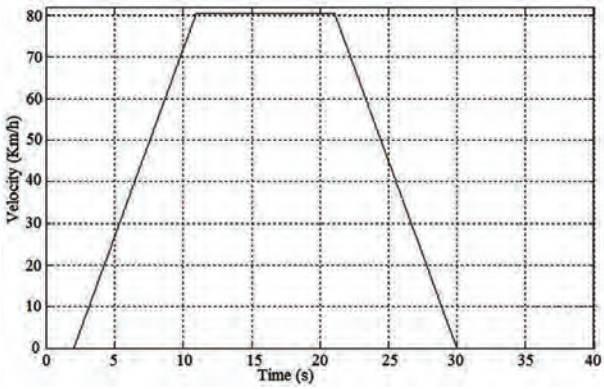


Figure 15. Linear speed of EV for cycle

- During a road drive cycle the power imposed by the EV remains almost constant and its value depends on the speed reached. Under these drive conditions, due to the few regenerative brakes, the consumed energy depends mainly on the speed of the EV.

- The city drive cycle imposes to the EV several repeated accelerations-decelerations (1.5m/s2), which allows to recover additional energy. During this drive cycle is important to impose an adequate control of the braking in order to maximize the available energy to be recovered.

- The proposed drive cycle is designed in order to be used under different driving conditions. The way to design this drive cycle is possible from an analysis of the EV dynamics based on the instant power, the average power, and the energy. From the above information, it is possible to define a hard or soft drive cycle in order to select the rated power of the motor.

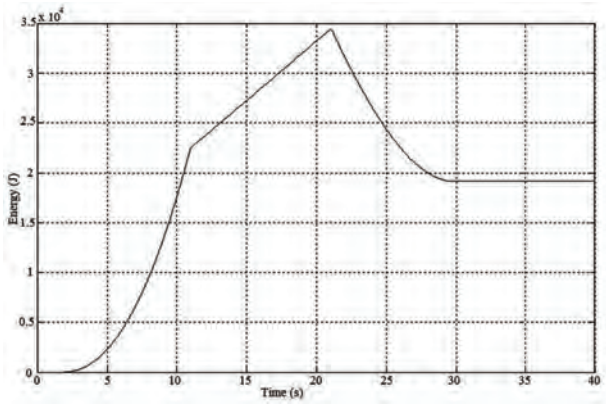


Figure 16. Energy consumed for the cycle proposed

5. EXERIMENTAL RESULTS

The experimental results have been obtained for the parameters given in Table 1; these parameters represent a typical EV of 1600Kg. An interesting test bench characteristic is the possibility of changing the parameters to correctly reproduce the behavior of the EV. Experimental results show the emulation of EV dynamics for different conditions. The phase current before a deceleration is shown in Figure 17. Figure 18 shows the phase current before and during EV acceleration. An increase of current can be observed due to the inertial mass of the vehicle (), which demands additional power during the acceleration or deceleration period. This requires an increase of the instant power, as a consequence the torque demanded from the induction motor (τ_L) is increased. During the deceleration period regenerative braking is possible, so an amount of energy can be recovered and stored to be used during the acceleration period. This can increase the autonomy of the EV.

Table 1. Parameter values of EV and Test bench

Parameters/Coefficient	Value
Electric vehicle (EV)	
m: EV mass	1600 Kg
C: Aerodynamic drag coefficient	0.42
R: Tire radius	0.32 m
f_r: Rolling resistance force	0.008
G: Transmission ratio	12.8
A: Area	2.2 m/s^2
J : EV equivalent inertia (or J_e)	0.2467 kg m^2
Motor	
P_m: Rated power of motor	3 hp (2.24 kW)
τ_e: Rated torque	12 N-m
ω: Rated angular speed	1800 rpm
V_s: Stator rated Voltage	230 V
I_s: Rated current	8 A
η_m: Efficiency	89.5 %
Drive	
η_i: Inverter efficiency	90 %
IGBT: Module CM100DU-24HB	1200 V/100 A
V_{dc}: dc voltage (26 batteries)	312 V
• Battery voltage	12 V
• Battery capacity	12 Ah
• Total Energy	13.48 MJ

Figures 17 and 18 show the correct dynamic emulation of the EV by the test bench. In Figure 19, it can be observed that the mechanical speed of the motor correctly follows the dynamic of the speed imposed by the reference control signal.

5. CONCLUSIONS

A methodology to analyze the EV dynamic based on simulation and experimentation has been proposed in this paper. The mathematical analysis allows a model to be developed and a test bench to be designed based on typical EV parameters. The test bench implemented can emulate a scaled EV. The equivalences between a real and the scaled EV, can be from the EV dimensions or from the capacity of the motor. The acceleration curves of the EV and the test bench motor have been presented; these curves show the demands of power, torque and energy to be taken into account to define the EV design considerations, such as specifications of the motor, converters, and others elements. Also, the necessary energy to handle the EV during a drive cycle can be used to design the battery bank.

The results obtained with the test bench and by simulation give a more complete and detailed information for the design. The tool proposed allows the EV dynamic to be studied by simulation varying the parameters and conditions with wide flexibility. These results can be validated with the experimental test bench tool. The EV has been correctly emulated by the test bench under different operating conditions (starting, cruising speed and braking). It is important to mention that the proposed test bench considered improves the EV behavior, which is close to real EV applications.

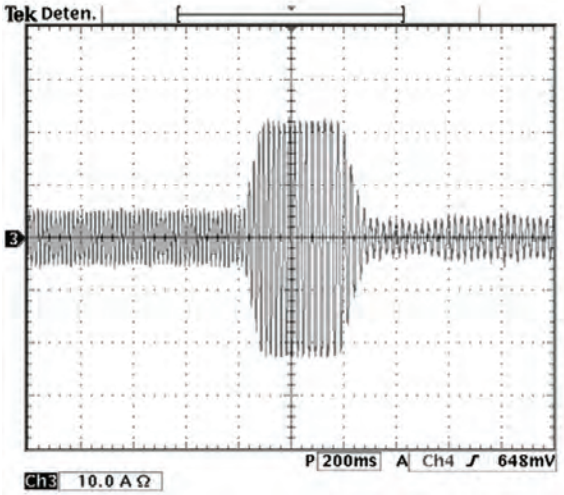


Figure 17. Phase Current during deceleration (1800 to 1200 RPM)

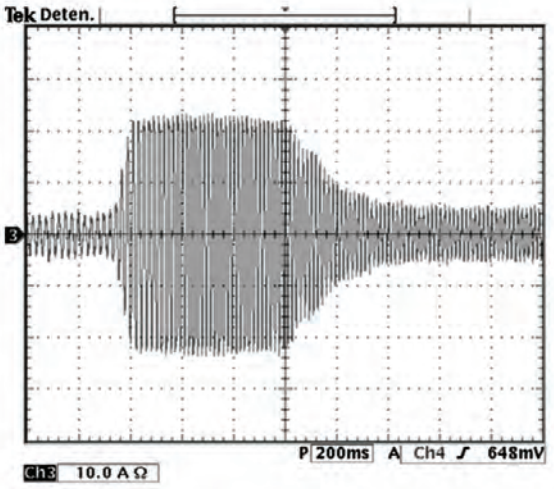


Figure 18. Phase Current during acceleration (1200 to 1800 RPM)

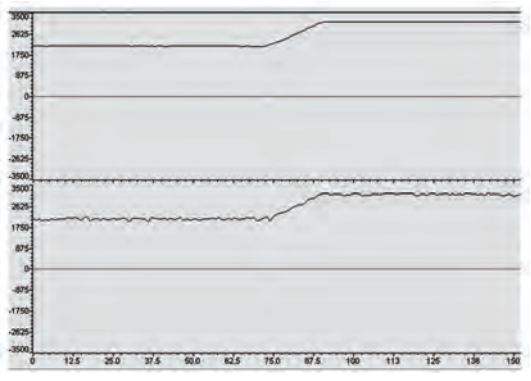


Figure 19. Motor speed and reference speed in acceleration

REFERENCES

[1] Popovic-Gerber, J. et al., "Power Electronics Enabling Efficient Energy Usage:Energy Savings Potential and Technological Challenges" IEEE Trans on Power Electronics, Vol. 27, No. 5, May 2012.

[2] Mejia, D.A., Lopez, J.M. and Gallego, L.A., "Energy Generation Expansion Planning Model Considering Emissions Constraints", Dyna, year 77, No. 163, pp. 7584.Medellín, September, 2010.

[3] Ramos, C.A., Jaramillo, A. A. and Pérez, E.A., "Design and implementation of a PEM fuel cell emulator for static and dynamic behavior", Dyna, year 78, No. 168, pp. 108-118. Medellín, August, 2011.

[4] Ehsani, M., Gao, Y., Gay, S. and Emadi, A., "Modern Electric; Hybrid electric, and Fuel Cell Vehicles", CRC Press, 2005.

[5] Thounthong, P., Pierfederici, S., Martin, J.P., Hinaje, M. and Davat, B., "Modeling and Control of Fuel Cell/ Supercapacitor Hybrid Source Based on Differential Flatness Control", IEEE Trans on Vehicular Technology, Vol. 59, No. 6, pp. 2700-2710.

[6] Chan, C., "The state of art of electric and hybrid Vehicles", IEEE Proc., Vol. 95, no. 4, pp. 704-718. 2007

[7] Dong, W., Choi, J.Y., Li, Y., Boroyevich, D., Lee, F.C., Lai, J. and Hiti, S., "Comparative experimental evaluation of soft-switching inverter techniques for electric vehicle drive applications", IEEE Industry Applications Conf., pp. 1469-1476, vol 3, 2001.

[8] Zandi, M., Payman, A., Martin, J.P., Pierfederici, S., Davat, B. and Meibody-Tabar, F., "Energy Management of a Fuel Cell/Supercapacitor/Battery Power Source for Electric Vehicular Applications", IEEE Trans on Vehicular Technology, Vol. 60, No 2, pp. 433 – 443, 2011.

[9] Kessels, J.T.B.A., Koot, M., De Jager, B., Van Den Bosch, P.P.J., Aneke, N.P.I. and Kok, D.B., Energy Management for the Electric Powernet in Vehicles With a Conventional Drivetrain", IEEE Trans on Control Systems Technology, Vol. 15, No. 3, pp. 494-505, 2007.

[10] Briat, O., Vinassa, J.M., Lajnef, W., Azzopardi, S. and Woirgard, E., "Principle, design and experimental validation of a flywheel-battery hybrid source for heavy-duty electric vehicles", IET Electric Power Applications, Vol. 1, No 5 pp.

665 – 674, 2007.

[11] Kessels, J.T.B.A., Koot, M., De Jager, B., Van Den Bosch, P.P.J., Aneke, N.P.I. and Kok, D.B., "Energy Management for the Electric Powernet in Vehicles With a Conventional Drivetrain", IEEE Transactions on Control Systems Technology, Vol. 15, No. 3, pp. 494 – 505, 2007.

[12] Camara, M.B., Gualous, H., Gustin, F. and Berthon, A., "Design and New Control of DC/DC Converters to Share Energy Between Supercapacitors and Batteries in Hybrid Vehicles" , IEEE Trans on Vehicular Technology, Vol. 57, No. 5, pp. 2721–2735, 2008.

[13] Haddoun, A., Benbouzid, M.E.H., Diallo, D., Abdessemed, R., Ghouili, J. and Srairi, K., "Modeling, Analysis, and Neural Network Control of an EV Electrical Differential", IEEE Trans. on Industrial Electronics, Vol. 55, No. 6, June 2008.

[14] Njeh, M., Cauet, S., Coirault, P. and Martin, P., "Hα control strategy of motor torque ripple in hybrid electric vehicles: an experimental study", IET Control Theory & Applications, Vol. 5, No. 1, pp. 131 – 142. 2011.

[15] Rodic, M., Jezernik, K. and Trlep, M., "A Feedforward Approach to the Dynamic Emulation of mechanical Loads", 2004 35th Annual IEEE Power Electronics Specialists Conference, Aachen, Germany, pp. 4595-4601, 2004.

[16] Hewson, C.R., Asher, G.M., Summer, M. "Dynamometer control for Emulation of Mechanical Loads", IEEE Trans. on lndustry Applications, pp. 1511-1518, 1998.

[17] Akpolat, Z.H., Asher, G.M. and Clare, J.C., "Dynamic Emulation of Mechanical Loads Using a Vector Controlled Induction Motor-Generator Set," IEEE Trans. on Industrial Electronics. vol. 46, no. 2, pp. 370-379, 1999.

[18] Akpolal, Z.H., Asher, G.M. and Clare, J.C., "Experimental dynamometer emulation of non-linear mechanical loads," IEEE Trans. on lndustry Applications, vol. 35, no. 6, pp. 1367-1273, 1999.

[19] Rodic, M. and Jezernik, K., "An approach to the dynamic emulation of mechanical loads" in Procs. of International conference on industrial technology IEEE ICIT 7003, Maribor, Slovenia, December 10-12, pp. 489-494, 2003.

[20] Rossini, W.M., Alvarenga, B., Chabu, I.E., Da Cruz, J.J., Cardoso, J.R. and Sales, R.M., "New Concept for Lifting in Onshore Oil Wells", IEEE Transactions on Industry Applications, Vol. 44, No. 4, pp. 951 – 961, 2008.

[21] Jinxin Fan, Chengning Zhang, Zhifu Wang, Yugang Dong, Nino, C.E., Tariq, A.R. and Strangas, E.G., "Thermal Analysis of Permanent Magnet Motor for the Electric Vehicle Application Considering Driving Duty Cycle", IEEE Transactions on Magnetics, Vol. 46, No. 6, pp. 2493 – 2496, 2010.

[22] Trzynadlowski, A.M., "Control of induction motor", Academic press, 2001.

[23] SIMULINK Simulation and Model-Based Design. Available: http://www.mathworks.com/products/ simulink [citado el 20 de agosto de 2013].

[24] Cholula, S. "Control Inteligente del frenado en propulsión eléctrica con regeneración". Tesis maestría CENIDET, Agosto 2005.

[25] Langaríca-Córdoba, D., "Control de un motor brushless Trifásico para aplicación a Vehículos eléctricos", Tesis maestría CENIDET, Octubre 2010.

BARRIER PROPERTY DETERMINATION AND LIFETIME PREDICTION BY ELECTROCHEMICAL IMPEDANCE SPECTROSCOPY OF A HIGH PERFORMANCE ORGANIC COATING

DETERMINACIÓN DE PROPIEDADES BARRERA Y DE PREDICCIÓN DE TIEMPO DE VIDA POR ESPECTROSCOPÍA DE IMPEDANCIA ELECTROQUÍMICA DE UN RECUBRIMIENTO ORGÁNICO DE ALTO RENDIMIENTO

JORGE ANDRÉS CALDERÓN-GUTIERREZ
DSc., Universidad de Antioquia - CIDEMAT, Medellín, Colombia. jacalder@udea.edu.co

FRANKY ESTEBAN BEDOYA-LORA
MEng., Universidad de Antioquia - CIDEMAT, Medellín, Colombia. frankybedoya@gmail.com

Received for review March 28 th, 2013, accepted September 20th, 2013, final version September, 24 th, 2013

ABSTRACT: The anticorrosion performance of an Epoxy-Mastic organic coating was evaluated during continuous immersion in saline solution using electrochemical impedance spectroscopy (EIS). The typical parameters of pore resistance and charge transfer resistance were determined employing an equivalent electric circuit. Constant phase elements (CPE) were used in order to determine fraction of water absorbed, mass diffusion, solubility and the swelling coefficients, as well as to predict the failure times of the coating. The results found by EIS measurements match very well with the high resistance to deterioration exhibited by the coating. It was also found that the excellent protection performance of the coating was mainly due to low water solubility and low permeability.

Key words: Organic coatings, electrochemical impedance spectroscopy, water diffusion, delaminated area, failure time prediction.

RESUMEN: El desempeño anticorrosivo de un recubrimiento orgánico tipo Epoxy-Mastic fue evaluado en condiciones de inmersión continua en solución salina usando espectroscopía de impedancia electroquímica (EIS). Se determinaron los parámetros típicos como la resistencia de poro y resistencia a la transferencia de carga usando un circuito eléctrico equivalente. Se usaron elementos de fase constante (CPE) para determinar la fracción de agua absorbida, coeficientes de difusión de masa, solubilidad y coeficientes de hinchamiento, así como también para predecir los tiempos de falla de dicho recubrimiento. Los resultados hallados por medio de medidas EIS concuerdan con la alta resistencia al deterioro que exhibe el recubrimiento. El excelente desempeño protector es debido principalmente a la baja solubilidad y permeabilidad de agua.

Palabras clave: Recubrimientos anticorrosivos, espectroscopía de impedancia electroquímica, difusión de agua, área delaminada, predicción de tiempos de falla.

1. INTRODUCTION

Two decades ago Haruyama [1] proposed a method for calculating the deterioration of organic coatings exposed to a corrosive environment using the equivalent circuit shown in Fig. 1. Haruyama's methodology phenomenologically explains the processes that occur inside organic coatings when they are exposed to an electrolyte and when there are minor defects or imperfections, such as small pores, in the coating. The Randles circuit (Fig. 2) has been used extensively to evaluate organic coatings that exhibit highly capacitive behavior or to evaluate pore-free coatings. However, the Randles circuit becomes useless after a few hours of continuous immersion because new time constants in the impedance spectra may appear. A complete description of each of the passive elements making up the two circuits shown in Fig. 1 and 2 can be found in numerous sources [2, 3]. The coating resistance (R_C) is closely related to the state of the coating, its additives or pigments, porosity and type of resin. The coating capacitance (C_C) is associated with the amount of water absorbed during the initial stages of exposure to the electrolyte [4]. The charge transfer resistance (R_{ct})

is directly related to the susceptibility to corrosion of the substrate and theoretically is the most appropriate parameter for measuring the protective properties of the coating [5]. The double layer capacitance (C_{dl}) is almost unanimously associated with the delaminated area of the coating [2]. When the diffusion process is shown in the impedance spectrum, the use of the Warburg impedance (Z_w) is normally accepted in order to give a physicochemical explanation to these processes [6].

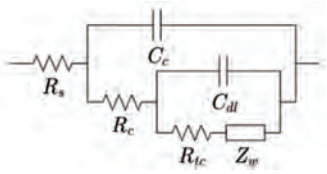


Figure 1. Equivalent electrical circuit used by Haruyama to explain the deterioration of a barrier organic coating using electrochemical impedance spectroscopy [1]

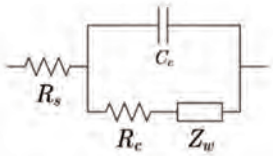


Figure 2. Randles circuit commonly used to simulate the electrochemical impedance of a defect-free barrier coating

2. THEORETICAL TREATMENT

The five EIS parameters described above have been used to evaluate the protective properties of organic coatings and even obtain empirical correlations [7-9]. However, when it is desired to adjust the EIS response using these passive elements of the equivalent circuit, in most cases, the fit is poor due to lack of uniformity in the coating or in the substrate [10]. Additionally, other phenomena can reduce the quality of the experimental EIS fit, for example geometrical effects such as polymer swelling [11], non-Fickian diffusion or variations in the time constants related to the corrosion processes. A commonly used alternative to reduce fit problems is to use constant phase elements (CPEs) instead of pure capacitances. It has been suggested that the existence of a distribution of relaxation times is the cause of CPEs [12]. A CPE can be thought of as an imperfect capacitor and allows a better fit of the experimental impedance when there is a flattening of the semicircle represented in the Nyquist plot [10, 13, 14]. According to Eq. 1, CPEs provides two terms: the pseudo-capacitance (Y_0) and an exponential constant (n).

$$Z = \frac{(j\omega)^{-n}}{Y_0} \quad (1)$$

In many cases the pseudo-capacitance can be represented as a simple capacitance (occurring when $n = 1$ since there is an absence of deviations from ideality in a perfect capacitor, *i.e.* there are no interferences caused by factors already cited). When equations and empirical relationships are used it is necessary to perform a conversion that enables the computation of the actual capacitance from a simulated CPE. This can be done with Eq. 2 [15].

$$C_c = Y_c(\omega_{\max})^{n-1} \quad (2)$$

Now, considering $\omega_{max} = 1/RC$, where ω_{max} is the frequency of the local maximum in imaginary region obtained from the Nyquist diagram, the actual capacitance can be calculated with Eq. 3.

$$C_c = \frac{(Y_c R)^{\frac{1}{n}}}{R} \quad (3)$$

Using the above relationships, several methods have been developed to evaluate the performance of organic coatings, as well as correlations and methods derived from the parameters obtained by the use of the circuit proposed by Haruyama (Fig. 1). One of the most popular methods for evaluating the performance of organic coatings, but also subject to substantial restrictions, is the breakpoint frequency method. This method, initially proposed by Haruyama [1], describes the relationship between the delaminated area and the total area of the sample, in accordance with Eq. 4 and 5.

$$f_b = f_b^0 \frac{A_d}{A} \quad (4)$$

$$f_b^0 = \frac{1}{2\pi\rho\epsilon\epsilon_0} \quad (5)$$

Where A_d and A are the delaminated area and the total area of the substrate, respectively. f_b^0 is the proportion constant, which can be obtained from Eq. 5. ρ is the electrical resistivity of the coating. ϵ is the relative electric permittivity and ϵ_0 is the vacuum electric permittivity (8.86×10^{-14} *F/cm*). The breakpoint frequency can be obtained by performing a scan at high frequencies and is located exactly at the point where the phase angle falls for the first time to 45° [1].

One of the first approaches when studying the lifetime of an organic coating using electrochemical impedance measurements is to determine the time to failure of the coating. This is the time after which the coating no longer offers significant protection against corrosion. Eq. 6, obtained by Bierwagen [16], makes it possible to establish the time after which a coating fails, according to an arbitrary failure impedance value.

$$t_{fail} = \theta \left[\ln \left(\frac{|Z|_0 - |Z|_m}{|Z|_{fail} - |Z|_m} \right) \right] \quad (6)$$

Where $|Z|_m$ is the impedance modulus of the substrate, $|Z|_0$ is the initial coating module and θ is considered as the characteristic decay time of the material. The difficulty of this method is in establishing a real value for $|Z|_{fail}$ from which it can be assumed that the coating has lost its protective properties. The impedance module used in this equation must be taken at a low experimental impedance frequency range, generally between 1×10^{-3} and 5×10^{-3} [16]. A coating is usually considered to have inadequate protection when the charge transfer resistance exhibits values below 10^6 ohm [10].

All organic coatings are somewhat permeable. This permeability may worsen with the addition of certain types of pigments, for example zinc powders, producing a significant reduction in the barrier property. However, there is an increase in the corrosion resistance of the coatings after prolonged periods when coating has a pigment volume concentration (PVC) greater than 60 [17]. The absorbed water content is of special interest since once the polymer has been saturated with water, even in the absence of electro-active species, the occurrence of corrosion phenomena, swelling, loss of adhesion and deterioration of mechanical properties of the coating are significantly enhanced. Furthermore, it has been found that the diffusion of water controls the start time of delamination and consequently the loss of protective capacity [18, 19].

The parameter most related to the amount of water absorbed is the coating capacitance, (C_C), expressed as:

$$C_c = \epsilon \epsilon_0 \frac{A}{d} \quad (7)$$

Coating capacitance is directly related to the electrical permittivity of the medium (ϵ), the permittivity of the vacuum (ϵ_0), the coating area (A), and the thickness of the coating (d). When the absorption process starts, coating capacitance changes due to the incorporation of water molecules. Water has a relative permittivity of 80, while the resins used in anticorrosive coatings have permittivity value typically in the range of 2.5 to 10. Therefore, the more water absorbed, the greater the capacitance.

Brasher and Kinsgbury [20], in their widely cited research, proposed a relationship to calculate the amount of water absorbed. This relationship is shown by Eq. 8.

$$X_v = \frac{\log\left(C_{c(t)} / C_{c(0)}\right)}{\log 80} \quad (8)$$

Where X_v is the fraction volume of absorbed water, and $C_{c(t)}$ and $C_{c(0)}$ are the capacitance of the coating at a given time and at the initial time, respectively. To make use of this expression it is necessary to fulfill several requirements, such as the absence of swelling in the coating, homogeneous distribution of the water, and absence of polar solvents. Additionally, the fraction of water absorbed must be relatively small [21]. Furthermore, it has been numerically shown that the values reported with this equation differ considerably from gravimetric data [22]. If the above conditions are satisfied, mathematical expressions may be obtained for calculating the diffusivity of the water using the relationship proposed by Brasher-Kingsbury. Eq. 9 for example, can be used to calculate water diffusivity [18], whereas more complex correlations can be used to obtain the swelling coefficients, such as Eq. 10 [11, 23].

$$\frac{\ln C_{c(t)} - \ln C_{c(0)}}{\ln C_{c(\infty)} - \ln C_{c(0)}} = \frac{2\sqrt{D}}{L\sqrt{\pi}}\sqrt{t} \quad (9)$$

$$\log C_t = \log\left(\frac{C_{c(\infty)}}{C_{c(0)}}\right) \left\{1 - \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\left[\frac{-(2n+1)^2 D\pi^2}{4L^2}t\right]\right\} + SC_c t + \log C_0 \quad (10)$$

As previously mentioned, $C_{c(0)}$ and $C_{c(t)}$ are the initial coating capacitance and the coating capacitance at a given time, respectively. $C_{c(\infty)}$ is the capacitance of water saturated coating, *i.e.* when the capacitance does not vary significantly with time. D is the diffusion coefficient of water in the coating. L and SC_c are the thickness and swelling coefficient of the coating respectively.

The solubility (S) and the permeation coefficient (P) can be calculated with simple relationships expressed in Eq. 11 and 12 respectively. These parameters are usually employed for comparison purposes.

$$S = \frac{V_\infty}{V_c + V_\infty} \times \rho_w = \frac{\log C_{c(\infty)} / C_{c(0)}}{\log 80} \times \rho_w \quad (11)$$

$$P = D \times S \quad (12)$$

Where ρ_w is the water density in $kg.m^{-3}$ and V_∞ and V_c are the volume occupied by the water to infinite time and the volume of the dry coating respectively.

2. MATERIAL AND METHODS

Carbon steel specimens were prepared by sandblasting to white metal grade SP-3, followed by washing with a mixture of ethanol/water and degreasing with acetone before applying the organic coating. The steel samples were painted with a commercial high solids epoxy mastic coating by spraying. In order to facilitate the coating application, 30% v/v of solvent was added. Table 1 shows general characteristics of the coating. Additionally, the coating was characterized by Micro-Raman Spectroscopy (Micro-Raman, HoribaJovinYvon), in order to corroborate the coating composition. As can be seen in Raman spectrum shown in Fig. 3, the characteristic vibration bands related with polydiphenylsiloxane [24], titanium dioxide (rutile phase) as pigment [25] and epoxy compounds [26] appear in the Raman spectrum. This is in agreement with information presented in Table 1.

Table 1. General characteristics of the studied coating

Coating	Epoxi Mastic Polysiloxane
Chemical nature	Epoxy-silicone hybrid
Catalyzer	Organosilane
Pigments	TiO_2, $CaSiO_3$, Silica
Solid Volume	93 % (Catalyzed)

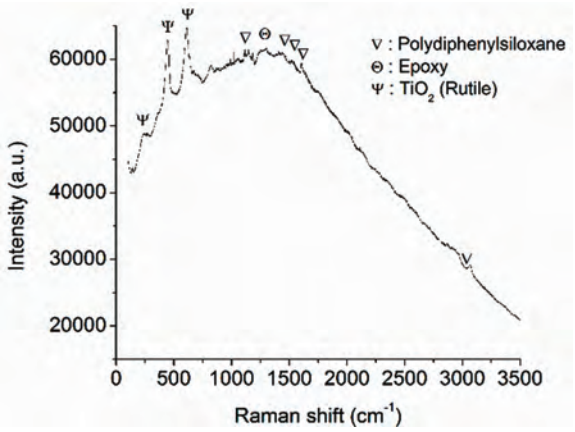


Figure 3. Raman spectra of mastic epoxy coating

Before the test, the painted specimens were kept in a natural dry room for 4 weeks to achieve a convenient curing and drying of the coating. The average thickness of the coatings dry film was 172 ± 19 microns. The painted steel coupons were exposed in continuous immersion in a 3% NaCl solution (0.5M NaCl) for about 440 days. The protective properties of the coating and corrosion phenomena were evaluated by measuring open-circuit potential and electrochemical impedance over time, following visual inspection. Electrochemical tests were conducted with a conventional three electrode cell, using the painted samples (38.5 cm^2 exposed area) as the working electrode. A saturated calomel electrode (SCE) and Pt plate were used as reference and counter electrodes respectively. The impedance measurements were performed at open circuit potential in a frequency range of 40 kHz to 3 mHz, using a perturbation amplitude of 10 mV. An AUTOLAB PGSTAT-30 potentiostat/galvanostat was used.

3. RESULTS AND DISCUSSION

Fig. 4 shows the Nyquist and Bode plots of impedance measurements of epoxy mastic coating taken during 441 days of continuous immersion in 0.5 M NaCl solution. The excellent performance and the efficient barrier protection at the beginning of the test can be clearly seen, as practically all diagrams consist of an open capacitive loop, with impedance values of 10^{10} ohm.cm^2. The capacitive loop tends to lean toward the real axis as the time of immersion increases, which means diminishing of the coating resistance and increases of capacitance due to electrolyte permeation.

Fig. 5 presents the evolution of the resistances related to the coating/metal system. The coating resistance (R_C) and charge transfer resistance (R_{ct}) tend to fall during the first days of immersion as a result of the absorption of water and the establishment of oxy-reduction reactions at the metal-electrolyte interface. However, this process seems to stop after 100 days of immersion since the charge transfer resistance tends to stabilize after this time. The modulus of the impedance follows the behavior of the charge transfer resistance more closely than the resistance of the coating. This suggests that the evaluation of the general anticorrosive performance of the coating should be done preferably by assessing the behavior of the (R_{ct}) during the evaluation time. It can be seen (R_{ct}) does not fall below a protection level value of 10^6 ohm.cm$^{2.}$ at any time and that (R_C) is also maintained high, although with a tendency to decrease. It was confirmed visually that there were no defects such as blistering, delamination, cracking or corrosion.

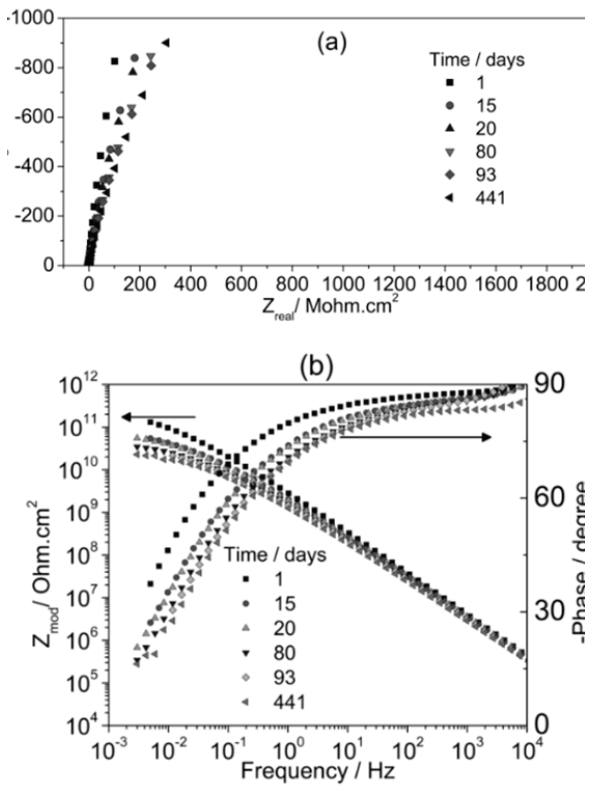


Figure 4. Electrochemical impedance spectra of epoxy mastic coating at different times of exposition in 3% NaCl solution: (A) Nyquist plots. (B) Bode plots

The open circuit potential (OCP) vs SCE values are presented in Fig. 6. A rapid evolution of (OCP) to anodic values was observed, from -250 mV in the first days of immersion up to -0.02 mV after 100 days, from then on the potential remained almost constant at -0.02 mV. The progressive increase of the OCP and stabilization in anodic values is consistent with that observed in the evolution of the (R_{ct}), which tends to remain constant after 100 days of exposure. The progressive increase of the OCP and its subsequent stabilization at positive potentials indicates the existence of an anodic control of the oxidation-reduction process on the metal substrate and the anodic area decreases as a result of the barrier imposed by the polymeric coating. Because of the high barrier property of the coating against the diffusion of oxygen molecules and the passage of the electrolyte, a cathodic control, which is commonly observed in the initial stages of the under-film corrosion process [27], is not present.

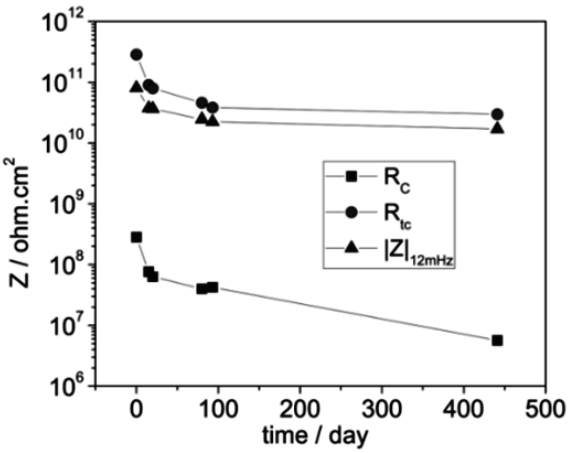


Figure 5. Evolution of coating resistance (R_C), charge transfer resistance (R_{tc}) and impedance module at 12 mHz ($|Z|_{12mHz}$) of epoxy mastic coating at different times of exposition in 3% NaCl solution

3.1. Delaminated Area

One of the most cited methods for delaminated area calculation is the breakpoint frequency method [1]. As mentioned in the introduction section of this work, to properly use this method the experimental measurements must match the impedance of the circuit of Fig. 1. Furthermore, the two time constants associated with the coating and the metal-coating interface must appear sufficiently spaced to allow the direct reading of the breakpoint frequency from the phase diagram. In this study, even though the experimental data fits

very well with the impedance of the electrical circuit using constant phase elements, the separation between the two time constants of the phenomena is not clear and the breakpoint frequency is not easily detected. For this reason the aforementioned analysis method is not suitable for the quantitative evaluation of the coating deterioration and coating delamination.

The inability to evaluate the anticorrosion performance of intact coatings by the breakpoint frequency method has previously been reported by Pistorius [28]. Pistorius establishes that intact coatings only show one time constant in the impedance spectrum or two time constants very close to each other. According to the Haruyama postulation, for a clear observation of the two time constants, the relationship $C_C.R_C << C_{dl}.R_{tc}$ must be satisfied. This situation prevents the accurate determination of the breakpoint frequency value.

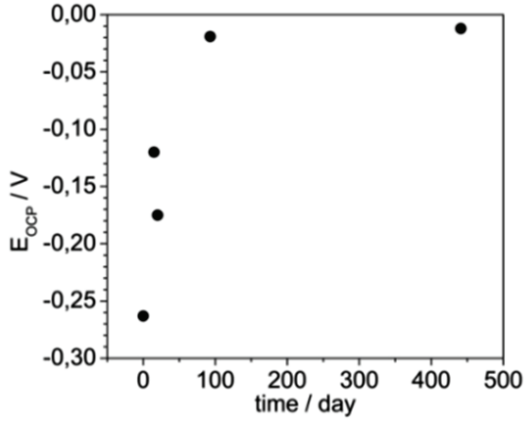


Figure 6. Open circuit potential vs SCE of epoxy mastic coating at different times of exposition in 3% NaCl solution

3.2. Failure time determination

The time to failure of the coating (t_{fail}) was calculated once parameter θ was known. The θ parameter can be easily obtained from Eq. 6 and by performing a linear regression of the impedance module (when $f \rightarrow 0$) versus immersion time. It was necessary to use a minimum frequency of 12 mHz to calculate the impedance module, since this is the lowest frequency that can be attained without data dispersion. The $|Z|_m$ value must be negligible compared to $|Z|_{fail}$. After completing the regression process, a slope of $1/\theta = 0.01058$ was obtained, which corresponds to $\theta = 94.5$ days. Upon defining a failure impedance of 9.9×10^6 *ohm.cm*^{2}, *i.e.* when the resistance of the system drops to the range of minimum protection (10^6 *ohm.cm*^{2}) [10, 29], the failure time found was t_{fail} = 851 days. This result is the predicted time from which the coating would not be expected to offer enough corrosion protection to the substrate under immersion conditions in 0.5 M NaCl solution, and coating damage signs would be observable. The experiment performed in this study was carried out for about 441 days. Neither coating damage nor signs of corrosion were observed during this time, as predicted by the Bierwagen expression [16] and experimentally corroborated using electrochemical impedances and visual inspection.

3.3. Fraction of water absorbed and diffusion coefficients for the Fick process

The calculation of the fraction of water absorbed X_v was performed using the Brasher and Kinsgbury expression [20], as can be seen in Eq. 8. The real coating capacitance (C_C) was calculated from the pseudo-capacitance (Y_C) and CPE values using Eq. 3. Fig. 7 shows the variation in the fraction of absorbed water (X_v) and the pseudo-capacitance (Y_C) during the immersion of metal coated samples in 0.5M NaCl solution. It can be seen that even after 441 days of immersion, no stabilization of the fraction of water absorbed is reached. Therefore, it can be said that complete water saturation of the epoxy mastic coating is not achieved. This corroborates the good performance of the barrier and anticorrosion properties of the coating, as the barrier properties are shown to persist longer than 400 days of immersion. The variation of the CPE exponential term for coating capacitance (n_c) over time is shown in Fig. 8. It can be seen that the value of n_c remains relatively constant during the first 120 days of immersion, indicating that there is little geometric variation of the polymer during immersion. This is due to the occurrence of low water absorption into the polymer during this period of time. After 120 days of immersion, a slight decrease of the n_c value can be observed, reaching a final value of 0.94 at the end of the immersion (441 days). This means that geometrical change, due to swelling of the polymer caused by water absorption, is only observable at longer immersion times. The swelling of the coating caused by water absorption could imply significant changes in the model of electrolyte diffusion through the polymer. Therefore, this must be taken into account

for the calculation of the diffusion coefficient for longer immersion times, as discussed in the next section.

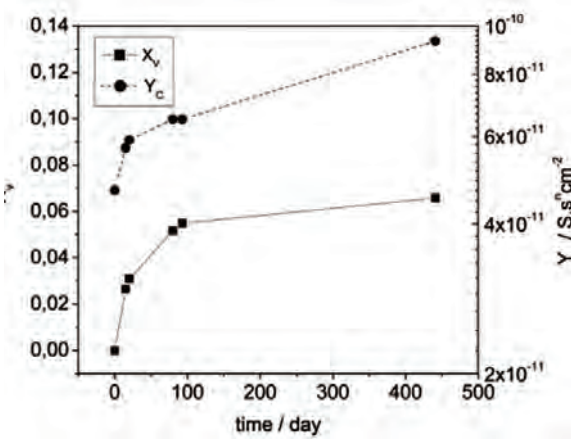


Figure 7. Coating pseudo-capacitance (Y_C), and water fraction (X_v) of epoxy mastic coating at different times of exposition in 3% NaCl solution

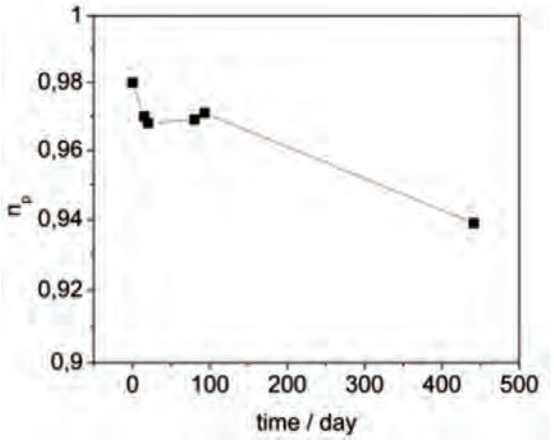


Figure 8. Exponential term n_c associated with the coating CPE at different times of exposition in 3% NaCl solution

The diffusion coefficient of water was calculated using Eq. 9. By plotting the left side of Eq. 9 vs the square root of the immersion time a linear behavior can be seen during the first 100 days of immersion, see Fig. 9. This behavior corresponds to a Fick type diffusion of the water through the coating. The calculated diffusivity value was $D_{AB} = 1.91\times10^{-15}\ m^2/s$. This diffusivity value is significantly lower compared to values reported for a conventional epoxy coating, which are in the range of $10^{-13}\ m^2/s$ [11, 18, 30, 31]. However, the low diffusivity value obtained for the epoxy mastic coating during the early stage of immersion is in agreement with the excellent anticorrosion performance showed by the coating, even after 400 days of continuous immersion. In order to verify the complete water diffusivity property of the epoxy mastic coating, a second calculation considering the film swelling due to water absorption was carried out. This calculation considered the nonlinear diffusion mechanism (non-Fickian diffusion process) shown by the coating during a longer immersion time. In addition, swelling and permeation coefficients of the coating can also be obtained.

3.4. Water diffusion coefficient for nonlinear behavior

The calculation of the non-Fickian diffusion and the swelling coefficients during water permeation into the coating was performed using Eq. 10. It was executed through a non-linear regression using the Levenberg-Marquardt algorithm and truncation at the 8th term when the summation counter was $n = 7$. It was also observed that truncation at lower terms leads to similar results. The swelling coefficient is an indirect measure of the interaction between the incoming electrolyte and the polymer. The swelling coefficient of the polymeric coating can be related to the variations of impedance parameters such as capacitance and the exponent (n_p) during the time of exposure to a corrosive media. The diffusion coefficient obtained by non-Fickian diffusion behavior was $D_{AB} = 2.55\times10^{-15}\ m^2/s$ and the swelling coefficient was $SC_C = 3.22\times10^{-12}\ s^{-1}$. It is important to highlight that the diffusion coefficients obtained by both linear and nonlinear methods give similar results. The low diffusivity and swelling coefficients obtained during the longer immersion period of the epoxy mastic coating is consistent with the excellent anticorrosion performance of the coating.

Using Eq. 11 and 12 it was possible to calculate the water solubility (S) and the permeation (P) coefficients of the epoxy mastic coating. The respective values were $S = 65.6\ kg.m^{-3}$ and $P = 1.67\times10^{-10}\ g.m^{-1}.s^{-1}$. The permeation coefficient is relatively low compared to that observed for other polymer materials, which exhibit P values in the range of $10^{-8}\ g.m^{-1}.s^{-1}$. Similarly, the water solubility coefficient of the epoxy mastic coating is also low compared to other polymers [32]. This may be the main reason why the coating has high barrier properties and high anticorrosion performance.

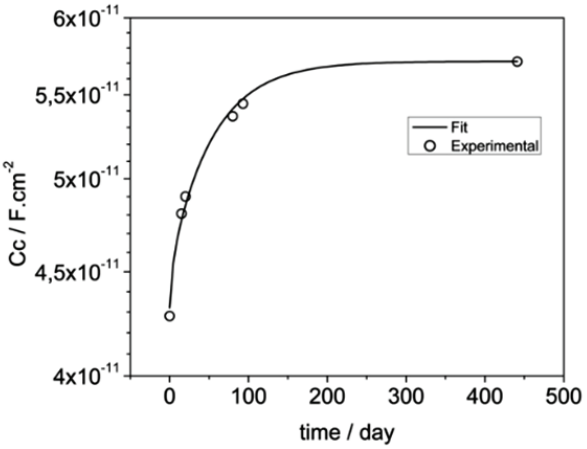


Figure. 9 Coating capacitance and fitting (Eq. 9)

3.5. Evaluation of an intentionally caused defect on delamination

During immersion of the epoxy mastic coating in 0.5M NaCl solution for 441 days, no visual signs of coating defects or corrosion phenomena were observed. However, due to the longer immersion time, some micro failures or micro blisters could have occurred, which if present would indicate the first stage of delamination. In order to verify if the electrochemical information of the metal/coating interface obtained from EIS measurements was correctly related with the anticorrosion coating performance, a small incision on the coating (0.2 mm x 5 mm) was made, as shown in Fig. 10. EIS measurements on coated samples with incisions were carried out during 14 h of immersion in 0.5M NaCl solution.

Fig. 10 shows photographs of coated samples with incision. As can be seen, no coating adhesion loss or increase in delamination area in the incision was observed. Fig. 11 shows Bode impedance diagrams of the coated samples with and without incisions at different immersion times. Unlike what was observed for samples without defects, which exhibited only one time constant, the impedance of the samples with incision exhibited two distinguishable time constants. The time constant at high frequencies can be related to the metal/electrolyte interface. The impedance modulus of this time constant is close to 10 kOhm.cm^2. This value is what is normally found for the charge transfer resistance of the metal/electrolyte interface in corrosion processes.

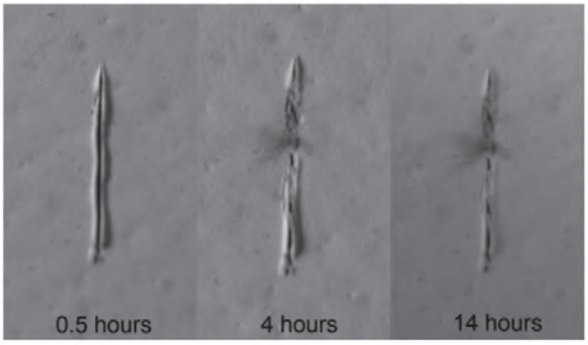


Figure 10. Coating photographs with incision at different times of immersion

The impedance modulus of the second time constant observed at a low frequency exhibits values close to 5 MOhm.cm^2. Clearly, these impedance modulus values cannot be related only with the coating but rather with metal corrosion processes under the coating near to the incision.

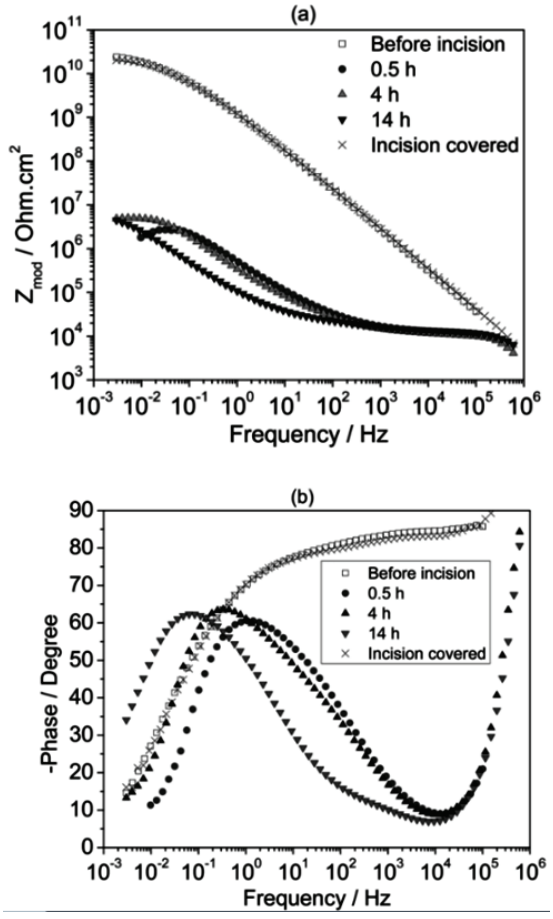


Figure 11. Bode (A) and Phase (B) diagrams for the coating before and after a 5 mm incision

When the coating has a defect, the overall impedance is related with the corrosion of the substrate, with impedance values 4 orders of magnitude lower than those observed in the absence of a defect. This result was confirmed when impedance measurements were taken of coated samples with the incision covered with paraffin. Paraffin completely covered the incision preventing electrolyte penetration in that area. The impedance values of the coating when the defect was covered with paraffin were similar to those observed in the defect-free coating, which indicates that the impedance actually follows the evolution of the coating defects. In addition, these results indicate that the separation of time constants in the impedance diagrams does not necessarily imply loss of adhesion of the coating or the onset of delamination. Therefore, the breakpoint frequency method for calculating the delaminated area from impedance diagrams, in those cases where it is possible to implement it, would be more suitable for calculating the percentage of the defective area than the delaminated area.

5. CONCLUSIONS

The high anticorrosion performance of the epoxy mastic coating is due to its superior barrier property, which is caused by low water solubility and low diffusivity. During the evaluation of the anticorrosion performance of a polymeric coating by EIS, the modulus of the impedance follows most closely the behavior of the charge transfer resistance than the resistance of the coating, which suggests that the behavior of this parameter should preferably be taken into account when looking at the overall anticorrosive coating performance.

The breakpoint frequency method, which is useful when the two time constants are well differentiated and used traditionally for calculating the delaminated area, could be used to estimate the percentage of the defective area rather than the delaminated area in a coated sample.

ACKNOWLEDMENTS

The authors wish to acknowledge the financial contribution of the “Estrategia de Sostenibilidad 2013-2014 de la Universidad de Antioquia”.

REFERENCES

[1] Haruyama, S., Asari, M. and Tsuru, T., Corrosion Protection by Organic Coatings. in Electrochemical Society (ECS) Proceedings. 1987. Pennington: The Electrochemical Society.

[2] Amirudin, A. and Thierry, D., Application of electrochemical impedance spectroscopy to study the degradation of polymer-coated metals. Progress in Organic Coatings, 26(1), pp. 1-28, 1995.

[3] Piratoba, U., Olaya, J.J. and Mariño, A., Impedancia Electroquímica - Interpretación de diagramas típicos con circuitos equivalentes. Dyna, 77(164), pp. 69-75, 2010.

[4] Popov, B.N., Alwohaibi, M.A. and White, R.E., Using Electrochemical Impedance Spectroscopy as a Tool for Organic Coating Solute Saturation Monitoring. Journal of Electrochemical Society, 140(4), pp. 947-951, 1993.

[5] Stern, M. and Geary, A.L., Electrochemical Polarization. Journal of Electrochemical Society, 104, pp. 56, 1957.

[6] Kending, M., Mansfeld, F. and Tsai, S., Determination of the long term corrosión behavior of coated steel with A.C. Impedance measuremends. Corrosion Science, 23, pp. 317-329, 1983.

[7] Scully, J.R., Electrochemical Impedance of Organic-Coated Steel: Correlation of Impedance Parameters with Long’Term Coating Deterioration. Journal Electrochemical Society, 136, pp. 979-990, 1989.

[8] Hirayama, R. and Haruyama, S., Electrochemical impedance for degraded coated steel having pores. Corrosion, 47, pp. 952-958, 1991.

[9] Deflorian, F. et al., Organic coating capacitance measurement by EIS: ideal and actual trends. Electrochimica Acta, 44, pp. 4243–4249, 1999.

[10] Singh, S.K. et al., Electrochemical impedance study of thermally sprayable polyethylene coatings. Corrosion Science, 51(3), pp. 595-601, 2009.

[11] Westing, E.P.M.V., Ferrari, G.M. and Wit, J.H.W.D., The determination of coating performance with impedance measurements - II. water uptake of coatings. Corrosion Science, 36, pp. 957-977, 1994.

[12] Barsoukov, E. and Macdonald, J.R., Impedance Spectroscopy Theory, Experiment, and Applications. 2 ed. 2005: John Wiley and Sons, Inc.

[13] Shreepathi, S. et al., Service life prediction of organic coatings: electrochemical impedance spectroscopy vs actual service life. Journal of Coatings Technology and Research pp. 191-200, 2010.

[14] Miszczyk, A. and Darowicki, K., Multispectral impedance quality testing of coil - coating system using principal component analysis. Progress in Organic Coatings, 69, pp. 330-334, 2010.

[15] Hsu, C.H. and Mansfeld, F., Concerning the Conversion of the Constant Phase Element Parameter Y0 into a Capacitance. Corrosion, 57, pp. 747-748, 2001.

[16] Bierwagen, G. et al., EIS studies of coated metals in accelerated exposure. Progress in Organic Coatings, 46(2), pp. 149-158, 2003.

[17] Parka, J.H. et al., The improvement of anticorrosion properties of zinc-rich organic coating by incorporating surface-modified zinc particle. Progress in Organic Coatings, 74(1), pp. 25-35, 2012.

[18] Hu, J.M., Zhang, J.Q. and Cao, C.N., Determination of water uptake and diffusion of Cl- ion in epoxy primer on aluminum alloys in NaCl solution by electrochemical impedance spectroscopy. Progress in Organic Coatings, 46, pp. 273-279, 2003.

[19] Deflorian, F. and Rossi, S., An EIS study of ion diffusion through organic coatings. Electrochimica Acta, 51, pp. 1736-1744, 2006.

[20] Brasher, D.M. and Kingsbury, A.H., Electrical measurements in the study of immersed paint coatings on metal. I. Comparison between capacitance and gravimetric methods of estimating water-uptake. Journal of Applied Chemistry, 4, pp. 62-72, 1954.

[21] Rammelt, U. and Reinhard, G., Application of electrochemical impedance spectroscopy (EIS) for characterizing the corrosion-protective performance of organic coatings on metals. Progress in Organic Coatings, 21, pp. 205-226, 1992.

[22] Castela, A.S. and Simoes, A.M., An impedance model for the estimation of water absorption in organic coatings. Part II: A complex equation of mixture. Corrosion Science, 45, pp. 1647-1660, 2006.

[23] Westing, E.P.M.V., Ferrari, G.M., and Wit, J.H.W.D., The determination of coating performance with impedance measurements. Electrochemical Acta, 7, pp. 899-910, 1994.

[24] Socrates, G., Infrared and Raman Characteristic Group Frequencies: Tables and Charts, ed. L. John Wiley and Sons. 2001, London.

[25] Goresy, A.E. et al., Natural shock-induced dense polymorph of rutile with α-PbO2 structure in the suevite from the Ries crater in Germany. Earth and Planetary Science Letters, 192, pp. 485-495, 2001.

[26] Vašková, H. and Křesálek, V., Quasi real-time monitoring of epoxy resin crosslinking via Raman microscopy. International Journal of Mathematical Models and Methods in Applied Sciences, 5, pp. 1197-1204, 2011.

[27] Kendig, M.W. and Henry Leidheiser, J., The Electrical Properties of Protective Polymer Coatings as Related to Corrosion of the Substrate. Journal of Electrochemical Society, 123(7), pp. 982-989, 1976.

[28] Pistorius, P.C., Electrochemical testing of "intact" organic coatings. in 14th International Corrosion Congress. 1999. Cape Town.

[29] Gray, L.G.S. and Appleman, B.R., EIS: Electrochemical Impedance Spectroscopy A Tool To Predict Remaining Coating Life? Journal of Protective Coatings & Linings, 30, pp. 66-74, 2003.

[30] Legghe, E. et al., Correlation between water diffusion and adhesion loss: Study of an epoxy primer on steel. Progress in Organic Coatings, 66(3), pp. 276-280, 2009.

[31] Krzak, M. et al., Water diffusion in polymer coatings containing water-trapping particles. Part 2. Experimental verification of the mathematical model. Progress in Organic Coatings, pp. 207-214, 2012.

[32] Thomas, N.L., The barrier properties of paint coatings. Progress in Organic Coatings, 19, pp. 101-121, 1991.

EFFECT OF THE ADDITION OF PHOSPHOTUNGSTIC ACID ON THE THERMAL STABILITY OF ZIRCONIUM OXIDE

EFECTO DE LA ADICIÓN DE ÁCIDO FOSFOTÚNGSTICO EN LA ESTABILIDAD TÉRMICA DEL ÓXIDO DE CIRCONIO

JUAN MANUEL HERNÁNDEZ-ENRÍQUEZ
Ph.D., Instituto Tecnológico de Ciudad Madero, Ciudad Madero Tamaulipas, México, jmanuelher@hotmail.com
RICARDO GARCÍA-ALAMILLA
Ph.D., Instituto Tecnológico de Ciudad Madero, Ciudad Madero Tamaulipas, México, rgalamilla@yahoo.com.mx
GUILLERMO SANDOVAL-ROBLES
Ph.D., Instituto Tecnológico de Ciudad Madero, Ciudad Madero Tamaulipas, México, guillermosandov@gmail.com
JOSÉ AARÓN MELO-BANDA
Ph.D., Instituto Tecnológico de Ciudad Madero, Ciudad Madero Tamaulipas, México, melobanda@yahoo.com.mx
LUZ ARCELIA GARCÍA-SERRANO
Ph.D., Instituto Politécnico Nacional (CIIEMAD), México D.F., México, draluzg@gmail.com

Received for review September 6 th, 2012, accepted November 1th, 2013, final version November, 25 th, 2013

ABSTRACT: Zirconium hydroxide [$Zr(OH)_4$] was synthesized by the sol-gel method using zirconium *n*-butoxide and 1-butanol as synthesis precursors. A part of the synthesized material was impregnated with phosphotungstic acid ($H_3PW_{12}O_{40}$) in order to evaluate the effect of this dopant agent on the thermal, textural and structural properties of zirconium oxide (ZrO_2). Pure and modified hydroxides were calcined in an air flow at 400, 500 and 600°C for 3 h and afterwards were characterized by thermal analysis, infrared spectroscopy, X-ray diffraction and nitrogen physisorption. Results showed that the interaction between the heteropolyanion and zirconium oxide surface produces a stabilizing effect of the crystalline structure and delays the drastic sintering of the material caused by the thermal treatment, obtaining high specific surface areas and small crystallite size for the modified zirconium oxides.

Key words: Zirconium oxide; Phosphotungstic acid; Thermal stability; Specific surface area; Crystallite size

RESUMEN: Se sintetizó por el método sol-gel hidróxido de circonio [$Zr(OH)_4$], utilizando como precursores de síntesis *n*-butóxido de circonio y 1-butanol. Una parte del material sintetizado se impregnó con ácido fosfotúngstico ($H_3PW_{12}O_{40}$) para evaluar el efecto de este agente dopante sobre las propiedades térmicas, texturales y estructurales del óxido de circonio (ZrO_2). Los hidróxidos puro y modificado se calcinaron en atmósfera dinámica de aire a 400, 500 y 600°C durante 3 h y posteriormente fueron caracterizados mediante análisis térmicos, espectroscopía de infrarrojo, difracción de rayos X y fisisorción de nitrógeno. Los resultados mostraron que la interacción del heteropolianión con la superficie del óxido de circonio produce un efecto estabilizador de estructura cristalina y disminuye la drástica sinterización del material causada por efecto del tratamiento térmico, obteniéndose altas áreas específicas y tamaños de cristal pequeño en los óxidos de circonio modificados.

Palabras clave: Óxido de circonio; Ácido fosfotúngstico; Estabilidad térmica; Área específica; Tamaño de cristal

1. INTRODUCTION

Metal oxides are one of the most common and outstanding categories of solid catalysts. Most heterogeneous catalysts are composed of a support and an active phase [1]. Many industrial applications require supports or catalysts that combine high specific surface areas, a suitable porous structure that allows good reactant diffusion throughout the inner particles of the catalyst, as well as high thermal stability [2].

One metal oxide with applications in diverse science and engineering fields is zirconium oxide (ZrO_2), one of the most important heterogeneous catalysts [3]. The great diversity of applications for ZrO_2 is related to its acid-base and oxido-reductive properties [4], which allow it to behave as a catalyst in alkilation, isomerization, esterification, alcohol dehydration, combustion, hydrogenation reactions, among others [5-9]. One of ZrO_2's disadvantages is the sudden decrease of specific surface area during calcination due to a sintering process which occurs slowly when

increasing the thermal treatment temperature, and results in the decrease in pore volume and increase of average diameter [10]. During thermal treatment, ZrO_2 is also subject to complete transformation from a tetragonal to a monoclinic structure when the calcination temperature reaches 600°C [11]. For zirconium oxide based materials it is important stabilize high specific surface areas and tetragonal structure, since these parameters are considered as promoters of the catalytic activity [12].

Since the early developments in the decade of 1970, chemical methods using the sol-gel route under different experimental synthesis conditions allowed suitable control of textural and structural properties in zirconium oxide [13]. During the last few years the application of new preparation methods as well as the combination of extreme critical drying techniques has resulted in high thermal stability, tetragonal structure and high specific surface area for this kind of materials [14-16]. In the same way it has been possible to stabilize the tetragonal structure of ZrO_2 with high specific surface areas by means of adding dopant agents like sulfate (SO_4^{2-}), phosphate (PO_4^{3-}) and borate (BO_3^{3-}) ions [17-19], as well as by the addition of metal cations especially: Mg^{2+}, Ca^{2+}, Sc^{3+}, Y^{3+}, La^{3+} and Ce^{4+} [2]. The amount of dopant agent required to stabilize the textural and structural parameters previously mentioned depends on the nature of the dopant agent and mainly on the preparation method.

In previous works we have reported successful preparation of ZrO_2 with high thermal stability, a defined tetragonal structure and high specific surface areas in the calcination temperature interval between 400 and 600°C, using the sol-gel method and 15% weight of PO_4^{3-} and BO_3^{3-} ions as structure-stabilizing agents. The synthesized materials were active in alcohol dehydration and light paraffin isomerization due to their excellent acid properties [20, 21].

The aim of this work is to study the thermal stability that can be induced in zirconium oxide by impregnation of precursor hydroxide with a heteropolyacid, phosphotungstic acid ($H_3PW_{12}O_{40}$), as well as its effect on specific surface area and crystalline structure.

2. EXPERIMENTAL METHOD

2.1. Material synthesis

The synthesis of zirconium hydroxide [$Zr(OH)_4$] was made by the sol-gel method as previously indicated [20]. The impregnation of $Zr(OH)_4$ with the acid agent was performed by the incipient wetness method using a $H_3PW_{12}O_{40}$/ethanol solution, adding the required volume of this solution to obtain a catalytic support with 15% by weight of the dopant agent. Pure and modified hydroxides were dried at 100°C for 24 h and afterwards they were calcined under dynamic air atmosphere during 3 h at 400, 500 and 600°C. Taking calcination temperature and acid agent as reference, Table 1 lists the nomenclature used throughout this work.

Table 1. Nomenclature of synthesized materials

Material	Calcination temperature (°C)	Acid agent
Z400	400	---
Z500	500	---
Z600	600	---
ZH400	400	$H_3PW_{12}O_{40}$
ZH500	500	$H_3PW_{12}O_{40}$
ZH600	600	$H_3PW_{12}O_{40}$

2.2. Characterization techniques

The characterization of synthesized materials was carried out by thermogravimetry, infrared spectroscopy, X-ray diffraction and nitrogen physisorption.

Thermal analyses were performed in the thermogravimetric instrument TA Instruments STD 2960 Simultaneous DSC-TGA. The samples were analyzed under dry air flow (10 mL/min) at a heating rate of 10ºC/min in the temperature interval from ambient up to 900ºC.

Infrared spectroscopy was carried out in a Fourier Transform Spectrometer (Perkin-Elmer Spectrum One) with transparent wafers containing the sample to be analyzed and KBr as a binder. Spectra were recorded at a resolution of 4 cm^{-1} and by co-adding 16 scans.

The crystalline structure of zirconium oxides was determined in a Bruker Advance D800 diffractometer which used Cu Kα radiation (λ = 1.5406 Å) and graphite monochromator in the secondary beam; the intensities of the diffraction lines were obtained in the range of 0-70° in the 2θ scale, with 0.02° steps and 2.4 s per point.

Nitrogen physisorption was used to determine the specific surface areas of the materials at the temperature of liquid nitrogen (-196°C) in a Quantachrome Autosorb-1 instrument. Prior to the measurements, samples were outgassed at 350°C for 2 h. The specific surface area was calculated using the BET equation and the BJH method was used to calculate the average pore diameter and the pore volume of the samples.

3. RESULTS AND DISSCUSION

3.1. Thermal analysis

By means of thermogravimetry, the physical and chemical processes developed in the synthesized zirconium hydroxides ($Zr(OH)_4$ and $Zr(OH)_4$-$H_3PW_{12}O_{40}$) versus thermal treatment temperature were studied.

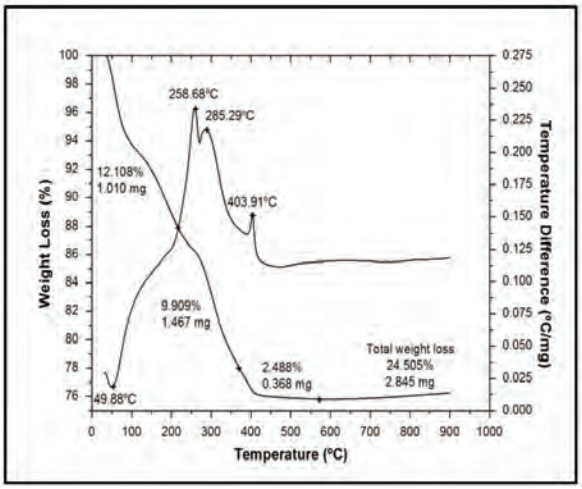


Figure 1. TGA-DTA profiles of $Zr(OH)_4$

In the thermogravimetric profile of Figure 1 that corresponds to $Zr(OH)_4$, multiple-stage decomposition of material can be seen, with three important weight losses. The first one is recorded in the interval of temperature from ambient to 200°C, and is related with the evaporation of water and alcohol that remain from synthesis, occluded up to that moment in the material's porous structure, in the curve of thermodifferential analysis such material releases are shown as endothermal signals near 50°C and 200°C [10]. The second weight loss is detected in the temperature range from 200°C to 370°C and is attributable to the combustion process of residual organic matter, related with the exothermal signals located at 258°C and 285°C in the thermodifferential profile. The last stage of weight loss recorded above 370°C could be related to de-hydroxylation of the material and the possible change of zirconium oxide from the amorphous to crystalline phase [10,22]; the thermodifferential analysis showed an exothermal change at 404°C during this weight loss. The transformation to zirconium hydroxide led to a total weight loss of 24%, which stabilized after 500°C.

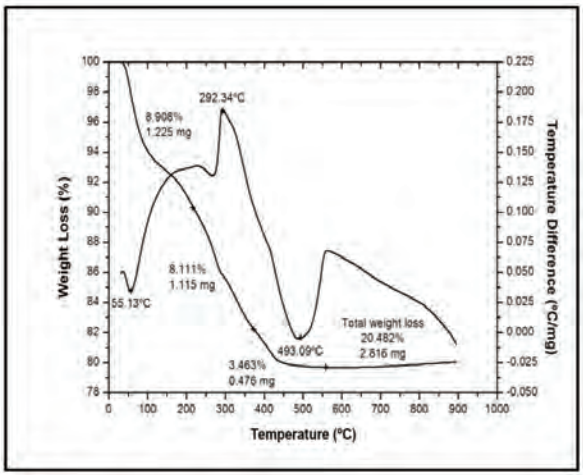


Figure 2. TGA-DTA profiles of $Zr(OH)_4$-$H_3PW_{12}O_{40}$

The thermograms of Figure 2 corresponding to sample $Zr(OH)_4$-$H_3PW_{12}O_{40}$, show a behavior similar to the thermograms of $Zr(OH)_4$, with a slight shift in temperatures at which the physical and chemical processes previously described occur, which indicates the presence of a different specie in the $Zr(OH)_4$. The broad endothermal change centered at 493°C is attributable to a slow transition from the amorphous state of the material to a crystalline state, probably caused by the interaction between the heteropolyanion ($[PW_{12}O_{40}]^{3-}$) and zirconium atoms.

Figure 3 shows the thermogravimetric and thermodifferential analysis for phosphotungstic acid ($H_3PW_{12}O_{40}$). The curve of the thermogravimetric analysis shows that the most important transitions of weight loss for this material are located in the range from ambient temperature to 300°C. The thermodifferential analysis profile shows endothermal signals near 68°C and 190°C that are result of the

elimination of physisorbed and structural water contained in the heteropolyacid. The exothermal peak located at 612°C is characteristic of this material and is related with the decomposition of its Keggin structure in oxides of tungsten and tungsten-phosphorus [23], this change takes place without weight loss.

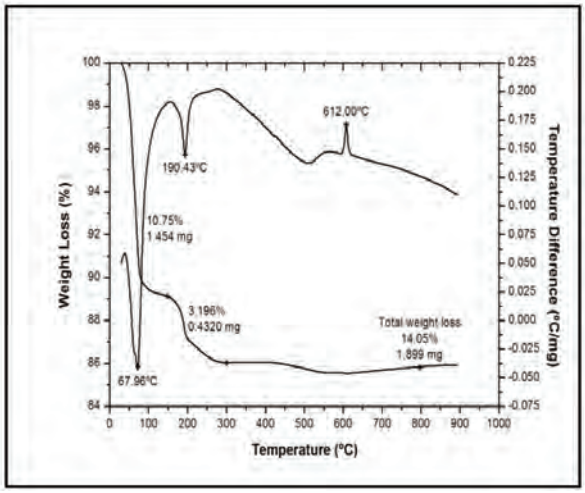


Figure 3. TGA-DTA profiles of $H_3PW_{12}O_{40}$

Thermal analyses of samples Z600 and ZH600 showed a very similar behavior between them. In the thermogravimetric analyses it can be noted that calcined materials have a fast initial weight loss characteristic of a desorption or drying process, this weight loss is not very significant and it was approximately 4% (Figure 4). The endothermal signal for these catalytic solids observed in the thermodifferencial analysis curve near the temperature of 60°C is associated with the evacuation of physisorbed water in the solid surface, after this change, the materials appeared to be thermally stable up to 900°C (Figure 5). Some studies reported that degradation or transformation of heteropolyacid ($H_3PW_{12}O_{40}$) supported in ZrO_2 occurs at temperatures greater than 500°C [24], nevertheless, in this work there were no evidence of changes related to this transformation, this fact could be related to the interaction between the heteropolyanion and the ZrO_2 surface. Thermal analyses performed by Ivanov et al. [23] to $H_3PW_{12}O_{40}$ and the $H_3PW_{12}O_{40}/ZrO_2$ system showed that the decomposition temperature of heteropolyacid is located around 612°C, however, for the $H_3PW_{12}O_{40}/ZrO_2$ system there were no changes attributable to heteropolyacid and with this can be corroborated the high stability of the material, which coincides with results obtained in the present work.

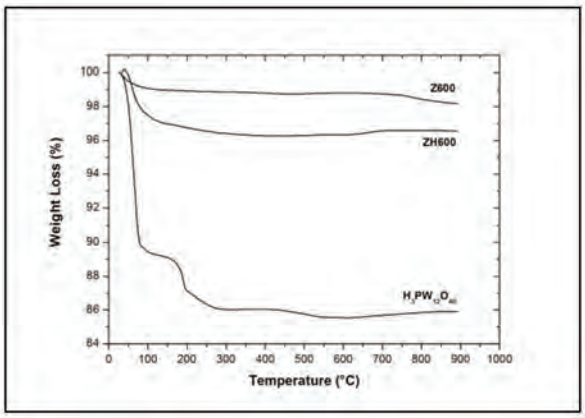


Figure 4. TGA profiles of materials $H_3PW_{12}O_{40}$, ZrO_2 and ZrO_2-$H_3PW_{12}O_{40}$

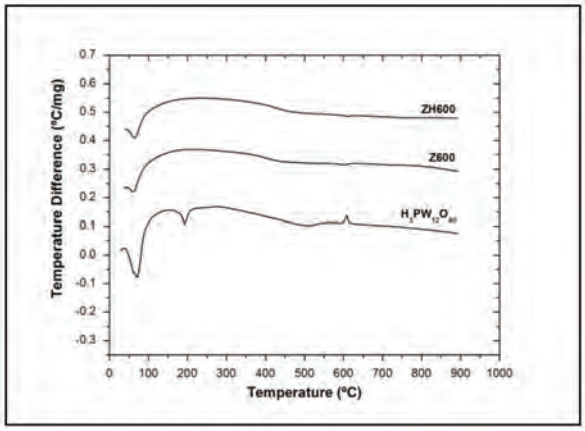


Figure 5. DTA profiles of materials $H_3PW_{12}O_{40}$, ZrO_2 and ZrO_2-$H_3PW_{12}O_{40}$

3.2. Infrared spectroscopy

Table 2 summarizes the most important data of infrared spectroscopy analyses carried out to heteropolyacid and to synthesized zirconium oxides. The infrared spectroscopy study of phosphotungstic acid with a Keggin structure ($H_3PW_{12}O_{40}$) consists of five characteristic bands in the zone 1100-400 cm^{-1}: 1081, 984, 891, 795, and 593 cm^{-1}. The position of these bands can be assigned, in turn, to the stretching vibrations associated to P-O, W=O, W-O-W bonds and to the bending vibration movement of in and out of the plane caused by P-O bonds, respectively [25]. These signals show a slight shift when heteropolyacid is introduced in the matrix of zirconium oxide, suggesting a strong interaction between the heteropolyanion coming from $H_3PW_{12}O_{40}$ and the oxide surface. Although the bands slightly change in position, they do not disappear from this spectrum, only those that coincide with the absorption bandwidths shown

by pure zirconium oxide in the region 750-500 cm^{-1} overlap. These signals associated to Zr-O stretching also slightly modify their position; this is possibly caused by material re-crystallization [26]. In the synthesized oxides also there are signals located at 3500 and 1620 cm^{-1}, related with some hydroxyl groups that exist after the calcination process [27]. The hydroxyl groups contained in $Zr(OH)_4$ at the moment of impregnation are responsible for the fixation and stabilization of the heteropolyanion in zirconium oxide.

3.3. X-ray diffraction

By means of X-ray diffraction the crystallization process of materials thermally treated in the temperature interval between 400°C and 600°C was studied. Figure 6 shows the development of crystalline structures as a function of thermal treatment temperature. In the diffraction patterns of pure zirconium oxide it can be noted that the material Z400 is mainly composed of tetragonal crystalline structure according to pattern card number JCPDS #80-2155. Also low intensity signals are seen at 2θ = 17.55, 24.52, 28.24, 31.53, 38.61, 41.22, 45.59, 54.10, 55.54 and 65.75° that characterize the existence of zirconium oxide with a monoclinic structure if compared to card number JCPDS #7-3430. These signals intensify as the calcination temperature increases, showing an almost completely monoclinic structure in the material Z600, and the full transformation of this structure at temperatures higher than 600°C can be predicted. With these results it can be concluded that the concentration of the crystalline phases in pure zirconium oxide depends mainly on the thermal treatment temperature.

The X-ray diffraction patterns of zirconium oxide modified with phosphotungstic acid are different to those of pure zirconium oxide. In these materials there could be seen only reflections at 30.25, 35.10, 50.41, and 60.02° in the 2θ scale that are characteristic of the formation of zirconium oxide with a tetragonal structure. The material ZH400 shows an X-ray diffraction pattern with incipient signals of very wide bases that are characteristic of materials with poor crystallinity. The increase in calcination temperature generates an intensification of signals shown in diffractograms of materials ZH500 and ZH600, which is related to an increase in crystallite size. The absence of diffraction lines associated to phosphotungstic acid in materials ZH400, ZH500 and ZH600 could be related to the good dispersion ability of the acid agent in zirconium oxide as well as with its low concentration.

The relationship between crystallite size and diffraction line width is given by the Scherrer formula. By means of equation (1) it was possible to estimate the crystallite sizes for the synthesized materials, knowing that λ is the wavelength radiation used (λ = 1.5406 Å), θ is the Bragg angle and β is the line width at half maximum height.

$$D = \frac{0.9\lambda}{\beta \cos\theta} \quad (1)$$

In Figure 6 the crystallite sizes obtained with ZrO_2 and ZrO_2-$H_3PW_{12}O_{40}$ calcined at different temperatures are shown. It was found that before the complete transformation from the tetragonal to the monoclinic structure take place, the crystallite size of pure zirconium oxide was between 4 and 16 nm. The Figure shows a sudden increase in the crystallite size for this material, which depends on the thermal treatment temperature. The crystallite size difference between pure oxide and the modified material with phosphotungstic acid is remarkable. The interaction between heteropolyanion and zirconium atoms limits solid state diffusion and delays crystal growth [28], obtaining values between 2 and 7 nm, making the influence of calcination temperature on this parameter less significant.

Table 2. Infrared absorption frequencies of materials $H_3PW_{12}O_{40}$, ZrO_2, and ZrO_2-$H_3PW_{12}O_{40}$

Material	Assignments and Wavenumbers (cm^{-1})					
	υ (P-O)	υ (W=O_t)	υ (W-O_c-W)	υ (W-O_e-W)	δ (P-O)	υ (Zr-O)
$H_3PW_{12}O_{40}$	1081	984	891	795	593	
Z400						754, 617 508
Z500						754, 615 505
Z600						747, 580 502
ZH400	1081	969	867	822		627, 519
ZH500	1081	969	867	820		628, 519
ZH600	1074	958	863	820		624, 515

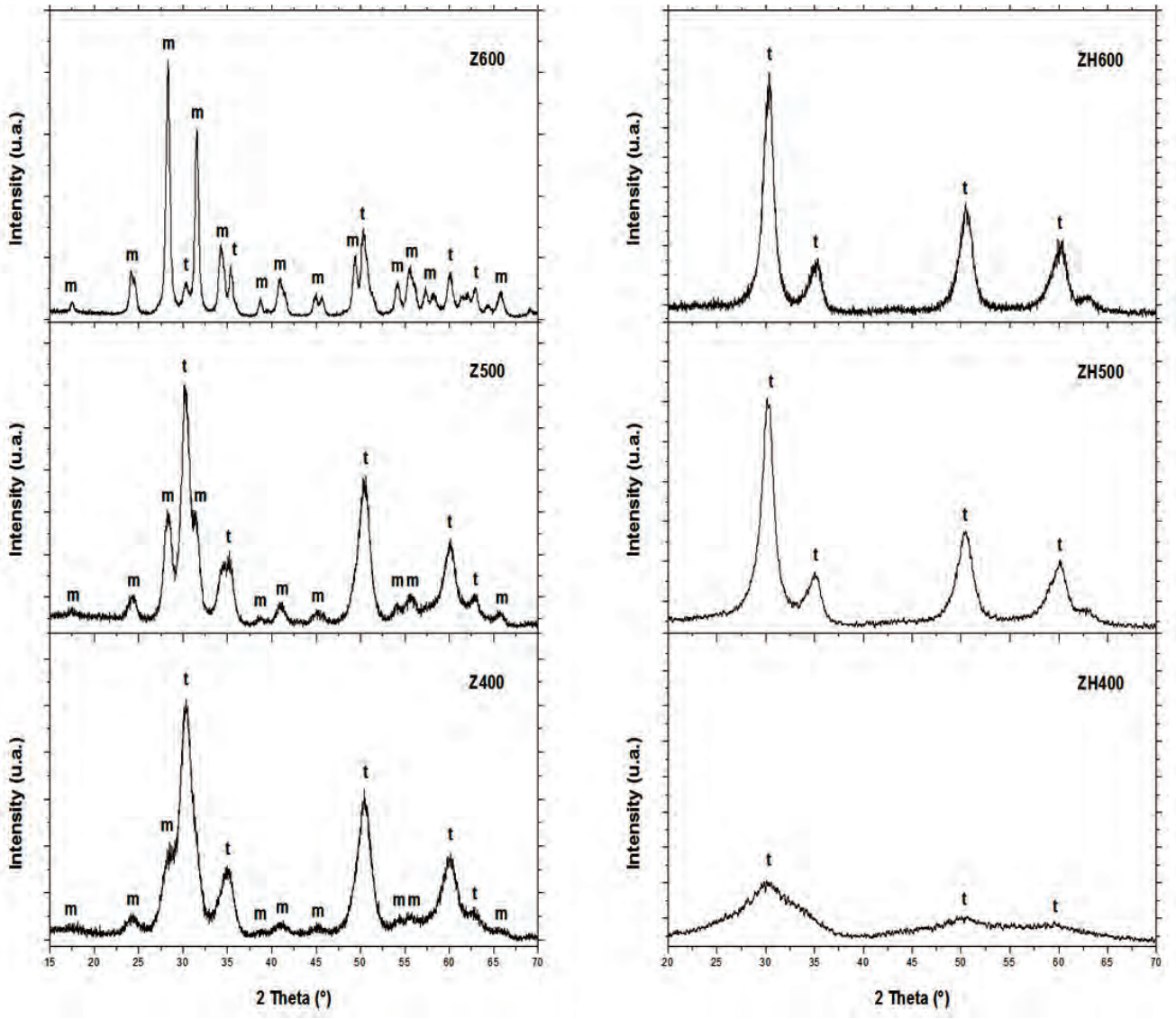


Figure 6. Crystalline structures of materials ZrO_2 and ZrO_2-$H_3PW_{12}O_{40}$: tetragonal (t) and monoclinic (m)

3.4. Nitrogen physisorption

Figure 7 shows specific surface areas of synthesized materials and their evolution according to thermal treatment temperature. Zirconium hydroxide [$Zr(OH)_4$] showed a specific surface area of 595 m^2/g, which decreases to 144 m^2/g when the calcination temperature of the material reaches 400°C. It is worth mentioning that major physical and chemical changes of the samples take place during thermal treatment, in the range between 100 and 400°C, which lead to a sudden decrease in the specific surface area.

The material's sintering process that occurs after 400°C causes a decrease of 88% in the specific surface area. Impregnation of $Zr(OH)_4$ with phosphotungstic acid showed a beneficial effect on specific surface areas obtained in the materials after the calcination process. The presence of a dopant agent reduces the loss of the material's specific surface area in the studied temperature range. Therefore, material modified with phosphotungstic acid and calcined at 600°C (ZH600) surpasses by up to six times the specific surface area value obtained with pure zirconium oxide calcined at the same temperature (Z600).

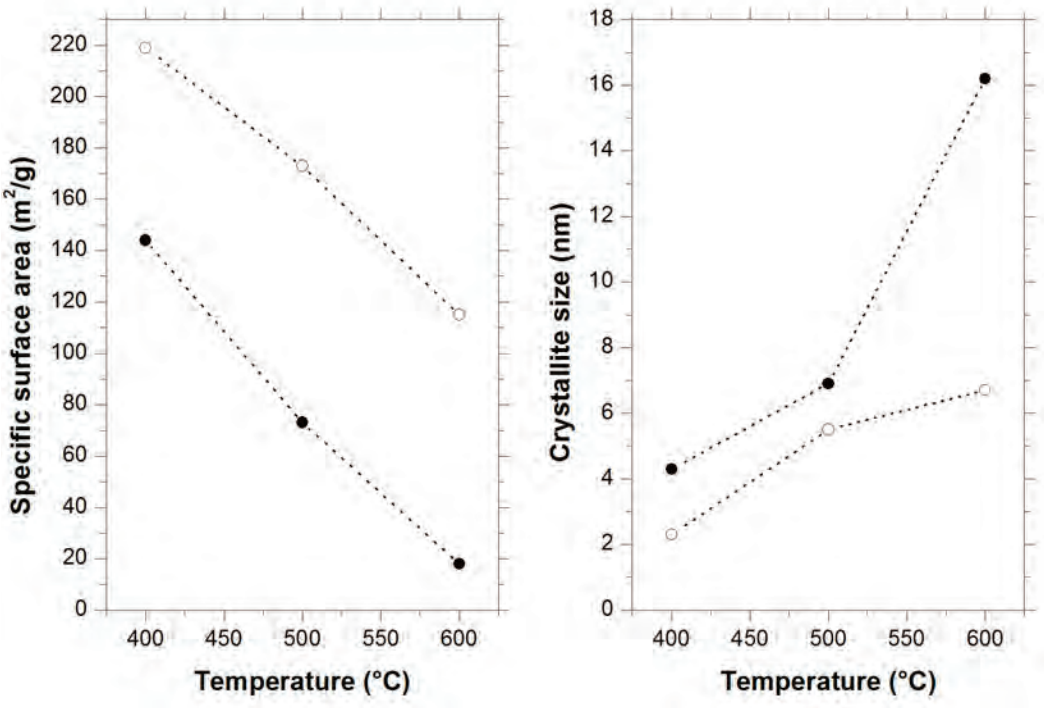


Figure 7. Effect of calcination temperature on specific surface area and crystallite size of materials (●) ZrO_2 and (○) ZrO_2-$H_3PW_{12}O_{40}$

4. CONCLUSIONS

With the addition of phosphotungstic acid in the phase of hydroxide precursor, the crystallization behavior of zirconium oxide was modified. The presence of heteropolyanion in zirconium oxide inhibited monoclinic structure formation, limited diffusion in the solid state, favored stability of the tetragonal structure in the interval of temperatures studied, delayed crystal growth and also encouraged development of nanocrystalline materials with high specific surface area values. The heteropolyanion coming from phosphotungstic acid caused a thermal stability effect similar to the one obtained with SO_4^{2-}, PO_4^{2-} and BO_3^{3-} ions and some metal cations like Y^{3+}, La^{3+} and Ce^{4+}.

ACKNOWLEDGEMENTS

The authors are grateful to División de Estudios de Posgrado e Investigación at the Instituto Tecnológico de Ciudad Madero for the support given during the stage of post-doctoral studies, from which this research work was taken.

REFERENCES

[1] Ready, B. M. and Khan, A., Recent advances on TiO_2-ZrO_2 mixed oxides as catalysts and catalyst supports, Catal. Rev., 47, pp. 257-296, 2005.

[2] Mercera, P. D. L., Van Ommen, J. G., Doesburg, E. B. M., Burggraaf, A. J. and Ross, J. R. H., Stabilized tetragonal zirconium oxide as a support for catalysts. Evolution of the texture and structure on calcination in static air, Appl. Catal., 78, pp. 79-96, 1991.

[3] Restrepo, G., Marín, J. M., Ríos, L., Macías, M. y Navío, J. A., Síntesis, caracterización y evaluación fotocatalítica de sistemas ZrO_2-SiO_2, Dyna, 73 (150), pp. 67-74, 2006.

[4] Chary, K. V. R., Seela, K. K., Naresh, D. and Ramakanth, P., Characterization and reductive amination of cyclohexanol and cyclohexanone over Cu/ZrO_2 catalysts, Catal. Commun., 9, pp. 75-81, 2008.

[5] Senso, N., Jongsomjit, B. and Praserthdam, P., Effect of calcination treatment of zirconia on W/ZrO_2 catalysts for transesterification, Fuel Process. Technol., 92, pp. 1537-1542, 2011.

[6] Lavrenov, A. V., Bogdanets, E. N. and Duplyakin, V. K.,

Solid acid alkylation of isobutane by butenes: the path from the ascertainment of the reasons for fast deactivation to the technological execution of the process, Catalysis in Industry, 1 (1), pp. 50-60, 2009.

[7] Sato, S., Takahashi, R., Yamamoto, N., Kaneko, E. and Inoue, H., Vapor-phase dehydration of 1,5-pentanediol into 4-penten-1-ol, Appl. Catal. A, 334, pp. 84-91, 2008.

[8] Brown, A. S. C. and Hargreaves, J. S. J., Methane combustion in the presence of zirconia-based catalysts, Top. Catal., 52, pp. 458-463, 2009.

[9] Yung, M. M., Zhao, Z., Woods, M. P. and Ozkan, U. S., Preferential oxidation of carbon monoxide on CoO_x/ZrO_2, J. Mol. Catal. A, 279, pp. 1-9, 2008.

[10] Hernández, J. M., García, L. A., García, R., Cortez, L. A. y Cueto, A., Síntesis, caracterización y evaluación catalítica de un ZrO_2 con fase monoclínica, Superficies y Vacío, 22 (2), pp. 1-9, 2009.

[11] Quiñones, J. y Lee, Y., Zirconia en polvo de alta área superficial: influencia de los parámetros de síntesis, Rev. LatinAm. Metal. Mater., 21 (1), pp. 12-17, 2001.

[12] Heshmatpour, F. and Aghakhanpour, R. B., Synthesis and characterization of superfine pure tetragonal nanocrystalline sulfated zirconia powder by a non-alcoxide sol-gel route, Adv. Powder Technol., 23, pp. 80-87, 2012.

[13] Colón, G., Navío, J. A., Macías, M., Hidalgo, M. C. y Sánchez, P. J., Procedimiento para la obtención de polvos nanométricos de óxido de circonio con modificación de su superficie específica y estructura tetragonal estable, Patente Española ES 2 190 300 B1, 2004.

[14] Kazemi, F., Saberi, A., Malek Ahmadi, S., Sohrabi, S., Rezaie, H. R. et al, A novel method for synthesis of metastable tetragonal zirconia nanopowders at low temperatures, Ceramics-Silikáty, 55 (1), pp. 26-30, 2011.

[15] Suh, Y. W., Lee, J. W. and Rhee, H. K., Synthesis of thermally stable tetragonal zirconia with large surface area and its catalytic activity in the skeletal isomerization of 1-butene, Catal. Lett., 90 (1-2), pp. 103-109, 2003.

[16] Cao, Y., Hu, J. Ch., Hong, Z. S., Deng, J. F. and Fan, K. N., Characterization of high-surface-area zirconia aerogel synthesized from combined alcohothermal and supercritical fluid drying techniques, Catal. Lett., 81 (1-2), pp. 107-112, 2002.

[17] Chen, F., Zhu, K., Huang, L., Chen, Y. and Kooli, F., Preparation of rodlike t-ZrO_2 nanoparticles by two-step calcination of 1,12-diaminododecane-zirconia mesostructured composites, Mater. Res. Bull., 41, pp. 10-18, 2006.

[18] Yadav, G. D. and Nair, J. J., Sulfated zirconia and its modified version as promising catalysts for industrial processes, Micropor. Mesopor. Mater., 33, pp. 1-48, 1999.

[19] Mekhemer, G. A. H., Characterization of phosphate zirconia by XRD, Raman and IR spectroscopy, Colloids Surf. A, 141, pp. 227-235, 1998.

[20] Hernández, J. M., Cortez, L. A., García, R., Castillo, A., Sandoval, G. et al, Synthesis and characterization of mesoporous and nano-crystalline phosphate zirconium oxide, J. Alloys Compd., 483, pp. 425-428, 2009.

[21] Cortez, L. A., Hernández, J. M., Castillo, A., Rivera, J. L., Sandoval, G. et al., Isomerización de *n*-pentano mediante óxidos de circonio dopados con ion borato, Rev. Mex. Ing. Quim., 5 (3), pp. 321-327, 2006.

[22] Shibata, K. and Kiyoura, T., Acid property of titania-zirconia, J. Res. Inst. Catalysis, 19 (1), pp. 35-41, 1971.

[23] Ivanov, A. V., Vasina, T. V., Nissenbaum, V. D., Kustov, L. M., Timofeeva, M. N. et al, Isomerization of *n*-hexane on the Pt-promoted Kegging and Dawson tungstophosphoric heteropoly acids supported on zirconia, Appl. Catal. A, 259, pp. 65-72, 2004.

[24] López, E., Hernández, J. G., Schifter, I., Torres, E., Navarrete, J. et al, Thermal stability of 12-tungstophosporic acid supported on zirconia, Appl. Catal. A, 193, pp. 215-225, 2000.

[25] Haber, J., Pamin, K., Matachowski, L. and Mucha, D., Catalytic performance of the dodecatungstophosphoric acid on different supports, Appl. Catal. A, 256, pp. 141-152, 2003.

[26] López, T., Tzompantzi, F., Hernández, J., Gómez, R., Bokhimi, X. et al., Effect of zirconia precursor on the properties of ZrO_2-SiO_2 Sol-Gel oxides, J. Sol-Gel. Sci. Technol., 24, pp. 207-219, 2002.

[27] Skoog, D. A. y Leary, J. J., Análisis Instrumental, Mc-Graw Hill, México, pp. 325-328, 2003.

[28] Sunita, G., Devassy, B. M., Vinu, A., Sawant, D. P., Balasubramanian, V. V. et al., Synthesis of biodiesel over zirconia-supported isopoly tungstate and heteropoly catalysts, Catal. Commun, 9, pp. 696-702, 2008.

LANDMINE DETECTION TECHNOLOGIES TO FACE THE DEMINING PROBLEM IN ANTIOQUIA

TECNOLOGÍAS PARA LA DETECCIÓN DE MINAS FRENTE AL PROBLEMA DE DESMINADO EN ANTIOQUIA

LORENA CARDONA
Ph.D.(c), Universidad Nacional de Colombia, lcardon0@unal.edu.co
JOVANI JIMÉNEZ
Ph.D., Universidad Nacional de Colombia, jajimen1@unal.edu.co
NELSON VANEGAS
Ph.D.(c), Universidad Nacional de Colombia, nvanegas@unal.edu.co

Received for review March 14 th, 2013, accepted August 22th, 2013, final version September, 20 th, 2013

ABSTRACT: This paper presents a review of existing landmine detection techniques. The review is made with an analysis of the strengths and weaknesses of each technique in relation to the landmine detection problem in Antioquia, which ranks first in Colombia by the number of victims from landmines. According to the uniqueness of landmines and terrains in Antioquia, this paper suggests some research topics that may help in the demining task for this affected department.

Key words: ammonium nitrate, APM, demining, explosive detection, IED, improvised explosive devices, landmine, landmine detection.

RESUMEN: En este artículo se hace una revisión de las técnicas existentes para la detección de minas antipersona. La revisión se acompaña de un análisis de las fortalezas y debilidades de cada técnica frente al problema de detección de minas en Antioquia, el departamento con mayor número de víctimas por minas antipersona en Colombia. De acuerdo a las características de las minas y de los terrenos antioqueños, se sugieren áreas de investigación que pueden ayudar a resolver el problema de desminado en este departamento tan afectado.

Palabras clave: AEI, artefactos explosivos improvisados, desminado, detección de explosivos, detección de minas, MAP, minas antipersona, nitrato de amonio.

1. INTRODUCTION

Landmines are weapons, usually buried, that explode when stepped on and are designed to injure or kill, leaving long-term psychological effects and posing a financial burden to the community [1]. Colombia ranks second in victims from landmines [2]. From 1990 to April 2011, Colombia recorded 9.277 victims, and 22.68% of them were in Antioquia [3], the most affected department. Colombia is seriously affected by this problem because landmines are currently used by groups outside the law to protect coca plantations and to counter the army [4].

Colombian mines usually include some elements that complicate the wounds caused by the explosion, as feces, glass and plastic scrap, which cause infections due to fragments that are not detected by x-rays [4]. In addition, these mines are enclosed in casings of various shapes, materials and sizes (e.g. wooden boxes, PVC pipes, glass and plastic bottles), that may contain more explosive: while conventional mines contain from 30 to 520 g, mines in Colombia usually contain from 250 g to 4 kg, and some have been found with more than 20 kg [4]. Due to government controls, the mines of illegal groups rarely contain military explosives, instead they contain ANFO (Ammonium Nitrate/ Fuel Oil) [5] to which some products, like coffee or paint, are added to avoid canine detection. Due to the characteristics of Colombian landmines mentioned above, these are called improvised explosive devices (IEDs) [6].

A crucial step to prevent further victims from landmines and IEDs is to detect them and destroy them. Today, detecting IEDs is a problem in Antioquia due to the characteristics of these devices and of the land where they are buried. Here, a review of landmine detection technologies is made, in order to orient future research in the topic.

2. EXPLOSIVE DETECTION

2.1. Explosive trace detection

These technologies detect traces of explosives on the soil surface or in the air. The explosive's transportation from the mine to the soil surface begins with the

diffusion of the explosive vapor through the mine casing or through its cracks and seals. About 95% is absorbed by the ground, and only 5% migrates towards the surface as vapor (although much is lost in degradation processes [7]). Several factors affect its arrival to the surface: diffusion is inversely proportional to mine casing density, casing paint absorbs chemical signals [7]; explosive degradation increases with temperature, moisture and organic matter; and finally, soil disturbances generated by mine burying can cause dispersion of chemical signatures [8]. Explosive residues in the soil are not uniform and the highest concentrations are not always at the mine location [9]. Additionally, the vapor pressure for most common explosives is very small, requiring highly sensitive sensors [10].

2.1.1. Biological methods

Dogs. Dogs outperform the best chemical sensors and can detect various byproducts from mine explosive degradation, however, they require constant training, and their performance decreases with time. In addition, they have mood and behavior variations difficult to predict [11], [12].

Rats. As an alternative to dogs, the use of giant African rats is being researched [13], [14]. They are trained with food to indicate the presence of explosives by scratching the ground with their feet. Rats have advantages over dogs: they are cheaper, can be deployed in large numbers, are lighter and resistant to tropical diseases [15].

Bees. Bees are trained to associate the odor of an explosive with food, by placing a mixture of sugar and the explosive nearby their hive. The use of bees is limited because they are difficult to track [16], can only work under limited climatic conditions, and their performance with multiple sources is unknown, being it possible that they only detect the strongest ones [15].

Ants. It has been found that ants can locate piles of explosives (TNT and RDX) and carried them to their ant hill. In [17], anecdotal evidence is discussed and in [18], a possible application is proposed for producing smart-landmines capable of self-passivation with time.

Bacteria. There are genetically modified bacteria that emit fluorescence in the presence of TNT [19]. A bacterial detection process involves spraying the bacteria in the minefield, letting them grow for several hours and then returning to search for fluorescence signals. It is possible to cover large areas rapidly with these bacteria [20], however they are highly sensitive to environmental conditions and false alarms of unknown cause have been found. Also, there is no known strain of bacteria capable of detecting other explosives than TNT.

Plants. *Aresa Biodetection* company has genetically modified a thale-cress plant called *Arabidopsis Thaliana* to change color when in contact with nitrogen dioxide emanating from explosives [21]. This technique has associated problems: nitrogen oxides are also formed by denitrifying bacteria, causing false alarms [22], these plants do not grow to be high, making it difficult to watch the results, and there is some concern about native plants contamination [21].

Human role. Humans play an important role in biological landmine detection, especially when dogs or rats are used. They are required for animal training and for interpretation of animal signals. A human and an animal comprise a unique team to perform the detection task together [23].

2.1.2. Chemical methods

Mass and ion mobility spectrometry. In mass spectrometry, a sample of air is introduced into a vacuum chamber where it is ionized. In ion mobility spectroscopy, ions are formed in a reactor. In both techniques ions are accelerated and separated according to their mass/charge ratio which is used for detecting explosives. However, these techniques have false alarms caused by non-explosive substances [24].

Infrared absorption spectroscopy. This technique relies on the fact that molecular vibrations have characteristic frequencies in the infrared spectrum. Resonant absorption of light by these vibrations is observed when the molecule dipole moment changes [24]. However, infrared spectra of large molecules can have broad bandwidths, leading to an indistinct spectrum. Also, many explosives decompose at the high temperatures needed to achieve the vapor pressure required for detection.

Opto-acoustic spectroscopy. This technique relies on the fact that optical energy absorbed by molecules is partially transformed into thermal energy by means of relaxation processes. Using pulsed radiation to excite the sample, pressure pulses are detected by sensitive microphones, yielding a photo-acoustic spectrum [25].

Raman scattering. Raman scattering is an inelastic scattering of light by molecules or atoms. During the interaction of light with the molecule, the energy of the vibrational and rotational states can be exchanged and a lower energy quantum of light is emitted. However, the Raman effect is extremely weak, having insufficient sensitivity for the landmine detection application [26].

Immuno-chemical sensors. Antigen-antibody complexes are formed for proteins containing the molecule of interest. When the substance is bound to an antibody, there is a change in some physical property which serves for detection. Although their selectivity, antibodies are not easily applied in field because they are not reusable [27].

Electronic noses. They are combinations of multiple sensors with pattern recognition algorithms, constituting an artificial olfactory system. They reduce the demand for highly selective sensors and put more emphasis on the algorithms to identify the sample [28]. Sensors used in electronic noses are: fluorescent polymers, optical fiber, polymeric films, gold nanoclusters, piezoelectric materials and micro-electromechanical systems [29].

2.1.3. Explosive trace detection techniques for IEDs detection in Antioquia

Most of the chemical and biological sensing research has focused on military explosives such as TNT, DNT, RDX, tetryl and PETN. Antioquia requires the development of new sensors capable of detecting Ammonium Nitrate (AN). However, AN is a common fertilizer and traces of it may be present around the minefield confusing such sensors. Moreover, vapor concentration of explosives depends on several factors that should be studied. It has been found that paint absorbs explosive vapors, preventing their escape to the surface, and it is usual to find a mixture of paint with the explosive in IEDs of Antioquia. The incidence of this mixture should be investigated, or perhaps exploited for IEDs detection. Also, explosive trace dispersion should be studied, since it is known that it could move within 10 meters from the mine location, and the current knowledge on vapor transportation from mines is insufficient for a reliable location [30].

Among biological systems, dogs and rats are the most suitable for IEDs detection in Antioquia, since bacteria are only able to detect TNT, plants suffer from high rates of false alarms and their effect on native plants has not been studied, and bees would be very difficult to track in the high vegetation lands of Antioquia; its use would depend on the development of new strategies for tracking bees in poor visibility conditions.

Currently dogs are used for IEDs detection in Colombia. Although reliable, dogs are slow and armed groups have begun to place anti-dog mines around IEDs, and also IEDs are being buried at greater depths to prevent explosive vapors from reaching the surface [6]. Alternatively, a research is being carried out on using *Wistar* rats for IEDs detection in Colombia [31]. According to the researchers, *Wistar* rats are better than African rats because African rats can damage the local habitat and they weigh three times more than *Wistar* rats, so they could trigger IEDs.

2.2. Bulk explosive detection

2.2.1. Neutron analysis methods

Thermal neutron analysis (TNA). In TNA neutrons are emitted and when absorbed by a nucleus, gamma rays are emitted with an energy that is specific to the nucleus. However, TNA cannot distinguish nitrogen oxides in the explosives from those in the surroundings, having a high rate of false alarms. In addition, its sensitivity is limited, its price is high and it is bulky and slow for use in field [32].

Fast neutron analysis (FNA). This technique uses high-energy neutrons to detect and differentiate gamma radiation at different energies. FNA technique is sensitive to almost all elements in explosives and allows their identification, but the equipment is usually complex and expensive [33].

Pulsed fast neutron analysis (PFNA). This technique applies the same concept of FNA, using a pulsed beam of neutrons. PFNA allows explosive composition to be

determined, as well as its location and concentration, however, it needs large particle accelerators, which are bulky and expensive [34].

Pulsed fast thermal neutron analysis (PFTNA). PFTNA applies long-lasting beams of neutrons. Its main advantages are high reliability and mobile construction [34]. Its main disadvantage is an insufficient sensitivity to effectively detect the sub-kilogram amounts of explosive in the mines.

2.2.2. Nuclear quadrupole resonance (NQR)

This technique sends radio frequency pulses that excite nitrogen nuclei in the explosive, which induces an electric potential at a receiver coil. The spectrum recorded from each explosive is unique and thus the detection is highly specific and less susceptible to false alarms [34]. Its major drawback is a low signal/noise ratio (SNR) [30]. Also, it is affected by radio frequency interference (RFI), by the distance between the detection coil and the explosive, and by temperature [35].

2.2.3. Bulk explosive detection techniques for IEDs detection in Antioquia

Bulk explosive detection has advantages over trace explosive detection because the last one is affected by factors that are difficult to characterize, cannot locate IEDs accurately and suffers from insufficient sensitivity and false alarms because the pressure and concentration of the explosive vapors are too low.

With neutron analysis methods, there is a concern about the radiation dose that a human operator can receive, and even if it is within safe limits, a perceived risk may hinder its acceptance. Furthermore, it has been found that hydrogenated materials, such as water, can produce false alarms, preventing its use when humidity is above 20% [30]. It has also been found that by using neutron sources of low intensity to get a portable equipment, the detection task takes longer [30]. In addition, there are some concerns about the high cost of both neutron sources and gamma rays detectors. It would be necessary to study the incidence of the factors mentioned above in the detection of IEDs in Antioquia, and to determine infrastructure requirements to support research on neutron techniques.

NQR technique has the advantage of detecting the landmine explosive without emitting harmful radiation and with high chemical specificity. Despite of some drawbacks of this technique, much research has been carried out to deal with them [36], [37], [38]. These research results can be exploited to create a system for IEDs detection in Antioquia. However, in the literature, few works are found about AN detection by NQR [39], [40], [41] and none of them is related to landmine detection. Despite this, for NQR, AN detection is easier than TNT detection, the most used explosive in landmines [41].

3. MINE CASING DETECTION

3.1. Electromagnetic methods

3.1.1. Metal detector (MD)

MD measures the disturbance of an emitted electromagnetic field, caused by the presence of metallic objects in the soil. A major disadvantage of this sensor is that each piece of metal scrap triggers an alarm. The large number of false alarms makes the landmine detection task to be slow and costly [42].

3.1.2. Ground penetrating radar (GPR)

GPR emits radio waves into the earth and analyzes the returning signals produced by reflections at the discontinuities of the dielectric constant. Ultra-Wide Band Synthetic Aperture Radar (UWB SAR) with the ground penetrating capability has become an alternative way to detect landmines over large areas from a safe standoff distance [43]. The main drawback of GPR is that irregularities in the soil (e.g. roots, rocks) can produce false alarms [44]. Additionally, its performance is sensitive to complex interactions between mine casing material, soil texture and moisture, wave frequency and soil surface roughness. Another limitation is the detection of small shallow mines because the reflection at the soil surface masks the response of such mines [30].

3.1.3. X-ray backscattering

This technique exploits the fact that mines and soil have densities and effective atomic numbers that differ by a factor of two [45]. However, it is non-portable because

of the size and weight of the x-ray generators, and of the power requirement to reach enough penetration [46]. Also, it emits radiation, hindering its acceptance.

3.1.4. Electrical impedance tomography (EIT)

EIT uses a two-dimensional array of electrodes placed on the ground to construct a map of conductivity distribution in which mines are seen as anomalies [47]. This technology is especially suited for wet soils, because water enhances conductivity. Also, EIT is cheap enough to be disposable and remotely inserted to improve safety [22]. The main drawback of EIT is that it requires physical contact with the soil. It cannot be used on very dry soils and it is sensitive to electrical noise. Its resolution does not reach that of a GPR [30].

3.1.5. Electromagnetic techniques for IEDs detection in Antioquia

Among the electromagnetic techniques, GPR is likely to be the most convenient for Antioquia. MDs are not suited because most IEDs have little or no metal content at all. X-ray backscattering equipment is not portable and in Antioquian minefields, it may be impossible to drive a bulky transportation platform. EIT requires physical contact with the ground and Antioquian lands are irregular, with high vegetation and IEDs can be activated with an uncertain weight.

GPR is regarded as one of the best subsurface investigation techniques and its potential application in Colombia has already been researched [48], [49]. As mine detection using GPR becomes very complex in the presence of clutter, its outcome depends on the knowledge of prevailing environmental conditions, soil texture and moisture [50]. These factors should be studied in the most affected areas of Antioquia (Tarazá, Guadalupe, Ituango, Toledo, Anorí and Campamento [3]), and tests should be made to validate GPR performance.

3.2. Mechanical methods

3.2.1. Seismic-acoustic techniques

Seismic-acoustic systems introduce sound and/or seismic waves into the earth that are partially reflected when they reach a landmine, causing vibrations at the soil surface. This technique has a low false alarm rate [51], however, sound waves are significantly attenuated by soil, so mines at depths greater than a mine diameter cannot be detected. Additionally, surface vibrations are small and difficult to measure when there is vegetation, turning it into a slow detection system. However the use of sensor arrays has proven to accelerate the detection [52].

3.2.2. Ultrasound

Emitted ultrasound waves are reflected at interfaces between materials with different acoustic properties, which could be used to create images of hidden internal anatomy. Its main advantage is the ability for good penetration into wet lands. However, the sound waves attenuate significantly at the soil-air interface [22].

3.2.3. Prodding

In prodding, a trained operator prods the ground with a stick of about 25 cm long, forming a small angle with the ground surface. Today, prodding the land is used to confirm the presence of a landmine and it is considered the only way to guarantee an exhaustive detection [53]. It is a very slow and dangerous technique because landmines can have any orientation due to soil movement or to nearby explosions and they can be detonated by the probe [46].

3.2.4. Mechanical techniques for IEDs detection in Antioquia

Ultrasound and acoustic/seismic systems have high probability of detection and low false alarm rate, however, they are not portable, seismic waves can detonate IEDs and they cannot detect deep objects. They also have issues with, dense vegetation, such as there is in Antioquia.

Prodding has several disadvantages, but it is still the quintessential confirmatory method for landmine detection. Although probes have been developed which can identify the landmine casing material, they are rarely used in field because, in the hands of a skilled operator, a normal probe may give enough information for landmine detection [46]. Moreover, instead of upgrading the probe for prodding, it would be ideal to have an alternative confirmatory method

without the hazards of prodding. In this sense, NQR can constitute a reliable landmine confirmation system because of its high specificity and because its SNR increases with sensing time.

4. INFRARED AND HYPERSPECTRAL DETECTION

Infrared and hyperspectral methods detect abnormal variations in the electromagnetic radiation reflected or emitted by landmines or by the ground above them.

4.1. Millimeter wave radar (MMWR)

This technique relies on the fact that in the millimeter-wave band, soil has high emissivity and low reflectivity, while metals behave in the opposite way [54]. Active MMWR uses an excitation source while passive MMWR does not, relying only on environmental temperature. MMWR is a good technique for detecting metallic objects, but not for detecting plastic ones.

4.2. Infrared cameras (IR)

This is a thermal detection method that relies in the fact that diurnal temperature variations in areas close to landmines differ from surrounding areas [55]. By using high power radiation (i.e. active IR), these temperature differences can be induced [56], but a landmine could be detonated. IR performance is affected by factors such as weather conditions, landmine size and composition. IR sensors have trouble detecting deep objects and are most suited to detect mined areas than individual mines [22].

4.3. Laser induced breakdown spectroscopy (LIBS)

LIBS uses a high intensity laser to generate a micro-plasma at high temperatures, dissociating the material into atoms and ions. Electron-ion collisions produce a continuous spectrum whose analysis allows element identification. LIBS has been used in identification of different types of anti-personnel and anti-tank landmine casings [57]. LIBS is capable of in-situ point probing and of chemical sensing for landmine detection. It is being directed towards a development of a probe for prodding the soil that emits laser light and collects plasma emissions through an optical fiber [58].

4.4. Visible light

Mine detection by this technique involves the capture of light in the visible wave range using an optical system for imaging [59]. The main limitation of this technique is that it can easily be blocked by camouflage or foliage, so it can only be used on flat lands with little vegetation [22]. It is more suited to detect surface-laid mines.

4.5. Optical light detection and ranging (LIDAR)

LIDAR detects polarization changes in the back-scattered light after illuminating an object with linearly polarized light. Landmines on soil surface can be detected because of their smooth nature compared with the rough background [60]. This method is not capable of subsurface imaging, and does not work well when there is vegetation.

4.5. Infrared/hyperspectral techniques for IEDs detection in Antioquia

Passive IR methods are slow and not suitable for close landmine detection, and active techniques can detonate IEDs and require high power supply, which is difficult to provide in a portable system [46]. MMWR gets weaker signals than infrared ones and it is ineffective in detecting plastic landmines.

There was a research about LIBS in Colombia [61] intended to identify the polymers used for IEDs manufacture and for explosive trace detection, but it was concluded that LIBS has a limited use in Colombia due to the wide variety of IEDs, and to the low sensitivity of the technique at the explosive concentrations. A recent work on AN and fuel oil detection by LIBS [62] showed its ability to identify AN, but not to identify fuel oil when mixed with soil; they also found that the technique had trouble with wet soils and mentioned potential false alarms produced by fertilizers.

LIDAR and visible light techniques do not work well when there is vegetation. Even so, in [63] the images taken by an on-board camera were used, looking for texture and color abnormalities at the regions where the mines were buried. Tests were conducted on soil with gravel and short grass, demonstrating the feasibility of this technique for landmine detection.

5. SENSOR FUSION

A general conclusion of the reviewed papers about landmine detection is that no technology is capable of detecting all types of landmines under all environmental conditions and that there is a need for developing multi-sensor detection systems in order to overcome each sensor limitations [30], [46], [50], [64], [65].

Sensor fusion for landmine detection has been mainly dedicated to the combination of GPR and MD technologies, and in some cases IR sensors [66]. In Colombia, the Universidad de los Andes together with Indumil developed a detection system that combines a GPR with a MD [67]. However, they have not reported any experimental results.

Further research in sensor fusion is needed to include other technologies besides MD, even more so in Antioquia where most IEDs do not contain metallic pieces. A good solution could be to include an explosive detection technology, since it would decrease false alarms from non-explosive objects in the soil.

6. FINAL RECOMENDATIONS

In this paper a review was made about the existent technologies for landmine detection, showing their advantages and disadvantages and analyzing their prospect to become IEDs detection systems for Antioquia. It is concluded that no single technique is capable by itself of solving the IEDs detection problem and that the integration of two or more sensors is required. For Antioquia, the combination of GPR and NQR technologies is suggested. The reason is that GPR is a mature technology, portable equipment is available and there is research that shows its feasibility for IEDs detection in Colombia; and NQR detects the explosive inside IEDs, complementing GPR. However, there is a need for research on AN detection inside IEDs by NQR.

IEDs detection in Antioquia is a complex problem because of the difficulty accessing the areas, because IEDs are non-metallic and because they are currently being placed by armed groups that tailor IEDs to the current demining technologies in order to hinder their detection. But demining of out-of-conflict areas is possible if further research is carried out on adapting the existing technologies to the local problem. However, there is a barrier because the Army and the government do not seem to understand that research costs a substantial amount of money and that results are uncertain and will only appear over the medium to long term. Landmine clearance in Antioquia is subjected to the amount of investment in research on this topic.

In the development of a technology for IED detection in Antioquia it is important to maintain communication with the demining groups to meet their needs. According to [46], many scientists have attacked the problem without an ongoing discussion with users, and as a result, they have obtained technologies that cannot be used in the field.

The way information is presented to the operator is also an important research area. Most existing systems give a sound signal difficult to interpret and to hear in certain conditions. The advanced landmine imaging system (ALIS) [68], which shows images from two sensors (MD and GPR) in real time, is one of the most important contributions in this area.

Knowing the research carried out in Colombia in the context of demining is important to avoid repeated work [46]. Besides the works already mentioned (research using rats to detect explosives, the fusion of a GPR and MD, research on LIBS for the characterization of materials, and mine detection by surface image analysis), there are other two works developed in robotics: a remotely operated robot [69] and a study on snake-type modular robots for use in landmine detection [70]. These works are important for two reasons: first, some researchers might be encouraged to develop detection technologies that are not suitable for hand-held use, but that would be effective onboard a robot; and second because ideally, the mine detection task should be carried out by autonomous or semi-autonomous systems to prevent exposure of people or animals to the hazards of landmine detection.

Finally, current detection technologies are unable to detect mines at great depths and a sensor fusion will not solve this problem. This is a major research focus since IEDs are currently being placed by armed groups seeking smarter ways to make these artifacts difficult to detect.

REFERENCES

[1] Duttine, A. and Hottentot, E., Landmines and explosive remnants of war: a health threat not to be ignored. Bulletin of the World Health Organization, 91, 160–160A, 2013.

[2] Efe, A., Colombia, segundo país con mayor número de víctimas de minas. Periódico El País, Sección Judicial, 2011.

[3] Duque, J.E., Informe de víctimas de minas antipersonal 1990 - Abril 2011. Medellín, Antioquia, Gobernación de Antioquia, 2011.

[4] Parra-Gallego, P.E., IEDs : A major threat for a struggling society. The Journal of ERW and Mine Action, 13, pp. 1–5, 2009.

[5] Pino, Y.A., Determinación de técnicas de detección de explosivos óptimas para el departamento de Antioquia [Undergraduate Thesis]. Medellín: Universidad Nacional de Colombia, 2009.

[6] Hendrickx, J.M.H., Molina, A., Diaz, D., Grasmueck, M., Moreno, H.A. and Hernandez, R.D., Humanitarian IED clearance in Colombia. Proceedings of SPIE, 6953, pp. 69530C–1–69530C–9, 2008.

[7] Phelan, J.M. and Webb, S.W., Mine Detection Dogs: Training, operations and odor detection, In: Odor detection : the theory and the practice, Geneva Centre for Humanitarian Demining, pp. 209–285, 2003.

[8] Gutiérrez, J.P., Padilla, I. and Sánchez, L.D., Transport of explosive chemicals from the landmine burial in granular soils. Revista Facultad de Ingenierías Universidad de Antioquia, 56, pp. 20–31, 2010.

[9] Göth, A., Mclean, I.G. and Trevelyan, J., How do dogs detect landmines?, In: Odor detection : the theory and the practice, Geneva International Centre for Humanitarian Demining, pp. 195–208, 2003.

[10] Östmark, H., Wallin, S. and Ang, H.G., Vapor Pressure of Explosives: A Critical Review. Propellants, Explosives, Pyrotechnics, 37, pp. 12–23, 2012.

[11] Sargisson, R.J., Mclean, I.G., Brown, J. and Bach, H., Environmental determinants of landmine detection by dogs : findings from a large-scale study in Afghanistan. The Journal of ERW and Mine Action, 16, pp. 74–81, 2012.

[12] Williams, A., The anomalies, perceptions and contradictions that exist in the use of dogs in demining. International Symposium on Humanitarian Demining, Šibenik, Croatia, pp. 27–30, 2010.

[13] Poling, A., Weetjens, B.J., Cox, C., Beyene, N.W. and Sully, A., Using giant African pouched rats (cricetomys gambianus) to detect landmines. The Psychological Record, 60, pp.715–728, 2010.

[14] Mahoney, A.M., Mine detection rats : effects of repeated extinction on detection rates [PhD Thesis], Kalamazoo, Michigan, Western Michigan University, 2012.

[15] Habib, M.K., Controlled biological and biomimetic systems for landmine detection. Biosensors & Bioelectronics, 23, pp. 1–18, 2007.

[16] Rains, G.C., Tomberlin, J.K. and Kulasiri, D., Using insect sniffing devices for detection. Trends in Biotechnology, 26, pp. 288–294, 2008.

[17] Mcfee, J.E., Achal, S., Faust, A.A., Puckrin, E., House, A., et al. Detection and dispersal of explosives by ants. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIV, Orlando, Florida, pp. 730302-1–730302-10, 2009.

[18] Achal, S.B., Mcfee, J.E. and Howse, J., Gradual dispersal of explosives by ants and its possible implication for smart-landmine production. International Symposium of Humanitarian Demining, Šibenik, Croatia, pp. 60–65, 2010.

[19] Yagur-Kroll, S., Lalush, C., Rosen, R., Bachar, N., Moskovitz, Y. et al. Escherichia coli bioreporters for the detection of 2,4-dinitrotoluene and 2,4,6-trinitrotoluene. Applied Microbiology and Biotechnology, pp. 1–11, 2013.

[20] Cross, E. and Osborn, S., An emerging remote sensing technology and its potential impact on mine action. International Symposium on Humanitarian Demining, Šibenik, Croatia, p. pp. 66–70, 2010.

[21] Deyholos, M., Faust, A.A, Minmin, M., Montoya, R. and Donahue, D.A., Feasibility of landmine detection using transgenic plants. Proceedings of SPIE, 6217, pp. 62172B-1–62172B-12, 2006.

[22] Kasban, H., Zahran, O., Elaraby, S.M., El-Kordy, M. et al. A comparative study of landmine detection techniques. Sensing and Imaging: an International Journal, 11, pp. 89–112, 2010.

[23] Ivanušić, M., Use of dog-handler teams in humanitarian demining in the republic of Croatia. International Symposium

on Humanitarian Demining, Šibenik, Croatia, pp. 28–34, 2010.

[24] Caygill, J.S., Davis, F. and Higson, S.P.J., Current trends in explosive detection techniques. Talanta, 88, pp. 14–29, 2012.

[25] Chen, X., Guo, D., Choa, F.S., Wang, C.C., Trivedi, S. et al. Standoff photoacoustic detection of explosives using quantum cascade laser and an ultrasensitive microphone. Applied Optics, 52, pp. 26–32, 2013.

[26] Piorek, B.D., Lee, S.J., Moskovits, M. and Meinhart, C.D., Free-surface microfluidics/surface-enhanced Raman spectroscopy for real-time trace vapor detection of explosives. Analytical Chemistry, 884, pp. 9700–9705, 2012.

[27] Mirasoli, M., Buragina, A., Dolci, L.S., Guardigli, M., Simoni, P. et al. Development of a chemiluminescence-based quantitative lateral flow immunoassay for on-field detection of trinitrotoluene. Analytica Chimica Acta, 721, pp. 167–172, 2012.

[28] Stitzel, S.E., Aernecke, M.J. and Walt, D.R. Artificial noses. Annual Review of Biomedical Engineering,13, pp. 1–25, 2011.

[29] Kong, D., Qi, Y., Zhou, L., Lin, B., Li, Z. et al. MEMS based sensors for explosive detection: development and discussion. Proceedings of the 3rd IEEE International Conference on Nano/Micro Engineered and Molecular Systems, Sanya, China, pp. 265–269, 2008.

[30] Macdonald, J., Lockwood, J.R., Mcfee, J., Altshuler, T., Broach, T. et al. Alternatives for landmine detection. Santa Monica, CA: RAND; 2003.

[31] Méndez-Pardo, L.F. and Pérez-Acosta, A.M., Del laboratorio al campo abierto: el uso de protocolos de adaptación y socialización en Rattus norvegicus, Suma Psicológica, 18, pp. 127–129, 2011.

[32] Mcfee, J.E., Faust, A.A., Andrews, H.R., Clifford, E.T.H. and Mosquera, C.M., Performance of an improved thermal neutron activation detector for buried bulk explosives. Nuclear Instruments and Methods in Physics Research A, 712, pp. 93–101, 2013.

[33] Sudac, D., Majetic, S., Kollar, R., Nad, K., Obhodas, J. et al. Inspecting minefields and residual explosives by fast neutron activation method. IEEE Transactions on Nuclear Science, 59, pp. 1421–1425, 2012.

[34] Bielecki, Z., Janucki, J., Kawalec, A., Mikolajczyk, J., Palka, N. et al. Sensors and systems for the detection of explosive devices - an overview. Metrology and Measurement Systems, XIX, pp. 3–28, 2012.

[35] Mikhaltsevitch, V.T., Techniques used for 14N NQR studies. Annual Reports on NMR Spectroscopy, 66, pp. 149–194, 2009.

[36] Gudmundson, E., Jakobsson, A. and Stoica, P., NQR-based explosives detection - an overview. International Symposium on Signals, Circuits and Systems (ISSCS), Iasi, Rumania, pp. 1–4, 2009.

[37] Somasundaram, S.D., Jakobsson, A. and Butt, N.R., Countering radio frequency interference in single-sensor quadrupole resonance. IEEE Geoscience and Remote Sensing Letters, 6, pp. 62–66, 2009.

[38] Jakobsson, A., Mossberg, M., Rowe, M.D. and Smith, J.A-S., Exploiting temperature dependency in the detection of NQR signals. IEEE Transactions on Signal Processing, 54, pp. 1610–1616, 2006.

[39] Prescott, D.W., Nuclear quadrupole spin dynamics : how weak RF pulses and double resonance cross-relaxation contribute to explosives detection [PhD Thesis], Fairfax, VA, George Mason University, 2010.

[40] Rudakov, T.N., Some aspects of the effective detection of ammonium nitrate-based explosives by pulsed NQR method. Applied Magnetic Resonance, 43, pp. 557–566, 2012.

[41] Mozzhukhin, G.V., Molchanov, S.V., Kupriyanova, G.S., Bodnya, A.V., Fedotov, V.V. et al. The detection of industrial explosives by the quadrupole resonance method: some aspects of the detection of ammonium nitrate and trinitrotoluene. NATO Science for Peace and Security Series B: Physics and Biophysics, pp. 231–244, 2009.

[42] Dula, J, Zare, A., H.O, D. and Gader, P., Landmine classification using possibilistic K-nearest neighbors with wideband electromagnetic induction data. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII, Baltimore, Maryland, pp. 87091F-1–87091F-12, 2013.

[43] Lou, J., Jin, T. and Zhou, Z., Feature extraction for landmine detection in UWB SAR via SWD and Isomap. Progress In Electromagnetics Research, 138, pp. 157–171, 2013.

[44] Sengodan, A. and Cockshott, W.P., The SIMCA algorithm for processing ground penetrating radar data and its use in landmine detection. 11th International Conference on Information Science, Signal Processing and their Applications, Montreal, Canada, pp. 983–988, 2012.

[45] Heuvel, J. and Fiore, F., Simulation study of x-ray backscatter imaging of pressure-plate improvised explosive devices. In: Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVII, Baltimore, Maryland, pp. 835716-1–835716-15, 2012.

[46] King, C., Blagden, P., Rhodes, G., Maresca, L., Wheatley, A. et al., Mine action : lessons and challenges. Geneva, Switerzalnd, 2005.

[47] Church, P., Mcfee, J.E., Gagnon, S. and Wort, P., Electrical impedance tomographic imaging of buried landmines. IEEE Transactions on Geoscience and Remote Sensing, 44, pp. 2407–2420, 2006.

[48] Lopera, O. and Milisavljevic, N., Prediction of the effects of soil and target properties on the antipersonnel landmine detection performance of ground-penetrating radar: A Colombian case study. Journal of Applied Geophysics, 63, pp. 13–23, 2007.

[49] Lopera, O., An integrated detection and identification methodology applied to ground-penetrating radar data for humanitarian demining applications [PhD Thesis]. Bogotá, Universidad de los Andes, 2008.

[50] Robledo, L., Carrasco, M. and Mery, D., A survey of land mine detection technology. International Journal of Remote Sensing, 30, pp. 2399–2410, 2009.

[51] Chi, W., Ying-Jie, Y. and Xing-Fei, L., An acoustic-to-seismic coupling based landmines detection system in lab-scale experimental environment. Journal of Tianjin University, pp. 160–166, 2011.

[52] Rajesh, K., Murali, R. and Mohanachandran, R., Realisation of ultrasonic Doppler vibrometer array for landmine detection. IEEE International Ultrasonics Symposium, Dresden, Saxony, pp. 1027–1030, 2012.

[53] Ishikawa, J. and Iino, A., A study on prodding detection of antipersonnel landmine using active sensing prodder. International Symposium on Humanitarian Demining, Šibenik, Croatia, pp. 15–23, 2010.

[54] Öztürk, H., Nazli, H., Yeğin, K., Biçak, E. and Sezgin, M. et al., Millimeter-wave detection of landmines. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII, Baltimore, Maryland, pp. 870913-1–870913-6, 2013.

[55] Thanh, N.T., Hao, D.N. and Sahli, H., Thermal infrared technique for landmine detection : Mathematical formulation and methods. Acta Mathematica Vietnamica, 36, pp. 469–504, 2011.

[56] Chakravarthy, P.K, Jogarao, C.S. and Varun, R.S.V.K., Advanced infrared technique for landmine detection. International Journal of Engineering and Technology, 2, pp. 906–912, 2012.

[57] Gottfried, J.L., Harmon, R.S. and Lapointe, A., Progress in LIBS for land mine detection. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIV, Orlando, Florida, pp. 73031F-1–73031F-24, 2009.

[58] Hauck, J.P-, Walker, M, Hamadani, S., Bloomhardt, N. and Eagan, J., Laser-induced breakdown spectroscopy based deminers' probe. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIV, Orlando, Florida, pp. 73031G-1–73031G-9, 2009.

[59] Staszewski, J.J., Hibbitts, C.H., Davis, L. and Bursley, J., Optical detection of buried explosive hazards: a longitudinal comparison of three types of imagery. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII, Baltimore, Maryland, pp. 870915-1–870915-9, 2013.

[60] Zhang, X., Jiang, Y. and Zhao, Y., Targets detection and discrimination using laser polarimetric imaging. Frontiers of Optoelectronics in China, 2, pp. 419–424, 2009.

[61] Díaz, D., Hahn, D.W. and Molina, A., Identificación de polímeros mediante espectroscopia de emisión de plasmas producidos por láser (LIBS), Dyna, 76, pp. 217–228, 2009.

[62] Díaz, D., Hahn, D.W. and Molina, A., Laser-induced breakdown spectroscopy (LIBS) for detection of ammonium nitrate in soils. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIV, Orlando, Florida, pp. 73031E-1–73031E-9, 2009.

[63] Coronado-Vergara, J., Aviña-Cervantes, G., Devy, M. and Parra, C., Towards landmine detection using artificial vision. IEEE/RSJ International Conference on Intelligent Robots and Systems, Alberta, 2005.

[64] Habib, M.K., Humanitarian demining : reality and the challenge of technology - the state of the arts. Advanced Robotic Systems, 4, pp. 151–172, 2007.

[65] Habib, M.K., Humanitarian demining mine detection and sensors. IEEE International Symposium on Industrial Electronics (ISIE), Gdańsk, Poland, pp. 2237–2242, 2011.

[66] Kolba, M.P., Torrione, P.A. and Collins, L.M., Fusion of ground-penetrating radar and electromagnetic induction sensors for landmine detection and discrimination. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XV, Orlando, Florida, pp. 76641S-1–76641S-7, 2010.

[67] MINISTERIO DE DEFENSA. INDUMIL y Universidad de los Andes desarrollan sistema detector de minas. Ejército Nacional : Informes y Noticias 2012:1.

[68] Sato, M. and Takahashi, K., Deployment of dual-sensor ALIS for humanitarian demining in Cambodia. Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII, Baltimore, 87090M, 2013.

[69] Rizo, J., Coronado, J., Forero, A., Otálora, C., Devy, M. and Parra, C., URSULA: robotic demining system. The 11th International Conference on Advanced Robotics, Coimbra, Portugal, 2003.

[70] Melo, K, Paez, L., Hernandez, M., Velasco, A., Calderon, F. and Parra, C., Preliminary studies on modular snake robots applied on de-mining tasks. In: EEE IX Latin American Robotics Symposium, Bogotá, Colombia, pp. 1 – 6, 2011.

ELECTRONIC, STRUCTURAL AND FERROELECTRIC PROPERTIES OF THE Ba_2ZrTiO_6 DOUBLE PEROVSKITE

PROPIEDADES ELECTRÓNICAS, ESTRUCTURALES Y FERROELÉCTRICAS DE LA DOBLE PEROVSKITA Ba_2ZrTiO_6

DAVID A. LANDÍNEZ-TÉLLEZ

PhD., Grupo de Física de Nuevos Materiales, Dep de Física, Universidad Nacional de Colombia, dalandinezt@unal.edu.co

CRÍSPULO ENRIQUE DELUQUE TORO

Ms.C., Grupo de Nuevos Materiales, Universidad Popular del Cesar, Valledupar, deluquetoro@gmail.com

JAIRO ROA-ROJAS

PhD., Grupo de Física de Nuevos Materiales, Dep de Física, Universidad Nacional de Colombia, Bogotá, jroar@unal.edu.co

Received for review November 08 th, 2012, accepted October 18th, 2013, final version December, 08 th, 2013

ABSTRACT: We report the synthesis, the structural characterization, the ferroelectric behavior and the electronic properties of complex perovskite Ba_2ZrTiO_6. Samples of Ba_2ZrTiO_6 were synthesized through the standard solid state reaction method. The crystalline structure was studied by means of X-ray diffraction experiments and Rietveld-like analysis. Results reveal that the material crystallizes in a rhomboidal structure, space group *R-3* (#148), with cell parameter *a*=5.8038(7) Å. The ferroelectric response of material was established from curves of polarization as a function of applied electric field. Our results reveal that the double perovskite Ba_2ZrTiO_6 has a ferroelectric hysteretic behavior at room temperature. The studies of the electronic structure show that Ba_2ZrTiO_6 behaves as a nonmetallic material with gap energy *2.32 eV*. The structural parameters obtained from energy minimization, through the Murnaghan equation state are 99.5% in agreement with the experimental data.

Key words: Double perovskite, ferroelectric properties, electronic structure

RESUMEN: En este trabajo reportamos la síntesis, la caracterización estructural, el comportamiento ferroeléctrico y las propiedades electrónicas de la perovskita compleja Ba_2ZrTiO_6. Las muestras fueron sintetizadas mediante el método estándar de reacción de estado sólido. La estructura cristalina se estudió a través de experimentos de difracción de rayos X y análisis de tipo Rietveld. Los resultados revelan que el material cristaliza en una estructura romboédrica, perteneciente al grupo espacial *R-3* (#148), con un parámetro de red *a*= 5,8038(7) Å. La respuesta ferroeléctrica del material se estableció a partir de curvas de polarización en función del campo eléctrico aplicado. Nuestros resultados muestran que la perovskita Ba_2ZrTiO_6 evidencia un comportamiento histerético ferroeléctrico a temperatura ambiente. Los estudios de la estructura electrónica muestran que esta cerámica se comporta como un material no metálico con brecha de energía *2,32 eV*. Los parámetros estructurales obtenidos a partir de la minimización de energía, a través de la ecuación de estado Murnaghan, son *99.5%* acordes con los datos experimentales.

Palabras clave: Perovskita compleja, propiedades ferroeléctricas, estructura electrónica

1. INTRODUCTION

Complex perovskites with formula $A_2BB'X_6$, where A represents an alkaline earth, B and B' are metal transition and X is the oxygen, have been studied because of their physical properties [1], which depend on structural distortions and the characteristic properties of B and B' cations. This chemical configuration results in multiple opportunities to combine different elements of the periodic table, generating the possibility of synthesizing new materials that involve a larger gamma of physical properties. One well-known simple perovskite, which is well known because it exhibits a ferroelectric character, is $BaTiO_3$ [2]. Partial substitutions of Ti by the Zr cation have been done by other authors, showing the evolution of the electric behavior from a ferroelectric response, for low concentration of Zr, to a ferroelectric relaxor [3] and a dielectric behavior for high concentrations of Zr [4]. However, the position of the substitutive cations in the crystallographic cell of the material is not usually reported. Location of all ions in the lattice is important for inferring the possibility of obtaining spontaneous polarization, in order to give rise to ferroelectric

behavior. In the case of polycrystalline samples, the ferroelectric behavior depends greatly on the grain size, because the effects of imperfection frequently dominate the ferroelectric response of small grains, where a significant fraction of the material volume may be influenced by grain boundaries. It is known that cationic disorder produces substantial changes of the ferroelectricity of complex perovskites [5]. The aim of this paper is to show the synthesis process for producing the double perovskite Ba_2ZrTiO_6; to analyze the crystallographic single phase; to study the electronic structure; and, the most important objective of the study being to investigate the ferroelectric response of the material at room temperature. Our results reveal that Ba_2ZrTiO_6 crystallizes in a rhomboidal double perovskite structure, which is a convenient lattice for determining the occurrence the ferroelectric character, observed by the occurrence of hysteretic curves of electric polarization as a function of variable applied fields.

2. EXPERIMENTAL PROCESS

Samples were synthesized by means of the solid state reaction recipe. The precursor powders $BaCO_3$ (Aldrich 99.9%), TiO_2 (Aldrich 99.99%) and ZrO_2 (Aldrich 99.99%) were stoichiometrically mixed according to the chemical formula Ba_2ZrTiO_6. The mixture was ground to form pellets of 9.55 mm diameter and 1.85 mm thickness. Then the material was annealed in a sequential thermal procedure at 850, 900, 950, 1000, 1100 and 1150 ºC for 144 hours, including five regrinding and pelletizing procedures to make six thermal steps of 24 hours each. An X-Ray diffraction (XRD) experiment was performed by means of a PW1710 diffractometer with $\lambda_{CuK\alpha}$=1.54064 Å. Rietveld refinement of the diffraction pattern was made by the GSAS code [6]. Electric polarization curves were obtained by means of a Radiant Ferroelectric Tester, which included a ±10 kV source for measurements in bulk-samples.

3. CALCULATION METHOD

We applied the Full-Potential Linear Augmented Plane Wave method (FP-LAPW) within the framework of the Kohn-Sham Density Functional Theory (DFT) [1], and adopted the Generalized Gradient (GGA) approximation for the exchange-correlation energy due to Perdew, Burke and Ernzerhof [7]. The self-consistent process was developed using the numeric package Wien2k [8]. Taking the experimental unit cell data as input, the structures studied in this work were fully relaxed with respect to their lattice parameters and the internal degrees of freedom compatible with the space group symmetry of the crystal structure. The resulting energies versus volume functions have been fitted to the equation of state due to Murnaghan [9] in order to obtain the minimum energy value, the bulk modulus, its pressure derivative and the equilibrium lattice parameters and associated volume. The muffin-tin radiuses used were 2.40, 1.74, 2.14 and 1.53 for Ba, Ti, Zr and O respectively, angular momentum up to l = 10 inside the muffin-tin sphere, a maximum vector in the reciprocal space of Gmax =12.0, $RMT*K_{max}$ = 7.0, and a mesh of 800 points in the first Brillouin zone (equivalent to a maximum of 85 k points in the irreducible Brillouin zone). Finally, the convergence criteria for the self-consistent calculation was 0.0001 Ry for the total energies, 0.0005 u.a. in the charge and 1.0 mRy/u.a. in the internal forces.

4. RESULTS AND DISCUSSION

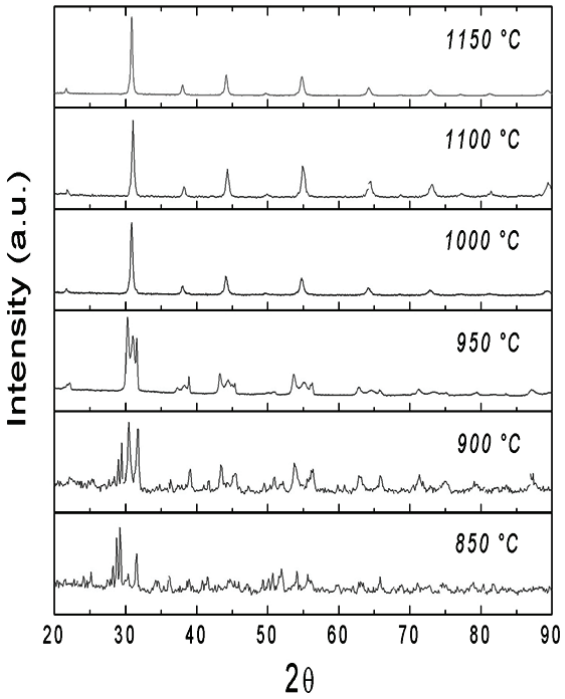


Figure 1. Diffraction patterns showing results of thermal treatment for obtaining rhomboidal structure for the double perovskite Ba_2TiZrO_6.

The sequence of crystallization obtained by the thermal process, described in the experimental setup, for Ba_2TiZrO_6 samples is shown in the diffraction patterns of Figure 1. As shown in Figure 1, the fourth thermal step, at 1000 *ºC*, evidences the perovskite phase of Ba_2TiZrO_6. However, some incipient peaks corresponding to impurities are observed close to 24º. These impurities vanish after the last thermal treatment at 1150*ºC*. A rigorous Rietveld refinement from experimental data of the diffraction pattern reveals that crystallization of material occurs in a rhombohedral double perovskite, which belongs to the space group *R-3* (#148).

The difference between simulated and experimental patterns is shown in Figure 2. The lattice parameters obtained from the refinement process were *a*=5.803(8) Å; primitive vectors angle α=59.9675º. This result is ~ 98.5% in agreement with the value supplied by the Structure Prediction Diagnostic Software [10], which gives *a*=5.7177 Å.

In Figure 2, the top plot represents experimental data, the middle plot is the simulated refined diffractogram, and the line at the bottom corresponds to the difference between experimental and simulated patterns. Indexes of crystallographic planes are indicated in figure 2. The parameters of refinement are: R_F^2=4.69 % and x^2=2.26.

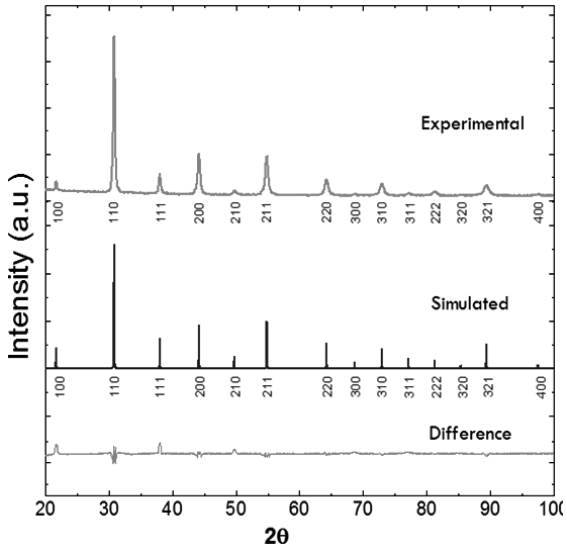


Figure 2. Experimental and simulated diffraction patterns of Ba_2TiZrO_6 material after Rietveld refinement. The bottom line corresponds to the difference between experimental and theoretical diffractograms.

Atomic positions and occupancies of ions in the structure are shown in table 1.

Table 1: Atomic position for the rhomboidal structure of Ba_2TiZrO_6, obtained by Rietveld refinement of diffraction patterns through the GSAS code.

Atom	*x*	*y*	*z*
Ba	*0.7558*	*0.7558*	*0.7558*
Ti	*0.0000*	*0.0000*	*0.0000*
Zr	*0.5000*	*0.5000*	*0.5000*
O	*0.7745*	*0.7745*	*0.2255*
Ba	*0.2442*	*0.2442*	*0.2442*
O	*0.2255*	*0.7745*	*0.7745*
O	*0.7745*	*0.2255*	*0.7745*
O	*0.2255*	*0.2255*	*0.7745*
O	*0.7745*	*0.2255*	*0.2255*
O	*0.2255*	*0.7745*	*0.2255*

One important parameter in perovskite materials is the tolerance factor, which is related to the probability of the formation of octahedral coordination of B and B‘ cations with the oxygen anion. The tolerance factor calculated for Ba_2TiZrO_6 material was 1.0312. This value, above 1.0, suggests the possibility of obtaining spontaneous polarizations from structural distortions that give rise to the ferroelectric character in the simple perovskite $BaTiO_3$ [11], for example.

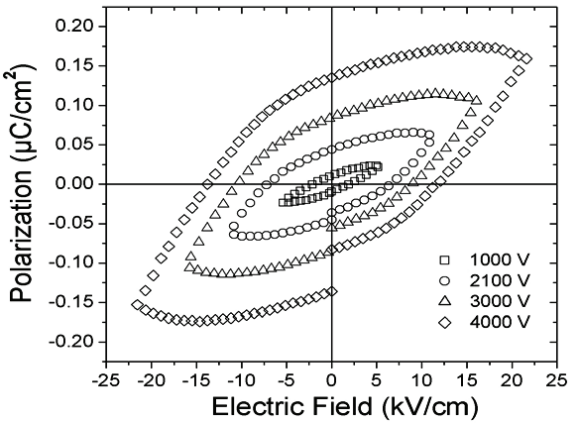


Figure 3. Ferroelectric hysteresis measured at ambient temperature for the perovskite Ba_2TiZrO_6, for several applied fields up to $2.16x10^4$ V/cm.

Another important characteristic for the possible application of the material in ferroelectric devices is the

response to high applied fields. In Figure 3 sequence of polarization for applied electric fields up to 22.5 kV/cm is shown. As observed in Figure 3, for the higher applied field we determine a saturation polarization of 0.175 μC/cm^2, in a curve with a remnant polarization of 0.3125 μC/cm^2 and a coercive field of 12.0 kV/cm. These values are in agreement with strong ferroelectric materials of the $BaTiO_3$ family [12,13] that evidence polarization saturation, remnant polarization and coercive fields, which are appropriate for technological applications in ferroelectric memories.

Figure 4 shows the optimization of energy as a function of volume. The minimum energy value is obtained for *-42368.114575 Ry*. The equilibrium volume is *945.7608 Bohr*^{3}, which corresponds with a lattice constant *a=11.0206 Bohr [a=5.831(8)]* with a volume modulus *Bo=155.9553 GPa*. The lattice parameter obtained from the GGA approximation is *99.5%* in agreement with our experimental results.

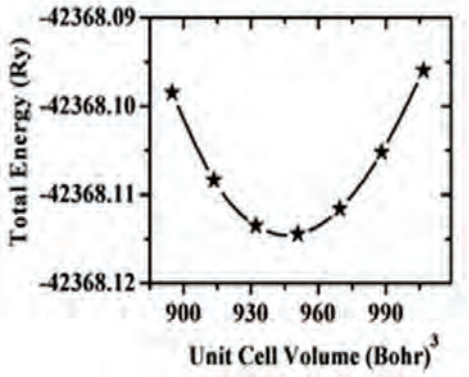


Figure 4. Calculated total energy as a function of volume and *c/a* factor for the Ba_2TiZrO_6 material

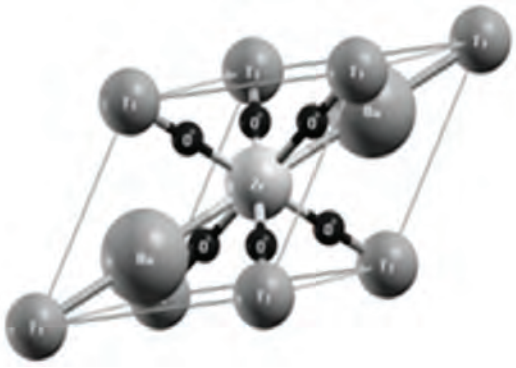


Figure 5. Rhombohedral structure of the Ba_2TiZrO_6 perovskite for the R-3 space group.

The difference between the two minimums is less than the convergence parameter. The minimums for the energy as a function of volume were obtained by adjusting Murnaghan's state equation to the square points [9]. The rhombohedral structure obtained by both procedures, Rietveld refinement of diffraction patterns and theoretical minimization of energy by the DFT method, is shown in figure 5.

Figure 6 corresponds to electronic properties of the perovskite Ba_2TiZrO_6. The Density of States (DOS) along the high symmetry directions in the first Brillouin zone and the band structure are shown. The energy of the electrons as a function of the wave vector *k* is also observed, which was taken along the Γ-Σ-M-K-Λ-Γ directions.

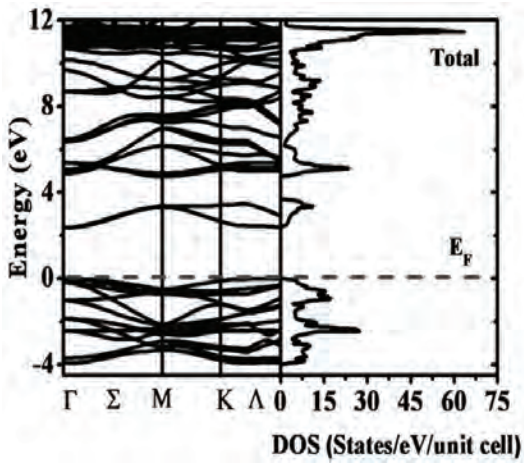


Figure 6. Band structure and total DOS for the Ba_2TiZrO_6 material. The Fermi level is the energy reference.

The DOS and band structure are calculated for the equilibrium configuration by using the lattice constant corresponding to the volume which minimizes the energy of the system.

The energy reference is selected to be the Fermi level in E=0. It is observed that the Ba_2TiZrO_6 material presents a nonmetallic behavior with a direct energy gap of 2,32 eV at the Fermi level. Furthermore, a valence sub-band sited in the regime between -4 eV and the Fermi level is observed in Figure 6, due to principal contributions of the *2s*-O, *3d*-Ti and *4d*-Zr orbital, as corroborated in Figure 7. A minority contribution of the *4s*-Ti and *5s*-Zr electrons is also observed. In the intermediate region

of the conduction band, between 2,32 eV and 3,7 eV, the majority contribution is due to the *3d*-Ti orbital.

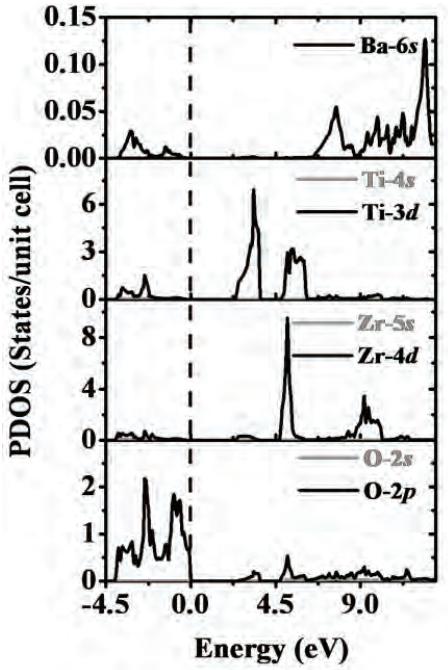


Figure 7. Partial DOS calculated for the Ba_2TiZrO_6 material.

On the other hand, from 3,7 eV up to 4,71 eV a gap of forbidden energies, internal to the conduction band, is determined. Likewise, the *6s*-Ba, *4d*-Zr, *3d*-Ti and *2p*-O electron contributions give rise to the regime between 4,74 eV and 12,0 eV, with a small share of *4s*-Ti, *5s*-Zr and *2s*-O orbitals.

5. CONCLUSIONS

In summary, we report the synthesis of polycrystalline Ba_2TiZrO_6 double perovskite. The complete thermal process is seen through a sequence of diffraction patterns. Rietveld refinement reveals that a single crystallographic phase of rhombohedral perovskite was obtained, space group *R-3* (#148), with lattice parameters *a=5.803(8)* Å. The tolerance factor was calculated to be 1.03. Measurements of polarization as a function of applied electric fields show a hysteretic behavior, which is characteristic of ferroelectric materials. Values of saturation and remnant polarization suggest that this material may be used for technological applications in devices for information storage. Lastly, calculations of band and electronic structure reveal the nonmetallic behavior of Ba_2TiZrO_6, with an energy gap of 2.32 eV and *Bo=155.9553 GPa* volume modulus.

6. ACKNOWLEDGMENTS

This work was partially supported by the *División de Investigaciones Sede Bogotá* (DIB) of the *Universidad Nacional de Colombia*.

REFERENCES

[1] Palkar, V.R., Malik, S.K., Observation of magnetoelectric behavior at room temperature in $Pb(Fe_xTi_{1-x})O_3$, Solid State Communications 134, pp.783-786, 2005.

[2] Landínez Téllez, D.A., Peña Rodríguez, G., Arbey Rodríguez, J., Fajardo, F., Roa Rojas, J., Structural, magnetic, multiferroic, and electronic properties of Sr_2TiMnO_6 double perovskite, Dyna 79, pp.111-115, 2012.

[3] Isupov, V.A., Nature of physical phenomena in ferroelectric relaxors, Physics of the Solid State 45, pp. 1107-1111 (2003).

[4] Baettig, P., Spaldin, N.A., Ab initio prediction of a multiferroic with large polarization, Applied Physics Letters 86, pp. 12505-12507, 2005.

[5] Ang, C., Guo, R., Bhalla, A.S., Cross, L.E., Dielectric dispersion at microwave frequencies, Journal of Applied Physics 87, pp. 3937-3940, 2000.

[6] Larson, A. C., Von, R. B., Dreele, *General Structure Analysis System (GSAS)*, Los Alamos National Laboratory Report LAUR, pp. 86-748, 2000

[7] Perdew, J.P., Burke, S., Ernzerhof, M., Generalized Gradient Approximation Made Simple, Physical Review Letters 77, pp. 3865, 1996.

[8] Blaha, P., Schwarz, K., Madsen, G. K. H., Kvasnicka, D. and Luitz J., WIEN2k, *An Aug-mented Plane Wave + Local Orbitals Program for Calculating Crystal Properties*, Karlheinz Schwarz, Techn. Universitat Wien, Austria 2001, ISBN 3-9501031-1-2.

[9] Murnaghan, F.D., The Compressibility of Media under Extreme Pressures, Proceedings of National Academy of Sciences, USA. 30, pp. 244-247, 1944.

[10] Lufaso, M.W., Woodward, P.M., The Prediction of the Crystal Structures of. Perovskites Using the Software Program SPuDS, Acta Crystallographic B 57, pp. 725-738, 2001.

[11] Choi, K.J., Biegalski, M., Li, Y.L., Sharan, A., Schubert, J., Uecker, R., Reiche, P., Chen, Y.B., Pan, X.Q., Gopalan, V., Chen, Q., Schlom, D.G., Eom, C.B., Enhancement of ferroelectricity in strained $BaTiO_3$ thin films, Science 306, pp. 1005-1009, 2004.

[10] Hohenberg, P., Khon, W., Density Functional Theory, Physical Review 136, B864-B871, 1964.

[12] D. Fu, M. Itoh, S. Koshihara, Lattice strain and polarization of the $BaTiO_3$–$CaTiO_3$ solid solution, Journal of Physics: Condensed Matter 22, 052204-052207, 2010.

[13] Iijima, Y., Remanent polarization of vacuum-deposited $BaTiO_3$ films, Electronics and Communications in Japan 68, pp. 27-35, 2007.

EFFECTS OF FLOTATION VARIABLES ON FELDSPATHIC SAND CONCENTRATION

EFECTOS DE LAS VARIABLES DE FLOTACIÓN EN LA CONCENTRACIÓN DE ARENAS FELDESPÁTICAS

ALEJANDRO ARGÜELLES - DÍAZ
PhD. Profesor, Universidad de Vigo, España, aargu@uvigo.es
JAVIER TABOADA - CASTRO
PhD. Profesor, Universidad de Vigo, España, jtaboada@uvigo.es
FERNANDO GARCÍA - BASTANTE
PhD. Profesor, Universidad de Vigo, España, bastante@uvigo.es
MARÍA ARAÚJO - FERNÁNDEZ
PhD. Profesora, Universidad de Vigo, España, maraujo@uvigo.es

Received for review December 26 th, 2012, accepted May 16th, 2013, final version June, 18 th, 2013

ABSTRACT: The aim of this research was to determine the influence of certain variables (e.g., collector addition, pulp dilution, conditioning times, pH and pulp agitation speed) on the yield for feldspar sand flotation aimed at separating quartz from feldspar and to estimate the values that provide an optimum yield. The feldspathic sands were obtained from a sedimentary deposit located in the Sarreaus region (Ourense province, NW Spain).

Flotation yield was based on two parameters: grade and recovery. Each of these parameters was determined in a series of experiments in which a single parameter was varied at a time while the remainder were kept constant. The results indicate that it is possible to obtain a high flotation yield using optimum process variables. With these optimum variables, the average concentrate grade and recovery were 95.1% and 25.6%, respectively.

Key words: Mineral processing, froth flotation, flotation reagents, flotation depressants, flotation collectors.

RESUMEN: El objetivo de este trabajo es determinar la influencia que determinadas variables (como son, concentraciones del colector, dilución de la pulpa, tiempos de acondicionamiento, pH y velocidad de agitación de la pulpa) ejercen en el rendimiento del proceso de flotación de arena feldespática, utilizado para separar el cuarzo del feldespato, así como estimar los valores que proporcionan un rendimiento óptimo. Las arenas feldespáticas provinieron de un depósito sedimentario ubicado en Sarreaus, provincia de Orense al noroeste de España.

El rendimiento de flotación se basa en dos parámetros: ley y recuperación. Cada uno de estos parámetros fue determinado en una serie de experimentos en los cuales se varió cada vez una de las variables manteniéndose fijos los valores del resto. Los resultados indican que es posible obtener un rendimiento elevado en la flotación habiéndose determinado las variables óptimas de proceso. Con dichas variables óptimas la ley media del concentrado fue del 95.1% y la recuperación media del 25.6%.

Palabras clave: Procesamiento de minerales, flotación por espuma, reactivos de flotación, depresores de flotación, colectores de flotación.

1. INTRODUCTION

Feldspathic materials for the glassmaking and pottery industries are obtained from feldspathic sands. However, in order to obtain an end product of suitable quality, a series of separation tasks need to be implemented, given that these minerals are - to a greater or lesser degree - generally rendered impure by ferromagnesian minerals. The flotation process to separate feldspar from quartz was developed in the 1930s [1]. Since then, many papers have been published on the application of different collectors, reagents, depressants and other substances to the flotation of feldspars [2-14].

Feldspar in Spain is generally recovered either by means of magnetic separation or in flotation circuits.

The lack of physical, chemical and size standardization in the feldspars sold in Spain - with proportions of quartz and iron that vary from one batch to another - have forced consumers to look elsewhere for suppliers, with negative repercussions for Spanish vendors. There is

also, moreover, a lack of investment in technologies and processes to improve the exploitation of mineral reserves.

Furthermore, the shortage of deposits rich in feldspathic material make it necessary to mine other types of local deposits (as described in this study), and this becomes more feasible in terms of recovery if suitable flotation techniques are used.

The main aims of this research were to determine the influence of specific parameters on the yields for feldspathic sand flotation aimed at separating feldspar from quartz (after removal of ferromagnesian minerals) and to define the values of the parameters which would optimize the flotation yield.

2. EXPERIMENTAL

2.1. Material

The mineral used for the purposes of this study was fine-grained arkosic sand with sodium-potassium feldspar, obtained from a quaternary deposit located in the Sarreaus region (Orense province, NW Spain).

Most of the feldspathic materials mined in Spain are obtained from sedimentary deposits, containing feldspathic sands composed mainly of quartz and feldspar along with other accessory minerals or impurities, such as muscovite, tourmaline, biotite, etc.

The sands used in our particular study contained biotite as an impurity, which had to be removed in a preliminary stage, prior to flotation, using high-intensity dry magnetic separators.

Fifty 5-kg samples were collected at the treatment plant after a drying process. The grain size analysis showed that 62% of the arkosic sand was less than 500 μm in diameter. The coarse sand fraction (greater than 500 μm) was 38% and the fine sand fraction (less than 63 μm) was 0.5%.

Samples were prepared and flotation tests were conducted at the mineralurgical laboratory of the University of Vigo (Spain).

The mineral was classified using an electric sieve shaker and the 63-μm to 500-μm fraction was collected. Sizes outside this range were discarded as unsuitable, since they tend to cause problems in the flotation process [15].

A high-intensity, 22000-gauss dry magnetic separator was used to separate the mica from the feldspar and quartz. This reduced the mica content of the initial sample by 95%. The mineral was then mixed and quartered.

The mineralogical composition of the 63-μm to 500-μm fraction was determined using X-ray diffraction. Chemical composition was determined using X-ray fluorescence.

Tables 1 and 2 show the mineralogical and chemical composition of the 63-μm to 500-μm fraction before and after the magnetic separation, respectively.

Table 1. Mineralogical and chemical composition of the 63-μm to 500-μm fraction before magnetic separation

Mica	*2.5%*	*Na_2O*	*1.292%*
Microcline	*29.0%*	*Al_2O_3*	*13.764%*
Albite	*10.0%*	*SiO_2*	*77.175%*
Quartz	*58.5%*	*K_2O*	*5.922%*
		CaO	*0.088%*
		TiO_2	*0.085%*
		Fe_2O_3	*0.152%*
		MgO	*0.088%*
		P_2O_5	*0.069%*
		Misc.	*1.365%*

Table 2. Mineralogical and chemical composition of the 63-μm to 500-μm fraction after magnetic separation

Non-magnetic fraction			
Mica	*0.0%*	*Na_2O*	*1.409%*
Microcline	*35.7%*	*Al_2O_3*	*13.949%*
Albite	*10.4%*	*SiO_2*	*76.653%*
Quartz	*53.9%*	*K_2O*	*6.406%*
		CaO	*0.098%*
		TiO_2	*0.058%*
		Fe_2O_3	*0.108%*
		MgO	*0.000%*
		P_2O_5	*0.074%*
		Misc.	*1.245%*

Magnetic fraction			
Mica	*46.3%*	Na_2O	*1.262%*
Microcline	*10.2%*	Al_2O_3	*19.731%*
Albite	*7.3%*	SiO_2	*65.717%*
Quartz	*32.3%*	K_2O	*8.183%*
		CaO	*0.100%*
		TiO_2	*0.752%*
		Fe_2O_3	*1.222%*
		MgO	*0.615%*
		P_2O_5	*0.062%*
		Misc.	*2.356%*

Analysis of the data shown in Tables 1 and 2 indicated that the micaceous minerals could be removed by high intensity magnetic separation. Magnetic separation alone decreased the magnetic mineral content of the feed to 0.1% Fe_2O_3 and 0.058% TiO_2.

2.2. Flotation experiments

The experiments consisted of a series of froth flotation concentration tests [16], in which a single parameter was varied at a time while the remainder were kept constant, with the objective being to determine each parameter's influence on flotation yield [17].

The parameters and conditions under which each experiment was conducted are summarized in Table 3. Italics indicates the range of variation for each parameter.

The flotation machine rotor, with an electronic variable speed drive, was 0.15 m in diameter. The flotation cell size was 13 L.

The collector used was an amine (trade name ARMAC T) — the collector typically used in feldspar flotation [18]. Hydrofluoric acid (HF) was used as a depressant regulator to prevent quartz flotation. This has the advantage of being very effective but, because this acid is corrosive, the lowest possible concentrations were used [19]. Sulphuric acid was used as the pH regulator for the pulp.

It was not necessary to use a frother, since the amine ensured a consistent foam in the tests conducted. In tests using pine oil as a frothing agent, the resulting foam was too consistent and not only impeded transport, but also prevented the eclosion necessary to free the mineral.

Table 3. Variable values in six experiments

	Amine (g/t)	Solids in weight (%)	Conditioning rotor speed (rpm)	Flotation rotor speed (rpm)	HF (g/t)	pH	Conditioning time (min)
Exp. 1	*700-1200*	35	400	600	750	2.2	6
Exp. 2	800	35	400	600	*700-1200*	2.2	6
Exp. 3	800	35	400	600	750	2.2	*2-12*
Exp. 4	**4a)** 700 **4b)** 800 **4c)** 900	*20-60*	400	600	750	2.2	6
Exp. 5	**5a)** 750 **5b)** 800	35	400	600	750	*1.8-2.8*	6
Exp. 6	800	35	*200-800*	600	750	2.2	6

For each flotation test a quantity of 2 kg was used, corresponding to the non-magnetic fraction obtained in the separator. The laboratory experimental procedure is outlined as follows:

1. Dilute the sample in water
2. Measure pH
3. Add the amine and HF
4. Agitate during the conditioning time
5. Measure pH and control with sulphuric acid
6. Open the air inlet valve to the flotation cell
7. Begin the flotation test until overflow stops

8. Dry and weigh overflow and underflow.

To analyse flotation results two parameters were defined:

1. Grade as a % of weight. Quotient between feldspar weight in the float and floated weight, indicating the richness of the floated feldspar and also the ideal concentration of the reagents.

2. Recovery as a % of weight. Quotient between float weight and sample assayed weight, indicating the process yield.

An analysis of these two parameters defined the process yield and and indicated the ideal values of the process variables.

3. RESULTS AND DISCUSSION

Below the results obtained for each experiment are described. The effect of each variable on grade and recovery was interpreted to assess the flotation yield and to establish optimum values for the variables.

Table 3 summarizes the flotation tests, showing the values maintained when particular variables were studied.

3.1. Effects of amine addition

In this experiment (Exp. 1 in Table 3), flotation tests were implemented with varied amine additions while holding the remaining parameters constant.

Figure 1 shows variations in grade and recovery according to changes in amine addition.

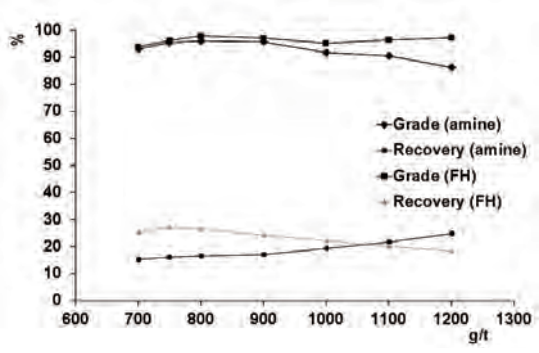


Figure 1. Influence of amine and hydrofluoric acid on grade and recovery [Exp. 1 and Exp. 2 in Table 3]

Recovery rose steadily with increasing amine addition, since greater amounts of collector caused more mineral to float. Grade, however, fell after peaking at 800 g/t because increasing quantities of quartz floated as well as feldspar. In a test conducted with 2200 g/t of amine, the entire sample floated, and no mineral was left in the tailings.

The optimum amine values fell within the range 750-850 g/t, with mean concentrate grade above 95%. It is not clear whether a major increase in recovery would compensate for a drop in grade due to higher amine usage.

The amine adsorption mechanism is not yet fully understood. At low collector addition and low pH, surface adsorption of the collector was high for feldspar and low for silicate (quartz). This high surface affinity of the collector for feldspar can be associated with the formation of inter-molecular complexes with a high hydrophobic capacity [20].

3.2. Effects of hydrofluoric acid

In this experiment (Exp. 2 in Table 3), flotations were implemented with different concentrations of HF, with the remaining parameters held constant. Figure 1 depicts the effect of varying HF concentration.

Recovery achieved a maximum value at a HF concentration of approximately 750 g/t, and then began to fall because HF has the effect of being a quartz depressant. HF also prevented the feldspar from floating, so the quantity of floating mineral fell as HF concentrations were increased.

Grade began to rise to a maximum, fell and then rose again. This happened because with low initial concentrations of HF, the quartz was depressed and did not float, with grade maintained almost constant for a HF concentration of 750-900 g/t.

A slight fall occurred between 900 and 1000 g/t, because the feldspar began to depress, but from 1000 g/t the grade began to increase again (concentration was sufficiently high to depress feldspar and quartz). Therefore, the flotation resulted in hardly any quartz, but there was also less feldspar and recovery became increasingly lower.

A range of 750-850g/t would seem to be the optimum HF concentration, since both grade and recovery would be close to a maximum.

3.3. Effects of pre-flotation pulp conditioning time variations

In this experiment (Exp. 3 in Table 3), the influence of pulp conditioning time on flotation yield was analysed. Flotations were implemented for five different time conditions, with all the other parameters held constant.

Figure 2 shows the results obtained for the different conditioning times. For shorter times, both grade and recovery were low since there was insufficient time for adsorption of the mineral by the collector. As time increased up to six minutes grade and recovery both improved, as the collector was gradually adsorbed and implemented its hydrophobic function with the feldspar.

From six minutes, grade and recovery appeared to stabilize so the optimum time was set at six minutes.

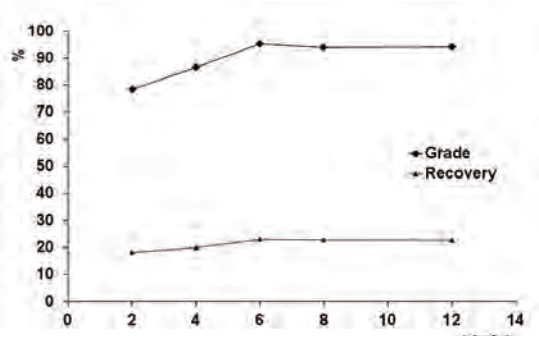


Figure 2. Variations in grade and recovery according to conditioning time [Exp. 3 in Table 3]

3.4. Effects of pulp dilution variations

Various dilutions were implemented in these experiments (Exps. 4a, 4b and 4c in Table 3), with the amine addition varied in each dilution. The proportions of solids by weight used were 20%, 30%, 35%, 40%, 50% and 60%. Amine concentrations were 700 g/t, 800 g/t and 900 g/t.

Figure 3 shows the results obtained. The interpretation of behaviour at each dilution is outlined below:

- For a dilution of 20% by weight, the recovery increased with collector concentration, as the quartz also floated; grade started to fall from above 800 g/t of amine, however, due to the greater presence of quartz in the float.

- For a dilution of 30% by weight, the grades and recoveries were greater than for the previous experiment, given that the proportion of solids in the sample was greater.

- For a dilution of 35% by weight, the graph obtained was similar to that of the previous case, with some small differences in recovery and grade indicating that the dilution difference was correspondingly small.

- For a dilution of 40% by weight, the graph obtained was similar to that for the 35%, with similar grades but with a gentle fall in recovery due to the fact that the solids concentration began to be significant, and the collector was therefore adsorbed with some difficulty by the mineral.

- For a proportion of 50% by weight, the grade was maintained at the preceding values; that is, it was little affected by an increase in the proportion of solids because the collector was not physically prevented from being adsorbed by the mineral. However, the collector was prevented from reaching further quantities of mineral, so recovery began to fall sharply.

- For a proportion of 60% by weight, grade was maintained at the above values for the same reasons as given above. Recovery underwent a notable fall, due to the fact that the proportion of solids was already high, so the probability of contact between the collector and mineral particles was reduced.

- A final assay conducted with 80% in solids resulted in a miniscule quantity of floating mineral. At these densities, particle clusters are very large, so the collector is adsorbed with great difficulty. Moreover, larger particles do not have the same mobility as in lower dilutions and the probability of contact with the collector is thus significantly reduced.

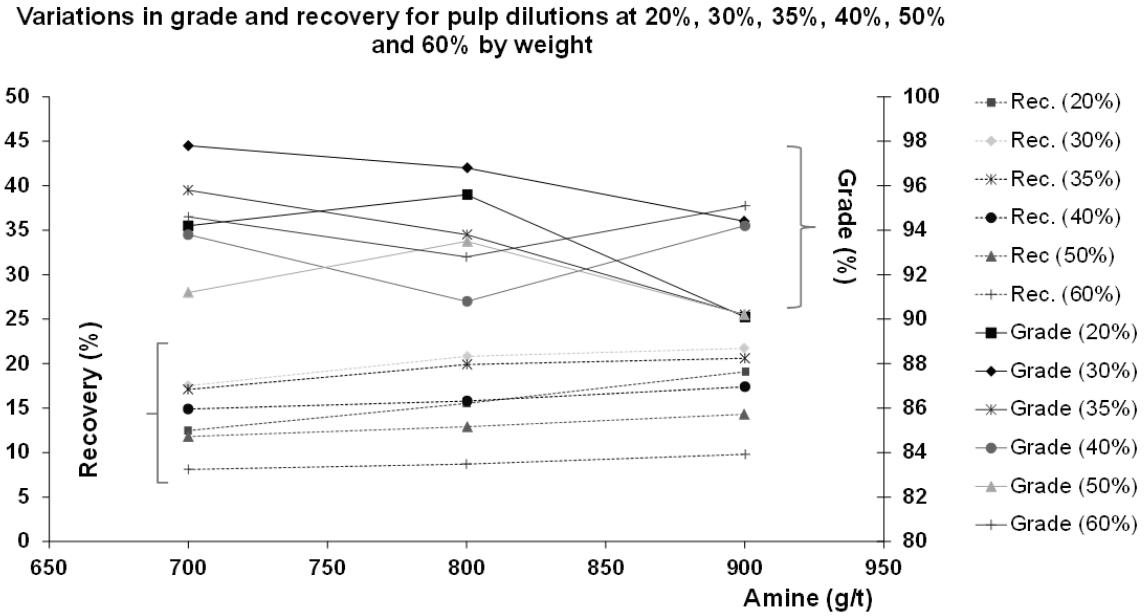


Figure 3. Variation in grade and recovery for sequential solid pulp dilutions (%) by weight [Exp. 4 in Table 3]

3.5. Effects of pulp pH variations

Two series of assays with pH variations were conducted, for two different additions of amine (750 g/t and 800 g/t). The results obtained are shown in Figure 4, where it can be observed that up to a value of approximately 2.2, the influence of pH on recovery was of little significance. Above this value, however, feldspar selectivity was reduced and, as a result, grade fell sharply; the recovery level also fell, but more gently.

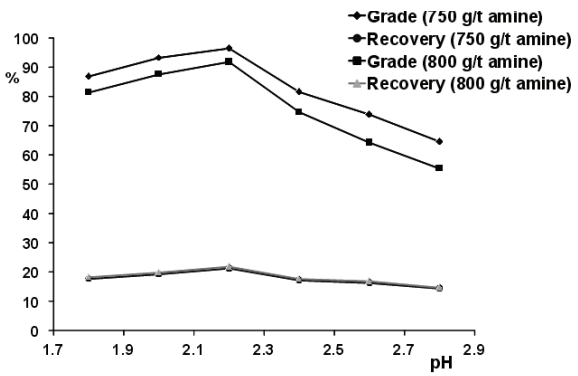


Figure 4. Variations in grade and recovery according to pH for amine additions of 750 g/t and 800 g/t [Exp. 5 in Table 3]

It can be observed that from pH values above 2.2, more quartz was floated and selectivity fell in line with grade. These results would indicate that selective feldspar flotation is only possible for a narrow range of pH conditioning. The loss of selectivity at high pH values can be attributed to a reduction in the Z potentials for the minerals. For pH values higher than 3, selectivity was lost totally.

These results are similar to those for the previous series. Recovery increased slightly because collector concentration rose, whereas grade decreased because more quartz floated due to the increase in collector concentration.

3.6. Effects of pre-flotation pulp conditioning rotor speed variations

In order to determine the effects of rotor speed, six flotations at different speeds were implemented, with all the remaining parameters held constant (Exp. 6 in Table 3). Results are shown in Figure 5.

At low rotor speeds, pulp agitation was low and most mineral particles were not in suspension. Therefore, the collector was not adsorbed, and recovery and grade were consequently low.

As speed was increased, recovery and grade gradually increased, with the maximum achieved for both in the interval 400-600 rpm.

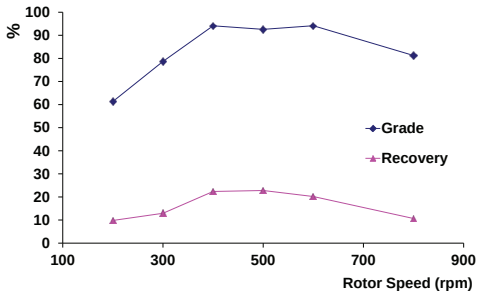


Figure 5. Variations in grade and recovery according to pulp conditioning [Exp. 6 in Table 3]

At speeds over 800 rpm, recovery and grade began to fall; the agitation of particles was high, causing some particles to break, resulting in a larger quantity of fine particles. These fines surrounded the other particles and prevented the collector from doing its work.

Moreover, a high level of agitation forced some of the pulp out of the cell and so a proportion of pulp was lost.

4. CONCLUSIONS

In all the assays conducted, recoveries were low, due to the fact that only a single flotation was implemented per sample. This recovery could be increased considerably by increasing amine addition, with the corresponding drawback: a large amount of quartz would be floated but the grade would be very low.

Sufficient feldspar remained for recovery in the tailings for each flotation. This could be easily recovered with successive flotations of the tailings, resulting in higher recovery rates, thereby justifying implementation in industry.

The most important factor is to ensure that variables are adjusted to obtain a high grade in a single flotation — as has been done in our research — since achieving high recovery rates over several successive flotations of the tailings is a less complex process.

An analysis of the results indicated that the optimum variables for the process were those summarized in Table 4.

With these optimum values the mean concentrate grade and mean recovery rates obtained for our flotation tests were 95.1% and 25.6%, respectively.

Table 4. Optimum flotation values for feldspathic sand

Amine addition	(750-850) g/t
Hydrofluoric acid	(750-850) g/t
Conditioning time	6 minutes
Pulp dilution	(30-40) %
pH	2.2
Agitation speed	(400-600) rpm

ACKNOWLEDGEMENT

The authors acknowledge the reviewer's comments and suggestions, which contributed greatly to improving the quality of the final version of this paper. Ailish M.J.Maher provided assistance with English usage in a version of the manuscript.

REFERENCES

[1] O'meara, R. G., Norman, J. E. and Hammond, W. E., Froth flotation and agglomerate tabling of feldspar, Bull. Am. Ceram. Soc., 18 (8), pp. 286-292, 1939.

[2] Manser, R. M., Handbook of Silicate Flotation, Warren Spring Lab. Stevenage, England, 1975.

[3] Rao, H. K. and Forssberg, K. S. E., Feldspar flotation: theory and practice, In: Selected Topics in Mineral Processing (Eds. A. Gash, R. Pradip Kumar and S. Venkatachalam), Port City Press, Baltimore, pp. 86-117, 1995.

[4] Shehu, N. and Spaziani, E., Separation of feldspar from quartz using EDTA as modifier, Minerals Engineering, 12 (11), pp. 1393-1397, 1999.

[5] Demira, C., Bentlib, I., Gülgönülc, I. and Çelikd, M. S., Effects of bivalent salts on the flotation separation of Na-feldspar from K-feldspar, Minerals Engineering, 16 (6), pp. 551-554, 2003.

[6] Bayat, O., Arslan, V. and Cebeci, Y., Combined application of different collectors in the flotation concentration of Turkish feldspars, Minerals Engineering, 19 (1), pp. 98-101, 2006.

[7] Gougazeh, M., Evaluation and beneficiation of feldspar from arkosic sandstone in South Jordan for application in the ceramic industry, American Journal of Applied Sciences 3 (1), pp. 1655-1661, 2006.

[8] Vidyadhar, A. and Rao, K. H., Adsorption mechanism of mixed cationic/anionic collectors in feldspar-quartz flotation system, Journal of Colloid and Interface Science, 306 (2), pp. 195-204, 2007.

[9] Mantilla, C., Pedraza, J. and La Verde, D., Application of zeta potential studies in the development of an alternative process for the flotation of feldspar minerals, DYNA, 75 (154), pp. 65-71, 2008.

[10] Aksay, E. K., Akarj, A., Kaya, E. and Cocen, I., Influence of fatty acid based collector on the flotation of heavy minerals from alkali feldspar ores, Asian Journal of Chemistry, 21 (3), pp. 2263-2269, 2009.

[11] Řepková, M., Botula, J. and Řepka, V., Application of the froth flotation for feldspar ores treatment, In proceeding of 11th International Multidisciplinary Scientific Conference, At Albena, Bulgaria, Vol. 1, pp. 1077-1082, 2011.

[12] Hacifazlioglu, H., Kursun, I. and Terzi, M., Beneficiation of low-grade feldspar ore using cyclojet flotation cell, conventional cell and magnetic separator. Physicochemical Problems of Mineral Processing, 48 (2), pp. 381-392, 2012.

[13] Heyes, G. W., Allan, G. C., Bruckard, W. J. and Sparrow, G. J., Review of flotation of feldspar. Transactions of the Institutions of Mining and Metallurgy, Section C: Mineral Processing and Extractive Metallurgy, 121 (2), pp. 72-78, 2012.

[14] Zhang, Z., Feng, Q., Wang, W., Xu, L. and Hunag, Y., Adsorption characteristics of feldspar and quartz in anion-cation collector. Zhongnan Daxue Xuebao (Ziran Kexue Ban)/Journal of Central South University (Science and Technology), 44 (4), pp. 1312-1318, 2013.

[15] Trahar, W. J., The selective flotation of galena from sphalerite with special reference to the effects of particle size, International Journal of Mineral Processing, 3 (2), pp. 151-156, 1976.

[16] Hadler, K., Aktas, Z. and Cilliers, J. J., The effects of frother and collector distribution on flotation performance, Minerals Engineering, 18 (2 SPEC. ISS.), pp. 171-177, 2005.

[17] Shimoiiza, J. and Nakatsuka, K., Study on the separation of feldspar from quartz by flotation on activation with FH, Journal of Mining and Metallurgical Inst. of Japan, 80 (918), pp. 1054-58, 1964.

[18] Sekulić, Ž., Canić, N., Bartulović, Z. and Daković, A., Application of different collectors in the flotation concentration of feldspar, mica and quartz sand, Minerals Engineering, 17 (1), pp. 77-80, 2004.

[19] Paulson, E. G., Reducing fluoride in industrial wastewater, Chemical Engineering (New York), 84 (22), pp. 89-94, 1977.

[20] Vidyadhar, A., Rao, K. H. and Chernyshova, I. V., Mechanisms of amine-feldspar interaction in the absence and presence of alcohols studied by spectroscopic methods, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 214 (1-3), pp. 127-142, 2003.

THE CONTENT OF MANUFACTURING STRATEGY: A CASE STUDY IN COLOMBIAN INDUSTRIES

EL CONTENIDO DE LA ESTRATEGIA DE MANUFACTURA: UN ESTUDIO DE CASO EN LAS INDUSTRIAS COLOMBIANAS

JORGE ANDRÉS VIVARES - VERGARA
PhD (c), Becario Colciencias, javivaresv@unal.edu.co
WILLIAM ARIEL SARACHE - CASTRO
PhD, Profesor Asociado Universidad Nacional de Colombia, wasarachec@unal.edu.co
JULIA CLEMENCIA NARANJO - VALENCIA
PhD, Profesora Asociada Universidad Nacional de Colombia, jcnaranjov@unal.edu.co

Received for review April 4 th, 2013, accepted June 14th, 2013, final version June, 19 th, 2013

ABSTRACT: Manufacturing strategy (MS) is a long-term plan for production and operations system aimed to support the company's corporate strategy. The content of MS addresses the goals and strategic decisions to face competition. Despite the number of contributions on this subject, few studies have been conducted in the Colombian context. Therefore, this article shows the results of a study undertaken in 36 Colombian companies addressing the three main components of the MS content: competitive priorities, strategic decision areas and the management's approach to manufacturing. The results allow two groups of companies with different performance level to be identified, as well as the relationship of this performance with strategic decision areas and the management's approach to manufacturing.

Key words: Manufacturing strategy, competitive priorities, strategic decision areas, manufacturing management approaches, Colombian industry.

RESUMEN: La estrategia de manufactura (EM) consiste en un plan a largo plazo para el sistema de producción/operaciones diseñado para apoyar la estrategia corporativa de la compañía. El contenido de la EM aborda las metas y las decisiones estratégicas necesarias para enfrentar la competencia. A pesar del creciente número de publicaciones sobre este tema, pocas investigaciones han sido realizadas en el contexto colombiano. Por tanto, el presente artículo expone los resultados de un estudio realizado en 36 empresas colombianas que aborda los tres componentes principales del contenido de la EM: las prioridades competitivas, las áreas estratégicas de decisión y los enfoques de gestión para la manufactura. Los resultados permitieron detectar dos grupos de empresas que exponen diferentes niveles de desempeño así como las relaciones de dicho resultado con las áreas de decisión estratégicas y los enfoques de gestión de manufactura.

Palabras clave: Estrategia de manufactura, prioridades competitivas, áreas de decisión estratégica, enfoques de gestión para la manufactura, industria colombiana.

1. INTRODUCTION

Due to the growth of global competition and its effects on international operations, greater efforts in strategic planning have been undertaken in order to ensure the long-term survival of companies. From the strategic planning point of view, a topic of great interest in the last four decades has been the new role of the manufacturing system in the company's competitive strategy.

Traditionally, the manufacturing system has been treated as a part of the company whose role is limited merely to address technical issues of production. However, due to the current competitive context, many companies have understood that the manufacturing system can become a powerful competitive weapon to improve their performance in the market. In this way, Manufacturing Strategy (MS) can be defined as a long-term plan for the production and operations system aimed to support the company's corporate strategy. Despite the fact that this topic has been widely addressed in many scientific investigations, few studies on the subject have been conducted in the Colombian context.

Therefore, this article shows the results of a study undertaken in 36 companies located in the Colombian central-western region, aimed at analyzing the content of their manufacturing strategies. Specifically, the

study addressed three main aspects: a) the competitive priorities; b) the strategic decision areas and c) the management's approach to manufacturing. According to the findings, two clusters of companies were identified. The first cluster (leader group) showed significant strengths in his MS and better performance. The second cluster (lagging group) showed important weaknesses in many areas and lower performance. Likewise, the results revealed that 8 strategic decision areas were directly related with the companies' performance levels; however, the assessment of management's approach to manufacturing shows that none of them had positive impacts on the companies.

This article is part of the research entitled "Impact of human resource management in operations strategy", undertaken at the Universidad Nacional de Colombia.

2. LITERATURE REVIEW AND HYPOTHESIS

The formal study of MS comes from the original contributions of Wickham Skinner [1, 2] and, since then, it has become a topic of growing interest in the scientific community over the past 40 years. The MS seeks to assign a strategic role for manufacturing systems beyond the traditional technical role given in the past [3]. Therefore, MS can be understood as a long-term plan for manufacturing systems containing decisions and strategic actions aimed to support the whole company's strategy [4].

The MS addresses two main elements: the formulation process and the content. The formulation process establishes how to proceed in order to strengthen and deploy the capabilities of manufacturing systems according to the company´s long-term strategy [5]. The content addresses the goals and strategic decisions to face the competition. Traditionally the content encompasses two main aspects: competitive priorities and strategic decision areas [3]. However, we consider that it is necessary to take into account a third aspect: the management's approach to manufacturing. These three topics, given the purpose of this article, are discussed in greater detail below.

2.1. Competitive priorities

The competitive priorities are the goals for manufacturing systems in order to increase the company's competitive advantage [6]. Competitive priorities have been named in different ways such as manufacturing goals, manufacturing tasks and manufacturing outputs, among others [7].

In the 80's, Miller [8], proposed a group of seven competitive priorities to guide the manufacturing system toward a better performance in the market: low cost, high quality, high service level, broad product portfolio, service attitude, product innovation and reaction to change. Due to the growth of international trade and the consequent increase of competitors, during the decade of the 90s new competitive priorities such as delivery, flexibility, and environmental responsibility became more relevant [9]. At present, there is a certain level of agreement on six main competitive priorities: cost, quality, flexibility, delivery, service and environmental responsibility [3,10-12].

In the Colombian context, few studies have been conducted regarding competitive priorities. According to the literature review, only one contribution in the metalworking sector [7,] and two in the apparel sector [13,14] were found. In the contribution of Sarache et al. [13] the so-called effectiveness indicator (*EI*) was developed. This indicator allows the performance assessment of manufacturing systems, based on their outcomes in competitive priorities. Because *EI* was applied in the present study, the mathematical expression is shown in equation 1.

$$EI_i = \sum_{j=1}^{n} W_{ij} C_{ij} \quad (1)$$

Where:

EI_i: Effectiveness of manufacturing system at Company *i*.

W_{ij}: Weight of competitive priority *j* at company *i*.

C_{ij}: Rating of competitive priority *j* at company *i*.

2.2. Strategic decision areas

The strategic decision areas for manufacturing systems greatly affect the company's survival. These are divided in structural and infrastructural decisions [3, 15, 16]. Structural decisions are characterized by their long-term impact not only because they require high investment but also because they significantly affect the manufacturing system's capabilities. In turn, infrastructural decisions address the management processes in diverse company areas in order to support

the manufacturing system [17]. Table 1 shows the structural and infrastructural decisions according to various contributions.

Based on the above, the performance and orientation of MS depends on two main aspects: the performance in competitive priorities and the way companies adopt to focus their strategic decision areas.

Table 1. Strategic decision areas in manufacturing

Structural decisions	Infrastructural decisions
Processes	Human resources
Capacity	Products
Facility location	Planning and control
Facility layout	Organization
Supply/distribution	Work study
	Quality management

Source: Author´s elaboration based on contributions of [3],[5],[16] and [17].

In this sense, the hypothesis 1 and 2 are as follows:

Hypothesis 1. There are different profiles of companies according to their EI and the orientation of to their strategic decision areas.

Hypothesis 2. There is a relationship between the performance in the strategic decision areas and the EI achieved by companies.

2.3. Management's approach to manufacturing

Companies have adopted various management approaches that must be taken into consideration as a part of the content of manufacturing strategy. From a broad perspective, these approaches are based on management philosophies aimed to improve effectiveness and performance of production systems. The most recurrent management's approach to manufacturing have been Just in Time (JIT) and Total Quality Management (TQM) [18,19], Total Productive Maintenance (TPM) [20], Theory of Constraints (TOC) [21,22], 5s and Kaizen [23,24]. Such approaches are not applied in an isolated way, but rather they act in an interconnected manner which commonly occurs between TPM, TQM and JIT or between TQM and Kaizen [25]. These considerations support the hypotheses 3 and 4.

Hypothesis 3. There is a relationship between the implementation level of management's approach to manufacturing and the EI achieved by companies.

Hypothesis 4. The management's approach to manufacturing adopted by companies are applied in a complementary way.

3. METHODOLOGY

3.1. Population and sample

The study was conducted in large and medium-sized industrial enterprises located in the Colombian central-western region. According to the government statistical reports, the population was composed of 48 companies. The survey was sent to production managers achieving a response rate of 75% (36 companies). Based on the contribution of [26], 11 semi-structured interviews were conducted in order to collect qualitative data to enrich the study outcomes.

3.2. Variables and measures

In this research, three groups of variables were addressed: competitive priorities, strategic decision areas and implementation of manufacturing management approaches. Table 2 summarizes the operationalization of variables.

Table 2. Variables, dimensions and measures used in the study

Variables	Dimensions	Measurement
Competitive priorities	Cost, quality, flexibility, delivery, service and environmental responsibility.	Performance assessment for each competitive priority regarding to the company's main competitor (Likert scale 1-5).
	Effectiveness indicator (EI)	Application of equation 1
Strategic decision areas	Processes, capacity, facility location, facility layout, supply/distribution, human resources, products, planning and control, organization, work study and quality management.	Level of performance in every decisión areas according to the context, business requirements and market expectations (Likert scale 1-5).
Manufacturing management approaches	JIT, TQM, TPM, TOC, 5s and Kaizen	0: not used. 1-5: according to the level of implementation/functionality.

3.3. Tests of validity and reliability

The survey content was structured according to contributions obtained from the literature review; also, two experts evaluated it. The internal consistency, tested by Cronbach's alpha coefficient, was 0.943, showing a high level of reliability [27]. By applying an analysis of variance among companies that responded to the survey and those that did not, the sample consistency was tested ($F = 0.004$, P-value ≥ 0.05 (0.950)). These results also were verified by the Mann-Whitney U test, repeating the process for subsets of medium and large enterprises ($U = 100$; P-value ≥ 0.05 (0.094); $F = 0.653$ for medium enterprises; $F=1.681$ for large enterprises; P-value = 0.429 for medium enterprises and 0,209 for medium enterprises).

Likewise, in order to improve the survey content, a pilot test in three companies was carried out. Finally, convergent validity was assessed by the principal component analysis factor with varimax rotation. The obtained solutions were suitable for all dimensions (KMO> 0.5; p-value <0.001 in all cases).

4. RESULTS AND DISCUSION

4.1. General description.

According to the results shown in Table 3, quality is the most important competitive priority, followed by service and deliveries. While cost was considered as the last, both flexibility and environmental responsibility were rated as satisfactory.

Table 3. Performance in competitive priorities

Competitive priority	Statistic		Percentage of companies by level of performance			
	Mean	**S**	**1-2**	**3**	**4-5**	**Total**
Quality	4.50	0.78	2.8%	8.3%	88.9%	100%
Service	4.14	0.80	2.8%	8.3%	88.9%	100%
Deliveries	3.97	0.74	2.8%	19.4%	77.8%	100%
Flexibility	3.83	0.94	8.3%	27.8%	63.9%	100%
Environmental Responsibility	3.81	0.98	8.3%	25.0%	66.7%	100%
Cost	3.58	1.03	13.9%	30.6%	55.5%	100%

Table 4 shows the results regarding strategic manufacturing decision areas. As is shown, 80.5% of companies consider that quality management is the best performing decision area. Although none of the analyzed decision areas showed a critical situation, a significant gap was detected in processes decisions, facility layout, work study and supply/distribution. This outcome clearly suggests the need of improvement programs in these decision areas to achieve a better performance in the manufacturing system.

Table 4. Performance in strategic decision areas

Strategic decision areas	**Statistic**		**Percentage of companies by score range**			
	Mean	**S**	**1-2**	**3**	**4-5**	**Total**
Quality management	4.08	0.84	5.6%	13.9%	80.5%	100%
Capacity	3.97	0.74	0.0%	27.8%	72.2%	100%
Products	3.97	0.85	5.6%	19.4%	75.0%	100%
Organization	3.97	0.77	5.6%	13.9%	80.5%	100%
Planning and control	3.86	0,72	5.6%	19.4%	75.0%	100%
Human resources	3.83	1.03	11.1%	19.4%	69.5%	100%
Facility location	3.64	1.02	11.1%	25.0%	63.9%	100%
Processes	3.64	1.07	13.9%	30.6%	55.5%	100%
Facility layout	3.64	1.05	13.9%	27.8%	58.3%	100%
Supply/distribution	3.58	0.77	8.3%	33.3%	58.4%	100%
Work study	3.42	1.05	16.7%	33.3%	50.0%	100%

Table 5. Effectiveness index (EI)

Statistic	Company size		Global
	Medium	Large	
EI	3.84	4.09	4.00
Maximum	4.66	4.75	4.75
Minimum	2.69	2.75	2.69
Median	3.84	4.10	4.07
Standard deviation	0.55	0.48	0.51
Coeff. of variation	14.2%	11.6%	12.7%

U de Mann-Whitney (99.500), *p-value* (0.100) > 0.05

By applying equation 1, the average *EI* for the group of surveyed companies was 4,0 ranging from 2.7 to 4.8 (See table 5). According to the scale proposed by Sarache [13], this result can be considered as satisfactory. On the other hand, no significant differences were found between medium and large enterprises. In general, the findings suggest that companies have a good level of performance in their competitive priorities that enable them to meet market needs adequately.

4.2. Hypothesis testing

By applying K-means cluster analysis with Ward's method, the hypothesis 1 was tested. Two groups of companies with significant differences in all variables were identified (see Table 6). The first cluster (named leader group), made up of 58% of the companies, showed better results not only in *EI* but also in the strategic decision areas compared with the second cluster (named lagging group).

Table 6. ANOVA results for cluster analysis

Variable	Cluster 1 Mean (S.D)	Cluster 2 Mean (S.D)	P-value
EI	4.21 (0.41)	3.71 (0.50)	0.002**
Capacity	4.38 (0.59)	3.40 (0.51)	0.000***
Facility location	4.05 (0.67)	3.07 (1.16)	0.003**
Processes	4.33 (0.66)	2.67 (0.72)	0.000***
Facility layout	4.33 (0.58)	2.67 (0.72)	0,000***
Supply/distribution	4.10 (0.44)	2.87 (0.52)	0.000***
Human resources	4.38 (0.67)	3.07 (0.96)	0.000***
Products	4.24 (0.77)	3.60 (0.83)	0.023*
Planning and control	4.14 (0.57)	3.47 (0.74)	0.004**
Organization	4.33 (0.58)	3,47 (0,74)	0.000***
Work study	3.86 (0.66)	2.80 (1.21)	0.002**
Quality management	4.48 (0.60)	3.53 (0.83)	0.000***

* Significant differences at 0.05; ** Significant differences at 0.01 *** Significant differences at 0.001.

Although the study did not address the management's approach to manufacturing to avoid missing data because some companies do not apply them, an additional assessment showed that companies in cluster 1 has a greater inclination towards implementing such approaches in their manufacturing systems (See Table 7).

Table 7. Management's approach to manufacturing applied for each cluster

Management approaches	Cluster 1				Cluster 2			
	Not used	Low	Medium	Good	Not used	Low	Medium	Good
JIT	10%	19%	19%	52%	40%	27%	26%	7%
TQM	5%	10%	18%	67%	40%	27%	26%	7%
TPM	10%	14%	33%	43%	47%	20%	33%	0%
TOC	29%	14%	19%	38%	47%	20%	20%	13%
5s	5%	10%	28%	57%	13%	53%	27%	7%
Kaizen	24%	10%	14%	52%	47%	27%	19%	7%

The results in Table 8 partially support the hypothesis 2. Among the eleven decision areas evaluated, only eight of them showed significant regression models that proved their direct relationship with *EI*. In the remaining decision areas (facility location, human resources and quality management) enough evidence was not found to establish some relationship with this indicator.

Table 8. Relationship between strategic decisions areas and EI

Variable	Spearman's rho		Regression			
	Coefficient	P-value	$\hat{\beta}_0$	$\hat{\beta}_1$	F-Test (p-value)	R^2
EI (dependent variable)	1.000	.				
Capacity	0.612***	0.000	2.244***	0.442***	0.000***	0.409
Facility location	0.305	0.070	3.440	0.154	0.068	0.095
Processes	0.592***	0.000	2.971***	0.282***	0.000***	0.355
Facility layout	0.510**	0.001	3.202***	0.217**	0.006**	0.200
Supply/distribution	0.600***	0.000	2.658***	0.374***	0.000***	0.321
Human resources	0.232	0.173	3.615***	0.100	0.236	0.041
Products	0.471**	0.004	3.045***	0.240*	0.016*	0.159
Planning and control	0.329*	0.050	0.315***	0.232*	0.050*	0.108
Organization	0.407*	0.014	2.772***	0.309**	0.004**	0.221
Work study	0.418*	0.011	3.143***	0.251**	0.001**	0.269
Quality management	0.285	0.092	3.317***	0.167	0.103	0.076

* Significant at 0.05. ** Significant at 0.01. *** Significant at 0.001.

On the other hand, the regression analysis exposed in Table 9, indicates that implementation of management's approach to manufacturing does not affect the *EI*. None of the analyzed management approaches showed significant results. Even more, the coefficient of determination (R^2) was very low in most cases; therefore, it was not possible to find statistical support for hypothesis 3.

Table 9. Relationship between EI and management's approach to manufacturing

Variable	Spearman's rho		Regression			
	Coefficient	P-value	$\hat{\beta}_0$	$\hat{\beta}_1$	F-Test (p-value)	R^2
EI (dependent variable)	1.000	.				
JIT	0.338	0.079	0.3869***	0.063	0.305	0.040
TQM	0.032	0.869	4.040***	0.010	0.900	0.001
TPM	0.074	0.715	3.986***	0.025	0.787	0.003
TOC	0.127	0.564	3.975***	0.033	0.692	0.008
5s	0.257	0.149	3.698***	0.103	0.172	0.059
Kaizen	0.208	0.330	3.939***	0.049	0.509	0.020

*** Significant at 0.001.

Regarding hypothesis 4, a factor analysis with varimax rotation that showed adequate results at 0.001 was carried out (KMO = 0.783 Test, P-value <0.001; communalities greater than 0.5). Through this analysis, two factors were extracted with a total explained variance of 85.13%. That is, the management's approach to manufacturing can be grouped in two factors: the first one made up of JIT, TQM, TPM and TOC; the second one made up of 5S and Kaizen (see Table 10). However, because both factors are located over one of the quadrants of the Figure 1, it is possible to conclude that all management's approach to manufacturing are closely related, complementing each other, whereby the hypothesis 4 is validated.

Table 10. Rotated Component Matrix

Management approaches	Component	
	1	2
JIT	0.240	0,858
TQM	0.635	0,664
TPM	0.652	0,664
TOC	0.128	0,893
5S	0.949	0,095
Kaizen	0.879	0,328

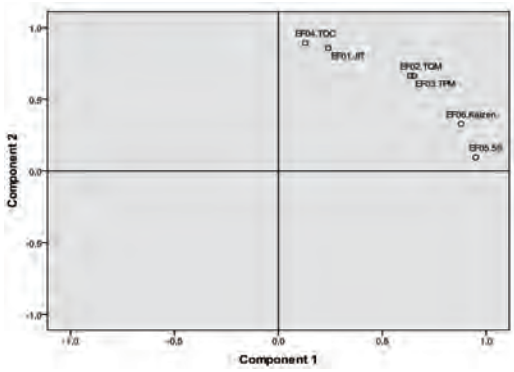


Figure 1. Rotated Component graphic

5. CONCLUSIONS

According to their competitive results, companies can be classified into a leader group or into a lagging group. Compared with the lagging group, the leader group showed better performance in its competitive priorities (average *EI* = 4.21) and a higher level of development in both strategic decisions areas and management's approach to manufacturing. Regarding the size, there was no significant difference between medium and large companies.

Also, the surveyed companies consider that the two of the most important competitive priorities are quality and cost. However, when the performance level was assessed, the cost showed the lowest rating. This finding indicates that companies must review their strategic decision areas and their management's approach to manufacturing in order to achieve a better performance level.

In addition, the study showed that a proper development of strategic decision areas positively affects the performance of competitive priorities. According to the results, capacity, processes, facility layout, supply/distribution, products, planning and control, organization and work-study are the decision areas that generate a greater effect on a firm's performance.

Furthermore, although the quality was the most prominent competitive priority, efforts regarding quality management are not generating the expected positive effects. Based on this result it is possible to infer that the quality management systems adopted by enterprises should be reviewed and improved.

Regarding management's approach to manufacturing two findings were significant. First, its level of implementation is still not adequate in most companies and many of them do not have a clear idea about its proper application. Second, there was no statistical evidence to establish that such approaches positively affect the firm's performance in their competitive priorities.

6. REFERENCES

[1] Skinner, W., Manufacturing Missing Link in Corporate Strategy, Harvard Business Review, 47 (july-August), pp. 136-145, 1969.

[2] Skinner, W., The focused factory. Harvard Business Review, May-June, pp. 113-121, 1974.

[3] Martín, M. Y Díaz, E., Posicionamiento estratégico de las empresas industriales en las prioridades competitivas de operaciones: desarrollo y aplicación de un indicador de medida. Cuadernos de Economía y Dirección de la Empresa, 39, pp. 59-94, 2009.

[4] Domínguez-Machuca, J.A., Álvarez-Gil, M., Domínguez-Machuca, M.A., García-González, S. Y Ruiz-Jiménez, A., Dirección de Operaciones: aspectos estratégicos en la producción y los servicios, McGraw-Hill, Madrid, 1995.

[5] Boyer, K., Swink, M. and Rosenzweig, E., Operations strategy research in the POMS journal. Production and Operations Management,14(4), pp. 442-449, 2005.

[6] Li, S. and Scullion, H., Developing the local competence of expatriate managers for emerging markets: A knowledge-based approach. Journal of World Business, 45 (2), pp. 190-196, 2010.

[7] Sarache, W., Cárdenas, D. Y Giraldo, J., Procedimiento para la definición y jerarquización de prioridades competitivas de fabricación. Aplicaciones en las pymes de la industria metalmecánica. Ingeniería y Competitividad, 7(2), pp. 84-91, 2005.

[8] Miller, S., Make Your Plant Manager's Job Manageable. Harvard Business Review, 1 (1), pp. 69-74, 1983.

[9] Olhager, J. and West, B.M., The House of Flexibility. Using the QFD approach to Deploy Manufacturing Flexibility. International Journal of Operations & Production Management, 22 (4), pp. 50-79, 2002.

[10] Ahmad, S. and Schroeder, R.G., The impact of human resource management practices on operational performance: recognizing country and industry differences. Journal of Operations Management, 21, pp. 19-43, 2003.

[11] Boyer, K.K. and Lewis, M.W., Competitive priorities: investigating the need for trade-offs in operations strategy. Journal of Production and Operations Management, 11 (1), pp. 1-13, 2002.

[12] Askar, M. and Mortagy, A.K., Assessing the relative importance of competitive priorities in Egyptian Companies. Advanced Management Journal, 72 (3), pp. 35-46, 2007.

[13] Sarache, W., Modelo con enfoque estratégico y procedimientos para contribuir al incremento del nivel de desempeño de las PyMEs de confección desde la función de producción. Tesis doctoral, Universidad Central de Las Villas, Santa Clara, Cuba, 2003.

[14] Sarache, W., Castrillón, O. and Giraldo, J., Prioridades competitivas para la industria de la confección: estudio de caso. Cuadernos de Administración, 24(43), pp. 89-110, 2011.

[15] Krajewski, L.J., Ritzman, L.P. and Malhotra, M.K., Administración de operaciones. Procesos y cadenas de valor. Octava edición, Pearson Educación, México, 2008.

[16] Ibarra, S. and Sarache, W., Dirección de la producción: su papel estratégico en la productividad. En: Colectivo de autores (Ed.). Gestión de la producción: una aproximación conceptual, Unibiblos. Bogotá, 2008.

[17] Miltenburg, J., Manufacturing strategy. How to formulate and implement a winning plan. Second edition, Productivity Press, New York, 2005.

[18] Zapata, A. and Sarache, W., Calidad y responsabilidad social empresarial: un modelo de causalidad. Dyna, No. 177, pp.31-39, 2013.

[19] Escanciano, C. and Iglesias-Rodriguez, F. Quality management and integrated total quality in Spanish mining: results of an empirical study. Dyna, No. 171, pp. 167-174, 2012.

[20] Kumar, R., Kumar, D. and KUMAR, P. Manufacturing excellence through TPM implementation: a practical analysis. Industrial Management & Data Systems, 106(2), 256-280, 2006.

[21] Inman, R., Lair, M. and Green JR, K., Analysis of the relationships among TOC use, TOC outcomes, and organizational performance. International Journal of Operations & Production Management, 29(4), pp. 341-356, 2009.

[22] Goldratt, E. Y Cox, J., La Meta. Un proceso de mejora continua, Ediciones Castillo, México, D.F., 1995.

[23] Bayo, A., Bello, A. and Merino, J., 5S use in manufacturing plants: contextual factors and impact on operating performance. International Journal of Quality & Reliability Management, 27(2), pp. 217-230, 2010.

[24] Magnier, R., Getting ready for kaizen: organizational and knowledge management enablers. The Journal of Information and Knowledge Management Systems, 41(4), pp. 428-448, 2011.

[25] Kaur, M., Singh, K. and Singh, I., An evaluation of the synergic implementation of TQM and TPM paradigms on business performance. International Journal of Productivity and Performance Management, 62(1), pp. 66-84, 2013.

[26] Valles, M.S., Técnicas cualitativas de investigación social. Reflexión metodológica y práctica profesional, Editorial Síntesis, Madrid, 1997.

[27] Hernández, R. Y Fernández, C., Baptista, P., Metodología de la investigación. 4ª edición, McGraw-Hill, México, D.F., 2006.

COMPUTER-AIDED DIAGNOSIS OF BRAIN TUMORS USING IMAGE ENHANCEMENT AND FUZZY LOGIC

DIAGNÓSTICO ASISTIDO POR COMPUTADORA DE TUMORES CEREBRALES UTILIZANDO MEJORA DE LA IMAGEN Y LÓGICA DIFUSA

JEAN MARIE VIANNEY-KINANI
M.Sc., Instituto Politécnico Nacional de México,jeanmariekinani@yahoo.com.mx

ALBERTO J. ROSALES-SILVA
Ph.D., Instituto Politécnico Nacional de México, arosaless@ipn.mx

FRANCISCO J. GALLEGOS-FUNES
Ph.D., Instituto Politécnico Nacional de México, fgallegosf@ipn.mx

ALFONSO ARELLANO
Specialist in Neurosurgery, Instituto de Neurología y Neurocirugía de México, alfonso_arellano@yahoo.com

Received for review January 21 th, 2013, accepted August 12th, 2013, final version September, 10 th, 2013

ABSTRACT: A robust medical image processing system depends upon a variety of aspects, including a proper image enhancement, and an optimal segmentation. An algorithm was proposed in this paper to facilitate the implementation of these two steps. First a Magnetic Resonance (MR) image is enhanced via spatial domain filtering and its contrast is improved, next, the image is segmented using *fuzzy C-mean* clustering, then the region of interest which might be the tumor or edema, is detected and delineated. The key advantage of this image processing pipeline is the simultaneous use of features computed from the intensity properties of the image in a cascading pattern which makes the computation self-contained. Performance evaluation of the proposed algorithm was carried out on brain images from different MRI's and the algorithm proved to be successful, comparing it with other dedicated applications.

Key words: MRI, Region of interest, Segmentation, Clustering.

RESUMEN: Un sistema de procesamiento de imágenes médicas robusto depende de una variedad de aspectos, incluyendo una mejora apropiada de la imagen, y una segmentación óptima. En este artículo se propone un algoritmo para facilitar la implementación de estos dos pasos. En primer lugar, una imagen de resonancia magnética (RM) se mejora vía filtrado en el dominio espacial y también se mejora su contraste, luego, la imagen se segmenta utilizando el agrupamiento difuso "fuzzy C-means" (FCM), posteriormente, la región de interés, que puede ser el tumor o edema, se detecta y delinea. La ventaja clave de esta canalización de procesamiento de imagen es el uso simultáneo de características calculadas a partir de las propiedades de intensidad de la imagen en un patrón en cascada que hace que el cálculo sea auto-contenido. La evaluación del rendimiento del algoritmo propuesto se llevó a cabo en imágenes cerebrales de diferentes sistemas de resonancia magnética, el algoritmo desarrollado probó ser exitoso en comparación a otras aplicaciones relacionadas.

Palabras clave: IRM (Imagen de Resonancia Magnética), Región de interés, Segmentación, Algoritmo de agrupamiento.

1. INTRODUCTION

For every person diagnosed with brain tumor, their life expectancy decreases by 22 years on average. Due to their location, i.e., at the center of thought, physical function and emotion, brain tumors are difficult to diagnose and treat, and as a result, their death toll both in adults and children is highly disturbing. In many case, brain tumors, like any other type of cancer are a result of uncontrolled growth of brain cells, which may spread across the brain if not taken care of. For this reason, timely diagnosis and steady monitoring are imperative. As part of oncologic imaging, computer-aided diagnosis of these unhealthy tissues is not only of high interest in serial treatment monitoring of disease burden (i.e.,the impact this health problem may have in terms of financial cost, mortality, morbidity etc), but is also getting popular in many image guided surgical approaches. Using T1 weighted MR images which are obtained after administration of a contrast agent (gadolinium), our proposed algorithm, focuses on identifying brain structures, i.e., white matter, gray matter, cerebral spinal fluid, and then, it detects any

abnormal region that always stands out due to its intensity spectrum; and with limited user interaction, the algorithm uses predefined parameters quantifying this region of interest (ROI) intensity to delineate it with higher precision thus facilitating subsequent treatment processes. Most of medical imaging modalities obtain images of gray scale intensities, including MRI's. It turns out that these images have noise, artifacts, poor resolution and contrast due to instrument and reconstruction algorithm limitations or even patient movement; this makes auto diagnosis a challenging task, and the algorithm's advantages and disadvantages may vary depending on the properties of the image under examination. Due to the image deterioration factors mentioned above, it's hard to develop a standard approach capable of working with all MR brain images [1]. For this reason, tradeoffs have always been present in computer-aided diagnosis systems. However, comparing our fuzzy clustering-based method to other methods like classifier, region growing, neural networks, deformable model-based systems, a big advantage of our approach is recognized especially when it comes to dealing with the adverse factors mentioned [2].

2. METHODOLOGY

2.1 Preprocessing

Medical images are often deteriorated by noise due to various sources of interference and other phenomena that affect the measurement processes in imaging and data acquisition systems. The nature of the physiological system under investigation may also diminish the contrast and the visibility of details [3]. Thus, pre-processing helps in generating enhanced versions of the original image that demonstrate certain features in manner that is better in some sense as compared to their appearance in the original image. Depending on the nature and quality of the original image, different methods are used to enhance the visibility of the image, among them, filtering, histogram equalization, intensity scaling, compensation for nonlinear characteristics, etc. In the methodology proposed, spatial filtering and histogram equalization were adopted because they turned out to be effective in dealing with most of the problems that show up in the post-acquisition phase, and this is the right way to tackle the issue since other types of artifacts are best taken care of during the acquisition process [4].

2.1.1. Filtering

In the proposed algorithm, filtering was carried out in the spatial domain using a Laplacian Filter. The Laplacian operator is an example of a second order method of enhancement. The Laplacian proved to be good at finding the fine detail in an MRI image in a much better way compared to other spatial filters like the median, Gaussian, gradient or fuzzy filters, because it sharpens the image in a manner that portrays the heterogeneity of the image components in a much clearer view, and this is thanks to the fact that it is a second derivative operator that works successfully on heterogeneous images like MR images [5]. This filter is implemented by applying a Laplacian to an image *f(x, y)* as follows:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}, \quad (1)$$

Commonly used digital approximations of the second derivative are:

$$\nabla^2 f = \left[f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) \right] - 4f(x,y), \quad (2)$$

After this transformation we obtain a sharper image with a higher degree of heterogeneity better suited for further processing.

2.1.2. Histogram equalization

Another way to improve the quality of MR images was through the enhancement of their contrast. This is a very important aspect for subsequent processing phases, because it produces images that have the same contrast, hence ensuring a consistent response from the detection algorithm. This adjustment was carried out using histogram equalization technique, because it is fully automatic, can cover the entire gray scale which means there is no loss of information, and is based on information that can be extracted directly from the given image without the need for further parameter specification.

Suppose the intensity levels are continuous quantities normalized to the range [0, 1] and let $p_r(r)$ denote the probability density function (PDF) of the intensity levels in a given MRI image, where the subscript is used for differentiating between the PDFs of the input and output images[5]. Suppose that we perform the

following transformation on the input levels to obtain output (processed) intensity levels *s*,

$$s = T(r) = \int_0^r p_r(w)\,dw, \tag{3}$$

where *w* is a dummy variable of integration and *r* is in the range [0, L-1], we define L being the gray scale intensity levels, with *r*=0 representing black and *r*=L-1 representing white. It can be shown that the PDF of the output levels is uniform, that is,

$$p_s(s) = \begin{cases} 1 & for \quad 0 \le s \le 1 \\ 0 & otherwise \end{cases}. \tag{4}$$

In other words, the preceding transformation generates an image whose intensity levels are equally likely, and in addition covers the entire range [0, 1]. The net result of this intensity-level *equalization* process is an image with increased dynamic range, which will tend to have a higher contrast, and when applied to the MR image, the sharpness increases and our detection stage will be more efficient since we will be dealing with images whose contrast has increased.

2.2. Segmentation

The principal goal of the segmentation process is to partition an image into regions (also called classes or subsets) that are homogeneous with respect to one or more characteristics or features [3]. In medical imaging, the segmentation step is important for feature extraction, image measurement and display, classification of image pixels into anatomical regions or pathological regions such as cancer (tumor in our case), tissue deformities, among others. There are a number of methods used for this classification task, and each of them maybe selected depending on the application and the type of image to be dealt with. In our case, the most preferable was fuzzy classification due to reasons that will become clear in the next sections.

2.2.1. Fuzzy C-mean Clustering

One of the well-established concepts in image segmentation is pixel classification. This concept assumes that the pixels in each subclass (tissue) have nearly constant intensities, which is true for anatomical structures with similar physiological properties [3]. This applies to structures like White Matter (WM), Gray Matter (GM), or the Cerebral Spinal Fluid (CSF) that are found in the brain. This is also true for the unhealthy tissues like tumors because the cells that make up these tissues present similar physiological properties, thus uniformly responding to the B_O field and the Radio-frequency system of the MRI. And for this reason, they always stand out in the classification process, since their intensity spectrum differs from the rest of other tissues; in this regard, any tissues in the MR image whose intensity spectrum differs from the WM, GM and CSF spectra and is spatially found in one of these tissues' locations will always be treated as a region of interest that catches the doctor's attention. This explains why our diagnosis methodology seeks to sort out all the tissues that show up in the MR brain image using an adaptive classification method, that is, a method that is capable of estimating the centroid and bounds of each tissue in an adaptive fashion. It was proved that the right tool for this identification process would be fuzzy classification as a clustering technique since it classifies pixels in their respective clusters [6] bearing in mind the variability of gray value along with pixel statistical uncertainty due to the randomness [7].

We can now define a family of fuzzy partition matrices M_{fc}, for the classification involving *c* classes (clusters) and *n* data points (pixels):

$$M_{fc} = \left\{ \underset{\sim}{U} \mid \mu_{ik} \in [0,\ 1];\ \sum_{i=1}^{c} \mu_{ik} = 1;\ 0 < \sum_{k=1}^{n} \mu_{ik} < n \right\}, \tag{5}$$

where $\underset{\sim}{U} \in M_{fc}$ is a fuzzy *c*-partition, μ_{ik} is the membership value that the *k*-th pixel has in the *i*-th cluster, with *i*= 1, 2,…,*c*, and *k* = 1, 2,…,*n; c* being the number of clusters, and *n* the number of pixels, and μ_{ik} must have the following restrictions:

$$\mu_{ik} = \mu_{A_i}(x_k) \in [0,1], \tag{6}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \quad for\ all\ \ k = 1, 2, \ldots, n, \tag{7}$$

with *A* being a family of fuzzy sets, and *x*, the data sample. It follows from the overlapping character of the classes and the infinite number of membership values possible for describing class membership that

the cardinality of M_{fc} is also infinity, that is, $\eta_{M_{fc}} = \infty$ [8].

To describe a method to determine the fuzzy *c*-partition matrix $\underset{\sim}{\mathrm{U}}$ for grouping a collection of *n* data sets (i.e.,

n pixels that quantify the corresponding intensities of our MR image) into *c* classes (i.e., *c* brain tissues including our ROI), we define an energy function J_m for a fuzzy *c*-partition [8],

$$J_m\left(\underset{\sim}{\mathrm{U}}, \mathbf{v}\right) = \sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_{ik})^w (d_{ik})^2, \tag{8}$$

where

$$d_{ik} = d(\mathbf{x}_k - \mathbf{v}_i) = \left[\sum_{j=1}^{m}(x_{kj} - v_{ij})^2\right]^{1/2}. \tag{9}$$

with d_{ik} representing the Euclidean distance between the *i*-th cluster center and the *k*-th pixel (data point in *m* space or *k*-th pixel's intensity); and the parameter $\mathbf{v}$

is a vector of cluster centers. Another parameter introduced in (8) is *w*, called a *weighting parameter* whose value has a range $w \in [1, \infty)$. This parameter controls the amount of fuzziness in the classification process.

Furthermore, in (9) appears $\mathbf{v}_i$ which is the *i*-th cluster center, and is described by *m* features (*m* coordinates) that can be arranged in vector form, that is, $\mathbf{v}_i = \{v_{i1}, v_{i2}, ...v_{im}\}$.

Each of the cluster centers mentioned above can be calculated in the following manner:

$$v_{ij} = \frac{\sum_{k=1}^{n} x_{ik}(\mu_{ik})^w}{\sum_{k=1}^{n}(\mu_{ik})^w}, \tag{10}$$

where *j* is a variable on the feature space, that is, *j* = 1, 2, . . .,*m*. and *i, k* define the cluster and pixel number respectively.

It should be pointed out that the function J_m can have a large number of values, the smallest one associated with the *best* clustering or in our case, the best detection. Because of the large number of possible values (now infinite due to the infinite cardinality of fuzzy sets) we seek to find the best possible, or optimum, solution without resorting to an exhaustive, or expensive, search.

The optimum fuzzy *c*-partition will be the smallest of the partitions described in (8), that is,

$$J_m^*\left(\underset{\sim}{\mathrm{U}}^*, \mathbf{v}^*\right) = \min_{M_{fc}} J\left(\underset{\sim}{\mathrm{U}}, \mathbf{v}\right). \tag{11}$$

As with many optimization processes, the solution to (11) cannot be guaranteed to be a global optimum because of its fuzziness in nature. What we seek is the best solution available within a prespecified level of accuracy. To get to this, an *iterative optimization* algorithm proved to be the best option [8].The steps of this algorithm are as follows:

1. Fix c ($2 \le c < n$) and select a value for parameter *w* depending on the degree of fuzziness of the images to be processed ($w>1$). Initialize the partition matrix, $\underset{\sim}{\mathrm{U}}^{(0)}$. Each step in this algorithm will be labeled *r*,

   where r = 0, 1, 2, ….It should be noted that *c* was chosen to be equal to 4, referring to Cerebral Spinal fluid, White matter, Gray matter and region of interest, i.e. tumor/edema.

2. Calculate the *c* centers $\left\{\mathbf{v}_i^{(r)}\right\}$ for each step.

3. Update the partition matrix for the *r*-th step, $\underset{\sim}{\mathrm{U}}^{(r)}$, as follows:

$$\mu_{ik}^{(r+1)} = \left[\sum_{j=1}^{c}\left(\frac{d_{ik}^{(r)}}{d_{jk}^{(r)}}\right)^{2/w-1}\right]^{-1}, \quad \text{for} \quad \mathrm{I}_k = \varnothing, \tag{12}$$

or,

$$\mu_{ik}^{(r+1)} = 0, \tag{13}$$

for all classes *i* where $i \in \mathrm{I}_k$, with,

$$\mathrm{I}_k = \left\{ i \,|\, \left(2 \le c < n \right) \;;\; d_{ik}^{(r)} = 0 \right\}, \tag{14}$$

and,

$$\underset{\sim}{\mathrm{I}}_k = \{1,\ 2,\ ...,\ c\} - \mathrm{I}_k, \tag{15}$$

and,

$$\sum_{i\in \mathrm{I_k}} \mu_{ik}^{(r+1)} = 1 . \quad (16)$$

If,

$$\left\| \underset{\sim}{\mathrm{U}}^{(r+1)} - \underset{\sim}{\mathrm{U}}^{(r)} \right\| \le \varepsilon_L , \quad (17)$$

then stop, otherwise set r=r+1 and return to step 2.

The parameters $\underset{\sim}{\mathrm{U}}^{(r)}$ and $\underset{\sim}{\mathrm{U}}^{(r+1)}$ represent the partition matrix for the rth and the (r+1)th steps respectively, as it was mentioned in step 1 and 3 of the algorithm. And the parameter ε_L represents a prescribed level of accuracy that is used to determine whether the solution is good enough to stop the algorithm. This comparison of a matrix norm || of two successive fuzzy partitions to a prescribed level of accuracy, ε_L, is due to the restrictions of (11).

In step 3, there is a considerable amount of logic involved in (12)–(16). Equation (12) is straightforward enough, except when the variable d_{jk} is zero. Since this variable is in the denominator of a fraction, the operation is undefined mathematically, and computer calculations are abruptly halted. So the parameters I_k and $\underset{\sim}{\mathrm{I}}_k$ comprise a bookkeeping system to handle situations when some of the distance measures, d_{ij}, are zero, or extremely small in a computational sense. If a zero value is detected, (13) sets the membership for that partition value to be zero. Equations (14) and (15) describe the bookkeeping parameters I_k and $\underset{\sim}{\mathrm{I}}_k$, respectively, for each of the classes. Equation (16) simply says that all the nonzero partition elements in each column of the fuzzy classification partition, $\underset{\sim}{\mathrm{U}}$ sum to unity [9], [10].

2.2.2. Level set methods

Level set methods (LSM) seek to define the active contours of ROI through the evolution of a numerical level set equation, which is actually a Hamilton-Jacobi equation [11], [12]:

$$\begin{cases} \dfrac{\partial \phi}{\partial t} + F\left|\nabla \phi\right| = 0 \\ \phi(0,x,y) = \phi_0(x,y) \end{cases} , \quad (18)$$

$\phi_0(x,y)$ where F represents comprehensive forces, including internal forces from the interface geometry like mean curvature, contour length, area and external forces from image gradient. ϕ is the level set function and, $\phi_0(x,y)$ is the initial contour. A complete level set equation can actually be written as:

$$\frac{\partial \phi}{\partial t} = \mu\zeta(\phi) + \xi(g,\phi),$$

where

$$\zeta(\phi) = \Delta\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right), \quad ,$$

and

$$\xi(g,\phi) = \lambda\delta(\phi)div\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + \nu g\delta(\phi) \quad (19)$$

with,

$$g = \frac{1}{1+\left|\nabla(G_\sigma * I)^2\right|} . \quad (20)$$

Equation (19) is an edge indication function that regularizes F in order to stop level set evolution near the optimal solution [13]. The constants $\mu, \lambda,$ and ν control the individual contribution of the terms above and they are tuned using *trial by error* method, and they vary from case to case. In our experiment these values varied from (0.01 to 0.3), (0.1 to 7), (-2 to 2) respectively.

Basically, the terms $\xi(g,\phi)$ attracts the level set function towards the variation boundary, and $\zeta(\phi)$ forces the function to approach the genuine signed distance function automatically. And the Dirac function is computed as follows [14]:

$$\delta_\alpha(x) = \begin{cases} 0, & |x| > \alpha \\ \dfrac{1}{2\alpha}\left[1+\cos\left(\dfrac{\pi x}{\alpha}\right)\right] & |x| \le \alpha \end{cases} , \quad (21)$$

where α is a constant regulating the Dirac function and is tuned using *trial by error* method . The interface defined as $\Gamma(t)$ can be determined by tracking the values of level set function according to the conditions

$$\begin{cases} \phi(t,x,y) < 0 \ldots.(x,y), \text{ inside } \Gamma(t) \\ \phi(t,x,y) = 0 \ldots.(x,y), \text{ at } \Gamma(t) \\ \phi(t,x,y) > 0 \ldots.(x,y), \text{ outside } \Gamma(t) \end{cases} . \quad (22)$$

The steps of the proposed algorithm are shown in figure 1 below.

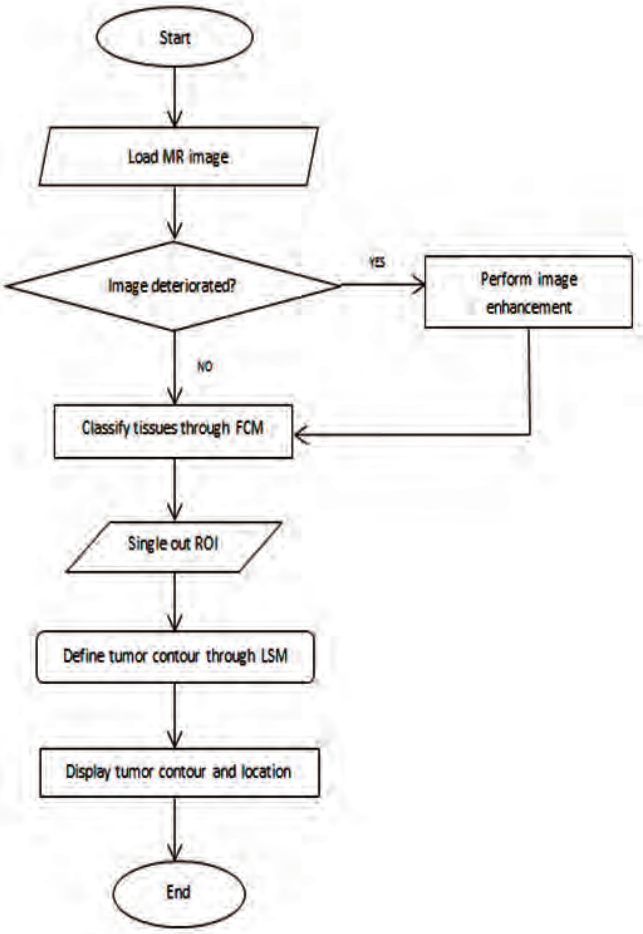


Figure 1. Flow diagram of the proposed computer-aided diagnosis algorithm.

3. EXPERIMENTAL RESULTS AND ANALYSIS

The experiments were carried out on MR brain images baring different resolutions varying between 1.5 to 4mm/pixel. The algorithm was implemented using MATLAB R2007b (Mathworks), on the Windows 7 operating system, and 3.00 GHz dual processor; its execution time was about 15 seconds which is quite acceptable. Another overwhelming advantage was the precision to track even the most negligible intensity change, which implies that our algorithm could carry out a quicker and more reliable diagnosis than a general clinician, especially in case of poor quality MR images because human eye is unable to trace these changes so easily.

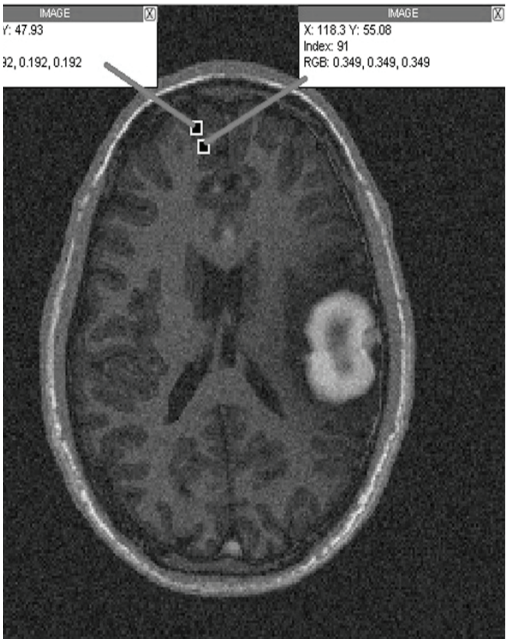


Figure 2. Patient 1.Original image.

As mentioned, the original image is preprocessed using the methods mentioned in Section 2, and then brain structures are categorized into white matter, gray matter, cerebral spinal fluid and the unhealthy tissue. Then, this tissue is selected and the algorithm works to define its contour. To carry out a much more reliable segmentation, it is necessary to find the region that has minimum energy, i.e., the region where the intensity shift is not that sharp. An example was taken and displayed in figure 2 where a gray matter region measuring 52 levels of intensity (index) is weighed against a white matter region in its vicinity measuring 91 levels of intensity. Taking into account the whole image's intensity spectrum, this intensity difference is not that big, and this is confirmed by the negligible difference of lightness of the two regions. Our fuzzy c-mean clustering should be able to segment this neighborhood into 2 different regions, so as to identify the expansion of the region of interest.

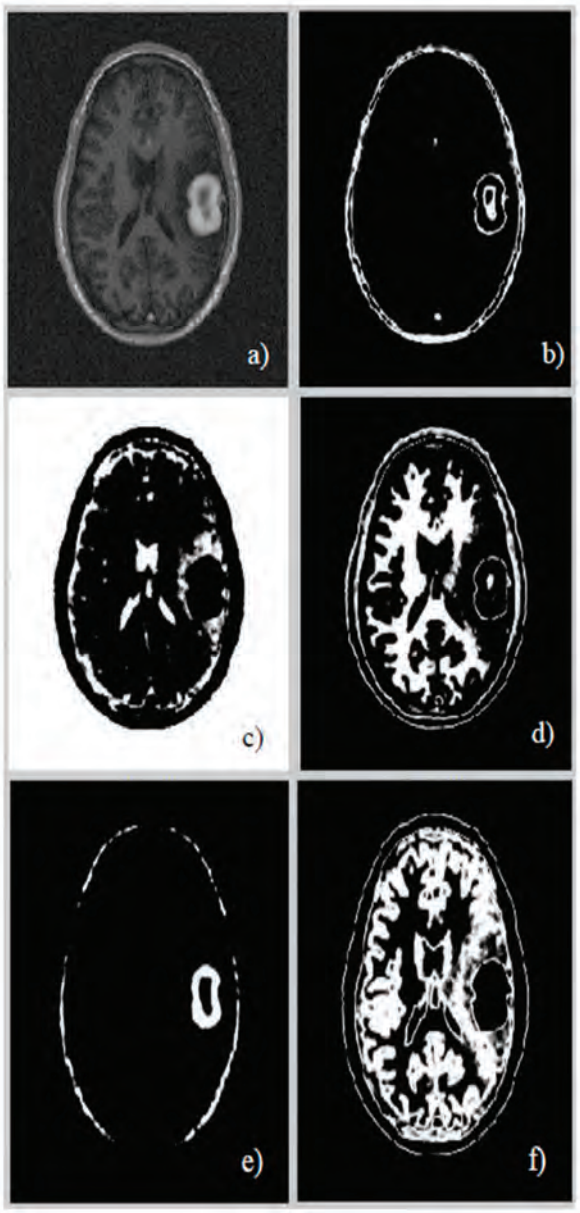


Figure 3. Processed images portraying FCM segmentation results of patient 1. (a) Original image, (b) brain tumor edges, (c) Cerebrospinal fluid (CSF) along with septum pellucidum, (d) White matter, (e) tumor area (f) gray matter.

As far as the tissues classification is concerned, one can see that in Figure 3, the fuzzy c-mean clustering routine takes in an MRI image (a) which was cleaned according to Section 2.1, processes it, using the classification scheme described in 2.2.1, then it outputs five different clusters (b)-(f) that correspond to five different types of regions, including the ROI. It is obvious that this fuzzy clustering phase is capable of detecting even very tiny shifts in intensity levels as portrayed on figure 2; this undoubtedly demonstrates the sensitivity of the system. This is followed by a ROI boundary definition that shows exactly the expansion of this region relative to the entire original image.

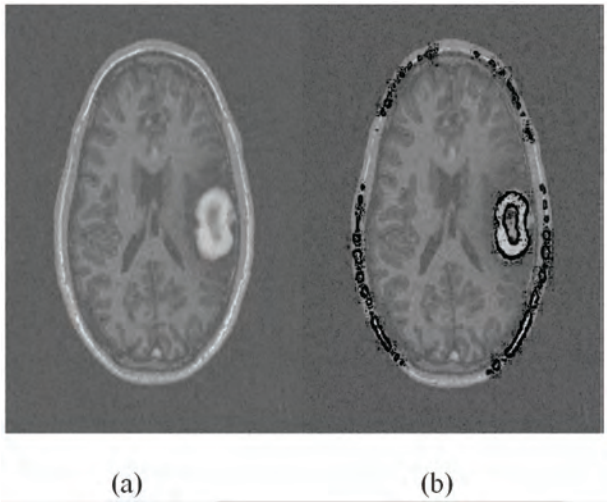


Figure 4. Comparison of the original image (a) with its corresponding processed image (b).

Figure 4 illustrates the final results (in green) of active contouring as described in equation (21). These results were obtained after 110 iterations. One can see that the image (b) portrays the contour of the tumor relative to the other tissues, and this delineation is obtained from the results of Figure 3(b) previously obtained from FCM segmentation.

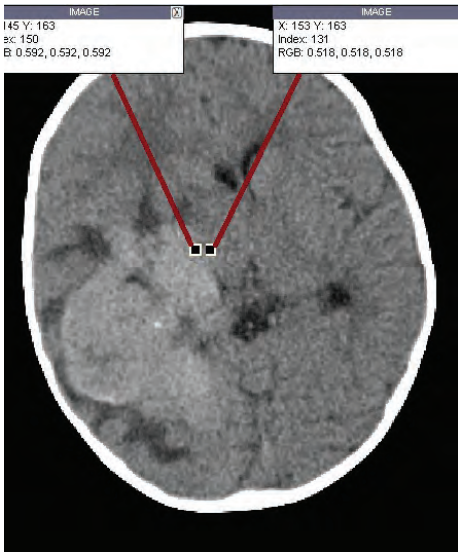


Figure 5. Patient 2. Original image

Figure 5 illustrates the intensity distribution of a different patient's MR image, with very tiny intensity

shift along the tumor edge. However, the FCM should detect this transition with accuracy.

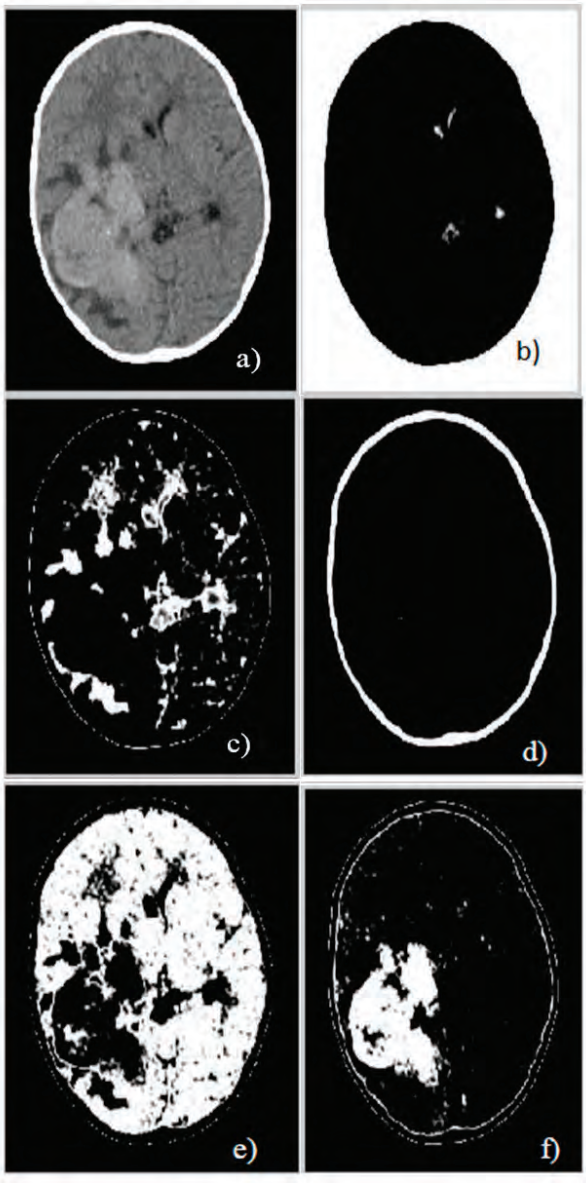


Figure 6. Patient 2 tissue classification, a) original image, b) septum pellucidum, c) white matter, d) cranium, e) gray matter along with edema, and f) tumor.

In figure 6, patient 2 who is in critical condition is diagnosed, as one can see his septum pellucidum was shifted upward (b) due to the expansion of edema that appears in (e). (c) Shows white matter tissues which were reduced because of the edema that spread over. Then (d) carries the cranium contour, whereas (e) portrays the gray matter intertwined with the edema. And (f) illustrates the tumor region. These results were obtained after FCM processing of figure 6(a) and, the specialist chooses his image of interest so that the level set method may give a final result of detection.

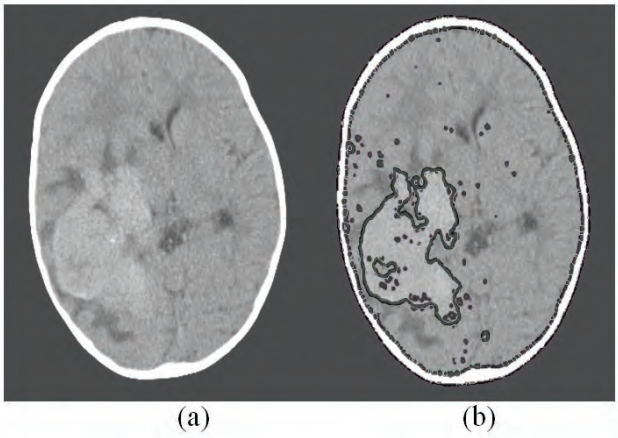


Figure 7. Patient 2 final results a) original image b) Tumor contour.

Figure 7 demonstrates two images, one being the input image, and the other one being the final processed image, where the tumor region is well outlined in spite of a cluttered neighborhood, and one can see clearly the high sensitivity of the system by comparing the two images. The tumor intensity spectrum is not that sharp, yet the algorithm can still trace its boundaries with high sensitivity.

Talking about the sensitivity, a quantitative evaluation of the performance and reliability of our system was carried out by numerically computing both the sensitivity and specificity using 75 images. Table 1, shows the results of the proposed method as compared to the neurologist's findings. The sensitivity of the system quantifies its ability to correctly identify subjects with the disease condition. In other words, it is the proportion of true positives (*TP*) that are correctly identified by the system, given by:

Table 1. Comparative results of the diagnostic test

Results of the proposed method	Results of Gold Standard test (as determined by the neurologist)		
	Positive	Negative	Row Total
Positive	32 (TP)	2 (FP)	34 (TP+FP)
Negative	1 (FN)	40 (TN)	41 (FN+TN)
Column Total	33 (TP+FN)	42 (FP+TN)	75 (TP+TN+FP+FN)

$$Sensitivity = \frac{TP}{TP + FN} \tag{23}$$

where *FN* represents the false negatives, and (*TP*+*FN*) represents the total number of subjects with the condition. On the other hand, the specificity is the ability of a system to correctly identify subjects without the condition. In other words, it is the proportion of true negatives (*TN*) that are correctly identified by the system:

$$Specificity = \frac{TN}{FP + TN} \tag{24}$$

where *FP* represents the false positives, and (*FP*+*TN*) represents total number of subjects without the condition. Based on the results presented in Table 1, the sensitivity and specificity of the system is 32/(32+1)=0.97 and 40/(2+40)=0.95 respectively.

Finally, another statistic that was used to measure the reliability of our system is the accuracy, which is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition, and is given by:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{25}$$

with (*TP*+*FP*+*FN*+*TN*) representing total number of subjects in study. Therefore the accuracy is (32+40)/(32+2+1+40)=0.96 in other words, 96% accurate.

4. CONCLUSION

The work mainly focused on the development of an algorithm that could promptly and accurately diagnose brain tumors even in their early stage of growth using both statistical and fuzzy methods. This was not an easy task due to the poor quality of some of the MR images; however, our preprocessing phase alleviated this issue significantly, thus facilitating the implementation of subsequent processing. A *fuzzy c-mean* clustering was implemented to classify different brain structures thereby allowing the detection of unhealthy tissues whose intensity span differed from the healthy ones i.e. white matter, gray matter, and cerebral spinal fluid. This method proved to be quite successful in the 75 cases studied; because it could accurately (96% of accuracy) process these images irrespective of their resolution. Using the controlling parameters generated by this clustering phase, a level set method could delineate these unhealthy regions, outlining their shape, position and expansion. This contour definition is normally the hardest task in manual segmentation due to changes in intensity as well as manual variations, (i.e., the manual segmentation's low degree of repeatability which is always present, even with the most meticulous surgeons). It is worth mentioning that an error of 4% i.e., (100-96% of accuracy) was due to artifacts generated by the brain tissues' heterogeneity, as it was proved by two false positives generated during the experiment.

5. ACKNOWLEDGEMENT

We would like to thank the Instituto Politécnico Nacional de México (National Polytechnic Institute of Mexico) for promoting our research, and the National Institute of Neurology and Neurosurgery of Mexico for providing us with medical guidance along with MR images, undoubtedly the support from these two institutions was vital in the accomplishment of this work.

REFERENCES

[1] Strickland, N. R., Image Processing Techniques for Tumor Detection. New York, Marcel Dekker, 2002.

[2] Bezdek, J. C., Hall, L. O. and Clarke, L. P., Review of MR Image Segmentation Techniques Using Pattern Recognition. Baltimore, Medical Physics, 1993.

[3] Bankman, N. I., Handbook of Medical Image Processing and Analysis. Burlington, Academic Press, 2009.

[4] Liang, Z., Principles of Magnetic Resonance Imaging; A Signal Processing Perspective. Bellingham, IEEE press, 2000.

[5] Gonzalez, C. R., Digital Image Processing. Upper Saddle River, Pearson, 2008.

[6] Bezdek, J. C., Pattern Recognition with Fuzzy Objective Function Algorithms. New York, Plenum Press, 1981.

[7] Tamalika, C. and Ajoy, K., Fuzzy Image Processing and Applications with MATLAB. Boca Raton, CRC Press, 2010.

[8] Bezdek, J., Keller, J., Pal, N. and Krishnapuram, R.,

Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. New York, Springer, 1999.

[9] Krishnapuram, R., Generation of Membership Functions via Possibilistic Clustering. Columbia, IEEE, 1994.

[10] Krishnapuram, R. and Keller, J. K., The Possibilistic C-Means Algorithm: Insights and Recommendations. Columbia, IEEE, 1996.

[11] Xu, C. and Prince, J. L., Snakes, Shapes and Gradient Vector Flow, IEEE Transaction on Image Processing 7. Baltimore, IEEE, 1998.

[12] http://www.cs.ubc.ca/~mitchell/ToolboxLS.

[13] Osher, S. and Fedkiw, R., Level Set Methods and Dynamic Implicit Surfaces. New York, Springer-Verlag, 2009.

[14] Li, C., Xu, C. and Andgui, C., Fox Level Set Evolution without Re-initialization, A New Variational Formulation in Proceedings of the 2005 IEEE Computer Vision and Pattern Recognition. Baltimore, IEEE, 2005.

FREE AND CONSTRAINED SINTERING OF 3-MOL % YTTRIA STABILISED ZIRCONIA

SINTERIZADO LIBRE Y CON RESTRICCIÓN DE ZIRCONIA PARCIALMENTE ESTABILIZADA CON ITRIA 3-MOL %

LUZ STELLA ARIAS-MAYA

M.Sc., Facultad de Ingeniería Mecánica, Universidad Tecnológica de Pereira, Pereira, Colombia, stellarias@hotmail.com

Received for review January 28 th, 2013, accepted July 29th, 2013, final version August, 20 th, 2013

ABSTRACT: Densification and shrinkage behaviour of ceramics or other powder materials can be predicted in a simple manner by using the master sintering curve. In this work, the densification data required to construct the master sintering curve of a 3-mol% yttria stabilised zirconia powder have been obtained. Bulk samples prepared by cold pressing, as well as thick films made of the same powder applied to a rigid substrate, were sintered. The free and constrained sintering experiments have been performed applying three different heating rates, with and without isothermal step at a maximum temperature of about 1450 °C. The shrinkage of the samples was measured *in situ* during densification using a push rod dilatometer. Also, a conventional box furnace was used, applying the same heating schedule, to compare the results. Microstructural analysis has been achieved by using scanning electron microscopy, environmental scanning electron microscopy, and optical microscopy. An acceptable agreement was obtained between similar samples sintered in the dilatometer and the box furnace, indicating that the dilatometer data could be used for both environments. The bulk samples reached lower densities than the films, probably due to their lower initial densities. For the samples free of constraint, a master sintering curve was achieved with an activation energy of 550 kJ mol^{-1}. For the constrained case, problems associated with the substrate and the uncertainties in the measurements hindered the characterisation of the film, i.e., a single master sintering curve could not be produced.

Key words: Yttria stabilised zirconia, Sintering, Master sintering curve, Restriction.

RESUMEN: La densificación y la contracción de cerámicos u otros materiales en polvo pueden ser predichos de una manera sencilla usando la curva maestra de sinterizado. En este trabajo se han obtenido los datos de densificación requeridos para construir la curva maestra de sinterizado de zirconia en polvo estabilizada con 3-mol% itria. Se sinterizaron muestras compactadas en frío, así como películas gruesas hechas del mismo polvo y aplicadas a un sustrato rígido. Las pruebas de sinterizado libre y con restricción fueron realizadas aplicando tres velocidades de calentamiento, incluyendo o no una etapa isotérmica a la máxima temperatura de ~1450 °C. La contracción de las muestras se midió durante la densificación usando un dilatómetro de varilla de presión. También se usó un horno convencional, aplicando la misma curva de calentamiento para comparar resultados. El análisis microestructural fue logrado mediante microscopía de barrido electrónico, microscopía de barrido electrónico ambiental y microscopía óptica. Se obtuvo una concordancia aceptable entre muestras similares sinterizadas en el dilatómetro y en el horno convencional, indicando que los datos del dilatómetro pueden ser usados para ambos ambientes. Las muestras que se sinterizaron sin restricción alcanzaron menores densidades que las películas, probablemente debido a que sus densidades iniciales eran menores. Para las muestras libres de restricción, se obtuvo una curva maestra de sinterizado con una energía de activación de 550 kJ mol^{-1}. Para las películas sinterizadas con restricción, algunos problemas asociados con el sustrato y con las incertidumbres de las mediciones impidieron la caracterización de la película, es decir, la obtención de una curva maestra de sinterizado con un solo valor de energía de activación.

Palabras clave: Zirconia estabilizada con ittria, Sinterizado, Curva maestra, Restricción.

1. INTRODUCTION

Sintering is a process by which crystalline powder particles reduce their free surface energy. The reduction of the total surface energy and grain boundary leads to densification and shrinkage of the body, due to the moving of atoms. The mass transport is possible over the free surfaces of particles and over the grain boundaries that form between the adjacent ones.

The finished ceramic piece's characteristics are affected by many variables and facts involved in its manufacturing. These characteristics are also affected by the properties of the powders and the additives used to prepare the piece. Due to this, it is difficult to predict the ceramic powder kinetics and behaviour during sintering, as well as the dimensions, properties and final appearance of the piece. Manufacturers usually have to use trial and error to determine the effects of sintering, and they have to adopt the composition and

appropriate heating routes. In consequence, the time and production costs increase.

For this reason, it is helpful for the sintering process to have tools that facilitate predictions of the piece change in size. The Master Sintering Curve (MSC) is a practical model developed to predict behaviour and shrinkage of metals and ceramics during sintering of compacted powders, regardless of the heating path. The MSC is obtained from data from various experiments, and it is valid for powders with the same microstructure and manufacturing process. Therefore, the model is vulnerable to the powder composition and physical characteristics, as well as the process conditions [1].

The MSC can also be used for comparing different powders, the additive effects and their ease of sintering [2]. In addition, it allows establishing the main sintering mechanism, through the activation energy [3]. The concept has been applied successfully to different systems free of constraint.

The axes of the MSC are relative density (RD) and the work of sintering (Θ), which depends on the sintering temperature (T) and time (t), representing the dominant diffusion mechanism. This dependent parameter is shown in Eq. 1, where Q_b is the activation energy for grain boundary diffusion and R is the gas constant.

$$\Theta = \int \frac{1}{T} \exp\left(-\frac{Q_b}{\mathrm{R}T}\right) dt \qquad (1)$$

Eq. (1) defines de MSC, where there is no assumption about the relationship between time and temperature [4]. All the terms affected by microstructure are on one side of the equation, and the time and temperature terms on the other side. Both sides are related experimentally [1,3]. The model is based on the linear contraction given in a concept developed for Hansen et al. [5], where the sintering process is given in one stage, instead of the traditional three stages. The model of Su and Johnson [1,3] assumes that the main diffusion mechanism is grain boundary diffusion.

The construction of a MSC requires the estimation of the activation energy (Q). The appropriate value can be obtained with the data gathered in a dilatometer using different heating rates. An initial value for Q is assumed, and then curves of density versus Θ, the work of sintering, are plotted for each heating rate. The curves must converge to a unique activation energy in a sigmoid curve [6].

Ceramic materials are widely investigated (e.g., [7-16]); in particular, zirconia is a ceramic recognized by its multiple applications in engineering, due to their exceptional properties, such as high hardness, toughness, high oxygen diffusivity and low thermal conductivity [12]. Yttria stabilized zirconia (YSZ) presents high resistance and toughness, high temperature resistance and also conducts oxygen ions at 800 °C or more. These qualities make YSZ a good candidate for applications such as electrolyte in solid oxide fuel cells (SOFCs) [11-14].

In SOFCs, a thin layer of YSZ is deposited as electrolyte on a rigid substrate. The film is sintered with the substrate, which impedes the free contraction in the film plane. Consequently, this contracts only in the direction perpendicular to the plane. This means that just the thickness dimension changes during sintering. This particular case is known as constrained sintering, where the densification process of the sample is affected due to the resulting stresses.

This work evaluates the possibility of obtaining a MSC for zirconia powders stabilized with yttria (YSZ) at 3% mol without restriction and of extending it to the case in which sintering is performed with restriction. Additionally, this work aims at comparing data taken from a conventional oven with the data given by a dilatometer, taking into account the difficulties to collect data mainly in the case where restriction is present. Some comparisons are possible through microstructural analysis.

2. METHODOLOGY

2.1. Samples

Two groups of samples were prepared for this study using 3-mol % yttria stabilized zirconia powders (Melox Chemicals 3Y XZ07078/14, batch number 04/316/01M). The powder and the screen-printed specimens were supplied by Rolls-Royce Fuel Cell Systems Ltd. More details about the powder characteristics, the formulation of the ink and the printing process can be found in [7]. Two groups of samples were prepared with the objective of analyzing the behaviour of the material during free

and constrained sintering, the latter one, when a layer of the material is placed on a rigid substrate.

2.2. Samples for free sintering

The ceramic powder was blended with 23 ± 2 wt% (standard deviation) polyvinyl alcohol (PVA) as binder. Each blend was used to prepare two green compacts with diameters of 20 mm and 2 mm height (Fig. 1). The discs were made in a Specac hydraulic uniaxial pressing machine at a pressure of 31 MPa. Each sample was cut in three parts; one of them was used in the dilatometer and at least other in the box furnace. The samples dimensions were measured using digital callipers with a precision of ±0.01 mm.

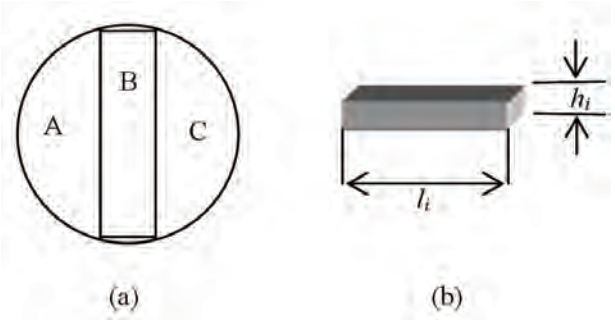


Figure 1. Schematic representation of the free sintering samples. (a) Sectioned disc (b) Sample for the dilatometer (the variables l_i and h_i correspond to the initial length and height, respectively)

2.3. Samples for constrained sintering

Each sample has a planar design. The electrolyte thick-film layers were placed using a Smetech Benchmark 90 screen-printing machine. The squeegee applied a force of 4.5 kgf to the ink during the printing operation. The mesh size was 165 apertures per inch with a nominal thickness of 100 μm ± 3μm and an open area of 44%. The electrolyte layer was formed as a laminate, with three applications to reduce the presence of defects. The substrate was a flat sided tube made from porous magnesia-magnesium aluminate spinel. Between the electrolyte and the substrate, a priming layer of about 30 μm was applied and pre-fired to smooth the surface. The electrolyte layer had a thickness of about 24 μm [7].

The samples were supplied as squares with 50 mm sides and 5 mm height. From them, smaller samples were cut with dimensions of 6×6×5 mm^3 for dilatometry and 6×10×5 mm^3 for the conventional oven (Fig. 2). Contamination of the samples was avoided, and then each cut was made with a diamond wheel without lubricant.

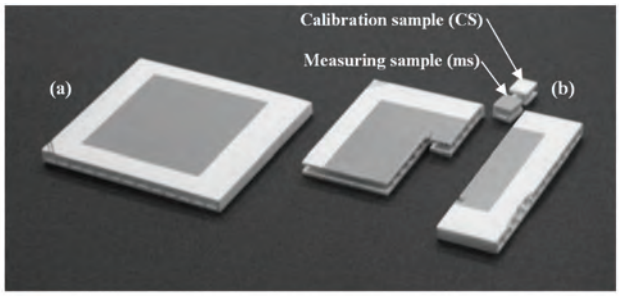


Figure 2. Samples for constrained sintering. (a) As supplied (b) Cut samples. The sample for the calibration does not contain the film under study

It was necessary to use two kinds of samples in the dilatometer (Fig. 2 (b)). The first one, named calibration sample, CS, lacked the electrolyte film and was used for performing the calibration of the dilatometer against which to measure the changes of thickness of the layer in study. The second type did include the electrolyte layer.

2.4. Sintering

Two ovens were used in this work: a conventional oven, Elite Thermal System Model BRF17/5-I-M with a Eurotherm 2416 programmer, and a dilatometer, Netzsch DIL 402C/41G with an "S" sample thermocouple and a pushrod that applies a force of only 30 cN to the sample. The pushrod transmits the signal to a Linear Variable Displacement Transducer (LVDT). The sample holder and calibration sample for the free sintering experiments are made of alumina. The gases released from the sample were purged using Argon at 44 ml min^{-1}.

In the dilatometer, it was necessary to correct the dimension changes in the system. For this purpose, a calibration sample (CS) was used. In the case of free sintering, the CS (alumina) was supplied by the Netzsch group. For constrained sintering the sample without the priming and electrolyte layers was used.

The samples were fired both in the dilatometer and the box furnace, which were programmed with the same schedule. This was done with the aim of comparing the results. Six sintering programs were performed, which are shown in

Fig. 3. The first step (1 °C min^{-1}) was the same for all the incursions and was used to burn out the binder and other organic additives in the samples (if higher heating rates are used, the vapours generated from the binders may damage the samples). For both free and constrained sintering, the calibration run had the isothermal step for each heating rate (3, 5, and 7 °C min^{-1}). Different heating rates and dwell times were used so as to obtain a generic master sintering curve, and were selected based on a previous work [7].

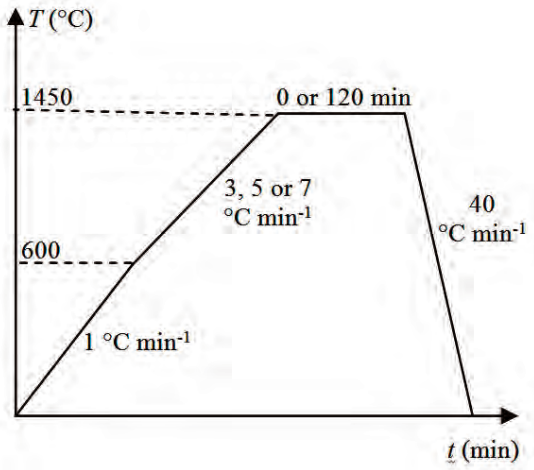


Figure 3. Sintering Schedule

2.5. Density

The final density of the free samples was measured by means of a digital balance, whose precision was of ±0.0001 g. This balance allowed the use of the Archimedes method. Deionised water was chosen as the fluid for the measurements. The samples were covered with a thin layer of nail polish, with the aim of avoiding water infiltration. The ASTM Standard test Method for Bulk Density and Volume of Solid Refractories by Wax Immersion (ASTM C914-95) [17] has been followed. The densities were given by the instrument with two decimal places, including the fluid temperature corrections.

Also, the final density for the samples sintered in the dilatometer was calculated assuming uniform contraction and constant mass:

$$\rho = \rho_i \left(1 - \Delta l / l_i\right)^{-3} \quad (2)$$

where, ρ_i is the initial density, $\Delta l/l_i$ is the linear contraction and Δl is the value given by the dilatometer. The final value was also measured by manual methods. The estimated data were compared with those measured before sintering to validate the results.

The density was determined from the surface of the electrolytic layer following the linear intercept method from the British standard EN 623-3:2001 method B [18]. The micrographs were obtained by means of a scanning electron microscope Hitachi S-3200N. For estimating the porosity of the samples, at least 100 pores were measured through lines drawn on the micrographs. The relative density is given by:

$$\mathrm{RD} = (1 - p_f) \times 100\% \quad (3)$$

where p_f is the final mean volume fraction of porosity:

$$p_f = \frac{\text{Total length of measured pores}}{\text{total length of line}} \quad (4)$$

Assuming constant area in the constrained layer, it is possible to find its initial density from the final density value and the relative change of length in the perpendicular direction to the layer ($\Delta l/l_{ic}$), being l_{ic} the initial thickness of the layer.

$$\frac{\rho}{\rho_i} = \frac{1}{1 - \dfrac{\Delta l}{l_{ic}}} \quad (5)$$

2.6. Layer thickness for the constrained films

The thickness of the electrolyte layer was measured before (l_{ic}) and after (l_{fc}) sintering. Images from the green layer were obtained by scanning electron microscopy (SEM) in both vacuum and variable pressure (ESEM) modes. Before using SEM, the samples were covered with gold (~20 nm) using a Turbo Sputter Coater (Emitech K575X High Vacuum). In this way, the charging effect due to poor conductivity of the samples is reduced. The ESEM does not require covering the samples, minimizing the risk of contaminating the samples before sintering.

Additionally, optical microscopy, using a digital camera, was used for capturing images from the films sintered with the isothermal step for 2 h at the maximum temperature. The mean of the thicknesses of the layers after sintering was estimated from the images. Sample preparation for light microscopy included mounting them in Epofix resin mixed with alumina powder. After settling each sample, they were polished, following the adequate procedure.

Using the mean of the final layer thickness (l_{fc}), and the change of length of the film measured by the dilatometer (including corrections), Δl_{dil}, the initial layer thickness, l_{ic}, was estimated:

$$l_{ic} = l_{fc} + \Delta l_{dil} \quad (6)$$

The average of the results for l_{ic} was considered as the general initial layer thickness assuming the same thickness for all of them.

3. RESULTS AND DISCUSSION

3.1. Free sintering

3.1.1. Density

The cold pressed sample initial densities are shown in Table 1. These were obtained from the data taken from the green bodies and were compared with the values estimated from the measurements after sintering. This fact proved the consistency in the sample fabrication process, which is important in finding an adequate MSC.

In general, the results are acceptable, taking into account that the samples could continue the densification process during cooling; mainly those that did not have a dwell step in their schedule. This affects the manual measurements, in the way that the RD tends to be slightly higher than the last density values given by the dilatometer. In addition, the RD tends to decrease with increasing heating rate, as it has been reported in the literature (e.g., [2,7]).

Table 1. Initial density for the free sintering samples. The standard error (e) is shown

Rate of heating (°C min^{-1})	3		5		7	
Isothermal time (h)	0	2	0	2	0	2
ρ_i (g cm^{-3}) (e = ±0.03)	2.12	2.23	2.04	2.04	2.02	2.14
Initial RD (%) (e = ±1%)	35	37	34	34	33	35

In a similar way, the results from the samples sintered in the dilatometer were compared with those densified in the conventional furnace. The results were consistent, which suggests that the process is similar in both cases. This consistency is important as, when the constrained samples are considered, there will be greater experimental difficulties in terms of measuring, e.g., the relative densities. The results from free sintering give confidence in the techniques.

3.1.2. Density and shrinkage curves

Figs. 4 and 5 show the relative change of length and relative density against time, given by the dilatometer. In the figures, "V3", e.g., means a heating rate of 3 °C min^{-1} and 2h corresponds to 2 h of soaking at T_{max}. For each heating program, with or without an isothermal step, both samples follow the same trajectory, as expected.

3.2. Constrained sintering

3.2.1. Layer thickness

Fig. 6 shows a micrograph of the electrolyte and priming layer, before sintering, that was obtained by ESEM. This type of micrograph was used for measuring the initial thickness of the YSZ film, even though, subsequent analysis suggested that this can have borders produced when the samples were cut, which could affect the results. Therefore, the initial thickness was estimated from the final thickness, as described below.

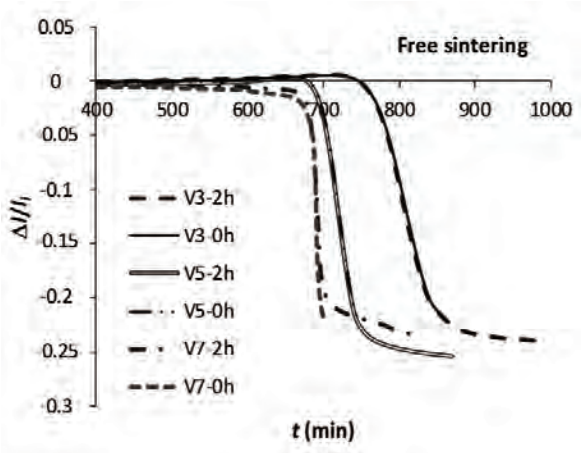


Figure 4. Relative shrinkage versus time for the dilatometer free sintering samples

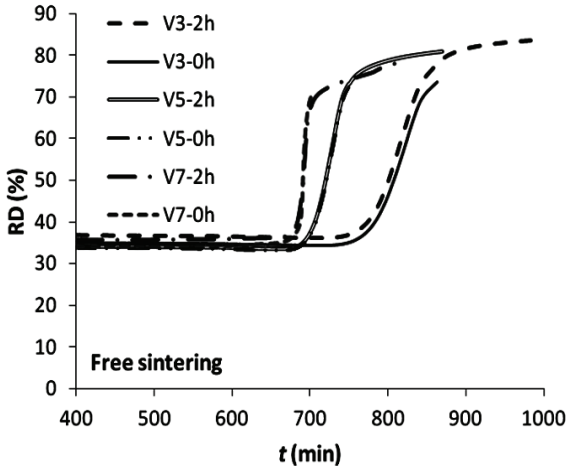


Figure 5. Relative density versus time for the dilatometer free sintering samples

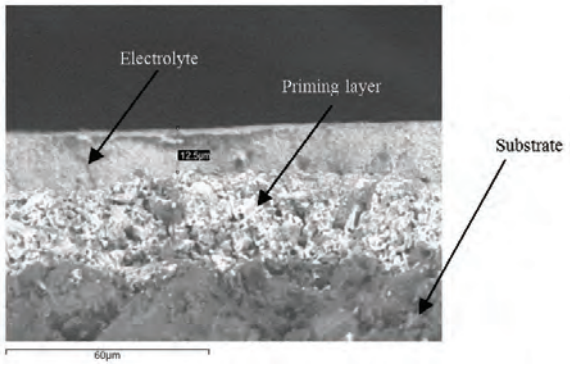


Figure 6. ESEM micrograph of the printed layer before sintering

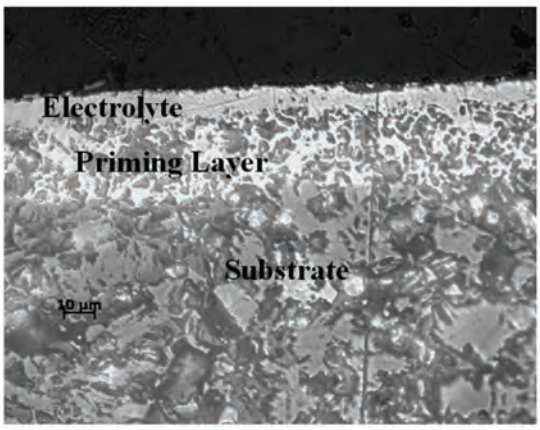


Figure 7. Cross section of a sample sintered in the dilatometer at 7 °C min^{-1} with 2 h holding at T_{max}

The cross section of one sample sintered with an isothermal step (two hours at about 1450 °C) is shown in Fig. 7. The image was achieved by light microscopy. The final thickness results are given in Fig. 8. There is an acceptable agreement in the data obtained for the samples sintered in both the dilatometer and the conventional oven. The bars correspond to the standard deviation.

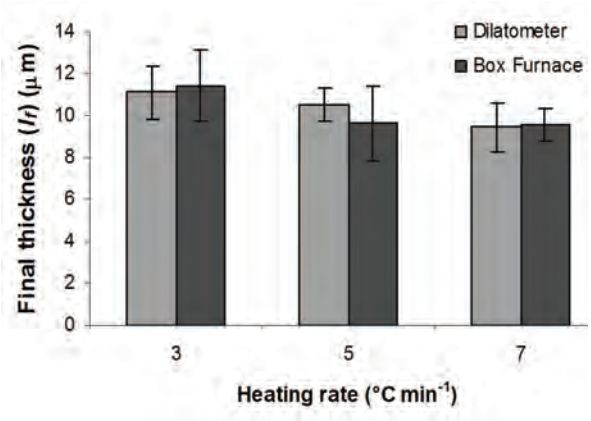


Figure 8. Thickness of the electrolyte layer after sintering for the samples in the box furnace and the dilatometer, with two hour dwell. Error bars indicate one standard deviation

The mean value for the final layer thickness (l_{fc}) and the change of length given by the dilatometer (Δl_{dil}) were used to estimate the thickness of the green layer (Eq. (6)). In this way, an average of 23 μm with standard error e = 0.2 μm (Table 2) was found. This value is in acceptable agreement with that reported by the manufacturer (24 μm).

3.2.2. Shrinkage of the film

Assuming 23 μm for the initial thickness of the samples, the shrinkage of the specimens sintered with soaking time in the box furnace was calculated using the data obtained by microscopy. The corresponding information is shown in Fig. 9. The results for the samples sintered in the dilatometer are similar to those sintered in the conventional oven. The estimated values by light microscopy are reported with the standard error. The average of the sintered film among the samples sintered in the dilatometer and the box furnace is 10 ± 1 μm. This value is similar to the sintered thickness reported by Wright [7]. The congruence in the results suggests confidence in the dilatometer measuring.

Table 2. Electrolyte thickness after and before densification for the samples with soaking time at 2 h a 1450 °C in the dilatometer

Heating rate (°C min^{-1})	Δl_{dil} (μm)	l_{fc} (μm)	l_{ic} (μm)
3	11.57	11.1 ± 1.0	22.7
5	12.94	10.7 ± 0.8	23.5
7	13.83	9.4 ± 1.0	23.3

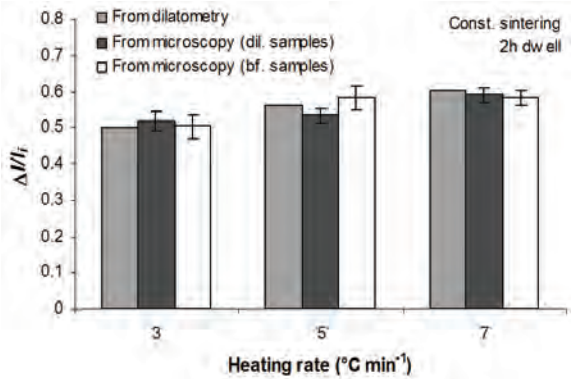


Figure 9. Relative change of layer thickness of the samples sintered with 2 h dwell, estimated assuming l_{ic} = 23 μm. The error bars correspond to one standard error. The dilatometer data do not present error bars because there was just one measurement

Assuming contraction of the layer only in the perpendicular direction to the substrate plane and with the data provided for the dilatometer, linear shrinkage versus time curves were built. The corresponding curves are given in Fig. 10.

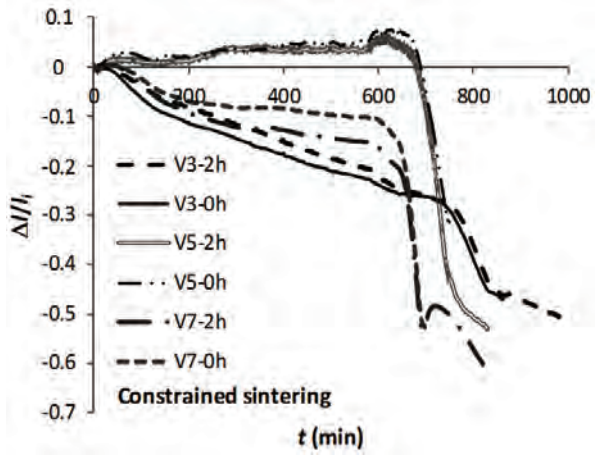


Figure 10. Relative change of thickness of the electrolyte film during sintering

The sintered samples with 5 °C min^{-1} of heating rate expanded during the first heating period and then the contraction begun. The other samples show contraction throughout the heating treatment. According to some observations the base layer may have suffered densification or the substrate was not sufficiently homogeneous. The latter could have affected the results.

3.2.3. Density

The relative densities are compared in Fig. 11. The standard deviation bars are longer for the samples with an isothermal step. This can be the result of the size, distribution, and shape variations of the microstructure, because at this degree of densification, the pores are just starting to link to each other, forming nets, and are taking a spherical shape. It means that there is not a homogeneous structure.

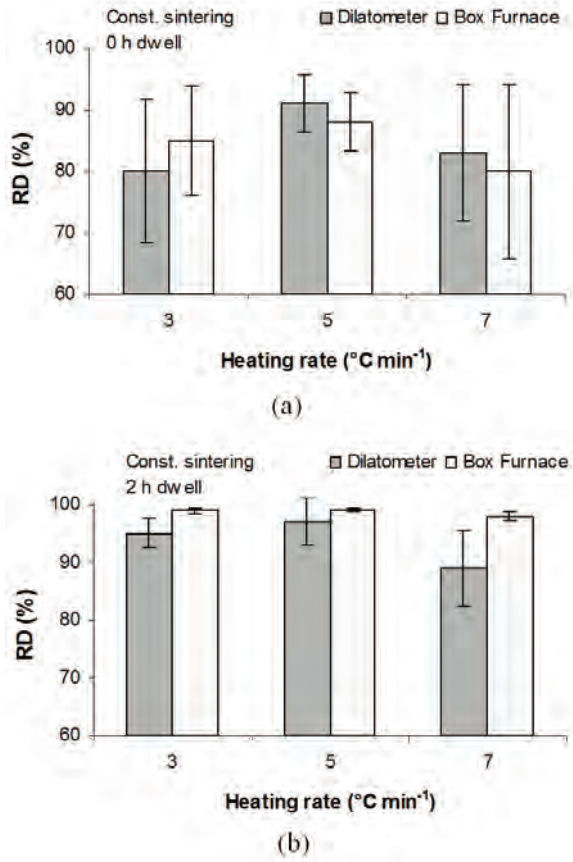


Figure 11. Relative density of the electrolyte layer sintered (a) without dwell, (b) with dwell. The error bars indicate one standard deviation

Even though the densities tend to be slightly higher for the samples from the conventional oven, the results are similar. A possible explanation is that in this furnace, the cooling step was bigger, leading to further densification and larger grain size. Additionally, the linear intercept method for estimating the density of the layer surface presents some difficulties. Furthermore, the restriction impedes to some degree the elimination of pores in a higher level in some zones [19] and hinders their reaccommodation. Similarly, internal stresses might contribute to the heterogeneous densification throughout the film [20].

In Fig. 12 the curves of relative density versus time are plotted. The uncertainty in the measurement of final density, taken from the micrographs, the heterogeneity in the layer, and the possible densification of the priming layer can be possible causes of the lack of convergence in the curves and for the variations in the initial density values.

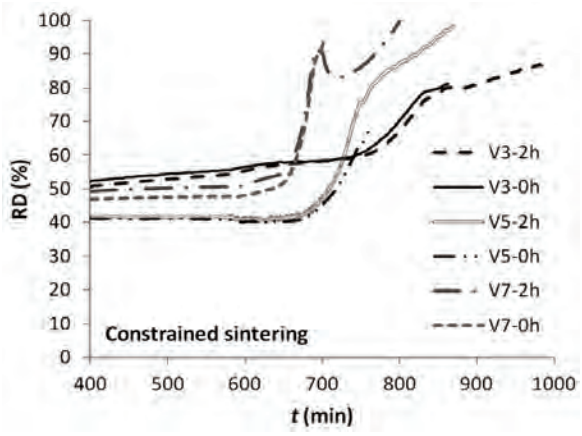


Figure 12. Variation of the relative density with time of the constrained layer sintered in the dilatometer at 3, 5, and 7 °C min^{-1} with zero and 2 h dwell

3.3. Master sintering curves

The data collected were used for trying to obtain the MSC. The curves were developed by Dr. Ruoyyu Huang from Leicester University, following the process given by [21], based on finite elements. The work of sintering (Θ) is determined by numerical integration and the activation energy is obtained when the curves converge.

The curve for the bulk samples is shown in Fig. 13. The data converge to a single curve, with an activation energy of 550 kJ mol^{-1}. This means that the densification kinetics is common to all the samples. There is a slight deviation for the rate of 7 °C min^{-1}, mainly above 72% relative density, at the threshold of the isothermal step. This tendency may be due to some fact that interferes with the densification kinetics and the activation energy for this particular powder.

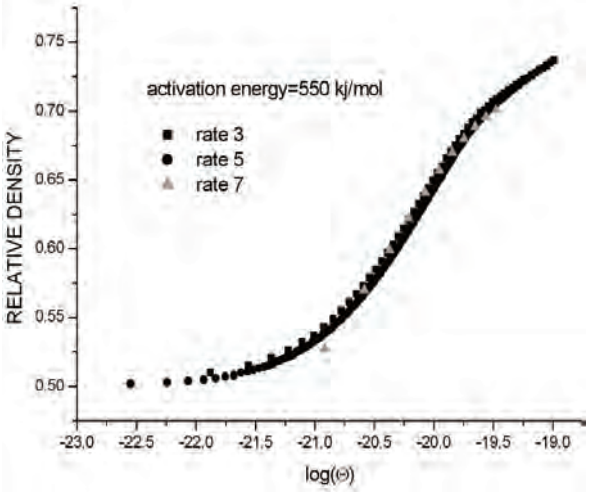


Figure 13. Master sintering curve for the 3-mol% YSZ free of constraint and no dwell

For the constrained layer, there was not a satisfactory convergence towards a universal curve during the process of densification. Two possible causes for this could be: there is no universal curve for this case, or the data are not reliable enough to produce the MSC. Given the problems related to the difficulty of subtracting the effect of the substrate, the latter reason must be eliminated by more experimentation before it could be concluded that a MSC cannot be achieved for this particular case.

3.4. Comparisons for free and constrained sintering

Apart from the master curves for both sintering cases (with and without restriction), it is also possible to compare the data, analyzing the relative density, as presented in Fig. 14.

The relative densities are, in general, higher for the samples sintered on the substrate. In previous works (e.g., in [22]), it has been found that the volumetric contraction tends to be lower for screen printed electrolytes than for bulk samples. However, in this

work, the initial densities of the free sintered samples were lower than those of the printed layers. The average relative density for the former was 35 ± 1%, while its counterpart for the latter samples was 43 ± 4%. This contributes to the differences in this study.

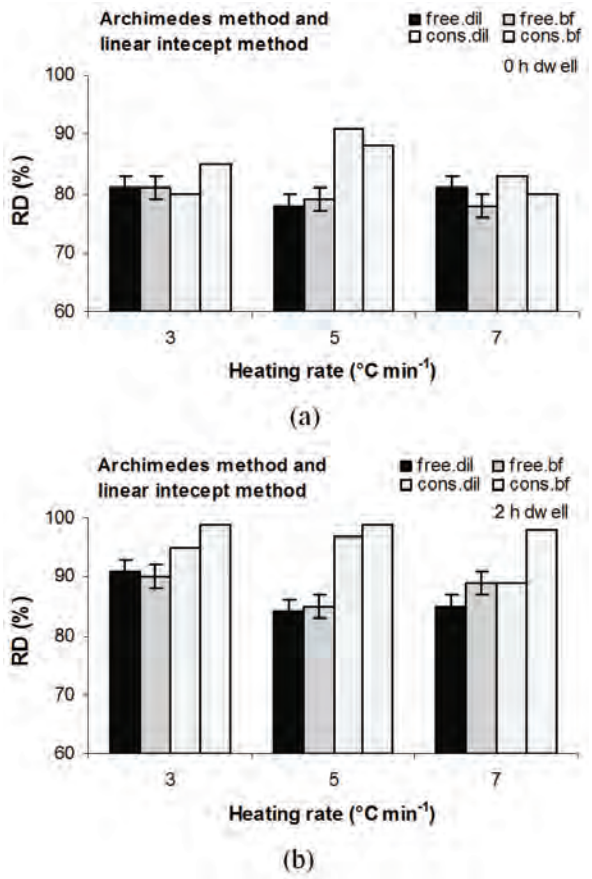


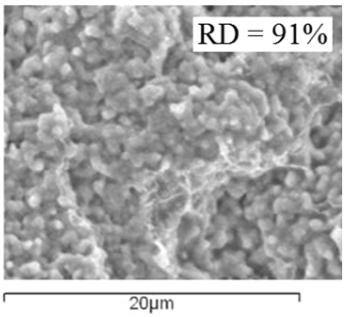
Figure 14. Relative density comparisons for the free and constrained samples sintered in the box furnace and in the dilatometer. (a) No dwell. (b) With dwell

Figure 15 provides examples of the bulk and electrolytic layer microstructures.

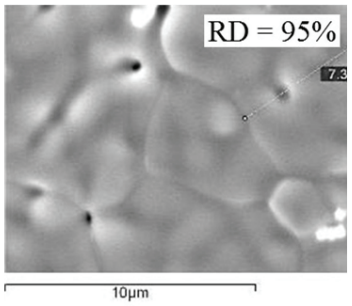
4. CONCLUSIONS

This work reported the results of experiments with 3-mol% yttria stabilized zirconia. The material was sintered free of constraint, as well as with restriction by applying it to a rigid substrate. Three different heating rates were used: 3, 5, and 7 °C min^{-1}. Half of the samples underwent an isothermal step at ~1450 °C; for the other half, this step was omitted. The samples relative length changes were measured in situ by a dilatometer.

For improving reliability, some experiments were performed using a conventional oven. The samples reached the same level of densification independently of the source utilized for sintering.



(a) Fractured surface - 3-mol% YSZ free of constraint



(b) Surface for the constrained samples

Figure 15. SEM micrographs; 3 °C min^{-1} with 2-h dwell in the dilatometer

A MSC was obtained for free sintering with an activation energy of 550 kJ mol^{-1}. This curve can be very useful for predicting the tendencies and behaviour during densification of this specific compacted powder.

In the case of the layer printed on a substrate, the latter gives some restriction to the stabilized zirconia film. In some of the SEM micrographs taken from the surface, non homogeneity in the porosity and accommodation of the microstructure was noticed, mainly at the intermediate rate of heating used in this work. This lack of homogeneity increases the uncertainty in the measuring. Other factors, such as the difficulty in subtracting the data from the layer itself and the contribution of the substrate to the dimensional changes, are considered as possible causes for the data not converging to a single activation energy and finding a unique curve.

REFERENCES

[1] Su, H. and Johnson, L., Sintering of alumina in microwave-induced oxygen plasma, J. Am. Ceramics Soc., 79(12), pp. 3199-3210, 1996.

[2] Kutty, T.R.G., Khan, K.B., Hegde, P.V., Banergee, J., Sengupta, A.K., Majumdar, S. and Kamath, H.S., Development of a master sintering curve for ThO_2, J. Nuclear Mat., 327, pp. 211-219, 2004.

[3] Su, H. and Johnson, L., Master sintering curve: a practical approach to sintering, J. Am. Ceramics Soc., 79(12), pp. 3211-3217, 1996.

[4] Johnson, D.L., Finding and utilising the master sintering curve. Proceedings of Sintering 03 - An International Conference on the Science, Technology and Applications of Sintering. Pennsylvania, 15-7, September, pp. 15-17, 2003.

[5] Hansen, J.D., Rusin, R.P., Teng, M. and Jonhson, D.L., Combined-stage sintering model, J. Am. Ceramics Soc., 75(5), pp. 1129-1135, 1992.

[6] Teng, M., Lai, Ch. and Chen, Y., A computer program of master sintering curve model to accurately predict sintering results, Western Pacific Earth Sciences, 2(2), pp. 171-180, 2002.

[7] Wright, G.J., Constrained sintering of yttria-doped zirconia: the relationship between microstructure and gas flow rate [PhD Thesis]. Guildford, UK: University of Surrey, 2006.

[8] García, C., García, C. and Paucar, C., Formación insitu de circonia en la síntesis de sustitutos óseos basados en matriz de alúminacircón infiltrada con hidroxiapatita, Dyna, 77(162), pp. 143-149, 2010.

[9] Ribero, D., Restrepo, R., Paucar, C. and García, C., Disminución de la temperatura en la síntesis de un material cerámico altamente refractario (mullita) a partir de hidroxihidrogeles, Dyna, 74(153), pp. 95-100, 2007.

[10] Mendoza, E.J., Zamora, L.A. and Galvis, F.V., Influencia del contenido de óxido de magnesio y del tratamiento térmico en la disolución de Ca+2 de un material cerámico bioactivo del tipo 31SiO2–11P2O5–(58-x)CaO–xMgO, Dyna, 73(150), pp. 167-173, 2006.

[11] Li, N., Mahapatra, M.K., and Singh, P., Sintering of porous strontium doped lanthanum manganite-yttria stabilized zirconia composite in controlled oxygen atmosphere at 1400 °C, J. Power Sources, 221, pp. 57-63, 2013.

[12] Lei, L., Fu, Z., Wang, H., Lee, S.W. and Niihara, K., Transparent yttria stabilized zirconia from glycine-nitrate process by spark plasma sintering, Ceramics Int., 38, pp. 23-28, 2012.

[13] Guoa, F. and Xiao, P., Effect of Fe2O3 doping on sintering of yttria-stabilized zirconia, J. Eur. Cer. Soc., 32, pp. 4157–4164, 2012.

[14] Wang, X., Kima, J.-S. and Atkinson, A., Constrained sintering of 8mol% Y2O3 stabilised zirconia films, J. Eur. Cer. Soc., 32, pp. 4121–4128, 2012.

[15] López-Honorato, E., Dessoliers, M., Shapiro, I.P., Wang, X. and Xiao, P., Improvements to the sintering of yttria-stabilized zirconia by the addition of Ni, Ceramics Int., 38, pp. 6777–6782, 2012.

[16] Mascolo, M.C., Mascolo, G. and Ring, T.A. Grain boundary evolution on sintering in yttria (8 mol%)-stabilized zirconia assisted by one or two driving forces, J. Eur. Cer. Soc., 32, pp. 4129–4136, 2012.

[17] ASTM Standard. C 914 – 95 Standard test method for bulk density and volume of solid refractories by wax immersion, 272-274, 1995.

[18] BSI-British Standards Institution. BS EN 623-3 Advanced technical ceramics-monolithic ceramics-general and textural properties-part 3: determination of grain size and size distribution (characterized by the linear intercept method), 1-22, 2001.

[19] Garino, T.J. and Bowen, K., Kinetics of constrained-film sintering, J. Am. Ceram. Soc., 5(2), pp. 251-257, 1990.

[20] Hsueh, C.H., Carneim, R.D. and Armstrong, T.R., Modelling of constrained sintering in multilayer systems, 2007. Available: http://www.ms.ornl.gov/ researchgroups/ SCG/PDF/Sintering.pdf.

[21] Kiani, S., Pan, J. and Yeomans, J.A., A new scheme to find the master sintering curve, J. Am. Ceramics Soc., 89(11), pp. 3393-3396, 2006.

[22] Carroll. D.R. and Rahaman, M.N., An initial stage model for the sintering of constrained polycrystalline thin films, J. European Ceramic Soc., 14(5), pp. 473-479, 1994.

IRON ORE SINTERING
PART 2. QUALITY INDICES AND PRODUCTIVITY

SINTERIZACIÓN DE MINERALES DE HIERRO
PARTE 2. ÍNDICES DE CALIDAD Y PRODUCTIVIDAD

JAVIER MOCHÓN
Ph. D., Centro Nacional de Investigaciones Metalúrgicas CSIC-CENIM (Spain), jmochon@cenim.csic.es
ALEJANDRO CORES
Ph. D., Centro Nacional de Investigaciones Metalúrgicas CSIC-CENIM (Spain), alcores@cenim.csic.es
ÍÑIGO RUIZ-BUSTINZA
Ph. D., Centro Nacional de Investigaciones Metalúrgicas CSIC-CENIM (Spain), irbustinza@cenim.csic.es
LUIS FELIPE VERDEJA
Ph. D. Escuela Técnica Superior de Ingenieros de Minas, Oviedo (Spain), lfv@etsimo.uniovi.es
JOSÉ IGNACIO ROBLA
Ph. D., Centro Nacional de Investigaciones Metalúrgicas CSIC-CENIM (Spain), jrobla@cenim.csic.es
FERNANDO GARCIA-CARCEDO
Ph. D., Centro Nacional de Investigaciones Metalúrgicas CSIC-CENIM (Spain), g.carcedo@cenim.csic.es

Received for review March 14 th, 2013, accepted August 27th, 2013, final version September, 6 th, 2013

ABSTRACT: Sinter plants have to process mineral mixes in order to obtain sinter of a suitable composition and quality to be loaded into the blast furnace. For this purpose a series of parameters need to be taken into account, such as the nature and composition of each component of the mineral mix and the conditions of the manufacturing process. Sinter is subjected to in-depth characterisation in terms of chemical and granulometric analysis, determination of the mineral phases in its structure and of quality indices such as reducibility, low temperature degradation, reduction degradation, and tumbler strength. It is also important to operate sinter plants with high productivity and to ensure a uniform sinter composition and quality so as to facilitate the steady state operation of the blast furnace.

Key words: Sintering, Quality indices, Productivity, Blast furnace

RESUMEN: Las plantas de sinterización tienen que procesar mezclas de mineral con el fin de obtener sinterizado de una composición y calidad adecuada para ser cargados en el alto horno. Para este fin se deben de tener en cuenta una serie de parámetros, tales como la naturaleza y la composición de los componentes de la mezcla de minerales así como de las condiciones del proceso de fabricación. El sinter es sometido a una caracterización en profundidad; análisis químico y granulométrico, determinación de las fases minerales en su estructura e índices de calidad, tales como reducibilidad, la degradación de la baja temperatura, la reducción de la degradación, y la resistencia - tumbler. También es importante que las plantas de sinterización funcionen con una alta productividad, garantizando una composición y calidad uniforme con el fin de facilitar la operación estable del horno alto.

Palabras clave: Sinterización, Índices de Calidad, Productividad, Horno Alto

List of abbreviations

AIME American Institute of Mining, Metallurgical and Petroleum Engineers, (USA)
ATS Association Technique de la Sidèrurgie Française
ASTM American Society for Testing Materials (USA)
BF blast furnace
BHP Broken Hill Propietary, (Australia)
BSE back scattered electron
CIT cahiers d'informations techniques
EDS energy dispersive spectrometry
FIMT fiber in metallic tube
ISF intensified sifting feeder
ISIJ Iron and Steel Institute of Japan
ISO International Standard Organization, (USA)
ISS Iron and Steel Society, (USA)
KSC Kawasaki Steel Corporation, (Japan)
NKK Nippon Kokan Keihin
QEMSCAN Quantitative evaluation of minerals by

scanning electron microscopy
R F Return fines
ROGESA Roheisengesellschaft Saar, (Germany)
SP sinter plant
SEM scanning electron microscope
TEM transmission electron microscopy
VDEh Verein Deutscher Eisenhüttenlente, (Germany)
XRD X-ray diffraction

1. INTRODUCTION

Improvements in the blast furnace process are achieved by technological improvements and a stable supply of high quality iron ores and sinters [1]. However, high quality iron ore resources are being depleted due to the heavy demand, and thus it is necessary to continue improving sintering technology in order to use lower quality iron ores in the raw mix. The blast furnace demands sinter with high strength, a low RDI, high RI, low fines content, good average calibrated sinter size and little variation in chemical composition in order to operate in a steady state regime. Efforts are being made to supply blast furnace operators with high quality sinter. Sinter quality control, by means of adequate sintering, is important in order to operate blast furnaces at a low fuel rate and stable operating rate.

Table 1. Data from the sinter plants of Western European countries.

	Minimum value	Maximum value
Coke consum, kg t^{-1} of sinter	39	54
Productivity, t m^{-2} per 24h	26	43
Fe total, %	51	61
FeO, %	4.0	11
Al_2O_3, %	0.6	1.8
MgO, %	0.7	2.2
RDI (>3 mm), %	27	33
Tumbler (>6.3 mm), %	63	79
Reducibility, R_{60}, %	49	78

In Western Europe, 9 countries operate 36 sinter plants which in 2004 manufactured 100.8 million tons of sinter. Table 1 sets out data on coke consumption, productivity, and sinter composition and quality indices in European sinter plants [2]. The figures show a low Al_2O_3 content, acceptable RDI values, and a large difference between the minimum and maximum parameter values.

2. QUALITY INDICES

The ongoing improvement of blast furnace operation is largely due to improvements in sinter quality. The blast furnace demands sinter with a high cold strength, low reduction degradation index and high reducibility index, in a very narrow band of chemistry variation, with the lowest possible fines content and a good average size.

2.1. Chemical composition

The chemical and structural composition are very important in sinter, and it is good for them to be stable so that both primary and final slags possess adequate characteristics in terms of softening and melting temperatures, liquid temperature and viscosity for the stable operation of the blast furnace.

It is important to have a high iron content, low gangue content, and basicity of the order of 1.6-2.1. Sinter reducibility, and sinter quality in general, improves with a higher level of hematite than magnetite, and its structure improves with a higher level of primary or residual hematite and ferrites than secondary or precipitated hematite.

2.1.1. FeO sinter

The FeO content is an important control parameter in the sinter plant. When the chemical composition of an ore mix is fixed, FeO can provide an indication of sintering conditions, in particular the coke rate [3]. A 2% increase in the FeO content in sinter has been found to lower (improve) the RDI by 8 points. However, a higher FeO content negatively affects reducibility. It is important to find an optimum FeO content in order to improve the RDI without altering other sinter properties [4].

2.1.2. Al_2O_3 sinter

The most harmful effect of alumina is to worsen the sinter RDI, which increases as the alumina content rises. Industrial experience with the blast furnace shows that within a 10-10.5% CaO content range an increase of 0.1%

in the alumina content raises the RDI by 2 points [4]. The strength and quality of sinter deteriorate as the alumina content rises. Alumina promotes the formation of SFCA, which is beneficial for sinter strength, but the strength of the ore components is lower, since a high alumina content in their lattice has been reported to be the main cause of the observed lower strength [5,6]. Alumina increases the viscosity of the primary melt that forms during the sintering process, leading to a weaker sinter structure with more interconnected irregular pores [7].

Sinter reducibility is determined by the chemical and mineralogical composition and by the pore structure. Due to the complexity of the effects of alumina on each of these factors, consideration of how alumina affects reducibility has produced contradictory results [8]. In work carried out in a sinter pot loaded with 65 kg of ore mixes with different alumina contents, an increase in the alumina content from 2 to 5.5% raised the sinter RI from 58 to 64% [9].

2.1.3. MgO sinter

MgO provides for an optimum blast furnace slag condition in terms of both good flowability and desulphurisation. It can be added to the blast furnace as raw flux in the form of dolomite or dunite, or as sinter. The addition of MgO to the raw mix improves the RDI, because MgO stabilises magnetite and thus decreases the hematite content, giving rise to less stress in the sinter during the hematite to magnetite reduction in the blast furnace stack [10].

It has been determined that replacing CaO with MgO in the form of dolomite for basicities of 1.6-1.9 leads to a slight reduction in sinter strength, reducibility and productivity [11,12]. In research carried out in a sinter pot with 65 kg of raw mix, the MgO content of four manufactured sinters was increased from 1.4 to 2.6% by the addition of dolomite to the mix. The iron ore used presents a low MgO content (0.01%) and a high Al_2O_3 content (2.99%). It was seen that raising the MgO content in the sinter, from 1.4 to 2.6%, increased the FeO content and decreased productivity and the RI, RDI and TI indices [13].

2.1.4. CaO sinter

CaO combines with the iron oxides to form compounds with a low melting point that favour the formation of the primary melt, a minimum level of which is needed in order to manufacture a strong sinter. These compounds are: $Fe_2O_3 \cdot CaO$ (1205 °C) and $FeO \cdot CaO$ (1120 °C). The properties of the melt formed during sintering determine the structure of the bonding phases originated in the sinter. The melt properties in the moments prior to solidification depend to a large extent on the chemical composition of the fines layer adhered to the granules and the assimilation of nucleus particles [14-16].

2.1.5. SiO_2 sinter

Silica combines with FeO and CaO to form compounds with a low melting point that favour the formation of the primary melt: $FeO \cdot SiO_2$ (1180 °C), $2FeO \cdot SiO_2$ (1205 °C), and $FeO \cdot SiO_2 \cdot CaO$ (1223 °C). Increasing the silica content and the basicity of the adherent fines causes the primary melt formation temperature to drop, which is favourable for the subsequent assimilation reaction at the liquid-solid interface between the fines and the nucleus particles [16,17].

2.2. Granulometric distribution

After being tipped from the pallets in the sintering machine, the sinter is hot screened. Its granulometric distribution is an important process parameter. The 12-35 mm fraction is sent directly to the blast furnace hoppers, the larger fraction is crushed to obtain smaller sized fractions, and the <5 mm fraction (return fines) is recycled to the sinter plant hoppers.

For the good operation of the process, it is important to keep a balance between the generation and recycling of return fines [2]:

B = RF generated / RF returned

$0.95 \leq B \leq 1.05$

The sinter is screened and each of the resulting fractions is weighed: >40 mm; 40-20 mm; 20-12 mm; 12-5 mm and <5 mm. The combined weight of all the fractions comprises the total cake weight. The useful sinter is the total cake minus the return fines generated (>5 mm fraction). The average grain size is calculated as a function of the kg of sinter corresponding to each fraction, and can vary over a broad interval between 25 and 45 mm.

To avoid excessive crushing of the product sinter as it

leaves the strand, some modifications have been made to the crusher. Kawasaki Steel Corporation (KSC) has improved productivity at its Kimitsu [18] and Chiba [19] plants (Japan) by using a smaller screen aperture for return fines. Kobe Steel has reduced the screen aperture for return fines in its Kakogawa plant (Japan) [20] from 5 to 4.5 mm, and has replaced the preskip 5 mm screens in the blast furnace with 3.5 and 4 mm screens. Changes to screen sizes and types have also been adopted at Chiba [19], as a consequence of which productivity has risen by 2%, and a rate of 80 kg fine sinter t^{-1} of HM can be used at No. 6 BF.

2.3. Reducibility index (RI)

Reducibility is an important characteristic of sinters which measures the ability to transfer oxygen during reduction in the blast furnace stack, giving an idea of fuel consumption needs in the furnace. The porosity and structure of the sinters and their mineral phases are intimately related with their reducibility. A heterogeneous structure is more reducible than a homogeneous structure [21-23]. It is also possible to predict reducibility behaviour from the concentration of each phase present [24]. The reducibility of mineral phases in decreasing order is:

$$Fe_2O_3 > Cao{\cdot}2Fe_2O_3 > Cao{\cdot}Fe_2O_3$$
$$2\,Cao{\cdot}Fe_2O_3 > Fe_3O_4$$

Hematite and magnetite are rapidly reduced to wustite (FeO), but the rates differ for subsequent reduction to metallic iron [25]. From hematite, wustite is quickly and homogenously reduced, although some wustite is surrounded by metal. From magnetite, the reduction is topochemical, following the sequence $Fe_3O_4 \rightarrow FeO \rightarrow Fe$, and almost all the wustite grains are surrounded by metallic iron, which delays the subsequent reaction. The reducibility of SFCA may be related with their morphology, porosity and whether or not they are coated with glass [26-28]. Acicular ferrite (<10 μm) formed at low temperature (<1300 °C) is more reducible, while columnar ferrite (>10 μm) formed at high temperature (>1300 °C, possibly coated with glass) is less reducible. Primary hematite is more reducible than secondary hematite because of its intrinsic porosity.

It has been researched how the mineralogy and porosity of iron ores affect reducibility [29,30]. Sinters have been manufactured in the laboratory (bench scale) and in a pilot plant and the relationship between porosity, reducibility and the tumbler index is determined. Higher porosity leads to greater reducibility, and the sinters with the largest surface area (open pores) present a more fragile structure and lower TI. Researchers have also investigated the behaviour of chlorine and alkalis in the blast furnace and their effect on sinter properties during reduction [31]. Studies have been carried out in the laboratory and in No. 5 BF operated by Roheisengesellshaft Saar (ROGESA) at Dillinger (Germany). The laboratory tests have shown that despite some differences the effects of chlorine, which combines to form KCl and NaCl, and alkalines on sinter are on the whole quite similar. Sinter reduction tests at up to 1100 °C show that the presence of alkalis favours the reduction of hematite to magnetite, due to the catalytic action of the alkali. The presence of chlorine compounds is unfavourable, as they are deposited on the sinter surface and inhibit its reduction. The presence of alkalis leads to an increase in the sinter stress, due to an increase in the reduction of hematite to magnetite, and cracks form which increase abrasion. By inhibiting the reduction reaction, chlorine compounds assure less abrasion up to 700 °C. At higher temperatures, the reduction reaction increases, with the corresponding rise in abrasion.

2.4. Low temperature degradation (LTD) index

The degradation of sinter is determined by the Low Temperature Degradation Index (LTD) and the Reduction Degradation Index (RDI). Sinter degradation during reduction at low temperature is determined by the dynamic LTD test, which is carried out at 600°C [32]. Degradation is originated, to a certain extent, in the transformation that takes place during the reduction of hematite to magnetite, accompanied by an increase in volume, giving rise to the presence of structural stresses in the sinter [33]. The degradation of sinter in the blast furnace occurs during reduction in the low temperature zone, and has a harmful effect on the burden strength in the furnace, with the consequent loss of permeability to reducing gases and an increase in coke consumption [34].

In research involving the addition of magnetite fines in a raw mix for sintering, a coke saving of 0.43% was seen for each 1% increase in magnetite in the raw mix, due to the fact that when hematite ore is replaced by

magnetite fines, the bed temperature rises as a result of the exothermic oxidation reaction of magnetite to hematite [35-37]. An increase of 5.1% was also seen in the LTD index for each 1% increase in hematite in the raw mix during the oxidation of magnetite, which is transformed into γ-Fe_2O_3 with the same cubic spinel lattice structure as magnetite. The TI and RI indices do not undergo any noticeable change when hematite is replaced by magnetite.

2.5. Reduction degradation index (RDI)

Sinter degradation during reduction at low temperature is more usually determined by the RDI static test [38], which is carried out at 550 °C. Low values are desirable for this index. The RDI is a very important parameter that is used as a reference in all sintering work and serves to predict the sinter's degradation behaviour in the lower part of the blast furnace stack.

Secondary hematite, also known as skeletal rhombohedral hematite, is the main cause of a poor sinter RDI. This is based on the frequent observation of cracks around the narrow neck regions of such hematite. On the other hand, it has been suggested that the cracks which form due to the volumetric change that accompanies the transformation of the crystalline phase from hematite to magnetite are responsible for the reduction degradation of the sinter [39]. Studies involving TEM, XRD and Vickers indentation tests have shown that secondary hematite is the most harmful sinter component for RDI [40]. Secondary hematite generally contains dissolved impurities like Al_2O_3, TiO_2 and MnO which raise the stress in magnetite by distorting the lattice. This magnetite forms during hematite reduction at 550 °C in the blast furnace.

Using XRD and dilatometry it has been found that the sinter structure depends on the maximum temperature reached in the bed, and that secondary hematite is present at higher temperatures [41]. Secondary hematite forms as a result of recrystallisation during the sintering of primary hematite. At lower temperatures a greater proportion of primary hematite (residual hematite) remains in the sinter composition. It has been observed that Al_2O_3 tends to be concentrated in the secondary hematite phase when the primary hematite → secondary hematite transformation takes place. Once again it is seen that an increase in the Al_2O_3 and TiO_2 concentration in sinter is harmful for the RDI. Continuing with this research [42], it was determined that the presence of a solid dissolution of Al_2O_3 and TiO_2 in hematite originates a 4% volume expansion during the reduction of hematite to magnetite at 550 °C, and causes distortion of the crystalline lattice of these phases and an increase in the magnitude of lattice stresses in the magnetite formed. The presence of cracks in the sinter structure after reduction at 550 °C are more frequent in the regions with a higher secondary hematite content, and harmful for the RDI, as has been noted.

The production rate and RDI have been researched in a sinter plant using neural networks [43]. The model considered 55 parameters and analysed a group of 695 RDI values recorded over a 3-year period. It was found that the production rate and the RDI depended on the same variables. A strong relationship was seen between the RDI and the outdoor ambient temperature at the plant. The RDI was also strongly dependent on the Ti content in the sinter, even when this was only very small. No relationship with alumina was found due to its low content (0.5%) and scarce variation in the tested period. The model found the coke ratio in the sinter mix to be the most important control variable with regard to the RDI.

To improve the operation of its No. 3 and No. 4 BFs at Fukuyama, Nippon Kokan Keihin (NKK) lowered the SiO_2 content in the sinters in its No. 4 and No. 5 SPs from 4.8 to 4.2%, taking into account pulverised coal injection rates of the order of 170 kg t^{-1} of HM. This led to an improvement in furnace permeability and reducibility, but worsened the RDI[44]. A relation was thus found between bed permeability and the RDI. NKK developed a totally new kind of optical fibre radiation thermometer named FIMT (fibre in metallic tube) which measures the temperature close to the tap hole immediately after tapping. With the combined actions of lowering the silica content in the sinter and improved melt temperature control, it has managed to lower the silica content in pig iron from 0.3 to 0.2% [44]. It has been reported that a 6% improvement in sinter RDI would lower the blast furnace coke rate by 14 kg t^{-1} of HM and increase blast furnace productivity by 3% [45].

2.6. Tumbler index (TI)

The cold strength of sinter is determined by the tumbler test [46], and depends on the strength of each individual ore component, the strength of the bonding matrix components and the ore composition. This test determines the size reduction due to impact

and abrasion of the sinters during their handling, transportation, and in the blast furnace process. Studies of the fracture strength of several mineral phases have allowed the following order to be established: primary (or residual) hematite > secondary hematite > magnetite > ferrites. Cold mechanical strength is directly related with the tendency for fines to form during transportation and handling between the sinter machine and the blast furnace throat.

The sinter strength depends to a large extent on the properties of the matrix formed by vitreous glass, silicates, olivines and ferrites. Vitreous glass presents a high degree of stress: the allotropic transformation, which starts at 697 °C, from β-$2CaO \cdot SiO_2$ to γ-$2CaO \cdot SiO_2$ is accompanied by a change in volume that causes the sinter strength to decrease. Ferrites have been identified as a strong bonding material that improves sinter strength [47,48]. The TI of sinter is dependent on critical flaws in the sinter and their propagation through sinter particles. Flaws are unavoidable because the different minerals and phases precipitate out of the melt at different times during the cooling cycle, and changes in volume almost always accompany the transformation of a liquid into a solid [49].

A research has been carried out in pilot plant and industrial scale, to enhance the strength of a sinter with high iron (58.8%) and low silica (4.38%) content [50]. By adding of serpentine and burnt lime into the ore mix to be sinterized and with a deeper sinter bed, there is a significant increase in the amount of magnesium ferrite and SFCA, which is associated with improved in sinter strength.

2.7. Sinter porosity

Sinter porosity is an important parameter that significantly affects its properties, in particular its reduction behaviour. Porosity (P) is calculated by determining the real density (d_R) and apparent density (d_A) of sinters before and after being subjected to the reducibility test:

$$P = (d_R - d_A) / d_R \quad (1)$$

Sinters experience a strong increase in porosity after undergoing the reducibility test.

In research carried out with hematite and goethite [51] ores, the changes caused to the initial pore structure during reduction tests at 550 and 950 °C were analysed. It was seen that the pore diameter needs to be larger than 0.01 μm for the reducing gas to have sufficient access to the pores to satisfactorily reduce the sinter. When the micropores coalesced to pores of a size of more than 1 to 5 μm, the specific surface area of the sinter decreased and so did its reduction [52].

Research has shown that eliminating the coalescence of micropores and increasing the number of small pores makes it possible to increase the surface area of the sinter and obtain a substantial improvement in its reducibility. Ferrites stabilise the micropores and lead to a rise in porosity, thus achieving higher reducibility [53,54]. The ferrite decomposition reaction to produce magnetite and silicates can be achieved at high temperature in a reducing atmosphere, and is the most important reaction to decrease sinter porosity.

Besides the increase in sinter porosity after being subjected to the reducibility test, there is also an increase in volume originated during the transformation of hexagonal hematite into cubic magnetite. The increase in volume that takes place due to this transformation is 25%.

The crystal structure of magnetite (Fe_3O_4) is of the spinel type, with a = 8.38 Å. It has a close-packed cubic lattice of O^{2-} ions with the smaller Fe^{2+} and Fe^{3+} ions distributed in the interstices. Hematite (α-Fe_2O_3) is of rhombohedral corundum type (a = 5.42 Å and x = 55° 14'). The O^{2-} ions are arranged in a close-packed hexagonal lattice and two thirds of the octahedral interstices are occupied by Fe^{3+} ions. The oxide has a small oxygen deficit, probably because of O^{2-} vacancies, but possibly also due to iron ions in additional interstitial positions [55].

2.8. Sinter structure

Given the diversity of the mineralogical components that comprise the raw mix, as well as the heterogeneity of the mix, it is understandable that the sinter structure will also be complex, being formed mainly by grains of iron oxide and calcium ferrites bonded by a gangue matrix [56-58]. The ferrites, whose amount increases with the basicity index, are easily reduced, and by increasing the mechanical

toughness of sinter to certain levels are considered very useful components. The ferrites are SFCA type and form by a solid-liquid reaction between hematite and the Fe_2O_3·CaO melt, with the subsequent assimilation of SiO_2 and Al_2O_3 in the melt. The gangue is composed of calcium, iron and magnesium silicates that are difficult to reduce, and come to form part of the slag in the blast furnace.

The structure and composition of sinter includes the presence of primary hematite (non-assimilated or residual), secondary hematite (precipitated), primary magnetite (non-assimilated or residual), secondary magnetite (precipitated) and ferrites as major phases, along with a smaller amount of gangue. There is sufficient porosity to favour the reducibility of the sinter, including micropores in many cases. The optimum structure for reducibility is formed by a nucleus of primary hematite surrounded by a lattice of acicular ferrites, as shown in Figure 1.

Figure. 1. Optimised sinter structure (x1100). H = hematite; F = acicular and columnar ferrite lattice

Anglo Research (South Africa) has developed a methodology for the characterisation of iron ore sinters and iron ore slime through the use of its QUESCAN, which uses EDS, SEM and BSE instruments. QUESCAN model results are based on an order of magnitude (approximately 10000 times) greater number of analysis points than those obtained by point counting [59,60].

3. PRODUCTIVITY

In a sinter plant, the onus is on achieving high productivity. This is done by assuring good bed permeability, and for this it is essential to optimise the granulation process. Moreover, for high sinter productivity it is necessary to maximise the sinter output. Various factors can influence output, such as: horizontal and vertical uniformity in the sinter bed, sinter bonding strength, crushing of product sinter, and selection of return fines screen aperture [19,20].

Non-uniform sintering usually results in part of the bed being more friable and can lead to high fines production. Where there is a lack of vertical uniformity, it is often necessary to increase the coke content in the top part of the bed. This is possible by segregation of the feed using devices such as an intensified sifting feeder (ISF) or a slit bar chute [61] (Fig. 2). Horizontal uniformity is improved by using multi-segment gates on the roll feeder outlet. The problem is most serious near the pallet walls where the air flow is highest. This can be reduced by compacting the top of the bed close to the side walls or installing a dead bar grate near the wall [62]. KSC has installed a new burner line to improve ignition at its Chiba works (Japan), thereby achieving more uniform sintering and consequently an increase in productivity [18]. ArcelorMittal has installed a more compact ignition hopper at its Fos-sur-Mer works (France), with 3 rows of burners under the hood [63]. Productivity has increased by 1 t m^{-2} per 24h as a result of more intense ignition, or because a larger surface area of the strand is now available for sintering.

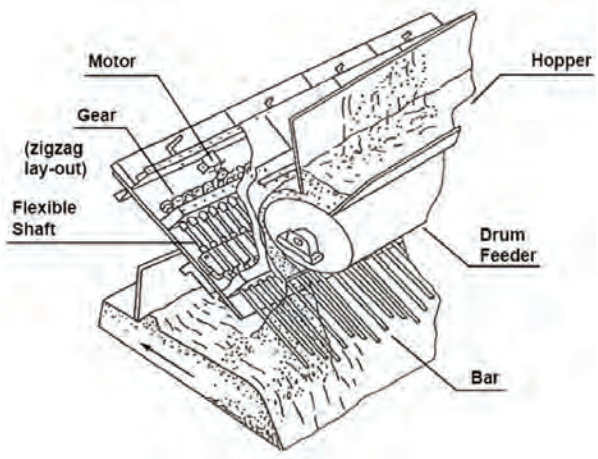


Figure. 2. Intensified sifting feeder (ISF)

In some plants higher production has been achieved by increasing the bed depth, generally together with a reduction in the strand speed. For this type of operation high permeability is essential and some improvements to granulation may be necessary, such as the addition

of (more) lime. At its Burns Harbor works (USA), Bethlehem Steel raised the bed depth from 406 mm to 635 mm and reduced the strand speed from 2.4 to 2 m min^{-1}, achieving a productivity increase of almost 30% [64]. KSC increased the bed depth at its Mizushima works (Japan) from 530 to 700 mm, achieving a 6% gain in productivity [65].

At its Kakogawa works (Japan), Kobe Steel enriches the air with oxygen [20]. Oxygen is injected below the hood that covers a large part of the strand, after the ignition hood. This improves coke consumption, with the result of operating with a narrower heating zone and a higher flame front speed. It is possible to improve production by 1 t h^{-1} with the use of a flow of 500 Nm^3 oxygen. Many plants produce sinter with a 1.5-3% MgO content by adding dolomite, serpentine or olivine in the feed [64]. Higher productivity is achieved with olivine and serpentine than with dolomite, a fact that may be attributed to the harmful effect of dolomite on sinter strength, and thus on output.

In research carried out at BHP Billington's Newcastle Technology Centre, productivity was seen to be the main challenge facing users of pisolitic ore [65]. It was widely observed that incorporating pisolitic ores in blends composed predominantly of Australian and Brazilian hematite ores causes sinter plant productivity to drop. The reason for this is a reduction in bed permeability caused by excessive melt formation. To improve productivity, water addition during granulation can be increased in order to compensate the fact that porous pisolite ore particles absorb a significant part of the added granulation water and thus reduce the amount of free water available on their surfaces for inter-particle adhesion, leading to a deterioration in granulation efficiency.

Research has been carried out in a laboratory pot grate, varying the MgO content in the raw material from 1.40 to 2.60% [66]. Dolomite and dunite are used as fluxes to add magnesia. Increasing the MgO content in the sinter mix means a higher temperature is required for melt formation, and the highly fluxed composition with MgO acts as a refractory phase, raising heat consumption and reducing productivity. On the other hand, it has been found that increasing the MgO content improves the RDI, due to the drop in hematite and ferrite phases and the rise in the magnetite phase, which presents lower degradation. Tata Steel published data recorded between April 1994 and March 2001 in its No. 2 SP at Jamshedpur (India) [67]. Raising the MgO content (range 1.75-3.25%) caused the plant productivity to decrease. The TI increased, but it was considered that for MgO contents of more than 4% the TI would decrease due to the formation of a vitreous matrix that exhibits a high degree of stress and a low formation of bonding phases. In contrast with the preceding work [66], it was seen that an increase in MgO also raised the RDI. This variation may be due to differences between experimental conditions and actual plant data.

REFERENCES

[1.] Cores, A., Verdeja, L. F., Ferreira, S., Ruiz-Bustinza, Í. and Mochón, J., Dyna, vol. 80, pp. 152- 171, 2013.

[2.] Cores, A., Babich, A., Muñiz, M., Ferreira, S. and Mochón, J., *ISIJ Int.*, vol. 50, pp. 1089-1098, 2010.

[3.] Loo, C. E., Proc. 2nd Int. Cong. on Sci. and Tech. Ironmaking. ISS Ironmaking Conf. Proc., 57, pp. 1229-1316, 1998.

[4.] Kumar, C. U., Ramana, R. V., Ali, S., Das, A. K., Kumar, A., De, A. K. and Mukherjee, T., ,Tata Search, pp. 20-25, 1995.

[5.] Sinha, M. and Ramna, R. V., ISIJ Int., 49, pp. 719-721, 2009.

[6.] Ostwald, J. and Pepper, M. D., BHP Tech. Bull., 26, pp. 46-48, 1982.

[7.] Loo, C. E. and Leung, W., ISIJ Int., 43, pp. 1393-1402, 2003.

[8.] De, S., Gupta, S. and Chatterjee, A., SEAISI Q., 21, pp. 35-46, 1992.

[9.] Umadevi, T., Deodar, A. V., Mahapatra, P. C., Prabhu, M. and Ranjan, M., Steel Res. Int., 80, pp. 686-692, 2009.

[10.] Lu, L., Holmes, R. J. and Manuel, J. R., ISIJ Int., 47, pp. 349-358, 2007.

[11.] Panigrahy, S.C., Riguad, M. A. J. and Dilewijins, J., Steel Res., 56, pp. 35-41, 1985.

[12.] Panigrahy, S. C., Verstraeten, P. and Dilewijins, J., Ironmaking Steelmaking, 11, pp. 17-22, 1984.

[13.] Umadevi, T., Nelson, K., Mahapatra, P. C., Prabhu, M. and Ranjan, M., Ironmaking Steelmaking, 36, pp. 515-520, 2009.

[14.] Napoleão, A., Pinheiro, P., Caparoli, L., Oliveira, D. L. A., Fujikawa, L. and Vervenee, R., Proc. 3rd Int. Conf. on Sci. and Tech. of Ironmaking, VDEh, Düsseldorf, Germany, pp. 127-132, 2003.

[15.] Hsieh, L. H. and Whiteman, J. A., ISIJ Int., 33, pp. 462-473, 1993.

[16.] Debrincat, D., Loo, C. E. and Hutchens, M. F., ISIJ Int., 44, pp. 1308-1317, 2004.

[17.] Bakker, T. and Heerema, R. H., ISS Ironmaking Conf. Proc., 55, pp. 487-491, 1996.

[18.] Saito, G., Motizuki, M., Amakawa, K., Kamiko, Y., Harada, T. and Yamada, H., Trans. Iron Steel Inst. Jpn., 28, B90, 1988.

[19.] Obata, H., Takahashi, H., Nakamura, M., Natsumi, T. and Komamura, K., ISIJ Int., 31, pp. 478-486, 1991.

[20.] Shibuta, K., Kuwano, K., Ito, R., Awaji, M., Ano, K. and Sugiyama, T., ISS Ironmaking Conf. Proc., 49, pp. 623-628, 1990.

[21.] Bhagat, R. P., Chattoraj, U. S. and Sil, S. K., ISIJ, 46, pp. 1728-1730, 2006.

[22.] Hessien, M. M., Kashiwaya, Y., Ishii. K., Nasr, M. I. and El-Geassy, A. A., Ironmaking Steelmaking, , 35, pp. 191-204, 2008.

[23.] Panigrahy, S. C., Rigaud, M. A. J. and Hegedus, G. S., Ironmaking Steelmaking, , 21, pp. 353-361, 1994.

[24.] Sakamato, N., Fukuyo, H., Iwata, Y. and Niyashita, T., Proc. 12th Mc Master Symp. on "Burden design for the blast furnace", Mc Master University, Hamilton, ON, Canada, pp. 137-150, 1984.

[25.] Maeda, T. and Ono, T., Trans. Iron Steel Inst. Jpn., 25, pp. 1191-1193, 1985.

[26.] Scarlett, N. V. Y., Pownceby, M. I., Madsen, I. C. and Christensen, A. N., Metal. Mater. Trans. B, , 35B, pp. 929-936, 2004.

[27.] Maeda, T., Nishioka, K., Nakashima, K. and Shimizu, M., ISIJ Int., 44, pp. 2046-2051, 2004.

[28.] Wang, S., Gao, W. and Kong, L., Ironmaking Steelmaking, , 25, pp. 296-301, 1998.

[29.] Busby, N. J. and Fray, T. A. T. "Pyrometallurgie 87". The Institution of Mining and Metallurgy in assotiation with The Institute of Metals. London, United Kingdom, pp. 141-166, 1987.

[30.] Busby, N. J., Fray, T. A. T. and Goldring, D. C., Ironmaking Steelmaking, 21, pp. 229-236, 1994.

[31.] Lectard, E., Hess, E. and Lin, R., Proc. 3rd Int. Conf. on Sci. and Tech. of Ironmaking, VDEh, Düsseldorf, Germany, pp. 424-429, 2003.

[32.] ISO Proposal, ISO/TC-102/SC-2/286-E, ISO, Geneve, Switzerland, 3rd Draft, 1974.

[33.] Roller, P. W., Trans Iron Steel Inst. Jpn., 26, pp. 834-835, 1986.

[34.] Kortman, H. A. and Burghardt, A., "Agglomeration 77", Ed. K. V. S. Sastry, Iron and Steel Soc. of AIME, New York, USA, pp. 219-231, 1977.

[35.] Panigrahy, S. C., Rigaud, M. A. J. and Hegedus, G. S., Ironmaking Steelmaking, 21, pp. 235-261, 1994.

[36.] Panigrahy, S. C., Rigaud, M. A. J., Chong, S. P. and Chanda, B., ISS Ironmaking Conf. Proc., 53, pp. 523-531, 1994.

[37.] Button, R. A. and Lundh, P. A., Ironmaking Steelmaking, 16, pp. 151-164, 1989.

[38.] ISO/TC-102/SC-3/285-E. A method for determining the low temperature disintegration behaviour of iron ores, pellets and sinters by cold tumbling after static reduction, ISO Proposal, ISO, Geneve, Switzerland, 3rd Draft, 1974.

[39.] Dawson, P. R., Ironmaking Steelmaking, 20, pp. 137-143, 1993.

[40.] Shigaki, I., Sawada, M. and Gennai, N., Trans. Iron Steel Inst. Jpn, 26, pp. 503-511, 1986.

[41.] Pimenta, H. P. and Seshadri, V., Ironmaking Steelmaking, 29, pp. 169-174, 2002.

[42.] Pimenta, H. P. and Seshadri, V., Ironmaking Steelmaking, 29, pp. 175-179, 2002.

[43.] Kinnunen, K. and Laitinen, P., Rev. Mètall. CIT, 102, pp. 361-371, 2005.

[44.] Rokugawa, S., Noda, H., Ichikawa, K., Sakamoto,

N., Sato, H. and Watanabe, T., Rev. Mètall. CIT, 96, pp. 1181-1190, 1999.

[45.] Murty, C., De, A., Chatterjee, A. and Rao, V. S., Tata Search, pp. 7-13, 1994.

[46.] "Standard method for tumbler test for iron ores, pellets and sinters". Designation: E279-69 (Reapproved 1979), ASTM, Pittsburg, PA, USA, 1980.

[47.] Dartnell, J. J., Iron Steel Inst., 207, pp. 282-292, 1969.

[48.] Barrie, D. J., Carmichael, I. F. and Kinloch, E. D. J., Iron Steel Inst., 207, pp. 563-569, 1969.

[49.] Lwamba, E. and Garbers-Craig, A. M., J. S. Afr. Inst. Min. Met., 108, pp. 293-300, 2008.

[50.] Zhu, D., Qiu, G., Fan, X., Hu, Y., Xiau, L., Tian, F., Chen, W. and Clout, J., Proc. 3rd Int. Conf. on Sci. and Tech. of Ironmaking, VDEh, Düsseldorf, Germany, pp. 133-137, 2003.

[51.] Restrepo, O. J., Vásquez, C. F. and Bastamente M. O., Dyna, vol. 75, pp. 163-170, 2008

[52.] Bristow, N. J., Goss, J. and Waters, A. G., Iron Steelmaker, 18, No. 9, pp. 61-79, 1991.

[53.] Bristow, N. J. and Loo, C. E., ISIJ Int., 32, pp. 819-828, 1992.

[54.] Bhagat, R. P., Chattoraj, U. S. and Sil, S. K., ISIJ Int., 46, 1728-1730, 2006.

[55.] Edström, J. O., "Symposium on Sinter". Special Report No. 53, The Iron and Steel Institute, London, United Kingdom, pp. 10-25, 1955.

[56.] Egundebi, G. O. and Whiteman, J. A., Ironmaking Steelmaking, 16, pp. 379-385, 1989.

[57.] Ahsan, S. N., Mukherjee, T. and Whiteman, J. A., Ironmaking Steelmaking, 10, pp. 54-64, 1983.

[58.] Mukherjee, T. and Whiteman, J. A., Ironmaking Steelmaking, 12, pp. 151-155, 1985.

[59.] Tonžetic, I. and Dippenaar, A., Miner. Engin., 24, pp.1258-1263, 2011.

[60.] Banerjee, A. and Mukherjee, A. K., ISIJ Int., 51, pp. 1392-1395, 2011.

[61.] Inazumi, T., Fujimoto, M. and Sato, K., Proc. 6th Int. Iron and Steel Cong., Nagoya, Japan, ISIJ, pp. 110-117, 1990.

[62.] Nakajima, R., Kurosawa, S., Fukuyo, H. and Yamaoka, Y., Proc. 6th Int. Iron and Steel Cong., Nagoya, Japan, ISIJ, pp. 163-170, 1990.

[63.] Huguet, C., Russo, P. and Soland, J., ISS Ironmaking Conf. Proc., 50, pp. 549-554, 1991.

[64.] Stubel, J. F. and Durko, D. P., ISS Ironmaking Conf. Proc., 47, pp. 657-669, 1988.

[65.] Matsumoto, K., Matsuda, K., Nakajima, Y. and Akizuki, H., ISS Ironmaking Conf. Proc., 47, pp. 613-619, 1988.

[66.] Umadevi, T., Roy, A. K., Mahapatra, P. C., Prabhu, M. and Ranjan, M., Steel Research Int., 80, pp. 800-807, 2009.

[67.] Yadav, U. S., Pandey, B. D., Das, B. K. and Jena, D. N., Ironmaking Steelmaking, 29, pp. 91-95, 2002.

ADVANCE OF FOUNDATION TECHNIQUES IN BRAZIL SINCE COLONIAL TIMES

PROGRESO DE TÉCNICAS DE CIMENTACIÓN EN BRASIL DESDE EL TIEMPO COLONIAL

RENATO P. CUNHA
Profesor Asociado, Ing. Civil, Ph.D., Universidad de Brasilia, Brasilia, Brasil, rpcunha@unb.br
PAULO J. R. ALBUQUERQUE
Profesor Asociado, Ing. Civil, D.Sc., Universidad Estadual de Campinas, Campinas, Brasil, pjra@fec.unicamp.br

Received for review August 1 th, 2012, accepted June 19th, 2013, final version July, 2 th, 2013

ABSTRACT: This paper presents a summary of the historical developments on foundation engineering that took place in Brazil since its early beginning until the present time. It relies on published information pinpointed from historical and reference books widely available today that were, on the other hand, written with basis on oral and written recounts and documents that survived through time. The intention is not to present a detailed discussion on the subject, but rather illustrate how, in common terms, foundation engineering technologies (and challenges) have been improved and approached throughout history in Brazil – shedding light as to what may be expected in this field for the near future of the new millennium.

Key words: History of foundations, usual foundation types, foundation techniques

RESUMEN: Este artículo presenta un resumen del desarrollo histórico que ocurrió en Brasil desde su principio hasta el tiempo presente. Está basado en información publicada y sintetizada de libros históricos y de referencia ampliamente disponibles actualmente que fueron, por otro lado, escritos con base en relatos orales y escritos y documentos que sobrevivieron a través del tiempo. No se tiene la intención de presentar una discusión detallada acerca del asunto, pero de ilustrar como, en términos generales, las tecnologías de la ingeniería de cimentaciones (y sus retos) fueron mejoradas y tratadas durante la historia de Brasil – reluciendo lo que se podrá esperar en este campo para el futuro inmediato del nuevo milenio.

Palabras clave: Historia de las cimentaciones, tipos usuales de cimentación, técnicas de fundación

1. INTRODUCTION

With the necessity of urban development, and the evolution of civil engineering, allied to nowadays a high level of demand for the construction of structures with increasingly slender and heavier buildings, foundation engineering in Brazil has advanced considerably since its early beginning.

In addition to the structural elements of the project, another recent factor of extreme importance for the development of the techniques for foundation execution and design is the limited (or lack of) urban space and the choice of location for new constructions in a quickly developing country. In many cities another aspect took place, i.e., previously disregarded regions of the "Brazilian continent" began to be populated and this fact imposed technical difficulties that necessarily needed to be overcome.

Therefore, the history of the foundation engineering in Brazil is briefly presented herein, concentrating on its most developed and populated cities, Rio de Janeiro and São Paulo (and surroundings). It does not intend to be a state of the art on the subject, but rather highlight the most interesting and relevant points in this evolution of design and construction techniques. Within this scope, as will be shown, some historical cases are emphasized together with the main techniques that have been employed in the country since the colonial times.

It shall be pointed out that the presented data comes from a compilation (and sometimes modified translation) of information submitted in the chapter 1 of a renowned book published in Brazil by the Soil Mechanics Society [1], combined with pertinent passages from another recently published book on the "History of Geotechnical Engineering in Brazil" [2]. Hence, the paper summarizes aforementioned data to

fulfill a broader information scope directed towards the international audience.

2. HISTORICAL DEVELOPMENTS

2.1. Colonial period (XV-XVIIIth. centuries)

In the colonial period of the XVI century, around 1530, the buildings could be classified into three groups according to the reference [3]. The first group corresponded to works/factories associated with the sugar industry and exporting port facilities, whereas in the second group one could find military buildings like forts and barracks, plus churches, monasteries, public buildings and related commercial chains, and finally the third group is comprised of civil works and houses of any dimensions plus business stores, warehouses, flour mills and so on. According to this reference, there are only a few reports on the foundation characteristics of these buildings, being possible to confirm a common knowledge that, since such remote period, foundations were made of excavated trenches filled in with "compacted" stones.

Around the XVII century, in the range of 1684, the reference [4] states that for the construction of the Monastery of St. Benedict, in the city of Rio de Janeiro, a "Statement of Construction" was written which contained specifications for the construction of the monastery. In one of the items presented by the document, one can find the techniques for the construction of the base foundation. In this statement the author demonstrates that concepts of "tension" and "resistance of the material' were already known at the time, being associated with the need of reaching a foundation level where a subsoil material of sufficient hardness to support the superstructure building exists. This document also mentioned that the presence of water in the subsoil could ruin the trenches excavated for the construction of this particular foundation.

Unfortunately the available information of this period is rather limited and further knowledge would require a more ample and thorough research, which could not be done.

2.2. Monarchic period (XIXth. century)

The monarchic period lasted around 1 century, and by its mid course, in the range of 1850, Brazil was ruled by an aristocratic and conservative society, driven by an agricultural economy that depended exclusively on slave labor. This was the period in which the coffee culture granted economic stability to the empire, allowing some few entrepreneurs, like Mr. Irineu Evangelista de Sousa, or the Baron de Mauá, to open a bank and to give financial incentives to large and diversified private enterprises, as, for example, the construction of the first Brazilian railway in Rio de Janeiro. The construction of this railroad was an impressive achievement for the time, since this construction was made up of bridges and tunnels, the largest of which with over 2000 meters in span. There are unfortunately few reports on the foundations of such bridges, nonetheless such engineering work can be highlighted as a true benchmark for the geotechnical engineering in Brazil. From then on, many other railroads and associated engineering works were built along the vast domains of the empire.

At the same time the construction of civil houses began to use panels made of bricks attached to beams and iron columns, due to the collapse of existing constructions built solely on "taipa" (which is a constructive technique imported from the Portuguese colonizers, still used nowadays in the poor inner rural areas of Brazil, based on the manual compaction of a mixture of local soil, predominantly clayey type, plus water and gravel).

Since then, the construction of industrial buildings, markets, warehouses, railway stations and associated works, such as the "Light Station" in São Paulo emerged. This construction, among others erected between 1884 and 1896 in the cities of São Paulo, Rio de Janeiro, Recife and Manaus, included the use of imported bricks, steel structural elements, and a metal roof. In view of such developments, particularly in the structural field, it became also necessary to leverage and improve current construction (and design) techniques for the foundations. The trenches filled with "compacted" stones were soon replaced by shallow foundations, strip foundations, or rigid blocks of brick masonry or carved rock, manually laid down at foundation base and compacted inside the ditches. According to reference [5], in the late XIX century, after the installation of the first engineering schools in the country, witnesses could observe that the Brazilian civil engineers held some understanding of the "technology of soil".

It is worthy to mention the interesting case of the water reservoirs from the Pedregulho station in Rio de Janeiro (Figure 1), where two tanks were constructed to store water for this (expanding) city, one upstream of the other with a differential topographic level of 5 meters.

After several days of tropical rain, the lower reservoir built near the slope base, had several cracks in the walls and had to be immediately emptied. According to [6], formal transcripts exist of a meeting dated July 12, 1880, i.e., over two months after the construction of the reservoir. The records reveal that the engineer Honório Bicalho (well known in Brazil) diagnosed the cause of the accident as due to differential settlements that were produced by "the lack of homogeneity of the terrain, related to the decomposition of the primitive rock, gneiss, which originated a residual soil of compressible and collapsible characteristics".

During the nineteenth century the interest in mining works, particularly iron mining, started to gain strength in Brazil. From then on, it started a large demand for geological surveys in such related works, and a close relationship between geology and (geotechnical) civil engineering was established in the country. Then, in 1907, the "Engineering Geology" field of expertise was initiated, whose early works lie in the geological survey for the implementation of railways and roads. Engineering geology also significantly contributed at such a time (as of present) to the development and implementation of the projects of the foundations of embankments and dams, or the special art works (bridges, tunnels) of road/railway constructions.

Figure 1. Historic picture of the Pedregulho reservoir (Public Archives - city of Rio de Janeiro)

With the expansion of the empire, enlargement of the cities and population growth, and associated natural evolution of the engineering problems, new challenges emerged by the mid to the final period of this century. Many problems were related to pier foundations designed for the construction of ports, in regions comprised of geologically characterized quaternary deposits, that means with a high frequency of sedimentary saturated layers of soft soil materials.

It is also reported that even though wooden piles were known and openly used at such time (as for "classical" example the Municipal Theater of Rio de Janeiro, erected in the early XX century), the lack of knowledge on the behavior of the subsoil, allied to the inexistence of settlement control techniques, led local contractors to avoid constructing new buildings in areas of soft deposits. On the other hand, it was precisely the construction of a foundation comprised of timber piles, driven by an innovative (for the time) steam powered hammer, which promoted the successful construction of deep foundations on such coastal sedimentary strata. The particular case of the design and construction of the Customs Docks in Rio de Janeiro, which started in 1866, is an example of that.

2.3. XXth century

In the early twentieth century the designers and contractors started to count with one of the best construction materials of the modern civil engineering industry: Reinforced concrete. It was known that concrete multi layer structures of this era could be built over shallow footings of reinforced concrete or on top of massive rigid blocks of non reinforced concrete. Besides, deep foundations could be represented by reinforced concrete types in addition to the wooden ones already in use. This era was also initially marked by the entry in the market of the first specialized company of seaport construction and pile foundations in Rio de Janeiro, and by the expansion of the port of Santos, where the first reinforced concrete pier supported by concrete piles and retained by a sheetpiling wall was constructed. Between 1926 and 1927 the dock of Rio de Janeiro was also expanded and the foundations consisted of pressurized piers made with driven shaft elements of reinforced concrete.

However, until the end of the '20s, empiricism was predominant in geotechnical engineering.

"Experimentation, however little, was carried out directly in the stonemasonry of the work and in a restrained fashion", according to the former ABMS president, the engineer, Antonio Nápoles Neto, in a lecture given in 1970 [2].

Possibly the first Brazilian geotechnical undertaking was described by Domingos da Silva Cunha in the Brazilian Journal of Engineering (1920) with the title "Experiments on terrain for the study of foundations". In this paper, this author warned of the variations in resistance and deformation of the soil as a function of chemical content, variations in the terrain's humidity and depth.

Nevertheless, it was precisely in the next decade that the first academic investigations were carried out by technological research institutes of the young Republic, and by the Polytechnic School of São Paulo (USP), the latter one more focused on the particular problems of foundation engineering. In 1934, the Laboratory for Material Testing of the city of São Paulo was transformed into the Institute of Technological Research of São Paulo (IPT), attached to the academic activities of the Polytechnic School. Early work in IPT started in 1938 under the coordination of engineer Odair Grillo (well known in Brazil) and aimed both the correct paving of earth roads and the study of foundations of bridges and buildings. In regard to foundation engineering, which is the scope of this paper, the first step to solving engineering problems was the development in 1939 [6] of a drilling/sounding equipment for the exploration and analysis of the subsoil. Thus, IPT standardized a soil exploratory device method, later published by Vargas (1945), which consisted of the introduction of a standard thin walled steel sampler by a percussion method with water circulation, simultaneously to the counting of the required number of strokes of a falling impact hammer of 60kg in weight, from a gravity drop height of 75cm. The number of blows necessary to penetrate this sampler to a standard length of 1 feet (30 cm) was termed as N(IPT). Besides this particular number, another standard was created by a private company for research and project in soils and foundations. This latter one also started to be largely adopted in the country, with a distinctive nomenclature of N(MG). Subsequently, the reference [7] suggested in 1948 the worldwide method commonly practiced today, known as Standard Penetration Test (SPT), with its corresponding N(SPT) number. The high variability of the testing results from both "Brazilian" techniques at the time, given differences in procedure and especially due to distinct "standard" N quantifications, posed serious problems of comparison and foundation design (via empirical rules, as still used nowadays). Therefore, over time, persuaded by regular discussions on technical congresses, the IPT´s number N variable was replaced by the international SPT quantification method of reference [7], this approach was later regulated and standardized for the whole country by the Brazilian normative NBR 6484 in its first version of 1980.

Among the numerous construction sites and foundation works carried out in the country during the 20th century, some are worth mentioning and deserve to be described, especially given their innovative solutions adopted and the engineering difficulties they faced with the available technology at the time.

A first case that deserves mention in the city of São Paulo, around 1925, is the construction of the first skyscraper in Latin America, the Martinelli building (Figure 2). The foundations were laid down using wooden piles and spread footings up to 16 m depth. It required lowering the water level, which resulted in structural damage to a neighboring building, leading to a temporary suspension of the work. The 25 storey, reinforced concrete MartineIli skyscraper was considered the highest in the world at the time.

Figure 2. Martinelli building in São Paulo [2]

Another case was the construction of the foundations of the State Bank Building. It shall be initially mentioned that, given the experience coming from previous geotechnical investigation campaigns in São Paulo, designers and contractors already had some knowledge beyond the local geotechnical parameters, i.e., they could also anticipate in some way the foundation behavior of the common buildings in the region. Nevertheless, in the present case, the construction process of the piles of this building encompassed the dynamic penetration of closed end tubes through a stiff clay stratum. This execution feature caused the lifting of the surrounding ground level by values around 70cm, leading to breakage of surrounding piles that had been (previously) molded in place. It was observed by load tests carried out in the site that such piles had compressive bearing capacities as low as 30% of the expected values. The problem was then solved by the installation of additional piles, summing up to a total of 400 piles in this particular work, all of them partially excavated around the shaft after the casting period to prevent damage during the driven stage of the new piles.

Another example that also deserves mention is the case of the Building of the Paulista Insurance Company in São Paulo. This is a 26 storey building that began construction in 1940, and in which the investigation of the subsoil was partially conducted inside the construction area given the existence of an old building that was about to be demolished (during further construction stages). Thus, the piles were designed based solely on one of the sides of the site, which allowed those foundations to behave as end bearing due to an existing shallow stiff clay stratum. However, it happened that on the other (non surveyed) side existed a layer of soft silt. This misconception in design was further aggravated by the (satisfactory) load tests solely carried out on piles that were installed on the investigated side. After three years, in 1943, the building was opened and soon afterwards, the IPT Institute detected in the monitoring records an accelerating tendency of column settlements from the building side located in the non investigated part of the site. Daily differences, or settlement increases of up to 1mm from morning to afternoon readings, were noticed, together with a steadily tilting of the building. This edification is presented in Figure 3.

According to what was presented by reference [8], the solution of the problem was achieved by an overall freezing process of the subsoil in 162 freezing steps, in which the final temperature was kept at -20° Celsius. It was followed by the drilling of injections wells, where galvanized metallic pipes were installed up to 14m in length. With the hardened ground and transitory suspension of settlements, injection of concrete was carried out through these pipes, transforming them into cast-in-situ reinforcement piles. The building was finally realigned in the vertical direction by the temporary insertion and pumping of hydraulic jacks onto existing columns.

Figure 3. Historic picture from the end-of-construction building of the Paulista Company ([2])

Still around the ‘40s two other studies conducted by the IPT stand out. One of them is related to the foundations of the bridge over the Rio Grande river during the construction of the railway Corumbá-Santa Cruz de la Sierra (in Bolivia), which crossed a terrain filled with “quicksands” and swelling clays. The other one refers to the design of the foundations of the Steel plant in Volta Redonda, where for the first time foundation problems associated with soils derived from rock alteration, or saprolites was noticed [9].

In the ‘50s, and with continuing studies of the previous decade, both IPT and the private company Geotécnica

SA monitored the settlements in tall buildings founded on shallow foundations at the coastal zone of Santos, an area typically known (later on) by its thick layers of saturated soft clays. The results of the investigation in Santos were henceforth published by references [10], [11] and [12], by reference [13] and by reference [14], the latter ones related to the settlement of foundations over soft sand deposits. Given such experience, reference [15] presented arguments and conclusions about the Brazilian expertise in forecasting and monitoring the settlements in buildings, during the state-of-art of an international conference on Soil Mechanics in 1977.

Within this same decade, the project of the Guaíba river crossing in Rio Grande do Sul deserves a special note, as this project highlights the use of bridges founded on deep foundations of the Franki type with relatively long lengths. In Rio de Janeiro the project and construction of the building Marquês de Herval, conducted by Prof. Costa Nunes (well known internationally), also calls for attention. This building was constructed by subsoil excavation with the simultaneous water level drawdown down to as much as 9 meters deep, in a region surrounded by older constructions founded on shallow foundations.

Figure 4. Construction of the foundations of the ministry buildings in Brasília (DF Public Archives)

It was also in the '50s that the construction of today's capital of the Federative Republic of Brazil, the city of Brasilia, had started. Figure 4 presents a photo of the construction phase of the ministry buildings in 1959, in the heart of the new city. Companies from the Brazilian capital (at the time), Rio de Janeiro, were responsible for conducting the site investigation of the subsoil, which is composed of sandstones and siltstones covered by a layer of residual soil, originated from the same typical metamorphic type rocks of the region. Faced with such distinctive geological conditions (from the coastal zone), the contractors opted for deep foundations with Franki type piles and uncased manually excavated caissons.

Later on in Rio de Janeiro, between 1960 and 1964, the Duque de Caxias oil refinery was built by employing several types of deep foundations, among them pressurized drilled caissons, which led to a total length that exceeded 400km of piles. Other relevant information of this work is the consumption of concrete and steel that reached 120 cubic meters and 10 tons (100 kN) respectively. Since then, one could affirm that the foundation engineering was finally consolidated in the country.

In 1963, the José Bonifácio steel plant, belonging to the Paulista Steel Group (Cosipa) was inaugurated close to Santos. The region is characterized by the presence of wetlands with soft clay deposits at surface extending to up to 25 m in depth. Faced with such geological conditions the solution by shallow foundations proved to be unattainable, allowing the exclusive use of deep foundations which totaled 110 thousand piles varying from centrifuged precast concrete types, metallic H shaped piles and mixed (composite) piles. The main problems encountered in implementing the foundations of the plant were the execution of the jointing sections of the long piles, the enhanced corrosion by the marine environment, the excess pore pressures generated during the dynamic insertion of the piles and negative friction between the shaft and the surrounding soft clay by consolidation phenomena. To understand and manage the behavior of the foundations, the IPT conducted over 300 pile load tests at this site.

Still in the '60s, the construction of the highway Piaçaguera-Guarujá, now known as Cônego Domênico Rangoni, had started close to Santos. About 12km of the road lies in a region of very soft organic soils of up to 40 meters in thickness. Six bridges were required to be built in this highway, with particular emphasis on the one crossing the Bertioga channel, that required 8 pressurized drilled caissons in the central span and dozens of metallic piles, with varying lengths from 10 to 60m on the edges. The road was opened in 1970 as a result of the implementation of pioneering techniques of geotechnical engineering.

The '70s was marked by several works of great technical challenge and financial impact for the

country, taking advantage of the good winds in which Brazil was sailing on at the time (nowadays recalled as the "economic miracle" era). Among several works it is highlighted the Transamazônica highway (1972) through the rainforest, The Lagoa-Barra highway (1971), the Steel railway (1973), the Immigrants highway (1976), the undergrounds in both Rio de Janeiro and São Paulo, and some of the astonishing in size, for the time, hydroelectric power plants (Itaipú, Tucuruí) that still serve the country nowadays.

Regarding the work of foundations, the literature highlights the Rio-Niterói Bridge (1974), where a tragic accident took place during a load test over the Guanabara Bay, causing the death of two workers. This bridge is 13.3km long and 26.6m wide, with six lanes of traffic, and is still considered nowadays a national landmark of engineering, given all the executive difficulties and the geotechnical particularities. As for the bridge foundations themselves, in the dry (land) sections metallic and Franki piles were used, whereas in the wet (sea) section pressurized 1.8m diameter drilled caissons were constructed with lengths that varied according to the thickness of the superficial organic soft clay of the bay. This was considered a pioneer solution for the time. Altogether 1138 caissons were built, of these 462 were pressurized by air, 199 were of mixed type and 477 used the "Bade-Wirth" technology. The pressurized caissons with enlarged base (similar example in Figure 5) were adopted whenever it was possible to support them in the sedimentary compacted sandy stratum located underneath the soft clay, up to a total length of 30 meters.

Figure 5. Example of pressurized drilled caisson platform and work site over a river in Brazil

The mixed caisson types were built with the dynamic insertion of a steel pile into a young residual soil and the subsequent submerged concrete molding of the caisson's shaft. Finally, the Bade-Wirth types were founded on rock and were mechanically drilled with the support of a metallic shaft, following the subsequent structural reinforcement with a steel cage that was lowered down just before the final molding stage.

The '80s began with works of great impact that in the vast majority have been started in the preceding decade. As key examples one can recall the first nuclear power plant in Brazil, whose construction started in 1970 (Angra dos Reis 1 – Figure 6) and the conclusion of the hydroelectric power plants. This decade is also marked by the construction of the Carajás railroad, the expansion of the Rio de Janeiro international Airport of Galeão (also known by Tom Jobim airport) and the construction of the international Airport of Guarulhos in São Paulo.

In 1982 the Port of Sepetiba in Itaguaí was opened, renamed in 2006 to Port of Itaguaí. The region of the implementation of the port has a subsoil typically composed by an organic silty clay of high compressibility and low bearing resistance, with thicknesses of up to 12 meters. Large diameter long piles were designed and implemented for the storage yard of the port, and the final solution also encompassed the dredging of the soft clay layer and its replacement by a hydraulically compacted sandy fill.

Figure 6. Angra dos Reis nuclear power plant (Portal Brasil, internet governmental site of free access)

In the '90s the magnificence of the Brazilian economy slowed down, given international external problems,

as the "petrol crises", and national internal financial calamities, such as hyperinflation and recession (nowadays known as the "lost era"). Nevertheless some milestone works can be recalled upon, such as the excavation and foundation reinforcement of the Sears store in São Paulo in 1989, to what would later be known as the Paulista Shopping Center. Located in an area of large commercial appeal and high acquisitive economic power, this work was marked by its great difficulty (see some of the few remaining pictures on Figure 7).

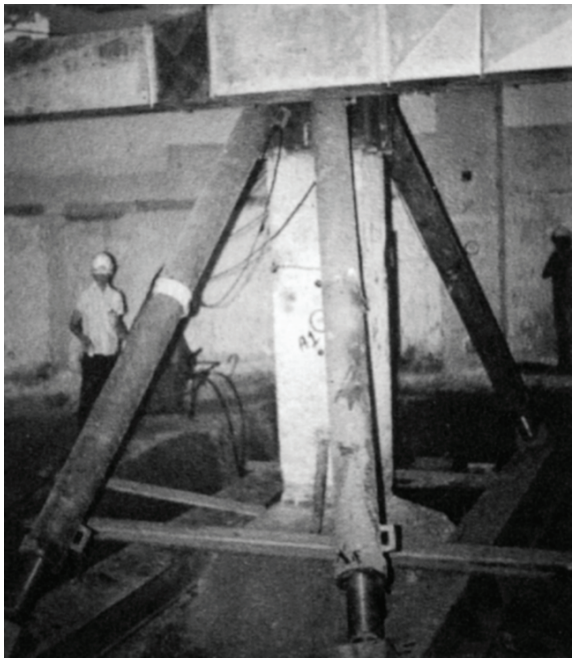


Figure 7. Foundation reinforcement at Paulista Mall ([2])

The original conception was the deployment of three additional basements to the original Sears Mall, built in 1949. It is reported that the original building consisted of three floors and a basement with a level difference of 4.3 m to the street. The first building block had shallow foundations founded at 2 meters below the basement level, and the second block was supported by drilled caissons of 14 meters long, which were constructed in 1972 during the initial expansion project of this mall.

The foundation solution adopted for the first building block was the substitution of the shallow foundations to deep ones, allowing the excavation of an extra 9 meters of soil, for the 3 additional required basements. For the second block, the strategy was the simultaneous reinforcement of the existing caissons along with the excavation. At the end, the caisson's bases were enlarged and concreted with additional steel reinforcement cages. Besides of the foundations, retaining structures made of (secant and non secant) pile curtain walls with caissons of enlarged (and non enlarged) bases were constructed to safely sustain the excavation process. It is worthy to mention that the commercial activities of the Sears Mall were not interrupted during the execution of such services.

2.4. XXIst century – new millennium

Similar to the readjusting tides, so are the economies of the world. After an era of depression and internal readjustments, which painfully demanded many personal sacrifices, political compromises and a new economic order, Brazil was able to turn over the table and reestablish itself in the path of sustained development and growth. Being part of what is called nowadays the BRICS group, the engineering possibilities, challenges and achievements for the new millennium are undoubtedly broad and large, as well as the encouraging mood and optimistic feeling of the people inside and outside the country. The coming Football World Cup event (in 2014) and the Olympic Games (in 2016) have called the attention for a general readjustment of key urban elements, such as transportation, airports, housing, hotel and entertainment venues, energy, security and, of course, stadiums and sport arenas of several types (for instance, see Figure 8).

Figure 8. Picture from the recently finished sports arena in Brasília during construction stage

Therefore, among the various geotechnical works which debuted together with the starting millennium,

one shall highlight the full renovation of the (tilted) buildings in Santos, the expansion of the Porto Alegre airport, the conclusion of the express South highway of Florianópolis, the construction of the piled embankment of the SESC, close to the future installations of the Olympic villa in Rio de Janeiro and the diverting beltway (ring road) around São Paulo.

Also, it is worthy to mention the constructions of refineries and pipelines in the northeastern and northern regions of Brazil, the expansion of the petrol industry by the discovery of unknown "pre-salt" oil basin in the southeast region, the renovation and construction of new sport arenas throughout the 12 base cities of the coming World Cup, and the (fast growing) residential urban expansion of interior cities all over the country, among other things.

The fast development of the country involves not only the use of more sophisticated techniques for foundations to support the increasing magnitude of the loads from superstructures that become more slender and tall everyday, but it also involves the formation of a skilled work force to safely deal with, the design and construction of such demanding structures. This point is valid for all fields of engineering, from foundation, to civil, mechanical, electrical, naval, energy and so on. Besides, the increasing demand for sustainable construction, space optimization and green areas have obliged urban constructions to use "less noble" subsoil areas for parking, transportation, storage and related items, leaving open environments for the more noble activities of living, entertaining and working. Besides, and regarding safety, the Brazilian standard obliges the execution of pile load tests in deep foundation projects with more than 100 piles. Such experimental validation must be done in even larger numbers in South America in order to advance the knowledge of this field throughout distinct soil sites [16].

Therefore, there is a potential challenge to develop and sustain an increasing demand for technologies capable of producing deep foundations with longer, slender and vibration/noise free piles, that would cause minimal intervention and disruption to the surrounding environment, and would eliminate the production of undesirable (environmentally unfriendly) by-products, such as contaminated soil, soil-bentonite mixtures, etc. It is with such spirit that one must realize that the foundation engineers of the future (those who are still working or studying today) will face a new reality in which the foundation problems will be engaged and solved altogether with other perspectives, i.e., from the geological and geotechnical point of view to the environmental restrictions, special construction characteristics, optimization procedures, and for sure the mandatory sustainable attitudes of the new age (as so well recognized in the recent international Conference "Rio+20" held in Brazil).

3. CONCLUSIONS

Despite the fact of being one of the South American countries that got independence from the colonial power at a later stage (1822), Brazil was able to accompany and even surpass the urban development of some of its neighbors.

The continuing investment in new machines and technology, the formation of a skilled work force at distinct levels (from technicians to Ph.D. Professors), the appreciation of the high education value, the national and international outsourcing of exceptional professionals, the free exchange of knowledge between multi-level institutions, and the urgent reconfiguration of many of the outdated Brazilian standards are some of the challenges for the future.

REFERENCES

[1] ABMS. Foundations: Theory and Practice. 2nd edition, PINI Editors, São Paulo. 751p, 1996. (In Portuguese).

[2] Sayão, A., History of Geotechnical Engineering in Brazil. 60 years of the Brazilian Association of Soil Mechanics and Geotechnical Eng. ABMS Ed., São Paulo, P. 252, 2010.

[3] Katinsky, J.R., Colonial Construction systems. History of Technique and Technology in Brazil. UNESP. CEETEPS, São Paulo, (In Portuguese), 1994.

[4] Silva Nigra, D. and Clemente, M., Three Benedictine artists. Typography. Ministry of Education and Culture. Salvador, (In Portuguese), 1950.

[5] Grillo, O., The Soil Mechanics at the Technical Road. Bulletin of the DER. Year V, Vol. 5, No. 4. São Paulo, (In Portuguese), 1939.

[6] Vargas, M., Exploring the Underground for the Purpose of Civil Engineering. Publication No. 242. Year 41, No. 149, October. Technology Research Institute (IPT), São Paulo, (In Portuguese), 1945.

[7] Terzaghi, K. and Peck, R.B., Soil Mechanics in Engineering Practice, John Wiley and Sons, New Cork, 1948.

[8] Villares, A., The Underpining of the 26 story Building Cia de Seguros - Sao Paulo, Brazil. Geotechnique, Vol. 6, Issue 1, March, pp. 1-14. 1956.

[9] Nápoles Neto, A.D.F and Lorraine, P., The Foundations of the plant at Volta Redonda. Publication No. 539. Technology Research Institute (IPT), São Paulo, (In Portuguese), 1956.

[10] Teixeira, A.H., Case History of Building underlain by Preconsolidated Unusual Condition of Clay Layer. Proceedings of 1st Panam. Conf. Soil Mechanics. Mexico, 1959.

[11] Teixeira, A.H., Contribution to the study of the Time-Settlement Relation of Actual Structures. Proceedings of 1st Panam. Conf. Soil Mechanics. Mexico, 1959.

[12] Teixeira, A.H., Typical subsoil conditions and settlement Problems In Santos, Brazi. Proceedings of 1st Panam. Conf. Soil Mechanics. Mexico, 1959.

[13] Machado, J., Settlement of Structures in the City of Santos, Brazil. Proceedings 5 th Int. Conf. Soil Mech. and Found. Eng., Paris, 1961.

[14] Vargas, M., Foundations of Tall Building on Sand in Sao Paulo, Brazil. Proceedings of the 5th International Conference on Soil Mechanics and Foundation Engineering, Paris, 1961.

[15] Burland, J. B., Broms, B. B. and De Mello, V. F. B., Behavior of foundations and structures. State of Art Report. Proceedings of the 9th International Conference on Soil Mechanics and Foundation Engineering, Tokyo, Vol. 2, pp. 495–546, 1977.

[16] Valencia, Y., Camapum, J. y Lara, L., Aplicaciones adicionales de los resultados de pruebas de carga estáticas em el diseño geotécnico de cimentaciones. DYNA, n.175, pp.182-190. 2012.

ANALYSIS OF CSP-1 UNDER INFALLIBLE AND FALLIBLE INSPECTION SYSTEMS

ANÁLISIS DE PLANES DE MUESTREO CSP-1 CONSIDERANDO SISTEMAS DE INSPECCIÓN CON Y SIN ERROR DE MUESTREO

RODRIGO A. BARBOSA-CORREA
Ph.D., Universidad Del Norte, rbarbosa@uninorte.edu.co
GINA M. GALINDO-PACHECO
Ph.D., Universidad Del Norte, ggalindo@uninorte.edu.co

Received for review April 10th, 2013, accepted July 29th, 2013, final version August, 21st, 2013

ABSTRACT: In this paper, we discuss the implementation of Continuous Sampling Plan (CSP)-1 under two scenarios: (i) infallible, and (ii) fallible inspection systems. For both cases, we develop an optimization model for designing a CSP-1 that minimizes the total expected cost. We use Markov theory to derive the expected results from the application of the CSP-1. A Bayesian approach is used to model the inspection system reliability. Based on the analyses for the two models, we offer a discussion on the adverse effects of disregarding inspection errors when implementing CSP-1.

Key words: CSP-1, quality, inspection sampling plan, Bayes, Markov, simulation, inspection error, optimization

RESUMEN: En el presente artículo, presentamos un análisis de las implicaciones relacionadas con ignorar errores de inspección cuando se implementa un plan de muestreo continuo del primer tipo (CSP-1 por sus siglas en inglés). Nuestro análisis cubre dos escenarios: (1) inspección perfecta o infalible, e (2) inspección imperfecta o falible. Para cada caso, presentamos un correspondiente modelo de optimización cuyo objetivo es el de minimizar el valor esperado del costo total. El comportamiento de los planes CSP-1 es modelado utilizando teoría Markoviana, mientras que la confiabilidad de los sistemas de inspección es modelada mediante un análisis Bayesiano. Las soluciones de ambos modelos son confrontadas para establecer comparaciones entre los dos escenarios.

Palabras clave: Calidad Saliente Promedio (CSP), Planes de Muestreo, Teorema de Bayes, Cadenas de Markov, Simulación, Error Muestral, optimización

1. INTRODUCTION

Quality is one of the most important factors considered by customers at the moment of selecting their suppliers. In an ideal world, customers would prefer suppliers whose products were absolutely perfect. However, in reality, customers usually agree to tolerate a certain proportion of defective units. Then, the task for the suppliers is to implement inspection policies that guarantee that the average outgoing quality level (AOQL) of their product, does not exceed a certain value, based on the customer's expectations regarding the acceptable proportion of defective units. An inspection policy can be defined as inspecting 100% of the products. However, due to financial limitations, a 100% inspection is not always viable. An alternative is to implement sampling inspection plans (SIPs), in which only a fraction of the total products is inspected.

The parameters involved in SIPs are often selected in such a way that the total expected cost is minimized [1-4]. An important assumption when computing the expected cost of a SIP is that related to the efficiency or reliability of the inspection system. Such inspection systems can be human inspectors or machines (from now on we will use indifferently the terms "inspector" and "inspection system" to refer to both human and machine-based inspection systems). In both cases, the occurrence of inspection errors is inevitable. There are two types of inspection errors: Type I, which refers to classifying a defective unit as non-defective; and Type II, which refers to classifying a non-defective unit as defective. Disregarding these errors, i.e. assuming perfect performance of the inspector, is unrealistic and it can lead to inaccuracies in the computations of the expected SIP cost and performance.

In this paper, we use Markovian and Bayesian analysis to develop two optimization models for designing a

SIP of the type CSP-1 [5], which applies for products that are manufactured through a continuous process. The implementation of CSP-1 can be summarized as follows:

1. Initially, inspect 100% of the units until *i* consecutive units are found as non-defective.

2. Once *i* consecutive non-defective units have been inspected, discontinue the 100% inspection and start to systematically inspect only a fraction *f* of the units. The fractional inspection continues as long as the inspected units are non-defective. If a defective unit is found, reestablish 100% inspection, i.e. return to step 1.

In our first model, we consider the minimization of expected costs assuming perfect performance of the inspector. This initial model sets the path for developing a more realistic model in which our objective is to minimize the expected costs when inspection errors are taken into consideration. Based on the results obtained for such optimization models, we offer a discussion on the effects of disregarding inspection errors in CSP-1.

The remaining of the paper is organized as follows: Section 2 contains our literature review. In section 3, we state the description of the problem, as well as the corresponding notation and assumptions. Section 4 presents an optimization model for CSP-1 when disregarding inspection errors, whereas section 5 contains the optimization model when taking into consideration such errors. In section 6, we discuss the economic impact of disregarding inspection errors, based on the results obtained on sections 4 and 5. In section 7 we extend our discussion by including the transition costs between total and partial inspection. Finally in section 8 we present some remarks, conclusions and future research directions.

2. LITERATURE REVIEW

In the literature, we can find some papers that analyze the effect of inspection errors in SIPs, such as that offered by [6], who use a Bayesian model to evaluate the efficiency of inspectors. Another interesting work that uses a Bayesian approach is that given by [7]. In this latter paper, the authors consider a type of SIP that accounts for the number of defects in each inspected product. The authors address the problem of analyzing the best *a priori* distribution to model the number of defects per unit under the presence of inspection errors. Another interesting paper that also considers inspection errors is that offered by [8] which establishes a set of results for matching Dodge-Romig single plans with Dodge-Romig plans under the presence of inspection errors.

As mentioned in the previous section, the design of SIPs is often subject to economic criteria such as in [9]. In that paper, the authors offer a mathematical model to design both 100% and single sampling plans considering potential inspection errors, while minimizing a loss function that accounts for deviations of quality characteristics from a certain target value. An earlier related work is offered by [10] in which the authors develop a model for locating inspection stations in an *n*-stage production system. The optimization criterion used in [10] is the cost per good unit accepted by the customer. Reference [11] also studies the impact of inspection error from an economical perspective. More specifically, in [11] the author develops a mathematical model and an algorithm for designing a SIP under inspection errors, while minimizing the expected associated cost. Reference [12] offers a model for minimizing inspection costs while imposing upper bounds on the inspection errors.

One fundamental difference between our paper and those mentioned above, is that they do not address CSP-1 in particular. Two papers that specifically investigate CSP-1 under inspection errors from an economic point of view are those given by [4], and [3]. In the first one, the authors present an optimal mixed policy of precise inspection and CSP-1 under the presence of inspection errors and return cost. In the second, the authors develop a model using a renewal reward process approach for selecting an economically optimal decision involving three alternatives: “do 100% inspection”, “do not inspect” and “do a CSP-1 inspection”. Their model accounts for both types of inspection errors. A key difference between the work given by [3] and ours is that we assume that as long as the customer’s acceptable quality level (AQL) is satisfied, there is no penalization for defective units delivered to the customer. With this reasonable assumption, our optimization models are simplified to finding the optimal number of inspected units as

the basis for designing an optimal CSP-1, instead on focusing on finding directly the parameters of the inspection plan (see sections 4 and 5).

Also, one of the main contributions of our work is that it highlights the impact of disregarding inspection errors when implementing CSP-1, which allows visualizing the importance of recognizing and measuring inspection errors. Similar analyses have been made for attribute SIPs, [13-17], but to our knowledge, this issue has not been yet addressed from the perspective of a supplier that implements CSP-1.

3. PROBLEM DESCRIPTION, NOTATION AND ASSUMPTIONS

In this section we present the description of our problem, as well as the corresponding notation and assumptions.

3.1. Problem Description

We focus our analysis on a supplier company that uses a continuous production scheme. The production process has an inherent defective fraction, which is greater than the AQL specified by the customer. In order to comply with the customer's AQL, the supplier implements an inspection plan of the type CSP-1 that guarantees that the AOQL, i.e. the expected proportion of defective units that are delivered to the customers, is lower than or equal to the customer's AQL. (The implementation of CSP-1 is as described in the Introduction).

The only two parameters involved in CSP-1 are *i* and *f*. The values for these two parameters are specified such that the resulting AOQL does not exceed the customer's AQL. Given the defective fraction of the process, different combinations of *i* and *f* can be used to achieve a desired AQL.

As mentioned before, the inspector participating in CSP-1 is often assumed to be infallible. However, in practice, inspection errors are likely to occur. In fact, according to the Second Law of Thermodynamics [18] and the Principle of Uncertainty of Heisenberg [19], it is not possible to have perfect inspectors, and if they were perfect, it would be impossible to prove it. Our objective is to perform an analysis of the implementation of CSP-1 under two scenarios: (1) assuming infallible inspection systems; and (2) considering the presence of inspection errors. We then study the impact of disregarding such inspection errors.

3.2. Assumptions

Our assumptions can be summarized as follows:

- The value of the defective fraction of the production process is deterministic and known.
- Throughout CSP-1, every rejected unit must be replaced by an acceptable one, according to the inspector criteria.
- Additional units produced to replace rejected ones, are also inspected.
- As long as the AOQL delivered by the company is at most equal to the customer's AQL, there is no penalization for defective units received by customers.
- The purpose of the supplier company is to design a CSP-1 that guarantees an AOQL lower than or equal to the customer's AQL.
- For simplicity, we perform our computations for a shipment of Q units delivered to the customer.

3.3. Notation

- θ_2 : event of a unit being defective
- θ_1 : event of a unit being non-defective
- $P(\theta_j)$: probability of occurrence of event θ_j, for $j = 1,2$. Note that $P(\theta_2)$ is the fraction defective of the process, and that $P(\theta_1) = 1 - P(\theta_2)$
- AQL : acceptable quality level specified by the customer. We assume that $AQL < P(\theta_2)$. Otherwise, there would be no need to implement CSP-1.
- $AOQL_A$: average outgoing quality level for a CSP-1 under infallible inspection systems
- $AOQL_B$: average outgoing quality level for a CSP-1 under inspection errors
- Q : size of the production batch to be analyzed

- S_1 : event of the inspection system classifying a unit as non-defective
- S_2 : event of the inspection system classifying a unit as defective
- $P(S_k)$: probability of occurrence of event S_k for $j = 1,2$. Note that $P(S_1) = 1 - P(S_2)$
- U_A : expected number of inspected units for a CSP-1 under infallible inspection systems
- U_B : expected number of inspected units for a CSP-1 under fallible inspection systems
- c_s : cost of inspecting one unit
- c_r : cost of rejecting one defective unit

4. ECONOMIC ANALYSIS FOR CSP-1 UNDER IDEAL INSPECTION PROCEDURES

In this section we present a discussion on the optimal expected cost of implementing CSP-1 under ideal conditions, i.e. infallible inspection systems.

As mentioned before, for CSP-1 to be useful, the AOQL must be at most equal to the consumer's AQL. A way to assign values to the parameters involved in the CSP-1 is by using Markovian analysis, in which CSP1- procedure is considered as an ergodic Markov chain. Two Markovian states are identified: (1) 100% inspection and (2) systematic inspection [18]. When analyzing CSP-1 as a Markov chain, the transition matrix would be as depicted in Table 1 [20]. Based on Table 1, we can define the steady-state probabilities $x^*_{100\%}$ and x^*_f respectively for the states of 100% and fractional inspection [20]:

$$x^*_{100\%} = \frac{P(\theta_2)}{P(\theta_2) + P(\theta_1)^i} \tag{1}$$

$$x^*_f = \frac{P(\theta_1)^i}{P(\theta_2) + P(\theta_1)^i} \tag{2}$$

Recall that Q represents the amount of products being inspected. Then, the expected number of inspected units in CSP-1 is given by $Q\left[x^*_{100\%} + f x^*_f\right]$ [20], which can be expressed as follows:

$$U_A = Q\frac{P(\theta_2) + fP(\theta_1)^i}{P(\theta_2) + P(\theta_1)^i}, \tag{3}$$

Table 1. Transition Matrix for CSP-1

	100% Inspection	Fractional Inspection
100% Inspection	$1 - P(\theta_1)^i$	$P(\theta_1)^i$
Fractional Inspection	$P(\theta_2)$	$P(\theta_1)$

The total expected number of rejected units would be equal to the number of detected defective units, i.e. $U_A P(\theta_2)$. The AOQL would be equal to proportion of the undetected defective units, which can be expressed as a function of U_A as follows [20]:

$$AOQL_A(U_A) = \frac{[Q - U_A]P(\theta_2)}{Q} \tag{4}$$

Two types of costs must be considered for the implementation of CSP-1: inspection cost, and defective units cost. The expected inspection cost would be $c_s U_A$, where U_A is given by expression (3).

On the other hand, the cost due to defective units involves the replacement of defective detected units by non-defective ones. This implies producing and inspecting additional units in order to replace the defective ones. Note that due to the defective fraction of our process, for producing an amount of *X* non-defective units, the expected number of units to be produced would be $\frac{X}{P(\theta_1)}$, since we should expect $XP(\theta_2)$ units to be defective. Let us denote c_p as the cost of manufacturing one unit of product. Then, the expected cost due to defective units would be $\left[c_p + c_s\right]\frac{U_A P(\theta_2)}{P(\theta_1)}$. Additionally, we need to include the cost of rejecting each defective unit, denoted by c_r.

Then the total expected cost for CSP-1, can be expressed in terms of U_A as:

$$C_A(U_A) = U_A \left[c_s + \frac{c_p P(\theta_2)}{P(\theta_1)} + \frac{c_s}{P(\theta_1)} + c_r P(\theta_2) \right] \quad (5)$$

The expected cost for CSP-1 under infallible inspection systems, given by equation (5), has U_A as its only variable. The minimization of (5) is restricted by the fact that the AOQL given in (4) must be at most equal to the customer's AQL. Then our optimization model can be expressed as:

Minimize (5) subject to:

$$\left[1 - \frac{U_A}{Q} \right] P(\theta_2) \leq AQL \quad (6)$$

We have that the AOQL is a decrement function of the proportion of inspected units. On the other hand, by inspecting (5) we notice that the expected cost is an increasing function of U_A. Therefore, the optimal, value of U_A, i.e. U_A^*, is the minimum U_A that allows complying the constraint given by (6). Any value lower than U_A^* would imply an AOQL greater than the customer's AQL, whereas a value greater than U_A^* would not be economically optimal. Then, U_A^* can be computed as follows:

$$U_A^* = U_A \mid \left[1 - \frac{U_A}{Q} \right] P(\theta_2) = AQL \quad (7)$$

When solving (7) we obtain:

$$U_A^* = \left[P(\theta_2) - AQL \right] \frac{Q}{P(\theta_2)} \quad (8)$$

Having determined the value of U_A^*, it is possible to compute the parameters of CSP-1 using equation (3) which relates $U_A, P(\theta_2), f, i$ and Q, where $U_A, P(\theta_2)$ and Q would be given. We can specify the value of one of the parameters and then solve for the other, based on the desired value of the AOQL. In this case, it is simpler to select a value of i and then solve for f, as follows:

$$f = 1 - \frac{AOQL\left[P(\theta_2) + P(\theta_1)^i \right]}{P(\theta_2) P(\theta_1)^i} \quad (9)$$

At this point, we would like to highlight two properties regarding the optimal expected cost and the optimal expected number of inspected units for CSP-1 under infallible inspection. First we have that expression (8) can be stated as $U_A^* = Q - \frac{[AQL]Q}{P(\theta_2)}$. Therefore, we can conclude that U_A^* increases with $P(\theta_2)$. This is a reasonable result since, intuitively, we would expect to have to inspect more units to guarantee a certain AQL for greater defective fractions. Second, by inspecting expression (5) we can notice that the optimal expected cost increases with $P(\theta_2)$. Again, this finding is also reasonable since the inspection, replacement and rejection costs increases with the defective fraction of the process.

5. EXPECTED COST FOR CSP-1 UNDER INSPECTION ERRORS

In the ideal case discussed in the previous section, an underlying assumption is that the inspection procedure is infallible. Therefore, it is assumed that whenever an inspected unit is classified as non-defective, such a unit is actually non-defective (an analogous analysis applies for units that are classified as defective). However, in practice we would expect inspection systems not to be infallible

When we drop the infallibility assumption, the transition matrix given in Table 1 does no longer apply for CSP-1, since the probability of an inspected unit being classified as defective is not equal to the actual defective fraction of the process. Also, recall that S_2 is the event of the inspection system classifying a unit as defective, and S_1 as non-defective. Anytime an event S_1 occurs, it means that the inspected unit has been accepted, whereas S_2 implies the rejection of the inspected unit. In our model we consider two types of conditional probabilities:

- Validity: $P(S_i \mid \theta_j)$ *for* $i = 1,2, j = 1,2$ gives the probability that the inspector makes a decision S_i given that the actual status of the inspected unit is θ_j. Note that such conditional probabilities are inherent to the inspector.

- Prediction: $P(\theta_j \mid S_k)$ *for* $j = 1,2, k = 1,2$ gives the probability that an inspected unit is θ_j, given that the inspector has classified it as S_k.

The two types of inspection errors discussed in the Introduction section can be now defined in terms of the conditional probabilities stated above, as follows:

$$P_I = P\left(Error\,Type\,I\right) = P\left(S_2 \mid \theta_1\right) \tag{10}$$

$$P_{II} = P\left(Error\,Type\,II\right) = P\left(S_1 \mid \theta_2\right) \tag{11}$$

By applying Bayes theorem and using the conditional probabilities discussed above, we can compute the probability of a unit being classified as defective, $P\left(S_2\right)$, as follows:

$$P\left(S_2\right) = P\left(S_2 \mid \theta_2\right)P\left(\theta_2\right) + P\left(S_2 \mid \theta_1\right)P\left(\theta_1\right) \tag{12}$$

In a similar way, we can compute $P\left(S_1\right)$. Then, the transition matrix for a CSP-1, when considering inspection errors, would be as shown in Table 1, but replacing θ_j by S_j. We can perform a statistical analysis similar to that offered in section 3.1 to compute an expression for the expected number of inspected units $\left(U_B\right)$, and the corresponding AOQL ($AOQL_B$). For U_B we have:

$$U_B = \frac{P\left(S_2\right) + fP\left(S_1\right)^i}{P\left(S_2\right) + P\left(S_1\right)^i}Q \tag{13}$$

To find the AOQL, we need to compute the expected proportion of defective units that would be delivered to the customers. First, we have the expected number of defective units that are inspected, and that are accepted due to error type II, i.e. $U_B P\left(\theta_2|S_1\right)P\left(S_1\right)$. Also, when performing the inspection, some units will be rejected and additional units would be produced and inspected to replace the rejected ones, which would be equal to $\frac{U_B P\left(S_2\right)}{P\left(S_1\right)}$. Among these additional units, we will also have defective units that are erroneously classified as non-defective and are sent to the customers, i.e. $\frac{U_B P\left(S_2\right)}{P\left(S_1\right)}P\left(\theta_2|S_1\right)P\left(S_1\right)$ or simply, $U_B P\left(S_2\right)P\left(\theta_2|S_1\right)$. Additionally, there will be some defective units delivered to the customers which come from those units that are never inspected during the fractional inspection. Therefore, the AOQL in this case can be expressed as:

$$AOQL_B\left(U_B\right) = \frac{U_B P\left(\theta_2|S_1\right)P\left(S_1\right) + U_B P\left(S_2\right)P\left(\theta_2|S_1\right)}{Q} + \frac{\left(Q - U_B\right)P\left(\theta_2\right)}{Q} \tag{14}$$

which can be simplified as follows:

$$AOQL_B\left(U_B\right) = \frac{QP\left(\theta_2\right) + U_{i_2}\left[P\left(\theta_2|S_1\right) - P\left(\theta_2\right)\right]}{Q} \tag{15}$$

Regarding the expected cost for CSP-1 under inspection errors, we have four different components: (1) cost of inspecting units, (2) cost of replacing rejected units, (3) cost due to rejecting defective units, and (4) opportunity cost due to inspection error Type I. The first cost is simply $c_s U_B$. For the expected cost of replacing rejected units we have that, as discussed in section 3.1, we need to produce and inspect additional units to replace those that have been rejected. Note that a proportion equal to $P\left(S_2\right)$ of such additional units would be also rejected. Then, the expected number of additional units that we must produce and inspect to replace $U_B P\left(S_2\right)$ rejected units would be $\frac{U_B P\left(S_2\right)}{1 - P\left(S_2\right)}$.

Therefore, the second component of the expected cost for CSP-1 under inspection errors would be $\left[c_s + c_p\right]\frac{U_B P\left(S_2\right)}{P\left(S_1\right)}$. The expected costs related to rejecting defective units would be simply $c_r B P\left(\theta_2|S_2\right)P\left(S_2\right)$.

Finally, let us denote as c_I the cost of erroneously classifying a non-defective unit as defective (Type I error). Then, the expected cost due to a Type I error would be $U_B c_I P\left(\theta_1|S_2\right)P\left(S_2\right)$. Recall that we have assumed that defective units sent to the clients do not generate an extra cost as long as the AOQL is lower than or equal to the customer's AQL. Therefore, we do not introduce any cost due to errors Type II. Then, the total expected cost of implementing CSP-1 under inspection errors is given by equation (16).

$$C_B(U_B) = U_B \left[\begin{array}{l} c_s + c_s \dfrac{P(S_2)}{P(S_1)} + c_p \dfrac{P(S_2)}{P(S_1)} \\ + c_I P(\theta_1 | S_2) P(S_2) + c_r P(S_2) \end{array} \right] \quad (16)$$

The expected cost for CSP-1 given by equation (16), has U_B as its only variable. The minimization of (16) is restricted by the fact that the AOQL given in (14) must be at most equal to the customer's AQL. Then our optimization model can be expressed as:

Minimize (16) subject to:

$$\frac{QP(\theta_2) + U_B\left[P(\theta_2|S_1) - P(\theta_2)\right]}{Q} \leq AQL \quad (17)$$

Note that by assuming that $P(\theta_2) > P(\theta_2|S_1)$, the AOQL is a decrement function of the proportion of inspected units. This is a reasonable assumption, since the role of the inspector is to reduce the number of defective units delivered to the customers (otherwise, it would be better not to use the inspector). On the other hand, the expected cost is an increasing function of U_B. Therefore, we can proceed as before to find the optimal value of U_B, i.e. U_B^*, which would be the minimum U_B that allows complying the constraint given by (17). Then, U_B^* can be computed as follows:

$$U_B^* = U_B \left| \frac{QP(\theta_2) + U_B\left[P(\theta_2|S_1) - P(\theta_2)\right]}{Q} = AQL \right. \quad (18)$$

Then, we have:

$$U_B^* = \frac{\left[P(\theta_2) - AQL\right]Q}{P(\theta_2) - P(\theta_2|\ S_1)} \quad (19)$$

Having determined the value of U_B^*, it is possible to compute the parameters of CSP-1 using equation (13). To do so, we can specify the value of one of the parameters and then solve for the other, based on the desired value of the AOQL. As mentioned in the previous section, it is simpler to select a value of i and then solve for f, as follows:

$$f = \frac{\left[P(\theta_2) - AQL\right]\left[P(S_2)P(S_1)^i\right]}{P(S_1)^i\left[P(\theta_2) - P(\theta_2|\ S_1\right]} - \frac{P(S_2)}{P(S_1)^i} \quad (20)$$

By inspecting expression (20) we have the following properties, which will be stated without proof due to their simplicity:

Property 1: the optimal expected number of inspected units, i.e. U_B^*, is directly proportional to the difference between the fraction defective of the process and the customer's AQL.

Property 2: the optimal expected number of inspected units, i.e. U_B^*, is inversely proportional to the difference $\left[P(\theta_2) - P(\theta_2|S_1)\right]$. Notice that, given that a unit has been delivered to the customer, if we did not implement an inspection plan at all, the probability of such a unit resulting defective is $P(\theta_2)$; whereas if we implemented a CSP-1, such probability would decrease to $P(\theta_2|S_1)$. Then, this property states that U_B^* is inversely proportional to the improvement obtained for having an inspection system, regarding the chances of delivering a defective unit to the customer.

Regarding the optimal expected cost, we have that it increases with the proportion of rejected units and with the probability of rejecting a non-defective unit. As it depends on U_B^*, it also increases with $\left[P(\theta_2) - AQL\right]$ and decreases with $\left[P(\theta_2) - P(\theta_2|S_1)\right]$.

6. CSP-1 UNDER INFALLIBLE INSPECTION VS. CSP-1 UNDER INSPECTION ERRORS

Let us consider the case in which we are interested in implementing a CSP-1 while minimizing the total expected cost. If we took into consideration the presence of inspection errors, we would compute the optimal value for U_B^* as discussed in section 5. However, if we neglected such inspection errors by

assuming infallible inspection systems, we would compute a suboptimal value U_A^* based on the analysis provided in section 4. The difference between U_B^* and U_A^* can be stated as:

$$U_A^* - U_B^* = Q\left[P(\theta_2) - AQL\right] \left[\frac{1}{P(\theta_2)} - \frac{1}{P(\theta_2) - P(\theta_2|S_1)}\right] \quad (21)$$

Expression (21) is always lower than or equal to zero, by our assumptions that $P(\theta_2) > P(\theta_2|\ S_1)$ and $P(\theta_2) > AQL$. Therefore, by implementing a CSP-1 with the expected number of inspected units equal to $U_A^* < U_B^*$, we should expect that in the long run we will not be able to comply with the customer's AQL. By inspecting expression (21) we can formulate the following properties:

Property 3: The difference $\left[U_A^* - U_B^*\right]$ is directly proportional to the difference between the fraction defective and the customer's AQL. This means that even though both U_B^* and U_A^* increases with $\left[P(\theta_2) - AQL\right]$, the rate at which such an increment occurs for U_B^* is greater than for U_A^*. This can be verified by examining expressions (8) and (19), from which we obtain that the rate at which U_B^* changes with $\left[P(\theta_2) - AQL\right]$ is $\frac{Q}{P(\theta_2) - P(\theta_2 \mid S_1)}$. This rate is greater than that for U_A^*, which is equal to $\frac{1}{P(\theta_2)}$.

Property 4: The magnitude of the difference $\left[U_A^* - U_B^*\right]$ decreases with the improvement obtained for having an inspection system, defined as $\left[P(\theta_2) - P(\theta_2|S_1)\right]$. Moreover, in the ideal case in which an accepted unit never happens to be defective, i.e. $P(\theta_2|S_1) = 0$, we would have $U_A^* = U_B^*$.

A reasonable assumption is that if a customer receives a shipment of units whose AOQL is greater than the customer's AQL, all units in such a shipment would be rejected and returned to the supplier. Therefore, if we design our CSP-1 disregarding inspection errors, in the long run we should expect the customer to return all of our shipments! Hence, we can conclude that using U_A^* as the optimal policy in the presence of inspection error, is not economically suboptimal, but simply unviable.

7. EXTENSION CONSIDERING TRANSITION COSTS BETWEEN TOTAL AND PARTIAL INSPECTIONS

So far we have only considered the costs related to inspection, rejection and classification errors. This approach is valid only when the cost due to the transition between total and partial inspection is insignificant. However, let us assume that we we would need to stop the production process and reprogram the inspection system when shifting from one type of inspection to the other. In that case, the assumption that the transition costs between total and partial inspections can be neglected would no longer apply. Let us first focus in the ideal case in which there are no inspection errors. Let us also introduce the following additional notation:

- c_1: cost for shifting from 100% to systematic inspection
- c_2: cost for shifting from systematic to 100% inspection
- U_1: number of expected units to be inspected during 100% inspection
- U_2: number of expected units to be inspected during systematic inspection

The expected number of times that we would shift from 100% to systematic inspection is given by the steady-state probability $x_{100\%}^*$ times the total number of units, Q, times the probability of shifting from 100% to systematic inspection which is $P(\theta_1)^i$, i.e. $Qx_{100\%}^* P(\theta_1)^i$. We also have that $U_1 = Qx_{100\%}^*$. Then, we have that the expected number of times that we would shift from 100% to systematic inspection would be given by $U_1 p(\theta_1)^i$. Similarly, the expected number of times in which we would shift from systematic to 100% inspection would be given by $U_2 p(\theta_2)$, where

$U_2 = Qfx_f^*$. Then the total expected cost under ideal inspection would be given by:

$$C_A(U_A) = U_A \left[c_s + \frac{c_p P(\theta_2)}{P(\theta_1)} + \frac{c_s}{P(\theta_1)} + c_r P(\theta_2) \right] + c_1 U_1 p(\theta_1)^i + c_2 U_2 p(\theta_2) \quad (22)$$

As mentioned before, the minimization of the cost given by (22) is subject to (6).

Note that from expression (22) we have the following properties:

- $U_1 + U_2 = U_A$.
- Our findings presented in section 4 regarding the optimal number of units to be inspected still apply. This means that the optimal number of units to be inspected is $U_A^* = U_A \left| \left[1 - \frac{U_A}{Q} \right] P(\theta_2) = AQL \right.$.

  To see why, let us first consider the case in which we inspected less than U_A^*. In that case we would violate the constraint given by (6) which is not allowed. For the other case, let us consider that we inspect $U_A^{'}$ units where $U_A^{'} > U_A^*$. Then, clearly the first term of (22) would be greater for $U_A^{'}$ than for U_A^*. Also, $C_A(U_A)$ is an increasing function of both U_1 and U_2. So if there is a combination of values $U_1 + U_2$ that allows complying with the constraint, such a combination would dominate any other that yields to a greater values of $U_1 + U_2$. Therefore, U_A^* is always a better solution than $U_A^{'}$.

$C_A(U_A)$ can be expressed as:

$$C_A(U_A) = U_A \left[c_s + \frac{c_p P(\theta_2)}{P(\theta_1)} + \frac{c_s}{P(\theta_1)} + c_r P(\theta_2) \right] + c_1 U_1 p(\theta_1)^i + c_2 (U_A - U_1) p(\theta_2) \quad (23)$$

By realizing that $U_A^* = U_A \left| \left[1 - \frac{U_A}{Q} \right] P(\theta_2) = AQL \right.$,

we just need to focus on optimizing $C_A(U_A)$ in terms of U_1. Therefore, our optimization problem becomes:

$$Min\, Z(U_1) = c_1 U_1 p(\theta_1)^i + c_2 (U_A^* - U_1) p(\theta_2) \quad (24)$$

By inspecting (24) we have that the optimal value of U_1, namely U_1^* is given by:

- $U_1^* = U_A^*$ if $c_1 p(\theta_1)^i - c_2 p(\theta_2) \le 0$. This means that we would remain in 100% inspection and we would avoid shifting to systematic inspection. This can be achieved by making $i = U_A^*$ and f = 0.
- $U_1^* = 0$ if $c_1 p(\theta_1)^i - c_2 p(\theta_2) \ge 0$. This would imply that we never enter 100% inspection. This can be achieved by starting directly the systematic inspection, making $i = 0$. The value of f must be such that:

$$U_2^* = Qfx_f^* = Qf \frac{p(\theta_1)^i}{p(\theta_1)^i + p(\theta_2)} = \left[P(\theta_2) - AQL \right] \frac{Q}{P(\theta_2)} \quad (25)$$

Then, the optimal value for f would be given by:

$$f = \left(\frac{\left[P(\theta_2) - AQL \right]}{P(\theta_2)} \right) \left(\frac{p(\theta_1)^i + p(\theta_2)}{p(\theta_1)^i} \right) \quad (26)$$

Now let us discuss the case in which we consider inspection errors. By a similar analysis to that applied for the case without inspection errors, we obtain that:

$$C_B(U_B) = U_B \left[c_s + c_s \frac{P(S_2)}{P(S_1)} + c_p \frac{P(S_2)}{P(S_1)} + c_I P(\theta_1 | S_2) P(S_2) + c_r P(S_2) \right] + c_1 U_1 p(S_1)^i + c_2 U_2 p(S_2) \quad (27)$$

It can be easily shown that, as in the previous case, our findings for section 5 still apply and therefore $U_B^* = \frac{\left[P(\theta_2) - AQL\right]Q}{P(\theta_2) - P(\theta_2 | S_1)}$. Then, our optimization problem becomes:

$$Min\, Z(U_1) = c_1 U_1 p(S_1)^i + c_2 (U_B^* - U_1) p(S_2) \quad (28)$$

By inspecting (28) we arrive to similar conclusions as those exposed for the previous case:

- $U_1^* = U_B^*$ if $c_1 p(S_1)^i - c_2 p(S_2) \le 0$. Therefore $i = U_B^*$ and f = 0.
- $U_1^* = 0$ if $c_1 p(S_1)^i - c_2 p(S_2) \ge 0$. This can be achieved by starting directly the systematic inspection, making $i = 0$. The value of f must be such that:

$$U_2^* = Qfx_f^* = Qf \frac{p(S_1)^i}{p(S_1)^i + p(S_2)} = \frac{\left[P(\theta_2) - AQL\right]Q}{P(\theta_2) - P(\theta_2 | S_1)} \quad (29)$$

Then, the optimal value for f would be given by:

$$f = \left(\frac{\left[P(\theta_2) - AQL\right]}{P(\theta_2) - P(\theta_2 | S_1)} \right) \left(\frac{p(S_1)^i + p(S_2)}{p(S_1)^i} \right) \quad (30)$$

8. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this paper we have presented two models for designing optimal CSP-1 from an economic perspective, under two scenarios: infallible and fallible inspection systems. Our work combines Markovian and Bayesian analysis to model the uncertain parameters of the system. We found that for both models, the optimal decision is to implement a CSP-1 whose expected number of inspected units is the minimum required to guarantee that the AOQL does not exceed the customer's AQL. From our findings, we can conclude that implementing the model for infallible inspection systems when actual inspection errors are present, results in a violation of the customer's AQL. We have also extended our analysis for the cases in which the costs for shifting between 100% and systematic inspections cannot be neglected. Our results show that the optimal answer in this case is either to remain whether in 100% or in systematic inspection, but never switch between the two styles of inspection within the same process .

As future research direction we recommend to consider the defective fraction as stochastic rather than deterministic. In this regard, the probability function of the fraction defective can be modeled using the Normal approximation to the Binomial distribution. Also, our analysis can be extended to other SIPs, such as CSP-2, CSP-3 and other related ones.

REFERENCES

[1] Lorenzen, T.J., Minimum cost sampling plans using bayesian methods, Naval Research Logistics Quarterly, 32(1), pp. 57-69, 1985.

[2] Schmidt, J.W., Case, K.E. and Bennett, G.K., The Choice of Variables Sampling Plans Using Cost Effective Criteria, AIIE Transactions, 6(3), pp. 178-184, 1974.

[3] Lin, T.-Y. and Yu, H.-F., An Optimal Policy for Csp-1 with Inspection Errors and Return Cost, Journal of the Chinese Institute of Industrial Engineers, 26(1), pp. 70-76, 2009

[4] Yu, H.-F., Yu, W-C. and Wu, W.P., A mixed inspection policy for CSP-1 and precise inspection under inspection errors and return cost, Computers & Industrial Engineering, 57, pp. 652-659, 2009

[5] Dodge, H.F., A Sampling Inspection Plan for Continuous Production, Annals of Mathematical Statistics, 14(3), 64-279, 1943

[6] Quinino, R.d.C., Ho, L.L. and Trindade, A.L.G., Bayesian judgment of a dichotomous inspection system when the true state of an evaluated item is unknown, Computers & Industrial Engineering, 49, pp. 591-599, 2005.

[7] Chun, Y.H. and Sumichrast, R.T., Bayesian inspection model with the negative binomial prior in the presence of inspection errors, European Journal of Operations Research, 182, pp. 1288-1202, 2007.

[8] Sengupta, S. and Majumdar, A., Some Considerations of Dodge and Romig Single Sampling Plans under Inspection

Error, Sankhyā: The Indian Journal of Statistics, Series B, 47(2), pp. 242-246, 1985.

[9] Ferrel Jr., W.G. and Chhoker, A., Design of economically optimal acceptance sampling plans with inspection error, Computers & Operations Research, 29, pp. 1283-1300, 2002

[10] Balllou, D. P. and Pazer, H.L., The impact of Inspector Fallibility on the Inspection Policy in Serial Production Systems, Management Science, 28(4), pp. 387-399, 1982

[11] Wang, C-H., Economic off-line quality control strategy with two types of inspection errors, European Journal of Operational Research, 179, pp. 132-147, 2007.

[12] Wu, H.Y., Chuang, C.L., Kung, Y.S. and Lin, R.H., Determination of optimal inspection sequence within misclassification error bound, Expert Systems with Applications, 38, pp. 5366-5372, 2011.

[13] Bennett, G. K., Case, K. E. and Schmidt, J. W., The economic effects of inspector error on attribute sampling plans, Naval Research Logistics, 21, pp. 431–443, doi: 10.1002/nav.3800210306, 1974

[14] Collins Jr. R.,D., Case, K.E. and Bennett, G.K., "Inspection Error and Its Adverse Effects: A Model with Implications for Practitioners," AIIE Transactions, 10 (1), pp. 2-9, 1978

[15] Tang, K. and Schenider H., The Effects of Inspection Error on a Complete Inspection Plan, IIE Transactions, 19(4), pp. 421-428, 1987.

[16] Greenberg, B. S., Attributes Sampling under Classification Error, Encyclopedia of Statistics in Quality and Reliability, doi: 10.1002/9780470061572.eqr147, 2008.

[17] Rios, J., Diseño de un Plan de Muestreo por Atributos en busca de un Óptimo Social, Dyna, 179, pp. 53-61, 2011.

[18] Ben-Naim, A., La entropía desvelada: El mito de la segunda ley de la termodinámica y el sentido común, Tusquetes Editores S.A., pp. 118 - 256, 2011

[19] Serway, R.A., y Jewett, J.W., Física para Ciencias e Ingeniería Vol. II, Thompson - Cengage Learning Editores, 6ª edición, P. 606, 2005

[20] Barbosa, R., Paternina, C. and Llinás, H., Revisión Bayesiana de Planes de Muestreo por Aceptación CSP para Procesos de Producción Continua y por Lotes, Barranquilla: Project COLCIENCIAS – Uninorte, 2007

TWO-DIMENSIONAL COMPUTATIONAL MODELING OF THE ELECTROKINETIC REMEDIATION OF A COPPER-CONTAMINATED SOIL

PART I: MODEL VALIDATION

MODELACIÓN COMPUTACIONAL EN DOS DIMENSIONES DE LA REMEDIACIÓN ELECTROCINETICA DE UN SUELO CONTAMINADO CON COBRE

PARTE I: VALIDACIÓN DEL MODELO

VIRGILIO RUBIO-NIEBLAS
MSc. Instituto de Ingeniería, Universidad Autónoma de Baja California. Mexicali, México, virgilio62@gmail.com

MANUEL PEREZ-TELLO
PhD. Departamento de Ingeniería Química y Metalurgia, Universidad de Sonora. Hermosillo, México, mperezt@iq.uson.mx

RICHARD A. JACOBS
Independent consultant. Houston, Texas, U.S.A., rajacobs@gmail.com

RONALDO HERRERA-URBINA
PhD. Departamento de Ingeniería Química y Metalurgia, Universidad de Sonora, Hermosillo, México, rherrera@guaymas.uson.mx

SERGIO A. MORENO-ZAZUETA
PhD. Departamento de Ingeniería Civil y Minas, Universidad de Sonora. Hermosillo, México, alan@dicym.uson.mx

Received for review January 26 th, 2013, accepted October 8th, 2013, final version November, 12 th, 2013

ABSTRACT: A two-dimensional computational model for the electrokinetic remediation of copper from a contaminated soil is presented. The model is an adaptation of the original code developed by one of the authors, in which the chemical composition and transport of the copper species in solution were incorporated. The model includes electromigration, electroosmosis, ordinary diffusion, and convection transport mechanisms, it assumes local chemical equilibrium and unsteady state conditions. An artificially contaminated soil from Cananea, Mexico was used as an example. The model predictions were compared with experimental data collected in a laboratory, circular-shaped soil in which a central anode and four cathodes in the periphery were inserted. A reasonable agreement between the predicted and the experimental values was obtained in terms of the spatial distributions of pH and copper concentration. The potential applications of the computational model are discussed.

Key words: Electrokinetic remediation, copper-contaminated soil, mathematical model.

RESUMEN: Se presenta un modelo computacional en dos dimensiones para la eliminación electrocinética de cobre de un suelo contaminado. El modelo es una adaptación del código original desarrollado por uno de los autores, al cual se incorporaron la composición química y el transporte de las especies de cobre en solución. El modelo incluye los mecanismos de transporte por electromigración, electroósmosis, difusión ordinaria y convección, se supone equilibrio químico local y régimen no estacionario. Un suelo natural de Cananea, México se utilizó como ejemplo. Las predicciones del modelo se compararon con datos experimentales obtenidos a nivel laboratorio en un suelo circular con un ánodo central y cuatro cátodos insertados en la periferia. Las predicciones del modelo concordaron razonablemente bien con las distribuciones espaciales de pH y concentración de cobre. Se discuten las aplicaciones potenciales del modelo computacional.

Palabras clave: Electrocinética, modelo matemático.

1. INTRODUCTION

The contamination of soils by heavy metals is a current concern worldwide. Metallic contaminants are typically retained within the soil components, but may be released and transported to the aqueous phase in the pores and other soil locations depending on both weather conditions and the spatial distributions

of pH and oxidizing potential within the soil [1]. The remediation of affected sites has been proposed by several techniques. The present work focuses on an extraction process usually referred to as electrokinetic remediation of soils [2-6]. In this process, the affected soil is water-flooded to promote the dissolution of contaminants as ionic species in solution. Next, a series of electrodes is inserted into the soil, and an electrical potential difference is applied. The electric field causes the motion of the ionic species towards the electrodes of opposite sign, where they react to form insoluble neutral compounds which precipitate rapidly. At the end of the process, the small volume of the contaminant-rich soil in the vicinity of the electrodes may be removed by mechanical techniques for further treatment and disposal. An alternative way of operation consists of injecting a continuous flow of fresh water to selected electrode wells while the electrical potential difference is applied. The bulk flow of water facilitates the transport of the contaminants through and out of the soil by convection. The electrokinetic remediation process is a novel technology which is currently under development. The literature shows a number of experimental studies [2-6] for different types of contaminants. Such a practical knowledge has been complemented with mathematical formulations [2,7-9] aimed at representing on a quantitative basis the relevant phenomena occurring in the system; namely, electroosmosis, electromigration, ordinary diffusion, and convection, which are responsible for the transport of the chemical species throughout the soil. Such phenomena occur simultaneously with homogeneous, liquid-phase chemical reactions, as well as the electrochemical decomposition of water at the surface of the electrodes. Adsorption of some of the species on the surface of the soil pores may also occur [2,8]**.** The present investigation was motivated by recent reports regarding soil contamination by copper in some areas of the city of Cananea, Mexico [10]. Based on the analysis of the soil samples [11], it was of interest to elucidate the potential of electrokinetic remediation as a candidate to treat such soil. Over the last years, significant advances on the elucidation of the mechanisms governing the electrokinetic remediation of soils have been made [2-9]. Whereas most of the literature has focused on one-dimensional systems in which one anode and one cathode are used, recent experimental works [12,13] have addressed two-dimensional systems in which a spatial distribution of electrodes is considered. From a theoretical standpoint, Jacobs and Probstein [9] reported on the first two-dimensional mathematical model for the electrokinetic remediation of soils. The model was verified by comparing its predictions with experimental data collected in a bed of kaolin particles. Recently, Vereda-Alonso et al. [14] reported on a simplified two-dimensional model in which the motion of the chemical species was mostly attributed to electromigration; thus, electroosmosis, diffusion and convection mechanisms were neglected. The model predictions showed good agreement with the experimental data collected in a square two-dimensional porous medium made up of kaolin particles. The present investigation is based on a previous study by Jacobs [15], who found that electrode arrangements in which the anodes are surrounded by a number of cathodes provide greater efficiencies than those in which the cathodes are surrounded by the anodes. To the best of the authors´ knowledge, such an electrode arrangement has yet not been studied experimentally in the literature. The goal of this investigation was twofold: (1) to test whether the original formulation by Jacobs and Probstein [9] was capable of predicting with reasonable accuracy the kinetics of electroremediation of an artificially contaminated soil, and (2) to analyze the main features of the process in a two-dimensional soil field by means of computer simulation. Parts I and II of this series are thus devoted to such goals, respectively.

2. MODEL FORMULATION

The goal of the mathematical formulation was to provide a quantitative representation of the relevant phenomena occurring during the electrokinetic remediation in a two-dimensional framework. The mathematical model developed by Jacobs and Probstein [9] was the starting point. Because a detailed description of the mathematical formulation was reported by the authors [9], only a brief description is presented here. The continuity equation for the *i*-th aqueous species within the porous medium is written as

$$\frac{\partial \psi C_i}{\partial t} + \nabla \bullet \psi N_i = \psi R_i \qquad (1)$$

where all the symbols are defined in the Nomenclature. The flux N_i includes the contributions by electroosmosis, electromigration, ordinary diffusion, and convection transport mechanisms. The numerical solution of

Equation (1) is difficult by conventional finite-difference methods such as Runge-Kutta and related algorithms [16]. This is because the chemical reaction terms R_i are several orders of magnitude higher than the transport terms appearing on the left-hand side of Equation (1), which causes numerical instabilities. An alternative way of solving Equation (1) is by assuming local chemical equilibrium throughout the porous medium. This allows the specification of a new set of independent quantities T_k defined by

$$T_k = \sum_{i=1}^{N} \alpha_k C_i \qquad k = 1...(N-M) \tag{2}$$

where α_{ik} is a stoichiometric coefficient representing the contribution of species i to the conserved quantity k, N is the total number of chemical species, and M is the number of chemical reactions occurring in the system. Symbols T_k are defined in such a way that they are not affected by the chemical reactions. Multiplying through Equation (1) by α_{ik} and taking the summation over all i values yields

$$\sum_{i=1}^{N} \alpha_{ik} \left(\frac{\partial C_i}{\partial t} + \nabla \bullet N_i \right) = \sum_{i=1}^{N} \alpha_k R_i \tag{3}$$

Because the amount of every element is conserved throughout the chemical reactions, the right-hand side of Equation (3) is equal to zero. Further substitution of Equation (2) into Equation (3) yields, upon rearrangement

$$\frac{\partial T_k}{\partial t} + \sum_{i=1}^{N} \alpha_k \nabla \bullet N_i = 0 \quad k = 1...(N-M) \tag{4}$$

Because the right-hand side of Equation (4) does not contain chemical reaction terms, this equation may be solved by conventional finite-difference and finite-element methods. The numerical solution of Equation (4) thus provides the values of the conserved quantities T_k as functions of both time and position within the porous medium. Once the T_k values are known, the concentration of the individual species in the solution C_i may be computed from Equation (2) coupled to the chemical equilibrium relationships:

$$K_j = \prod_{i=1}^{N} (C_i)^{\gamma_{ij}} \qquad j = 1...M \tag{5}$$

Equations (2) and (5) is a set of N nonlinear algebraic equations with N unknowns; namely, C_1 through C_N. In this study, the set of equations was solved by means of the multivariable Newton-Raphson algorithm [16].

The specification of the chemical system was computed according to the procedure described by Butler and Cogley [17], and was based on the equilibrium diagram for the copper-water system reported by Fuerstenau and Palmer [18]. The resulting diagram is shown in Figure 1.

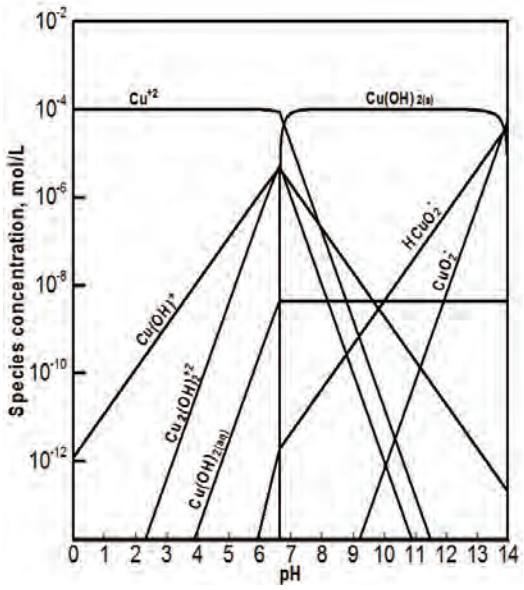


Figure 1. Distribution of chemical species in aqueous solution at 25 °C. Total copper concentration: 10^{-4} mol/L. Adapted from Fuerstenau and Palmer [18].

The resulting chemical system included: N = 13 chemical species and M = 5 conserved quantities, and is shown in Table 1. Also shown are the mobilities and effective diffusivities of the chemical species. Four continuity equations for the conserved quantities: T_1 through T_4 were solved within the porous medium. Quantity T_0 is the net electrical charge of the aqueous solution, which must be zero at all times. Thus, it is a restriction for the local values of the species concentrations. Table 2 shows the chemical reactions considered in this study, as well as their equilibrium constants computed from the HSC software [21]. The specification of such reactions satisfies thermodynamic consistency. In addition to the homogeneous reactions occurring in the aqueous phase, the heterogeneous decomposition of water was assumed to occur at the anode surface: $2H_2O_{(aq)} \rightarrow 4H^+_{(aq)} + O_{2(g)} + 4e^-$. At the cathode surface, water was assumed to react according to: $2H_2O_{(aq)} + 2e^- \rightarrow 2OH^-_{(aq)} + H_{2(g)}$. From a numerical standpoint, such reactions represent boundary conditions to Equation (4). The model also includes the

equations to compute the density of total electric charge, the electrostatic potential, and the pressure distribution throughout the soil. Details of such equations and the overall numerical strategy are discussed elsewhere [9] and thus are not repeated here. For numerical purposes, the system boundaries were established beyond the vicinity of the electrode wells so that mass transfer across the boundaries was assumed to be zero. The model equations were solved in rectangular coordinates by means of a finite-element algorithm, and the results were analyzed within the framework of the Tecplot visualization software [22].

3. EXPERIMENTAL WORK

The goal of the experimental program was to provide data to validate the mathematical model described above. Figure 2 shows a schematic representation of the experimental set up used in this study. A circular, 0.4 m diameter, 0.35 m-height acrylic cell was used to contain a two-dimensional soil field. The cell included five graphite electrodes.

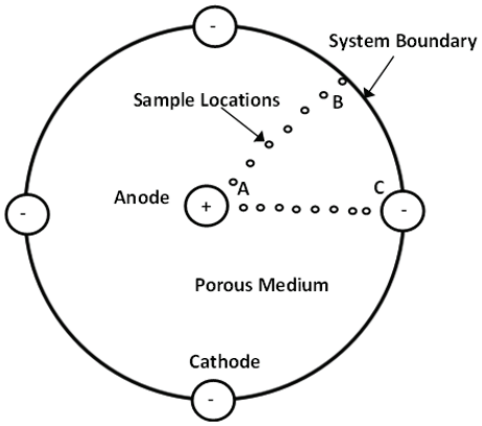


Figure 2. Schematic diagram of the two-dimensional soil used in this study.

One electrode working as the anode was placed at the center of the arrangement, whereas four electrodes working as the active

Table 1. Definition of chemical species and conserved quantities

Chemical species	Symbol for species concentration	Mobility* mol m/(N·s)	Effective diffusivity† m^2/s	Conserved quantity	Symbol
H^+	C_1	$3.7x10^{-12}$	$9.3x10^{-9}$	Solution net electrical charge	T_0
OH^-	C_2	$2.1x10^{-12}$	$5.3x10^{-9}$	Na^+	T_1
Na^+	C_3	$5.4x10^{-13}$	$1.3x10^{-9}$	NO_3^-	T_2
NO_3^-	C_4	$7.7x10^{-13}$	$1.9x10^{-9}$	SO_4^{-2}	T_3
SO_4^{-2}	C_5	$4.3x10^{-13}$	10^{-9}	*Total Cu*	T_4
HSO_4^-	C_6	$5.5x10^{-12}$	$1.3x10^{-8}$		
Cu^{+2}	C_7	$3.0x10^{-13}$	$7.5x10^{-10}$		
$CuOH^+$	C_8	$3.3x10^{-12}$	$8.3x10^{-9}$		
$Cu_2(OH)_2^{+2}$	C_9	$1.7x10^{-12}$	$4.1x10^{-9}$		
$HCuO_2^-$	C_{10}	$4.9x10^{-12}$	$1.2x10^{-8}$		
CuO_2^{-2}	C_{11}	$2.9x10^{-13}$	$7.1x10^{-10}$		
$Cu(OH)_{2(s)}$	C_{12}	0	0		
$Cu(OH)_{2(aq)}$	C_{13}	$3.7x10^{-12}$	$9.3x10^{-9}$		

*Computed by the method of Vanýsek [19]

†Computed from [20]: $D_i = D_i^0 \psi / \tau$

Table 2. Chemical reactions and equilibrium constants computed from the HSC software [21]

Reaction number *j*	Chemical reaction	Equilibrium constant K_j
1	$H_2O \leftrightarrow H^+ + OH^-$	10^{-14}
2	$Cu^{+2} + H_2O \leftrightarrow CuOH^+ + H^+$	$1.2\text{x}10^{-8}$
3	$2Cu^{+2} + 2H_2O \leftrightarrow Cu_2(OH)_2^{+2} + 2H^+$	$3.9\text{x}10^{-11}$
4	$Cu^{+2} + 2H_2O \leftrightarrow HCuO_2^- + 3H^+$	$2.7\text{x}10^{-28}$
5	$Cu^{+2} + 2H_2O \leftrightarrow CuO_2^{-2} + 4H^+$	$1.6\text{x}10^{-19}$
6	$Cu(OH)_{2(s)} \leftrightarrow Cu^{+2} + 2OH^-$	$4\text{x}10^{-42}$
7	$Cu(OH)_2 \leftrightarrow HCuO_2^- + H^+$	10^{-10}
8	$SO_4^{-2} + H_2O = HSO_4^- + OH^-$	$9.7\text{x}10^{-13}$

cathodes were placed at the border of the cell at right angles with respect to the center anode. The electrode wells consisted of five PVC cylindrical shells cut in half and glued to the inside of the acrylic cell. The half shells were perforated on their curved side to allow for the aqueous solution from the surrounding soil to flow inside the wells. The soil used in the experiments was from the copper mine at the city of Cananea, Mexico. This soil was free of copper contamination, it was previously sieved to eliminate particles larger than 0.635 cm (1/4 inch), and stored in plastic bags. Table 3 summarizes the properties of the soil after sieving [11]. According to the criteria established by the Unified Soil Classification System, this is a sandy clay soil. The unit dry weight of the soil was found to be 14.5 N/m^3, with a porosity of 44.8%. The chemical composition of the soil was determined by atomic absorption spectroscopy (AAS). From this analysis, the following amounts in weight percent were determined [11]: 54.7% SiO_2, 3.2% Fe_2O_3, 2.6% CaO, 2.2% K_2O, 1.47% MgO, and 0.15% Mn_2O_3. The soil also contained Zn (122 mg/kg), elemental Cu (60 mg/kg) and Pb (35 mg/kg). The native copper present in the soil did not come from a contaminant source. Preliminary experiments with water showed that the copper initially present in the soil was motionless under the presence of the electric field. Because the native copper is not intended to be removed by electroremediation, this was a good experimental result. In a typical electrokinetic experiment, 2 kg of sieved soil was placed inside the acrylic cell and distributed uniformly by a systematic shaking procedure. The soil was artificially contaminated by flooding it with an aqueous solution of copper sulfate. The solution was prepared so that a pre specified copper concentration was obtained (100 and 600 mg/L). The solution was poured onto the soil bed until saturation, and the system was left at rest for two days covered with a plastic cover. A direct-current power source was used to generate the electric field. For a given initial concentration of copper in the soil, replicated experiments were carried out for 24, 48 and 72 hours. Before and after the experiments, 2 mL samples of the aqueous solution were collected with a plastic syringe at the sample locations shown in Figure 2. The samples were analyzed to determine the pH by a potentiometer, and the copper concentration by AAS. The sample locations were placed every 4 cm along two radial directions as shown in Figure 2. The first radial direction connected the surface of the central anode to the surface of the cathodes in the periphery (line AC), and is referred to as the anode-to-cathode distance. The second direction followed a path forming a 45-degree angle with respect to line AC, and will be referred to as anode-to-border distance (line AB). Five equally-spaced sampling locations along lines AC and AB were established. For line AC, sample locations 1 and 5 were placed at the surfaces of the anode and cathode, respectively. For line AB, sample location 1 was 2 cm away from the anode surface, whereas sample location 5 was set to match the surface of the acrylic cell border.

4. DISCUSSION OF RESULTS

Table 4 shows the experimental conditions and model parameters used for the validation runs. Figure 3 shows the predicted and experimental values of total copper concentration T_4 along lines AC (left plots) and AB (right plots) as functions of time and initial copper concentration in the soil (lower and upper plots). For a given initial concentration of copper, a comparison of left and right plots shows that copper species traveled faster along the AC direction than the AB direction, as the experimental values in the vicinity of the cathode (left plots) are higher than their respective values in the vicinity of the cell border (right plots). This was an expected result because the intensity of the electric field reaches its maximum along line AC, which in turn causes the metallic ions to move at the fastest rate. Similarly, for a given AC or AB direction, a comparison of lower and upper plots indicates that the higher the initial content of copper in the soil, the slower the motion of the copper species. As a result, more copper was left behind when 600 mg/L was used (upper plots) as compared to the experiments conducted with 100 mg/L of initial copper (lower plots). A general trend observed in Figure 3 consists of a sudden change in the experimental values of copper concentration within a few centimeters across the soil. This change was up to two orders of magnitude for samples collected 4 cm apart. In this paper, the location at which this phenomenon occurs is referred to as the reaction front. Along line AC, the reaction front is observed in the proximity of the cathode surface, whereas for line AB it occurred within 4 cm of the cell border. During the experiments, the reaction front was observed to move outwards as time progressed. The motion of the reaction front involves the formation of an isoelectric zone and is explained later in this paper.

Table 3. Properties of the Cananea soil [11].

Property	Value	ASTM Method
Gravel content, wt. %	10	D422-63
Sand content, wt. %	48	D422-63
Fines content, wt. %	42	D422-63
Liquid limit, %	31	D422
Plastic limit, %	16	D422
Specific gravity, dimensionless	2.62	D854
Maximum unit dry weight, N/m^3	18.2	D698
Optimum moisture content, wt %	22	D4643
Hydraulic conductivity coefficient, m/s	$8.77x\ 10^{-8}$	D2434
pH	7.7	D4972

Table 4. Experimental conditions and model parameters for the validation runs

Parameter	Value
Cell diameter, m	0.45
Anode-to-cathode distance, m	0.2
Electrode diameter, m	0.0381
Anode-to-cathode potential difference ϕ, V	20
Soil porosity ψ	0.45
Soil tortuosity τ	1.1
Hydraulic permeability κ_h', m^2	10^{-15}
Zeta potential ζ, V	$-1.54x10^{-2}$
Electrical permittivity ε, F/m	$7x10^{-10}$
Solution viscosity μ, Pa•s	10^{-3}
Effective hydraulic permeability κ_h, $m^2 \bullet V^{-1}s^{-1}$	2.4×10^{-9}
Soil electrical conductivity σ_s, S/m	0.02875

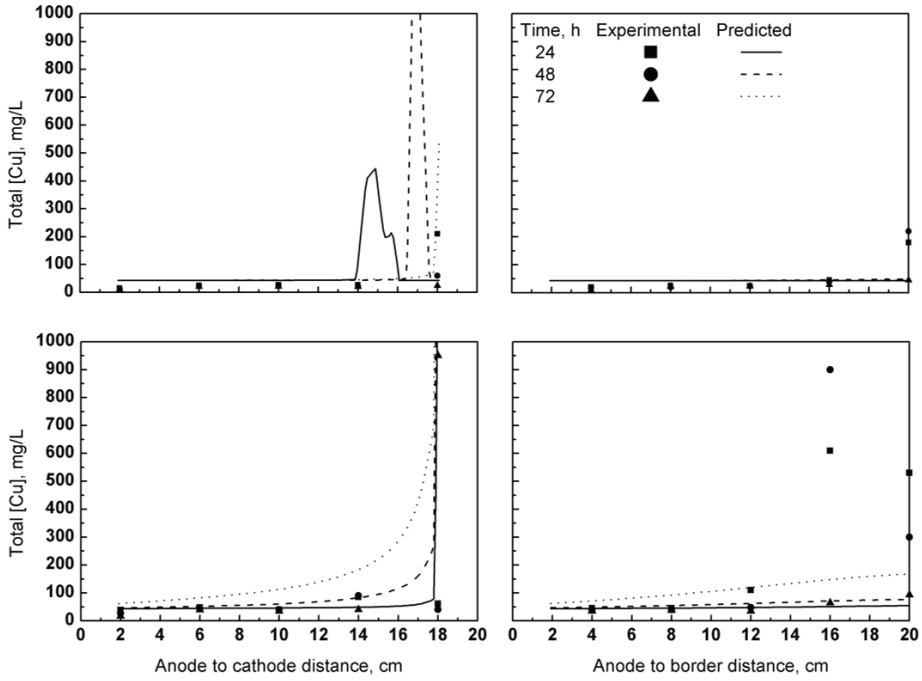


Figure 3. Experimental and predicted values of copper concentration T_4 along lines AC (left) and AB (right) in Figure 2. Initial copper concentration: 100 mg/L (lower plots), and 600 mg/L (upper plots).

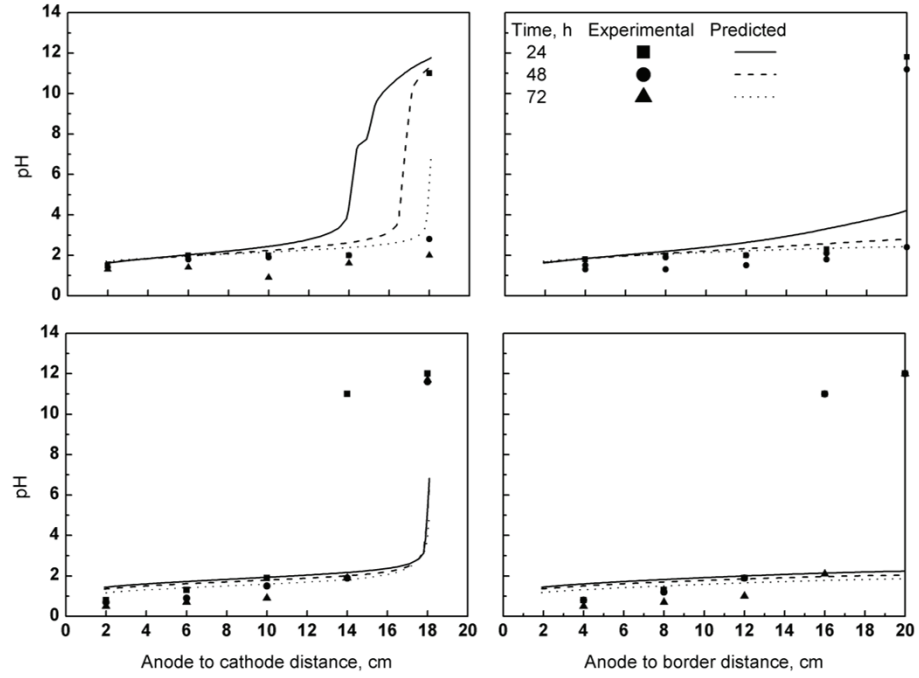


Figure 4. Experimental and predicted values of pH along lines AC (left) and AB (right) in Figure 2. Initial copper concentration: 100 mg/L (lower plots), and 600 mg/L (upper plots).

As far as the mathematical model is concerned, the presence of an abrupt reaction front makes the prediction of the system characteristics difficult. Figure 3 shows that the concentration profiles computed by the model along line AC (left plots) are smooth compared to the experimental profiles. Also, the rate of transport of the copper species within the time frame of zero to 48 h was underpredicted. For 72 hours, the computed values agreed well with the experimental data. Along line AB (right plots), the calculated values showed a reasonable agreement with the experimental data, with the exception of the experiment conducted with 600 mg/L of initial copper, in which case the model overpredicted the rate of copper motion. Despite the complex behavior of the experimental system, the present model was capable of predicting the general trends observed in the experiments. Overall, the model predictions in the proximities of the central anode agreed well with the experimental data in all cases studied. The copper concentration in this region was typically less than 50 mg/L, whereas near the cathodes and the cell boundaries it increased by up to three orders of magnitude.

A relevant feature of the present formulation is its capability to predict the concentration peaks of copper within the soil, as shown in Figure 3. The peak in copper concentration occurs simultaneously with a sudden change in pH across the reaction front. This behavior was also observed experimentally.

Figure 4 shows the predicted and the experimental values of pH for all cases studied. The trends are similar to those observed for copper concentration. Overall, the calculated values agreed well with the experimental values in the vicinity of the anode, whereas the prediction of the reaction front was less accurate. The behavior of the reaction front can be explained as follows. The electrolysis reactions cause significant changes in pH in the vicinity of the electrodes. At the reaction front, the Cu^{+2} ions react with OH^- ions to form $Cu(OH)_2$ according to reaction 6. Because $Cu(OH)_2$ is a neutral molecule, both the solution and the soil in this area increase their electrical resistance. As a result, an isoelectric region appears, and the transport of all the ionic species in this area is hindered.

5. CONCLUDING REMARKS

The two-dimensional mathematical model developed by Jacobs and Probstein [9] was adapted to represent the electrokinetic remediation of an artificially copper-contaminated soil from the city of Cananea, Mexico. The model predictions showed reasonable agreement with experimental data collected in a laboratory cell in terms of copper concentration and pH. For the circular arrangement, the model predicted the formation of an isoelectric region which moves from the central anode towards the cathodes in the periphery. This region is characterized by the presence of high gradients of both copper concentration and pH.

An overall evaluation of the present formulation indicates that the phenomena occurring during the electrokinetic remediation of a natural soil are complex, and require further investigation from both experimental and theoretical perspectives. Despite these difficulties, the present formulation was capable of representing the main features of the process in a two-dimensional soil. It also clarified the potential of the electrokinetic remediation to treat a copper-contaminated soil. Investigation on both experimental and theoretical aspects of this process is currently in progress in this laboratory, and will be the subject of a future publication.

6. NOMENCLATURE

Symbol	Description
C_i	Concentration of species *i,* mol•m^{-3}
D_i	Effective diffusivity of species *i,* m^2•s^{-1}
D_i^0	Diffusivity of species *i* at infinite dilution, m^2•s^{-1}
k	*k*th conserved quantity
K_j	Equilibrium constant for the *j*th chemical reaction, various units
M	Number of chemical reactions
N	Number of chemical species
N_i	Net flux of species *i*, mol•m^{-2}•s^{-1}
R_i	Generation rate of species *i* due to homogeneous chemical reactions, mol•m^{-3}•s^{-1}
t	Time, s
T_k	Concentration of conserved quantity *k,* mol•m^{-3}
T_0	Net electrical charge of the aqueous solution, *C*
α_{ik}	Contribution of species *i* to conserved quantity *k*

ε	Solution electrical permittivity, $F \bullet m^{-1}$
ζ	Soil zeta potential, V
τ	Soil tortuosity, dimensionless
κ_h'	Soil hydraulic permeability, m^2
κ_h	Soil effective hydraulic permeability, m^2
μ	Solution viscosity, Pa•s
ϕ	Electrical potential, V
ψ	Soil porosity, dimensionless
σ_S	Soil electrical conductivity, $\Omega^{-1} \bullet m$
∇	Nabla operator, m^{-1}

REFERENCES

[1] Hernán, S.L., González, L.M. y Espinosa, A., Modelación de elementos traza en el horizonte de suelos, plancha 170 (Vélez, Departamentos de Santander y Boyacá), (in spanish). Dyna, 156, pp. 157-164, 2008.

[2] Shapiro, A.P., Renaud, P.C. and Probstein, R.F., Preliminary studies on the removal of chemical-species from saturated porous media by electroosmosis, Phys. Hydrodyn. 11(5-6), pp. 785-802, 1989.

[3] Probstein, R.F. and Hicks, R.E., Removal of contaminants from soils by electric fields. Science, 260(5107), pp. 498-503, 1993.

[4] Hamed, J., Acar, Y.B. and Gale, R.J., Pb(II) removal from kaolinite by electrokinetics. J. Geotech. Eng.-ASCE, 117(2), pp. 241-271, 1991.

[5] Denisov, G., Hicks, R.E. and Probstein, R.F., On the kinetics of charged contaminant removal from soils using electric fields. J. Coll. Interface Sci., 178(1), pp. 309-323, 1996.

[6] Page, M.M. and Page, C.L., Electroremediation of contaminated soils. J. Env. Eng.-ASCE, 128(3), pp. 208-219, 2002.

[7] Shiba, S., Hirata, Y. and Seno, T., Mathematical model for hydraulically aided electrokinetic remediation of aquifer and removal of nonanionic copper. Eng. Geol., 77(3-4), pp. 305-315, 2005.

[8] Jacobs, R.A., Sengun, M.Z., Hicks, R.E. and Probstein, R.F., Model and experiments on soil remediation by electric fields. J. Env. Sci. Health Part A-Environmental Science and Engineering & Toxic and Hazardous Substance Control, 29(9), pp. 1933-1955, 1994.

[9] Jacobs, R.A. and Probstein, R.F., Two-dimensional modeling of electroremediation. AIChE J., 42(6), pp. 1685-1696, 1996.

[10] Gómez-Álvarez, A., Meza, F.A., Villalba, A., Valenzuela, J., Ramírez, J. and Almendáriz, J., Estimation of potential pollution from mine tailings in the San Pedro river (1993-2005), Mexico-US border. Env. Geol. 57(7), pp. 1469-1479, 2009.

[11] Moreno-Zazueta, S.A., Restauración por electroadsorción de suelos contaminados (In Spanish), [M.S. Thesis], Universidad de Sonora, Mexico, 1999.

[12] Almeira, J., Peng, C. and Wang, Z., Effect of different electrode configurations on the migration of copper ions during the electrokinetic remediation process. Asia-Pac. J. Chem. Eng., 4(5), pp. 581-585, 2009.

[13] Kim, D. H., Jo, S.U., Choi, J.H., Yang, J.S. and Baek, K., Hexagonal two dimensional electrokinetic systems for restoration of saline agricultural lands: a pilot study. Chem. Eng. J., 198-199, pp. 110-121, 2012.

[14] Vereda-Alonso, C., Rodríguez-Maroto, J.M., García-Delgado, R.A., Gómez-Lahoz, C. and García-Herruzo, F., Two-dimensional model for soil electrokinetic remediation of heavy metals. Aplication to copper spiked kaolin. Chemosphere, 54, pp. 895-903, 2004.

[15] Jacobs, A. R., Removal of contaminant material from a soil site. US Patent No. 5415744, 1995.

[16] Carnahan, B., Luther, H.A. and Wilkes, J.O., Applied Numerical Methods, John Wiley & Sons, Inc., 1969.

[17] Butler, J.N. and Cogley, D.R., Ionic equilibrium solubility and pH calculations, John Wiley & Sons, Inc., 1998.

[18] Fuerstenau, M.C. and Palmer, B.R., Anionic flotation of oxides and silicates. In: Flotation – A. M. Gaudin Memorial Volume–Volume 1, (M.C. Fuerstenau, Editor), Chapter 7, American Institute of Mining, Metallurgical, and Petroleum Engineers Inc., New York, pp. 148-196, 1976.

[19] Vanýsek, P., Ionic Conductivity and Diffusion at Infinite Dilution. In: CRC Handbook of Chemistry and Physics (D. R. Lide, editor), CRC Press, Boca Raton, FL, pp. 5-90, 1993.

[20] Hines, A. L. and Maddox, R.N., Mass Transfer: Fundamentals and Applications, Prentice Hall, 1985.

[21] HSC Chemistry for Windows version 5.0 Finland, 1999. http://www.outotec.com

[22] Tecplot Inc., Tecplot 360, R 2012. http://www.tecplot.com/ Bellevue, WA 980062011.

TWO-DIMENSIONAL COMPUTATIONAL MODELING OF THE ELECTROKINETIC REMEDIATION OF A COPPER-CONTAMINATED SOIL PART II: SENSITIVITY ANALYSIS FOR A TRIANGULAR SOIL FIELD

MODELACION COMPUTACIONAL EN DOS DIMENSIONES DE LA REMEDIACION ELECTROCINETICA DE UN SUELO CONTAMINADO CON COBRE PARTE II: ANALISIS DE SENSIBILIDAD PARA UN CAMPO DE SUELO TRIANGULAR

VIRGILIO RUBIO-NIEBLAS
MSc., Universidad Autónoma de Baja California.Mexicali, México, virgilio62@gmail.com

MANUEL PEREZ-TELLO
PhD., Universidad de Sonora. Hermosillo, México, mperezt@iq.uson.mx

RICHARD A. JACOBS
PhD., Independent consultant. Houston, Texas, U.S.A, rajacobs@gmail.com

RONALDO HERRERA-URBINA
PhD., Universidad de Sonora. Hermosillo, México, rherrera@guaymas.uson.mx

Received for review January 26 th, 2013, accepted October 8th, 2013, final version November, 12 th, 2013

ABSTRACT: The computer model described in the first paper of this series was used to perform a sensitivity analysis for the electrokinetic remediation of a triangular, copper-contaminated soil field in which one anode and two cathodes are placed in the vicinity of the triangle vertices. The input variables included the initial concentration of copper in the soil, the electrical potential applied between electrodes, and the absence or presence of wash water. The output variables included the cleanup efficiency and the fraction of copper eliminated from the soil f_{Cu} . Overall, the larger the electrical potential applied between electrodes, the shorter the time to achieve steady state. When no wash water is used, the cleanup efficiency may fluctuate over time before reaching steady state. When wash water is used, the final f_{Cu} obtained depends on the initial concentration of copper in the soil.

Key words: Two-dimensional, mathematical model, electrokinetics, sensitivity analysis

RESUMEN: El modelo computacional descrito en el primer artículo de esta serie se usó para realizar un análisis de sensibilidad del proceso de eliminación electrocinética de un campo de suelo triangular contaminado con cobre, en el cual se colocan un ánodo y dos cátodos en la proximidad de los vértices del triángulo. Las variables de entrada incluyeron la concentración inicial de cobre en el suelo, el potencial eléctrico aplicado entre los electrodos, y la ausencia o presencia de agua de lavado. Las variables de salida incluyeron la eficiencia de limpieza y la fracción de cobre eliminado del suelo f_{Cu} . En general, a mayor potencial eléctrico aplicado entre los electrodos, menor el tiempo requerido para alcanzar el estado estacionario. Cuando no se usa agua de lavado, la eficiencia de limpieza puede fluctuar en el tiempo antes de alcanzar el estado estacionario. Cuando se usa agua de lavado, el valor final de f_{Cu} depende de la concentración inicial de cobre en el suelo.

Palabras clave: Dos dimensiones, modelo matemático, electrocinética, análisis de sensibilidad

1. INTRODUCTION

Heavy metals in soils are important nutrients for plants in low concentrations, but at high concentrations they become toxic [1]. The contamination of soils by heavy metals is nowadays a major concern worldwide. In Part I of this series, a two-dimensional computational model for the electrokinetic remediation of a copper-contaminated soil was presented. The model is an adaptation of the original code developed by Jacobs and Probstein [2], and incorporates electromigration, electroosmosis, ordinary diffusion, and convection transport mechanisms. Based on a previous study reported by Fuerstenau and Palmer [3] on the

distribution of copper species in aqueous solution, the chemical system was represented by 13 chemical species which participate in 8 chemical reactions. The computer model was validated by comparing their predictions with experimental data collected in a laboratory-scale system containing a soil collected at the city of Cananea, Mexico. A reasonable agreement between the predicted and the experimental values was obtained in terms of the spatial distributions of pH and copper concentration when the initial copper content in the soil was 100 and 600 mg/L . Once the computational model was verified, it can be used to analyze the main features of the electrokinetic remediation process. The potential applications of a two-dimensional formulation were pointed out by Jacobs [4], who conducted a computational study in which a number of electrode configurations was tested. Jacobs [4] found that electrode arrangements in which the anodes are surrounded by a number of cathodes provide greater efficiencies than those in which the cathodes are surrounded by the anodes. The simplest configuration obeying this criterion is a triangular arrangement in which one anode and two cathodes are placed in the vicinity of the triangle vertices. In a recent experimental study at laboratory scale, Almeira et al. [5] found that a triangular arrangement consisting of one anode and two cathodes forming three 60° angles between the lines connecting the electrodes was the most efficient in terms of both copper elimination and energy efficiency. Such experimental results [5] and further observations by Almeira et al. [6] and Peng et al. [7] experimentally confirmed the validity of the model predictions reported by Jacobs [4]. Based on the previous results shown in Part I of this series and the analysis of the relevant literature [4-7], the goal of this investigation was to test the effects of the main operating variables on the overall performance of the electrokinetic remediation for a large, triangular, two-dimensional soil field. For that purpose, a number of computer simulations were conducted with the mathematical model described in Part I of this series, as described below.

2. SIMULATION STRATEGY

Figure 1 shows the system under study. An equilateral triangular soil field was assumed in which one anode and two cathodes are placed in the vicinity of the triangle vertices. The test variables included the initial concentration of copper in the soil pores (100 and 600 mg/L) and the electrical potential applied to the electrodes (500 and 1000 V).

The specification of the latter values was based on the criteria reported in the literature [4], which suggests the application of 1 V per centimeter of separation between electrodes. Table 1 shows the numerical parameters used for the simulation runs, in which the properties of the Cananea soil reported in Part I of this series were assumed.

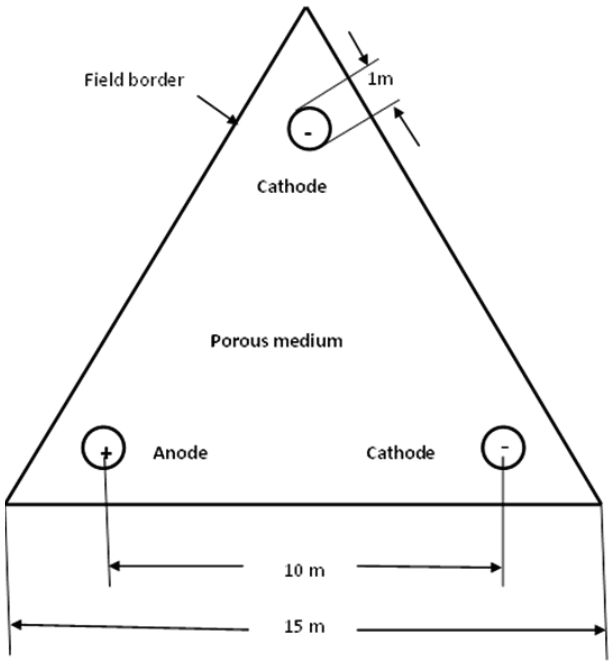


Figure 1. Equilateral triangular soil field for the simulation runs

For every combination of the input variables, computer simulations were done for the two alternative cases in which pure wash water is either injected or not injected through the anode well. Because the mechanisms governing the motion of the chemical species in the soil are the same with and without water injection, the mathematical formulation can be used to predict the performance of the system under both types of operation strategies. For each case, a corresponding indicator of the process yield was defined as follows. In the simplest case, no wash water enters the soil, and thus the copper species are retained within the soil. At sufficiently long times, the copper is expected to concentrate within a localized area in the vicinity of the cathodes.

Table 1. Numerical parameters for the simulation runs

Parameter	Value
Electrode well diameter, m	1
Anode-to-cathode distance, m	10
Anode-to-cathode potential difference ϕ, V	500, 1000
Soil porosity ψ	0.45
Soil tortuosity τ	1.1
Hydraulic permeability κ_h', m^2	10^{-15}
Zeta potential ζ, V	-1.54×10^{-2}
Electrical permittivity ε, F/m	7×10^{-10}
Solution viscosity μ, Pa•s	10^{-3}
Effective hydraulic permeability κ_h, $m^2 \bullet V^{-1}s^{-1}$	2.4×10^{-9}
Soil electrical conductivity σ_s, S/m	0.02875
Anode-to-cathode pressure difference (wash water only)	10^5 Pa

This soil will eventually be collected for further treatment or disposal. A feasible response variable for this case is the fraction of the total area in the arrangement which has been cleaned up to a given time t. In this study, this quantity is referred to as cleanup efficiency (CE); thus

$$\mathrm{CE} = \frac{\left\{\begin{array}{l}\text{Soil area in which} \\ [\mathrm{Cu}] < 50 \text{ mg/L at time } t\end{array}\right\}}{\text{Total soil area}} \qquad (1)$$

In this definition, the 50 mg/L limit appearing in the numerator was established based on local regulations. Although Equation 1 suggests that the CE may take on values in the range of zero to unity, in practical terms it may not reach the value of unity. This is because copper does not leave the system when no wash water is used. The calculation of the numerator in Equation 1 was accomplished within the framework of the Tecplot visualization software [8]. The procedure involves the calculation of the two-dimensional contours of copper concentration from the local concentration values within the computational domain, followed by the numerical calculation of the areas enclosed by those contours which comply the criterion shown in Equation 1.

An alternative way of operation consists of injecting pure wash water through the anode to maintain a continuous flow of the aqueous solution throughout the soil. This flow gradually takes the copper out of the system through the cathode wells. From a computational point of view, this mode of operation was simulated by setting a pressure difference between electrodes, as shown in Table I. In this case, an indicator of the process efficiency is the fraction of the initial copper in the soil which has been eliminated from the soil up to a given time *t*. This quantity is represented by the symbol f_{Cu} and was computed from the following expression:

$$f_{Cu} = \frac{\sum_{j=1}^{2} \int_0^t Q_{wj} C_{wj} dt}{V C_{Cu,0}} \qquad (2)$$

in which the numerator accounts for the amount of copper that left the soil through the cathode wells up to time *t*, and the denominator is the initial amount of copper in the soil field.

3. DISCUSSION OF RESULTS

Figure 2 shows the predicted values of the CE in the absence of wash water in the triangular soil field. For all cases considered, the maximum CE values were predicted to be 0.55 and 0.59 when the electrical potential difference was set to 500 and 1000 V, respectively. Overall, the effect of the electrical potential difference is mostly on the process kinetics. The larger the electrical potential difference, the faster the initial response of the CE, and thus the shorter the time to achieve steady state. This behavior

may be explained in terms of the electroosmosis and electromigration transport mechanisms, which are both dependent upon the electrical potential gradient $\nabla\phi$ according to the following expressions [2]

$$u_{eo} = \frac{\varepsilon\varsigma}{\tau^2\mu}\nabla\phi \quad (3)$$

and

$$u_{ei} = -\upsilon_i z_i F \frac{\nabla\phi}{\tau^2} \quad (4)$$

where u_{eo} and u_{ei} are the electroosmotic and electromigration velocities, respectively, and other symbols are defined in the nomenclature. It is noted that u_{eo} represents the velocity of the aqueous solution adjacent to the pore walls; thus, it affects all species in solution at a given location. On the other hand, the electromigration velocity u_{ei} is only attributed to those chemical species in solution carrying an electrical charge z_i

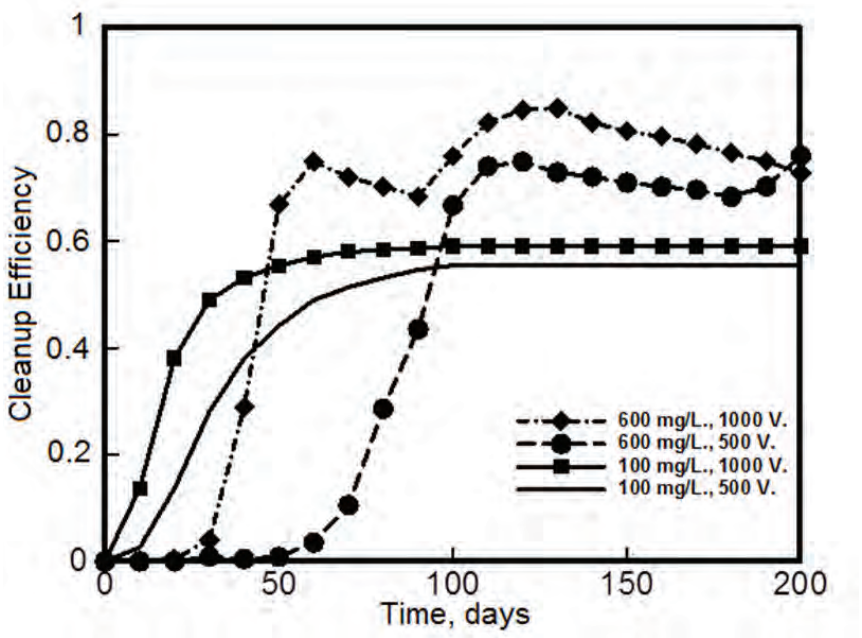


Figure 2. Predicted time evolution of the cleanup efficiency (CE) in the triangular soil field in the absence of wash water injection

z_i . From a macroscopic standpoint, any change in the electrical potential difference between the electrodes will affect the local electrical potential gradient $\nabla\phi$, and thus the kinetics of the overall process described by Equations (3) and (4). On the other hand, it is of interest to note the values of CE at long times, i.e., once steady state was achieved. Such CE values increased as the initial concentration of copper in the soil also increased. Therefore, as far as the cleanup efficiency is concerned, highly contaminated soils are expected to be cleaned up more efficiently than moderately contaminated soils. A complementary response variable is the time to reach steady state conditions. This quantity was defined as the time at which the cleanup efficiency CE varied by less than 0.01 percent from one time step to the next. In this study, this quantity is represented by the symbol: t_∞ . Figure 2 shows that t_∞ was in the range of 90-100 days when the initial copper content was 100 mg/L, whereas it was longer than 200 days when the initial copper content was 600 mg/L.

The initial copper concentration in the soil also affected the kinetics of the CE. Thus, when it was set to 100 mg/L, the CE increased with time up to an asymptotic value which depended upon the electrical potential difference applied to the soil. In contrast, when the initial copper concentration was set to 600 mg/L, the CE values showed significant fluctuations with time before reaching steady state conditions.

Computer visualizations of the concentration field in this time period provided an explanation for this behavior. A simplified diagram of such calculations is shown in Figure 3, which shows that the local values of pH are strongly related to the distribution of the chemical species throughout the soil. At *t*=24 days (Figure 3a) the H^+ ions produced at the anode moved to the cathodes, thus creating an acid zone.

At the same time, the OH^- ions produced at the surface of the cathodes moved toward the anode, thus forming a basic zone. The locations at which both species meet in the soil is characterized by a sudden change in pH values as a result of the neutralization reaction: $H^+ + OH^- \rightarrow H_2O$ The locations at which this reaction occurs is referred to in this paper as the reaction front, and depends upon the mobilities of both H^+ and OH^- ions [2], as well as the concentrations of the anionic and cationic copper species in solution.

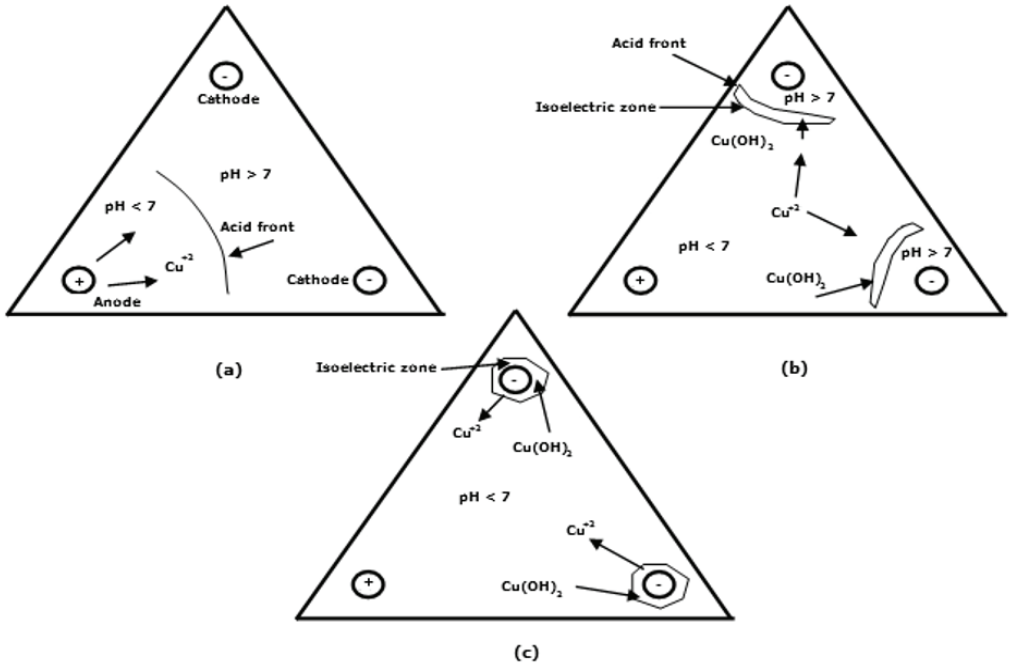


Figure 3. Predicted time evolution of the reaction front in the triangular soil field in the absence of wash water injection; (a) 24 days, (b) 48 days, (c) 72 days

At these locations, both solid and aqueous $Cu(OH)_2$ are formed. Because aqueous $Cu(OH)_2$ carries no electrical charge, it causes the electrical conductivity of the aqueous phase to decrease substantially, thus forming an isoelectric zone. The coexistence of both solid and aqueous $Cu(OH)_2$ occurs in the range of: $6.5 < \text{pH} < 13$, whereas in the range of $4 < \text{pH} < 6.5$ only the aqueous form is present [9]. As the acid front advanced at *t*=48 days (Figure 3b), the reaction front and the isoelectric zone became narrow. This is because the H^+ ions are continuously produced at the anode surface and react to form other cations such as $Cu(OH)_2^+$ and $Cu(OH)_2^{+2}$, which also moved toward the cathodes. On the surface of the cathodes, the OH^- ions are also formed and react to produce other anions such as $HCuO_2^-$ and CuO_2^- which move toward the anode. This process repeats over time (t=72 days, Figure 3c) until steady state is achieved. The fluctuating motion of the copper ions throughout the soil thus explains the fluctuating behavior of CE in Figure 2.

It is noted that the present formulation assumes that the transport of the chemical species occurs in the aqueous phase only, i.e., the solid phase is assumed to be motionless. It is also assumed that the soil is isotropic and its properties do not vary with time. Therefore, any change in the rate of transport of the chemical species may be attributed to changes occurring in the aqueous phase only. The incorporation of anisotropies, time-dependent soil properties, and other phenomena such as electrophoresis was beyond the scope of this investigation.

The results shown in Figure 3 qualitatively agree with the experimental observations by Almeira et al. [5,6] and Peng et al. [7] regarding the formation of three distinctive areas within the soil; namely, and acid region, a basic region, and a narrow area with low electrical conductivity. Overall, the results shown in Figures 2 and 3 indicate that the behavior of this type of a system in the absence of wash water injection is complex, and cannot be generalized. Detailed simulations for specific field dimensions, soil characteristics, and operating conditions must be done in order to analyze the process performance prior to its optimization.

Figure 4 shows the predicted behavior of the fraction of copper eliminated f_{Cu} computed from Equation (2) when pure wash water was used. Under the test conditions, the maximum values of f_{Cu} were predicted to be in the range of 0.8-0.95. At long times, all the curves approach an asymptotic value of f_{Cu} which is dependent upon the initial concentration of copper in the soil, and is independent of the electrical potential difference applied to the electrodes. The asymptotic values of f_{Cu} at sufficiently long times decreased as the initial copper concentration increased. Therefore, in terms of the fraction of copper eliminated from the soil f_{Cu}, the more dilute the copper in the original soil, the more efficient the electrokinetic remediation process.

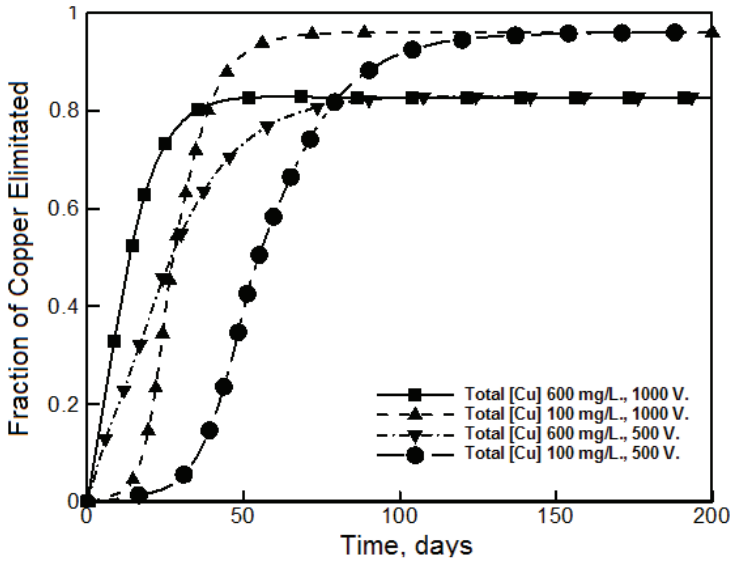


Figure 4. Predicted time evolution of the fraction of copper eliminated in the triangular soil field in the presence of wash water injection

This behavior contrasts with that observed for the cleanup efficiency in the absence of wash water injection shown in Figure 2. The time to reach steady state conditions throughout the soil was defined as the time at which the f_{Cu} value varied less than 0.01 percent from one time step to the next. Figure 4 shows that this quantity is strongly dependent upon the electrical potential difference applied to the electrodes. Thus, the higher the electrical potential, the shorter the time to reach steady state conditions and vice versa. As was discussed previously, this is attributed to the electroosmotic and electromigration velocities, which are dependent upon the electrical potential gradient according to Equations 2 and 3, respectively. The behavior of the system under the presence of wash water injection shows a simpler behavior than that observed when no wash water is injected to the soil. However, further studies are necessary to clarify whether the trends observed in this study for a triangular soil field can be generalized to other electrode arrangements.

4. CONCLUDING REMARKS

The computer simulation of the electrokinetic remediation of a copper-contaminated soil field with a triangular arrangement of electrodes was performed. The results indicated that the electrical potential difference applied between electrodes mostly affects the process kinetics due to the electroosmotic and electromigration velocities.

Also, the initial copper content in the soil determines the final process efficiency. When no wash water is injected to the soil, the final cleanup efficiency increased as the initial content of copper in the soil was increased, and it may fluctuate over time before reaching steady state conditions. Such fluctuations may be explained in terms of the chemical reactions occurring at both sides of the acid front moving throughout the soil during the operation.

On the other hand, when wash water is injected, the final fraction of copper eliminated from the soil depends upon the initial concentration of copper in the soil only. The results obtained in this study illustrate the potential use of the present formulation for the further optimization of this process.

NOMENCLATURE

Symbol	Description
CE	Cleanup efficiency, dimensionless
C_{wj}	Concentration of copper at *j*th well, mol•m^{-3}
$C_{Cu,0}$	Initial copper concentration in soil, mol•m^{-3}
f_{Cu}	Fraction of the initial copper eliminated from soil, dimensionless.
Q_{wj}	Volumetric flow rate of solution at *j*th well, m^3• s^{-1}
t	Time, s
t_∞	Time to achieve steady state, s
u_{ei}	Electromigration velocity of species *i*, m• s^{-1}
u_{eo}	Electroosmotic velocity of solution, m• s^{-1}
V	Total volume of solution in soil, m^3
z_i	Charge number of species *i*, dimensionless
ε	Solution electrical permittivity, F•m^{-1}
ζ	Soil zeta potential, V
τ	Soil tortuosity, dimensionless
υ_i	Mobility of species *i*,
κ'_h	Soil hydraulic permeability, m^2
κ_h	Soil effective hydraulic permeability, m^2
μ	Solution viscosity, Pa•s
ϕ	Electrical potential, V
ψ	Soil porosity, dimensionless
σ_S	Soil electrical conductivity, Ω^{-1}•m
∇	Nabla operator, m^{-1}

REFERENCES

[1] Hernán, S.L., González, L.M. y Espinosa, A., Modelación de elementos traza en el horizonte de suelos, plancha 170 (Vélez, Departamentos de Santander y Boyacá), (in spanish). Dyna, 156, pp. 157-164, 2008.

[2] Jacobs R.A., Probstein, R.F. Two-dimensional modeling of electroremediation. AIChE J., 42(6), pp. 1685-1696, 1996.

[3] Fuerstenau, M.C. and Palmer, B.R., Anionic flotation of oxides and silicates. In: Flotation – A. M. Gaudin Memorial Volume–Volume 1, (M.C. Fuerstenau, Editor), Chapter 7, American Institute of Mining, Metallurgical, and Petroleum Engineers Inc., New York, pp. 148-196, 1976

[4] Jacobs, A. R., Removal of contaminant material from a soil site. US Patent No. 5415744, 1995.

[5] Almeira, J., Peng, Ch. and Wang, Z., Effect of different electrode configurations on the migration of copper ions during the electrokinetic remediation process. Asia-Pac. J. Chem. Eng., 4, pp. 581-585, 2009.

[6] Almeira, J., Peng, Ch. and Li, P., Effect of electrode configuration on the distribution of Cu during electrokinetic soil remediation. J. Kor. Soc. Urb. Env., 10(2), 2010.

[7] Peng, Ch., Almeira, J. and Gu, Q., Effect of electrode configuration on pH distribution and heavy metal ions migration during soil electrokinetic remediation. Env. Earth Sci., 69, pp. 257-265, 2013.

[8] Tecplot Inc., Tecplot 360, R 2012. http://www.tecplot.com/ Bellevue, WA 980062011

[9] Jacobs, R.A., Sengun, M.Z., Hicks, R.E. and Probstein, R.F., Model and experiments on soil remediation by electric fields. J. Env. Sci. Health Part A-Environmental Science and Engineering & Toxic and Hazardous Substance Control, 29(9), pp. 1933-1955, 1994.

EVALUATION OF STRESS CRACKING ON GEOMEMBRANES AFTER ACCELERATED TESTS

EVALUACIÓN DE FISURACIÓN BAJO TENSIÓN DE GEOMEMBRANAS POS ENSAYOS ACELERADOS

FERNANDO LUIS LAVOIE
MSc, Civil Engineering, University of São Paulo (USP), Brazil, fllavoie@yahoo.com.br
BENEDITO DE SOUZA BUENO
PhD, Full Professor, University of São Paulo (USP), Brazil, bsbueno@sc.usp.br
PAULO CÉSAR LODI
PhD, Associate Professor, UNESP - São Paulo State University, Brazil, plodi@feb.unesp.br

Received for review October 29 th, 2012, accepted November 1th, 2013, final version December, 9 th, 2013

ABSTRACT: This paper presents results of stress cracking tests performed in high density polyethylene (HDPE) geomembranes (GM). Stress cracking tests were performed in accordance to ASTM D5397: Notched Constant Tensile Load Test (NCTL) and Single Point-Notched Constant Tensile Load Test (SP-NCTL). Tests were conducted to the fresh sample at 50ºC (standard test) and at 70ºC (accelerated condition) in order to compare the SC values. Results from accelerated tests (NCTL) showed, for instance, a total economy of 390 hours (comparing load stages of 25% yield stress) to perform the tests.

Key words: HDPE geomembranes, stress cracking, accelerated tests.

RESUMEN: Este trabajo presenta resultados de fisuración bajo tensión realizados con geomembranas (GM) de polietileno de alta densidad (HDPE). Las pruebas de tensión fueron realizados de acuerdo a la norma ASTM D5397: Notched Constant Tensile Load Test (NCTL) and Single Point-Notched Constant Tensile Load Test (SP-NCTL). Estos ensayos fueron realizados en muestras de control a 50 º C (ensayo padrón) y a 70 º C (condición acelerada) para comparar valores de fisuración bajo tensión. Los resultados de estos ensayos acelerados (NCTL) mostraron una economía total de 390 h (comparando etapas de carga de 25% de la tensión de ruptura) para los ensayos realizados.

Palabras clave: Geomembranas de polietileno de alta densidad, fisuración bajo tensión, pruebas aceleradas.

1. INTRODUCTION

Stress cracking (SC) is an external or internal cracking in plastic induced by tensile stress less than its short-term mechanical strength [1]. Stress cracking occurs in a brittle manner with little or no elongation near to the crack surface [2].

Halse et al. [3] and Peggs and Carlson [4] claim that failure due to stress cracking is associated with defects or imperfections that cause the stresses to be enhanced to higher values with up to a 6-fold magnification of tensile stress (relative to the average global stress) depending on the geometry of the defect. The defects may be of various types and shapes and generally include surface scratches, grinding gouges, patches, and seams. In addition, the presence of an external chemical environment such as detergents, surfactants, leachate, polar vapor, or liquid, may accelerate stress cracking. Stress cracking in the presence of chemicals is called "environmental stress cracking" [5, 6].

Polymers used in the fabrication of geosynthetics, such as polyethylene (PE), polyester (PET), unplasticized and plasticized poly vinyl chloride (PVC), are subjected to environmental stress cracking [7]. Among polyolefin polymers, polypropylene (PP) is less sensitive to SC than PE when it is associated with a very aggressive chemical environmental. The deformation limit that activates the phenomenon in PP is not yet well-known. The other polymer properties that affect susceptibility to stress cracking include the molecular weight and the co-monomer content [8, 9].

Higher molecular weight corresponds to longer chains [10], resulting in more tie molecules and more effective

tie molecule entanglements [11]. Similarly, high co-monomer content and longer co-monomer short-chain branches provide better cracking resistance, most likely because portions of the long-branch chains cannot be folded into the lamellae and therefore contribute to the amorphous tie molecules [11, 12].

High density polyethylene (HDPE) is a widely used polymer for manufacturing geomembranes used in liquid and waste containment facilities and/or used as a part of liner systems in modern landfills [13]. The primary function of GM is to provide a barrier to advective and diffusive migration of contaminants [14]. The relatively high crystallinity (40 to 50%) of the material provides both high chemical resistance and low diffusion rates, which are required in most containment facilities [6, 15]. However, despite its good chemical resistance, one of the concerns raised in using HDPE geomembranes is their susceptibility to stress cracking (SC) which, in turn, is a consequence of their high crystallinity [2, 16, 6, 13, 15].

The evaluation of stress cracking is performed according to ASTM D5397 [17]. The test is called Notched Constant Tensile Load (NCTL) and uses notched dumbbell shaped specimens placed under various tensile stresses. The tensioned specimens (usually 20) are immersed in a bath containing 10% Igepal / 90% water solution at 50ºC to accelerate the crack growth. A notch is introduced at the central constant-width section on the face of the specimen. The depth of the notch is such that the ligament thickness is equal to 80% of the nominal sheet thickness. The applied stresses typically range from 20% to 50% of the room temperature yield stress (σ_{yield}) in increments of 5%. Three replicate specimens are tested at each stress level, and the failure time of each individual specimen is recorded to an accuracy of 0.1 hour. When a specimen fails, its failure time is recorded by a timer. The test data is presented by plotting the applied stress versus average failure time on a log-log scale. Unfortunately, the full test takes a long time to complete (generally over 10,000 hours). Thus, the Single Point Notched Constant Tensile Load (SP-NCTL) test was developed and is included in ASTM D5397 [17] as an appendix (to be used as a quality control or conformance test). The concept is to select a stress level near, but slightly lower than the transition stress, and to specify the minimum failure time at that stress. A single applied stress of 30% yield stress is utilized with a minimum failure time in GRI-GM13 [18] (Specification for HDPE Geomembranes). In 2003, the specification was revised by increasing the minimum failure time from 200 hours to 300 hours to further enhance the SCR of HDPE geomembranes [6, 19].

Rowe and Sangam [12] concluded that stress cracking is important because: (a) even short cracks can allow excessive leachate through the geomembrane that may readily move laterally in areas of poor contact between the geomembrane and the underlying clay; and (b) short cracks can grow with time eventually allowing excessive leakage through the geomembrane even in areas of good contact with the clay. In either case, once the leakage increases substantially, the geomembrane ceases to perform the barrier function for which it was designed as discussed by Rowe et al. [20].

Several investigators have reported the vulnerability of HDPE geomembranes to stress cracking: Fisher [21], Peggs and Carlson [4], Hsuan et al. [2], Hsuan [6], Rowe and Sangam [12] and Rowe et al. [15].

The transition from ductile-to-brittle failure requires the knowledge of stress level, stress concentration factor, temperature and surrounding environment. However, the fundamental governing factor is the polymer's characteristics, among which crystallinity and molecular weight are the most important. In this sense, it is important to evaluate the SCR of an HDPE GM to assess its long-term performance.

As previously mentioning the standard test is performed in a bath containing 10% Igepal / 90% water solution at 50ºC to accelerate the crack growth. However, the effect of higher temperatures are not yet well know.

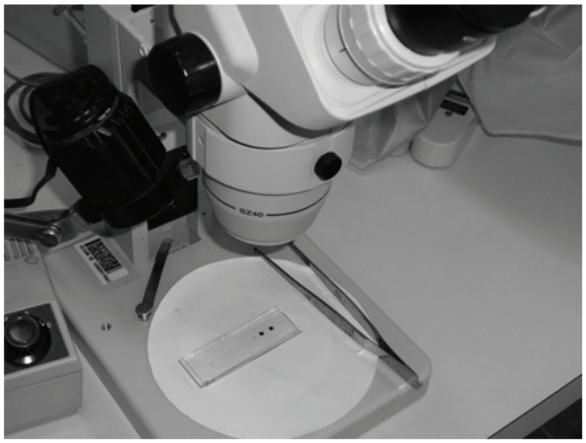
This paper presents results of SC tests performed in HDPE samples. The tests were conducted at 50ºC (standard test) and at 70ºC (accelerated condition) in order to verify the effect of the temperature and compare the SC values in both conditions. For this purpose, equipment was developed to process 20 specimens simultaneously. This equipment includes electronic acquisition of the failure times.

2. MATERIAL AND METHODS

Smooth HDPE geomembranes of 2.0-mm nominal thickness were used. Tests of SC (NCTL and SP-NCTL) were conducted at 50ºC and at 70ºC (accelerated

condition) to compare the SC values according to ASTM D5397 [17]. Additionally, the dispersion of carbon black (Fig.1) was evaluated in accordance with ASTM D5596 [22] to verify the degradation of the material.

(a)

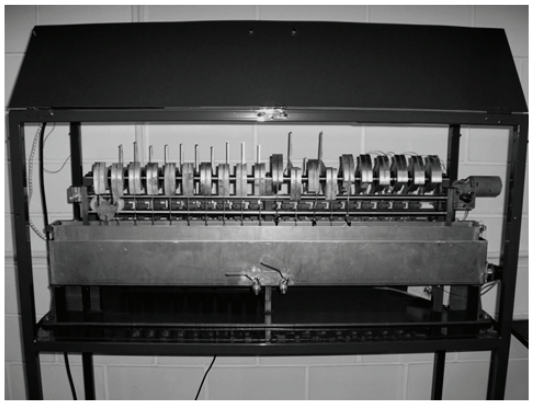


(b)

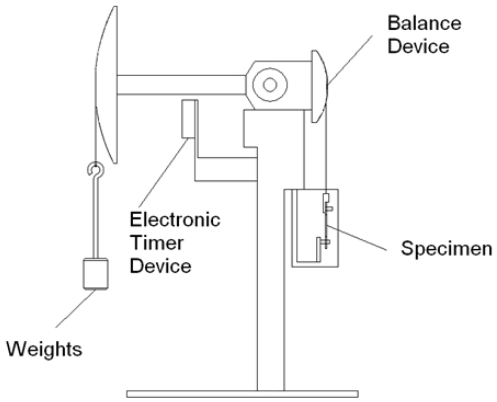
Figure 1. Preparation of the dispersion of carbon black test (a) cut specimens (b) microscopic evaluation.

2.1. Developed stress cracking test equipment

The equipment used in the SC tests was developed to process 20 specimens simultaneously. This equipment includes electronic acquisition of the failure times (Fig. 2). Force is applied to the specimen by a lever with metallic weights in its extremity. This lever applies a force equal to three times the load. The incubation system allows the mechanical control of temperature.



(a)



(b)

Figure 2. Test Setup

3. RESULTS AND DISCUSSION

3.1. Carbon black dispersion

Carbon black is the most common type of UV protection for polymeric products, and it consists of very fine particles (the primary particles) fused together to form the primary aggregates (Fig. 3). The UV absorbing efficiency of carbon black is governed by the average prime particle size. Primary aggregates composed of finer prime particles present a greater surface area for optical absorption than the primary aggregates composed of larger prime particles.

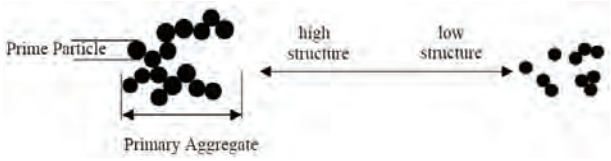


Figure 3. Schematic drawing of carbon black particles [23]

Thus, UV absorption increases as prime particle size decreases. However, with prime particles below 20 nm, the UV stabilizing efficiency tends to level off as light scattering becomes more significant with a further decrease in particle size. The carbon black particle size used for the UV protection of polymers used for geosynthetics is typically in the range of 22–25 nm [24, 23, 19]. The results obtained show that all specimens belong to the category I, which contains carbon black particles of circular geometries with diameters less than or equal to 35 μm. In this research, the results of carbon black tests were not efficient in detecting degradation in the samples.

3.2. Accelerated NCTL and SP-NCTL tests

Fig. 4 presents the NCTL curves and the failure times of the SP-NCTL tests. Table 1 and 2 presents the parameters of the NCTL tests and the SP-NCTL results concerning the samples at 50ºC and 70ºC, respectively.

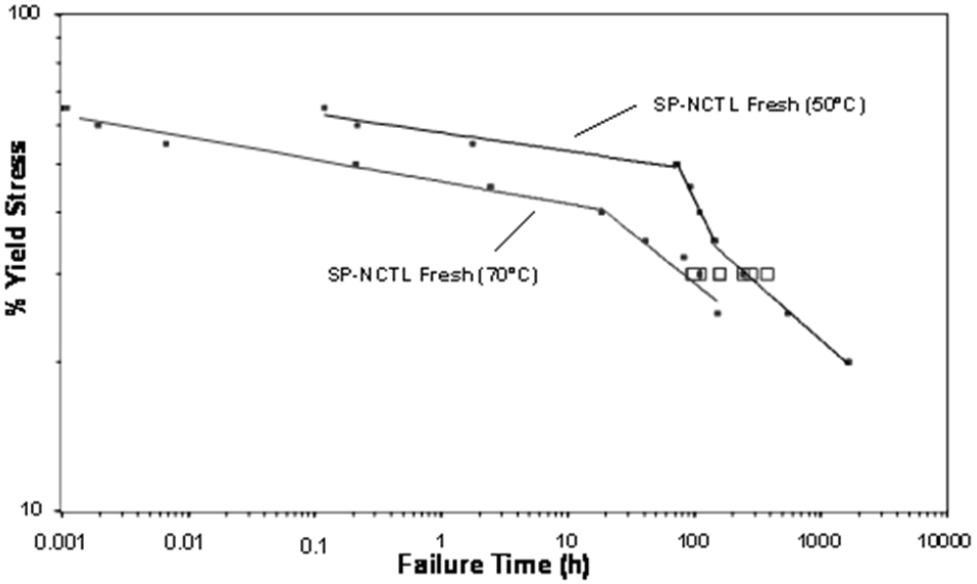


Figure 4. Results of NCTL tests at 50ºC and 70ºC and failure times (SP-NCTL)

Table 1. Parameters of the NCTL tests

Sample	Ductile Slope (%/h)	Brittle Slope (%/h)	T_t (h)	σ_t (%)
HDPE (50ºC)	-0.038	-0.23	143	35
HDPE (70ºC)	-0.045	-0.20	18	40

T_t = Transition time; σ_t = stress transition

Table 2. Average failure times (T_f) of the SP-NCTL tests

Sample	T_f (h)	CV (%)	Variation (%)
HDPE (50ºC)	243	38	-
HDPE (70ºC)	101	6	58 (decrease)

CV = coefficient of variation

The accelerated test (70ºC) presented a bi-linear curve (different than the test at 50ºC). The highest loading stages at 70ºC presented 6 stages of loading in the ductile region and only 4 stages of loading at 50ºC. The slope of the straight ductile accelerated test showed a slight increase when compared to the sample at 50ºC. The transition time (T_t) obtained in the accelerated test presented an 87% decrease when compared to the test at 50ºC and had a higher stress transition (σ_t). The brittle region of the curve for the accelerated test presented 5 stages of loading and 4 stages for the test at 50ºC. The slope of the accelerated tests decreased

when compared to the tests at 50ºC due mainly to the loading stage of 32.5% σ_{yield}.

Accelerated test results have shown time savings when tests with high temperature are used: a total economy of 390 h, comparing load stages of 25% σ_{yield}. However, the behavior of the sample under high temperature was completely different when it was compared to the test at 50°C (the type of curve and in the stress transition, σ_t, obtained).

The increase in the incubation temperature (70ºC) decreased the average failure times by 58%. Accelerated test results provided more homogeneous results , resulting in lower coefficients of variation as well.

Regarding the accelerated tests, the average time to failure (SP-NCTL) was close to the failure time of the NCTL (30%).

The results show that the sample maintains the same trend verified in the NCTL tests. There was a total economy of 265 h compared to the standard test (50°C).

4. CONCLUSION

From the data presented in this paper the following conclusions can be drawn:

- The SC equipment developed to register the failure times worked well producing precise failure times;
- The carbon black tests were not efficient to detect degradation in the GM samples;
- Results from accelerated tests (NCTL) showed time savings (390 h less compared to a load stage of 25% σ_{yield}). However, the behavior of the GMs under high temperatures was completely different when compared to the standard test at 50°C. This fact is discussed and evaluated by Hsuan and Koerner [25] who say the transition time changes in a systematic manner with changes in test temperature. The recommended test temperature for performing the control NCTL test (and the associated SP-NCTL test) will be material dependent and must be decided upon accordingly.
- Results from accelerated SP-NCTL tests showed a total economy of 265 h (compared to the standard test at 50°C), and the same trend was verified in the NCTL tests.

REFERENCES

[1]. ASTM D883, Terminology Relating to Plastics, Annual Book of ASTM Standards, vol. 8.01, American Society for Testing and Materials, West Conshohocken, Pennsylvania, USA, 2012.

[2]. Hsuan, Y.G., Koerner, R.M. and Lord Jr, A.E., Stress-Cracking of High-density Polyethylene Geomembranes, Journal of Geotechnical Engineering, ASCE, 119, pp. 1840-1855, 1993.

[3]. Halse, Y.H., Koerner, R.M., Lord Jr, A.E., Laboratory evaluation of stress cracking in HDPE geomembrane seams. In: Koerner, R.M. (Ed.), Durability and Aging of Geosynthetics. Elsevier Applied Science Publisher, pp.177–194, 1989.

[4]. Peggs, I.D. and Carlson, D.S., Stress cracking of polyethylene geomembrane: field experience. In: Koerner, R.M. (Ed.), Durability and Aging of Geosynthetics, Elsevier Applied Science Publishers Ltd., Amsterdam, Netherlands, pp.195–211, 1989.

[5]. Kulshreshtha, A.K., Chemical degradation. In: Hamid, A.H., Amin, M.B., Maadhah, A.G. (Eds.), Handbook of Polymer Degradation. Marcel Dekker, New York, USA, pp.55–94, 1992.

[6]. Hsuan, Y.G., Data Base of Field Incidents Used to Establish HDPE Geomembrane Stress Crack Resistance Specifications, Geotextiles and Geomembranes, Elsevier Science Publishers Ltd, 18, pp. 1-22, 2000.

[7]. Bright, D.G., The Environmental Stress Cracking of Polymers Used In Geosynthetic Products, Geosynthetics ‘93, IFAI, Vancouver, Canada, 1, pp. 925-933, 1993.

[8]. Lu, X. and Brown, N., The transition from ductile to slow crack growth failure in a copolymer of polyethylene. Journal of Materials Science 25, pp. 411–416, 1990.

[9]. Lu, X. and Brown, N., Unification of ductile failure and slow growth in an ethylene–octene copolymer. Journal of Materials Science 26, pp. 612–620, 1991.

[10]. Apse, J. L., Polyethylene resins for geomembrane

applications. In: Koerner, R.M. (Ed.), Durability and Aging of Geosynthetics.Elsevier Applied Science Publisher, Amsterdam, Netherlands, pp. 159–176, 1989.

[11]. Lustiger, A. and Rosenberg, J., Predicting the service life of polyethylene in engineering applications. In: Koerner, R.M. (Ed.), Durability and Aging of Geosynthetics. Elsevier Applied Science, Amsterdam, Netherlands, pp. 212–229, 1989.

[12]. Rowe, R.K. and Sangam, H.P., Durability of HDPE geomembranes. Geotextiles and Geomembranes, 20, pp. 77–95, Elsevier Science Publishers Ltd, 2002.

[13]. Rowe, R.K., Quigley, R.M., Brachman, R.W.I. and Booker, J.R., Barrier Systems for Waste Disposal Facilities, 2nd ed. Spon Press. P. 587, 2004.

[14]. Rowe, R.K., Long-term performance of contaminant barrier systems. Geotechnique 55 (9), pp. 631–678, 2005.

[15]. Rowe, R.K., Rimal, S. and Sangam, H., Ageing of HDPE geomembrane exposed to air, water and leachate at different temperatures. Geotextiles and Geomembranes 27, pp. 137-151, 2009.

[16]. Wright, D.C., Environmental Stress Cracking of Plastics, 1st ed., Rapra Technology Limited, Shawbury, UK, 1996.

[17]. ASTM D5397, Standard Test Method for Evaluation of Stress Crack Resistance of Polyolefin Geomembranes Using Notched Constant Tensile Load Test, American Society for Testing and Materials, West Conshohocken, Pennsylvania, USA, 2012.

[18]. GRI-GM13, The Stress Crack Resistance of HDPE Geomembrane Sheet, Geosynthetic Research Institute, Philadelphia, USA, 2006.

[19]. Hsuan, Y.G., Schroeder H. F., Rowe, K., Müller, W., Greenwood, J., Cazzuffi, D. and Koerner, R.M., Long-term performance and lifetime prediction of geosynthetics, In: EuroGeo 4 - 4th European Geosynthetics Conference, Edimburgo, Scotland, pp. 01 – 40, 2008.

[20]. Rowe, R.K., Hsuan, Y.G., Lake, C.B., Sangam, P. and Usher, S., Composite liner for a lagoon after 14 years of use.Proceedings of the Sixth International Conference on Geosynthetics, Vol.1, Atlanta, March.Industrial Fabric Association International, St Paul, MN, USA, pp.191–196, 1998.

[21]. Fisher, G.E., Controlling thermal damage in flexible membrane liners. Geotechnical Fabrics Report 7 (2), pp. 39–41, 1989.

[22]. ASTM D5596, Standard Test Method for Microscopic Evaluation of the Dispersion of Carbon Black in Polyolefin Geosynthetics. Philadelphia, American Society for Testing and Materials, West Conshohocken, Pennsylvania, USA, 2009.

[23]. Suits, L.D. and Hsuan, Y.G., Assessing the photo-degradation of geosynthetics by outdoor exposure and laboratory weatherometer. Geotextiles and Geomembranes 21 (1), pp. 111-122, 2003.

[24]. Cabot, Carbon blacks for protection of plastics exposed to ultraviolet light. CABOT Technical Report S–114, Cabot Corporation, Billerica, MA, P. 7, 1990.

[25]. Hsuan, Y.G. and Koerner, R.M., The single point-notched constant tension load test: a quality control test for assessing stress crack resistance, Geosynthetics International 2 (5), pp. 831-843, 1995.

SMART METERS ADOPTION: RECENT ADVANCES AND FUTURE TRENDS

ADOPCIÓN DE MEDIDORES INTELIGENTES: AVANCES RECIENTES Y TENDENCIAS FUTURAS

NATALIA CASTAÑO JARAMILLO
Ingeniera de Sistemas e Informática, Universidad Nacional de Colombia, ncastanj@unal.edu.co
CARLOS JAIME FRANCO CARDONA
Profesor Titular, Universidad Nacional de Colombia, cjfranco@unal.edu.co
JUAN DAVID VELÁSQUEZ HENAO
Profesor Titular, Universidad Nacional de Colombia, jdvelasq@unal.edu.co

Received for review July 30th, 2013, accepted December 20th, 2013, final version January, 30th, 2014

ABSTRACT: High growth in electricity demand and peaks in the load curve, caused mainly by households, require big investments in infrastructure that are used for short periods. Because of this, it is necessary to look for new developments that allow meeting the needs of users as well as using the electricity system resources efficiently. This is possible through the Smart Grid (SG), which additionally allows users to have autonomy in the electricity supply chain. Our focus of investigation is with households because they can monitor their demand and help to reduce the peaks of the load curve. To do this, users must use Smart Meters, because these devices allow consumers to obtain the information necessary to control their demand. This paper presents a systematic analysis of published literature related to the study of the SG from the demand side, analyses the current situation of this topic and the impact of Smart Meter penetration in households.

KEYWORDS: Demand side management; households; peak demand; Smart Grid; Smart Meters.

RESUMEN: El alto crecimiento en la demanda de electricidad y los picos en la curva de carga, causados principalmente por hogares, hacen necesarias grandes inversiones en infraestructura que solo es usada para periodos cortos. Esto ocasiona la búsqueda de desarrollos que permitan suplir las necesidades de los usuarios y usar los recursos del sistema eficientemente. Esto es posible por medio de las Redes Eléctricas Inteligentes, las cuales adicionalmente brindan a los usuarios autonomía en la cadena de suministro eléctrico. El foco de esta investigación son los hogares, ya que estos pueden monitorear su consumo y ayudar a reducir los picos en la curva de carga, Para esto los usuarios deben usar Medidores Inteligentes, los cuales le permiten obtener información necesaria para controlar su demanda. Este artículo presenta un análisis sistemático de la literatura publicada relacionada con el estudio de las redes eléctricas inteligentes desde el lado de la demanda, analiza la situación actual sobre este tema y el impacto de los medidores inteligentes en los hogares.

PALABRAS CLAVE: Gestión de la demanda; hogares; pico de demanda; Redes Eléctricas Inteligentes; Medidores Inteligentes.

1. INTRODUCTION

Several reasons justify the well-known necessity of reducing daily electricity consumption and flattening the peaks in the daily power curve with the aim of reducing investments and operational costs. First, the power grid constantly needs new investments because the demand for electricity is growing rapidly, and it is necessary to build new generating plants [1] and the infrastructure to connect them; moreover, installed generating capacity grows at a lower rate than electricity consumption, so that it is always necessary to make new investments in generation plants and power grid infrastructure. Second, inefficiency, maintenance and upgrade costs are growing with the size of the grid [2]. Third, growing environmental costs caused by new investments in generating plants and in the expansion of the power grid to connect them; also, it is necessary to account for the environmental impact of greenhouse emissions caused by generating plants, especially for coal-fired plants that are considered to be the major sources of pollution [3]. Fourth, the pronounced peaks and flattened troughs in the hourly load curve causes the inefficient use of generating plants so that it is necessary to install additional capacity just to satisfy the peaks in the power curve during short time periods [4]. Such peaks and low flat regions are illustrated in a typical load curve for the Colombian electricity market in Figure 1. As a consequence, electricity costs are inflated and the technologies in use are less efficient.

In addition, newly installed capacity is required only for attending the daily peak demand.

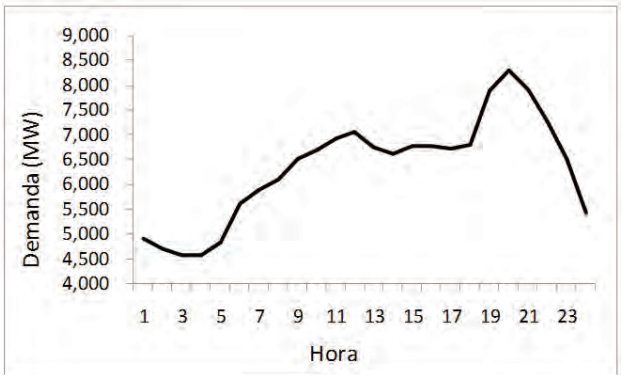


Figure 1. Typical Colombian load curve [5]

The Smart Grid (SG) paradigm is a current and very important technological trend with a strong impact throughout the electricity market affecting it's future development. The Smart Grid (SG) is defined as the full integration of automatic meter reading in the current power grid, and the associated infrastructure, which is conformed by sensors and the proper support software [6]. Additionally, there are many new technologies that are being developed such as controllable appliances, distributed generation, systems to store energy, among others [7].

In the SG, the behavior of the connected users is integrated in an intelligent way in order to use electricity in an efficient, sustainable, economic and secure manner [8, 10]. Moreover, business processes, objectives, and needs are met efficiently [4]. SG promotes and improves the efficient use of all available resources and provides more information and autonomy to the agents in the electricity supply chain [9].

By using intelligent devices (provided by SG technology), electricity utilities will be able to know the consumption, instantaneous condition of distribution networks, electricity prices and to control the loads with complete control and real-time system status; as a consequence, electricity utilities will be able to optimize the use of the network, to make better decisions, and to control customers consumption [1,11].

Users, by using intelligent devices [1], will be able to have more information about the system, its status, and characteristics [12]; their consumption habits; and electricity costs in real time [13,14]. With this information, they are able to reduce a part of the demand or to shift it to another period when the cost of electricity is cheaper [15] with the aim of mitigating costs and saving on their bills [16,17,18]. This is a financial incentive that encourages changes in consumption habits [1]. These changes in consumption habits due to economic signals are known as Demand Response (DR) [1,19,20].

By promoting investments in SG technology as an alternative to installing more generation capacity to attend the load curve peaks [21], the regulator enables energy price reduction, promoting a more efficient and competitive market, and also allows the system to support the use of renewable energy resources, distributed generation and advanced metering [11].

SG promotes the reduction of the high rates of growth in electricity demand and the flattening of the peaks in the power curve [18]. This is explained by, first, the effects of the electricity utilities decisions in order to optimize network use; and second, by the reduction of peak height and depth of the valleys in the load curve [22]. Several facts support the previous conclusions; according to the Electric Power Research Institute (EPRI), the use of electricity in the residential sector is equivalent to 38% of the total electricity consumption in the USA; thus, changes in consumption habits will modify the load curve. The feedback to users can result in a 9% reduction in household consumption [13,23] and a 10–20% reduction during peak demand periods [24,25]. In the European Union it has been detected that when 80% of the consumers reduce their demand in peak hours because of changes in the price, the reduction in associated capacity and transmission costs would be € 67 billion [26].

These facts explain the penetration of smart meters; previously Capgemini, forecasted that in 2012, from 25 to 40% of European households would have smart meters [27]. Other improvements derived from the use of SG are listed in Table 1.

Despite all the above benefits, SG technology is young and new developments are made rapidly so that for the researcher and practitioner it is very difficult to obtain a complete view of the advances in this area.

Table 1. Smart Grid vs. current system [4].

Current System	Smart Grid
Electromechanical	Digital
One-Way Communication	Two-Way Communication
Centralized Generation	Distributed Generation
Few Sensors	Sensors Throughout
Limited Control	Pervasive Control
Manual Restoration	Self-Healing
Failures and Blackouts	Adaptive and Islanding
Few Customer Choices	Many Customer Choices

Due to the importance of this topic, it is necessary to identify, organize, and evaluate the published manuscripts related to SG technology in order to establish: the current state of development, the current research lines, the main contributions of smart grid technology, the most appropriate research directions to be taken, and the most important new contributions that could be implemented.

The objective of this paper is to answer the following questions by means of an exhaustive literature research using the methodology of Systematic Literature Review:

Q1: Which have been the most influential articles in the area?

Q2: What has been implemented regarding SGs in Colombia?

Q3: How do SGs affect the electricity demand?

Q4: What SG technologies are installed in households?

Q5: How have the SG penetrations been modeled?

Q6: Why is it important to study consumer behavior?

This paper is organized as follows. A description of the research process is described first, followed by the systematic review of results. Then the research questions are answered.

2. REVIEW METHODOLOGY

In our research, we employ the methodology of Systematic Literature Review (SLR) that is rooted originally in the field of evidence-based medicine, and adapted to the fields of software engineering by Kitchenham and others [28,29] and to energy policy by Sorrell [30]. SLR is a scientific research technique characterized by a systematic and ordered procedure for searching, identifying and organizing written evidence with the aim of answering several research questions in an unbiased way.

3. RESEARCH PROTOCOL

In this section, our research protocol, based on the guidelines proposed in [28-30], is described.

- Terms used in the search string: smart grid, adoption, demand peak, modeling, smart houses, smart devices, technologies for home, Demand Side Management (DSM).
- Subject areas: engineering, energy, computer science, business, management and accounting, decision sciences, economics, econometrics and finance.
- Source: Scopus.
- Search period: available information until December of 2012
- Inclusion and exclusion criteria: after obtaining the search results that meet the criteria defined above, a manual filtering process was performed in which the articles that meet any of the research questions proposed are selected. The criteria used are articles that speak about:
  - Changes in the consumer habits due to Smart Meter utilization.
  - Demand Peaks and approaches to solve them.
  - Smart Meters penetration and their improvement of the electricity system.
  - Demand Side Management and household response.
  - The Smart Grid in Colombia.
  - SG technologies implemented in homes.

Other articles related to the topic of study were included at the suggestion of experts.

4. RESULTS

By using the automatic search procedure described in the previous section, we obtained 851 articles; next, inclusion and exclusion criteria were applied and the quantity of selected manuscripts was reduced to 49. In addition, eleven new manuscripts were added at the suggestion of experts. Finally, we obtained 60 articles.

By analyzing the selected articles, we found the following characteristics of the body of knowledge: First, research on smart grid is a young area of rapid development and growth. There are no papers before 2008. Figure 2 presents the number of articles published by year.

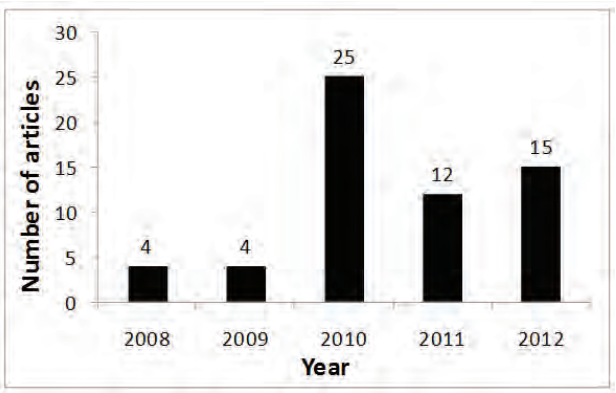


Figure 2. Distribution of articles - inclusion and exclusion criteria and experts suggestions

Second, we found six papers with more than 50 citations; they are presented in Table 2. Note that, all of them are related to methodological aspects of the smart grid.

Three, the most cited article was written by Ipakchi and Albuyeh in 2009 [11] and it is focused on explaining some of the causes of the high investment being made in SG and some possible improvements that are obtained with this approach.

Fourth, in Table 3 we present the conference proceedings and journals where the 60 selected manuscripts were published. The most important journal in the area is the IEEE Power and Energy Magazine with seven published articles and a total of 672 citations, followed by the IEEE Transactions on Smart Grid with 189 citations and seven published articles.

5. DISCUSSION

In this Section, we answer the research questions.

5.1. Q1: Which have been the most influential articles in the area?

Table 2 shows the articles with the highest number of citations (more than 50), so these have been the ones that have influenced research in this area. In Table 3 it can be seen the journals in which the selected articles are published.

5.2. Q2: What has been implemented regarding Smart Grids in Colombia?

We found only three research papers with the aim of studying the Smart Grid in Colombia [3234]. In [32], there is an overview of the technologies that are sought to implement the SG and a study of publications relevant to the development and use of SG technology and current research regarding this topic.

Table 2. Most cited articles.

Author	Title	Year	Research type	References	Citations
Ipakchi & Albuyeh.	Grid of the future. (IEEE Power and Energy Magazine) [11]	2009	Methodology	2	275
Farhangi.	The path of the smart grid (IEEE Power and Energy Magazine) [4]	2010	Methodology	1	206
Vojdani.	Smart integration (IEEE Power and Energy Magazine)[24]	2008	Methodology	1	80
Rahimi & Ipakchi	Demand Response as a Market Resource Under the Smart Grid Paradigm. (IEEE Transactions on Smart Grid)[19]	2010	Methodology	3	70
Conejo, Morales & Baringo.	Real-time demand response model (IEEE Transactions on Smart Grid)[31]	2010	Methodology	26	61
Garrity.	Getting Smart. (IEEE Power and Energy Magazine)[9]	2008	Methodology	1	58

Table 3. Journals and conference proceedings where the selected papers were published.

Journal or conference	Number of articles	Citations
IEEE Power and Energy Magazine	7	672
IEEE Transactions on Smart Grid	7	189
Energy Policy	4	41
Applied Energy	3	5
IEEE PES Innovative Smart Grid Technologies Conference Europe	3	1
Energy	2	13
IEEE Innovative Technologies for an Efficient and Reliable Electricity Supply	2	5
Power and Energy Engineering Conference, Asia-Pacific, 2010	2	1
The Electricity Journal	2	4
IEEE Transactions on Industrial Informatics	1	46
IEEE Transactions on Industrial Electronics	1	22
IEEE Transactions on Consumer Electronics	1	17
Annual Conference - Rural Electric Power Conference	1	10
IEEE Network	1	8
36th Annual Conference on IEEE Industrial Electronics Society, 2010	1	7
3rd European Modelling Symposium on Computer Modelling and Simulation, 2009	1	7
International Journal of Electrical Power and Energy Systems	1	6
Energy and Buildings	1	6
ACM Transactions on Intelligent Systems and Technology	1	6
IEEE/IFIP Network Operations and Management Symposium Workshops, 2010	1	4
International Journal of Energy Sector Management	1	4
IEEE Potentials	1	2
ABB Review	1	1
E-Energy 2010 - 1st Conference on Energy-Efficient Computing and Networking	1	1
Engineering & Technology	1	1
Electric Power Systems Research	1	1
IEEE/PES Transmission and Distribution Conference and Exposition: Latin America, 2008	1	0
Energy, Utilities & Chemicals	1	0
IEEE PES Innovative Smart Grid Technologies Conference Latin America, 2011	1	0
IEEE Innovative Smart Grid Technologies Conference, 2010	1	0
National Energy Technology Laboratory	1	0
2nd International Conference on Computational Intelligence, Modelling and Simulation, 2010	1	0
9th Conference Telecommunications Internet and Media Techno Economics, 2010	1	0
Tesis, Universidad Nacional de Colombia	1	0
45th International Universities Power Engineering Conference, 2010	1	0
Asia-Pacific Power and Energy Engineering Conference, APPEEC	1	0
Colombia Inteligente - http://www.colombiainteligente.org	1	0

In [33], a study was made by electricity utilities with the aim of determining the current state of the SG in Colombia and to build a reference based on the knowledge of the participants. In this work a reference map of the SG implementation in the Colombian electricity sector was defined. This study concludes that there is no specialized commercial network of Smart Grids in Colombia and that all research and development in this area focuses on the U.S. and Europe.

In [34], "Colombia Inteligente" was defined as a strategic framework that defines the guidelines and metrics toward a sustainable and efficient energy sector. The aim of this project is to establish guidelines, policies and strategies for the implementation of the REI in Colombia and to allow the country to have good practices in energy efficiency.

5.3. Q3: How do SGs affect the electricity demand?

SGs allow consumers to receive information about their consumption habits [15] and the associated electricity prices [35,36]. With this information, they can modify their consumption patterns by load shifting, shaping or cutting [37,38], consuming electricity in time periods with lower prices [1,3,15,21,39,40] in order to mitigate costs, save on their bills [16,17,38] or receive incentives [3,8,41,42].

As a consequence of changes in the consumption patterns, it is expected the following improvements are expected: more efficient use of energy [36], reduction of the peak demand [42,43], and increase of the grid sustainability by reducing the overall cost and carbon emissions [44,40,8]. With these improvements it is expected to reduce electricity demand between 10% and 20% during peak hours [1,20,24,25,26] and to reduce the depth of the valleys in the demand profiles [22]. The reduction of demand during peaks can result in a 9% reduction in household consumption [23,13]. In addition, the European Smart Metering Industry Group (ESMIG) believes that the greenhouse gas emissions for 2020 can be reduced by 20% [27].

5.4. Q4: What SG technologies are installed in households?

SG technologies installed in homes are:

- Smart meters for managing household electricity usage [16,45]. They are a key part of the communication infrastructure with the aim of making the users aware of their consumption [46,8] and receiving signals to help in the decision process [47,48]. With smart metering and home-based micro-storage, the demand will be more flexible to changes and will react to signals from the grid [49].
- Intelligent controllers and devices that cooperate in an intelligent way and that allow any appliance to be configured [50], so that when energy price is at some level (usually off-peak) they start working automatically [16,23,13]. These devices can administrate the consumer's demand based on the electricity market price[22].
- Advanced measurement systems, so called AMI (Advanced Metering Infrastructure), which are systems for the measurement, collection, storage, analysis and application of information from users [51].
- Online services like Google PowerMeter and Microsoft Hohm, which show the total consumption of a house using a friendly web interface and Smart Sockets, which control the energy consumption of the appliance [52]. However, Google PowerMeter and Hohm services were retired on September 2011 and May 2012 respectively.
- Smart sensors that offer a variety of technological functions like electricity controls, home interfaces, web interfaces, fire monitors, control appliances, gas leak monitors among others [53].

5.5. Q5: How have the Smart Grid penetrations been modeled?

Smart Grid penetration has been modeled from different points of view and different objectives:

- In [16], the authors develop a system to support users' intelligent decisions to increase the efficiency of energy use in the Smart Grid.
- In [20], an agent-based model is developed to simulate an electricity market with DR from different types of commercial buildings.
- In [31], an optimization model is developed to allow consumers to adjust their electricity demand level in response to the electricity price.
- In [54], the authors develop a methodology for modeling the prosumers' behavior (prosumers are the consumers who are themselves producers) with the aim of supporting policy makers decisions.
- In [55], an agent-based model is developed for simulating the behavior of a dynamic smart city in the future infrastructure of SG.
- In [56], an agent-based model is developed for simulating restructured electricity markets and exploring the impact of the price elasticity of consumer demand.
- In [57], a benchmark is proposed for evaluating management methodologies of domestic energy using different criteria in order to compare SG management solutions.

- In [20], an AHP (Analytic Hierarchy Process) methodology is developed for quantifying consumer preferences related to the use of appliances in peak periods when the price is higher. Then, a dynamic programming approach is used to achieve the optimal solution for managing appliances.
- In [58], a tool is designed to simulate a smart home, show the operation of DSM for customers, and estimate the home electricity consumption minimizing the customer cost.
- In [59], a model that simulates the response of consumers to dynamic pricing is developed. This model is integrated with an agent model that simulates deregulated markets.
- In [12], the authors design an in-home energy management (iHEM) application that uses a wireless sensor home area network to communicate all the appliances and the grid. This application accommodates consumers demand depending on the electricity price.
- In [43], a convex programming (CP) DR optimization framework is proposed for the automatic load management of different appliances in a smart home.
- In [7], the authors propose a model for energy use planning to optimize the consumption, generation and storage of residential electricity in a dynamic pricing environment.
- In [42], a model for demand response is developed; this work models consumer behavior in different scenarios and levels of consumers' rationality considering real time pricing.

5.6. Q6: Why is it important to study consumer behavior?

It is important to study consumer behavior because:

- It is necessary for consumers to modify their consumption patterns in order to reduce electricity costs and investments, to avoid price spikes [22] and to use appliances and energy more efficiently [14,26,60,23,13,54].
- Consumers can bring energy saving, energy efficiency and peak load shifts by using smart meters and information about the system [61].
- It is necessary to analyze and understand how consumers decide how much and how they want to consume energy [1,62].
- It is necessary to understand how to encourage the reduction of consumption or how to transfer it to periods with lower demand [63].
- It is necessary to determine how the most relevant information would be presented to the consumers in order to achieve changes in the consumption profile [3,64].
- It is necessary to understand how to encourage the use of advanced technology such as smart meters and other smart devices [47,65].
- In a scheme where consumers can directly participate in demand management it is necessary to find the best way to forecast the electric loads of individual consumers [66].
- Demand Response can provide competitive pressure to reduce energy prices, increasing awareness of energy use and providing a more efficient functioning of markets [11,19].
- With Demand Response, consumers can make the system more efficient, and with a 5% increase in network efficiency the equivalent of 42 coal thermal plants in the U.S. would not be needed, which would bring benefits to the system and the community [67].

6. CONCLUSIONS

Through the above study and the selected publications, the state of the art in Smart Grid and user behavior has been shown.

The publications of Kitchenham [28], Software Engineering Group [29] and Sorrell [30], propose the methodology of the systematic literature review, which was performed. These methodologies give as a result an analysis of the studies that some authors have made about the investigation subject, which allows further research into the area and will complement the work done.

The study of consumer behavior regarding SG is important because these users will become active entities in the electricity supply chain and their behavior can influence the improvement of the system.

Finally it can be said that in the literature there is no methodology to model and define policies for the entry of a SG and the Smart meters in an electricity system.

BIBLIOGRAPHY

[1] Lui, T. J., Stirling, W., & Marcy, H. O., Get Smart. Power and Energy Magazine, IEEE, 8 (3), 2010.

[2] Werbos, P. J., Computational Intelligence for the Smart Grid-History, Challenges, and Opportunities. IEEE Computational Intelligence Magazine, 6 (3), 14–21, 2011.

[3] U.S. Department of Energy, Environmental Impacts of Smart Grid. National Energy Technology Laboratory, 2011.

[4] Farhangi, H., The path of the smart grid. IEEE Power and Energy Magazine, 8 (1), 18–28, 2010.

[5] XM, Informe anual 2010, 2010.

[6] Rahman, S., Smart grid expectations. IEEE Power and Energy Magazine, 7 (5), 88, 84–85, 2009.

[7] Hubert, T., & Grijalva, S., Modeling for Residential Electricity Optimization in. IEEE Transactions on Smart Grid, 3 (4), 2224–2231, 2012.

[8] Clastres, C., Smart grids: Another step towards competition, energy security and climate change objectives. Energy Policy, 39 (9), 5399–5408, 2011.

[9] Garrity, T. F., Getting Smart. Power and Energy Magazine, IEEE, 6 (2), 38–45, 2008.

[10] Harris, A., Smart grid thinking. Engineering & Technology, 4 (9), 46, 2009.

[11] Ipakchi, A., & Albuyeh, F., Grid of the future. IEEE Power and Energy Magazine, 7 (2), 52–62, 2009.

[12] Erol-Kantarci, M., & Mouftah, H. T., Wireless Sensor Networks for Cost-Efficient Residential Energy Management in the Smart Grid. IEEE Transactions on Smart Grid, 2 (2), 314–325, 2011.

[13] Stragier, J., Hauttekeete, L., & Marez, L. De, Introducing Smart Grids in Residential Contexts : Consumers ' Perception of Smart Household Appliances. 2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply, CITRES 2010135–142, 2010.

[14] Santacana, E., Husain, B., Pinnekamp, F., Halvarsson, P., Rackliffe, G., Tang, L., Feng, X., Smart grids: The next level of evolution. ABB Review, 463 (7277), 10–15, 2010.

[15] Ajaja, A., Reinventing electric distribution. IEEE Potentials, 29 (1), 29–31, 2010.

[16] Sianaki, O. A., Hussain, O., Dillon, T., & Tabesh, A. R., Intelligent Decision Support System for Including Consumers' Preferences in Residential Energy Consumption in Smart Grid. 2010 Second International Conference on Computational Intelligence, Modelling and Simulation154–159, 2010.

[17] Collier, S. E., Ten steps to a smarter grid. IEEE Industry Applications Magazine, 62–68, 2010.

[18] Cecati, C., Mokryani, G., Piccolo, A., & Siano, P., An Overview on the Smart Grid Concept. IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society-, 3322–3327, 2010.

[19] Rahimi, F., & Ipakchi, A., Demand Response as a Market Resource Under the Smart Grid Paradigm. IEEE Transactions on, 1 (1), 82–88, 2010.

[20] Sianaki, O. A., Hussain, O., & Tabesh, A. R., A Knapsack Problem Approach For Achieving Efficient Energy Consumption in Smart Grid for End- users ' Life Style. 2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply, CITRES 2010159–164, 2010.

[21] Zhou, Z., Zhao, F., & Wang, J., Agent-Based Electricity Market Simulation With. IEEE Transactions on Smart Grid, 2 (4), 580–588, 2011.

[22] Moshari, A., Yousefi, G. R., Ebrahimi, A., & Haghbin, S., Demand-Side Behavior in the Smart Grid Environment. IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT Europe1–7, 2010.

[23] Hauttekeete, L., Stragier, J., Haerick, W., & De Marez, L., Smart, smarter, smartest... the consumer meets the smart electrical grid. Telecommunications Internet and Media Techno Economics (CTTE), 2010 9th Conference on2010.

[24] Vojdani, A., Smart Integration. IEEE Power and Energy Magazine, 6 (6), 71–79, 2008.

[25] Chassin, D. P., What Can the Smart Grid Do for You? And What Can You Do for the Smart Grid? The Electricity Journal, 23 (5), 57–63, 2010.

[26] Faruqui, A., Harris, D., & Hledik, R., Unlocking the

€53 billion savings from smart meters in the EU: How increasing the adoption of dynamic tariffs could make or break the EU's smart grid investment. Energy Policy, 38 (10), 6222–6231, 2010.

[27] Shargal, M., From Policy to Implementation : The Status of Europe's Smart Metering Market. Energy, Utilities & Chemicals, 2009.

[28] Kitchenham, B., Procedures for Performing Systematic Reviews. Keele University Technical Report TR/SE2004.

[29] Software Engineering Group, K. U., Guidelines for performing Systematic Literature Reviews in Software Engineering. Technical Report EBSE-2007-012007.

[30] Sorrell, S., Improving the evidence base for energy policy: The role of systematic reviews. Energy Policy, 35 (3), 1858–1871, 2006.

[31] Conejo, A. J., Morales, J. M., & Baringo, L., Real-Time Demand Response Model. IEEE Transactions on Smart Grid, 1 (3), 236–242, 2010.

[32] Perez Miranda, D., Estado y desarrollo de la tecnologia Smart Grid en ColombiaUniversidad Nacional de Colombia, Facultad de Minas, 2008.

[33] Aldana, A., Cespedes, R., Parra, E., Lopez, R., & Ruiz, M. E., Implementation of Smart Grids in the Colombian Electrical Sector2011.

[34] XM, CIDET, COCIER, CNO, CAC, & CINTEL, Colombia Inteligente. 2012[Online]. Available: http://www.colombiainteligente.org.

[35] Palensky, P., & Dietrich, D., Demand Side Management : Demand Response , Intelligent Energy Systems , and Smart Loads. IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, 7 (3), 381–388, 2011.

[36] Bouhafs, F., Mackay, M., & Merabti, M., Links to the Future. IEEE power & energy magazine(february), 24–32, 2012.

[37] Jia, W., Kang, C., & Chen, Q., Analysis on demand-side interactive response capability for power system dispatch in a smart grid framework. Electric Power Systems Research, 90, 11–17, 2012.

[38] Koutsopoulos, I., Tassiulas, L., & Hellas, T., Challenges in Demand Load Control for the Smart Grid. IEEE Network(October), 16–21, 2011.

[39] Zhu, N., Bai, X., & Meng, J., Benefits Analysis of All Parties Participating in Demand Responsein Asia-Pacific Power and Energy Engineering Conference, APPEEC 2011, 2011, 11–14.

[40] Blumsack, S., & Fernandez, A., Ready or not, here comes the smart grid! Energy, 37 (1), 61–68, 2012.

[41] Lujano-Rojas, J. M., Monteiro, C., Dufo-López, R., & Bernal-Agustín, J. L., Optimum residential load management strategy for real time pricing (RTP) demand response programs. Energy Policy, 45, 671–679, 2012.

[42] Venkatesan, N., Solanki, J., & Solanki, S. K., Residential Demand Response model and impact on voltage profile and losses of an electric distribution network. Applied Energy, 96, 84–91, 2012.

[43] Tsui, K. M., & Chan, S. C., Demand Response Optimization for Smart Home Scheduling Under Real-Time Pricing. IEEE Transactions on Smart Grid, 3 (4), 1812–1821, 2012.

[44] Logenthiran, T., Srinivasan, D., & Shun, T. Z., Demand Side Management in Smart Grid Using Heuristic Optimization. IEEE Transactions on Smart Grid, 3 (3), 1244–1252, 2012.

[45] Xin-wei, D., & Qiang, Y., Review of Smart Grid and its Development Prospect in Sichuan. Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific1–4, 2010.

[46] Benzi, F., Anglani, N., Bassi, E., & Frosini, L., Electricity Smart Meters Interfacing the Households. IEEE Transactions on Industrial Electronics, 58 (10), 4487–4494, 2011.

[47] O' Malley, G., Wu, J., & Jenkins, N., Technical Requirements of Smart Electric Power Distribution Networks in the UK. Universities Power Engineering Conference (UPEC), 2010 45th International2010.

[48] Misra, A., & Schulzrinne, H., Policy-driven distributed and collaborative demand response in multi-domain commercial buildings. Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking - e-Energy '10119, 2010.

[49] Ramchurn, S. D., Vytelingum, P., Rogers, A., & Jennings, N. R., Agent-based homeostatic control for green energy in the smart grid. ACM Transactions on Intelligent Systems and Technology, 2 (4), 1–28, 2011.

[50] Kofler, M. J., Reinisch, C., & Kastner, W., A semantic representation of energy-related information in future smart homes. Energy and Buildings, 47, 169–179, 2012.

[51] Zhang, J., Chen, Z., Yang, X., Chen, K., & Li, K., Ponder over Advanced Metering Infrastructure and Future Power Grid. Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific1–4, 2010.

[52] Park, S., Kim, H., Moon, H., Heo, J., & Yoon, S., Concurrent Simulation Platform for Energy-Aware Smart Metering Systems. IEEE Transactions on Consumer Electronics, 56 (3), 1918–1926, 2010.

[53] Daim, T. U., & Iskin, I., Smart thermostats: are we ready? International Journal of Energy Sector Management, 4 (2), 146–151, 2010.

[54] Lampropoulos, I., Vanalme, G. M. A., & Kling, W. L., A methodology for modeling the behavior of electricity prosumers within the smart grid. IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT Europe1–8, 2010.

[55] Karnouskos, S., & De Holanda, T. N., Simulation of a Smart Grid City with Software Agents. 2009 Third UKSim European Symposium on Computer Modeling and Simulation424–429, 2009.

[56] Thimmapuram, P. R., Kim, J., Botterud, A., & Nam, Y., Modeling and simulation of price elasticity of demand using an agent-based model. Innovative Smart Grid Technologies Conference, ISGT 20101–8, 2010.

[57] Bosman, M. G. C., Bakker, V., Molderink, A., Hurink, J. L., & Smit, G. J. M., Benchmarking set for domestic smart grid management. IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT Europe1–8, 2010.

[58] Gudi, N., Wang, L., & Devabhaktuni, V., A demand side management based simulation platform incorporating heuristic optimization for management of household appliances. International Journal of Electrical Power & Energy Systems, 43 (1), 185–193, 2012.

[59] Valenzuela, J., Thimmapuram, P. R., & Kim, J., Modeling and simulation of consumer response to dynamic pricing with enabled technologies. Applied Energy, 96, 122–132, 2012.

[60] Kadar, P., Understanding customer behaviour. 2008 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America(1), 1–4, 2008.

[61] Mah, D. N., Van der Vleuten, J. M., Hills, P., & Tao, J., Consumer perceptions of smart grid development: Results of a Hong Kong survey and policy implications. Energy Policy, 49, 204–216, 2012.

[62] Lineweber, D. C., Understanding Residential Customer Support for – and Opposition to – Smart Grid Investments. The Electricity Journal2011.

[63] Davies, S., Grid looks to smart solutions. Engineering and Technology(May), 2010.

[64] Vassileva, I., Wallin, F., & Dahlquist, E., Understanding energy consumption behavior for future demand response strategy development. Energy, 46 (1), 94–100, 2012.

[65] Saffre, F., & Gedge, R., Demand-Side Management for the Smart Grid. Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP300–303, 2010.

[66] Javed, F., Arshad, N., Wallin, F., Vassileva, I., & Dahlquist, E., Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting. Applied Energy, 96, 150–160, 2012.

[67] Hamilton, B., & Summy, M., Benefits of the smart grid. Power and Energy Magazine, IEEE, 9 (1), 2011.

INVITED ARTICLE

A FIRST COURSE IN SOFTWARE ENGINEERING METHODS AND THEORY

UN CURSO INICIAL SOBRE TEORÍA Y MÉTODOS DE LA INGENIERÍA DE SOFTWARE

CARLOS ZAPATA

Ph.D., Profesor Asociado, Facultad de Minas, Universidad Nacional de Colombia, sede Medellín, cmzapata@unal.edu.co

IVAR JACOBSON

Ph.D., Chairman of Ivar Jacobson International, ivar@ivarjacobson.com

Received for review December 10th, 2013, accepted February 04th, 2014, final version February, 10 th, 2014

ABSTRACT: Even though the software engineering curriculum has been designed many times, some problems still remain: the gap between academy and industry, the failure to continuously update the courses, the difficulties for combining theory and practice, and the lack of a sound, widely accepted theoretical basis. The SEMAT (Software Engineering Methods and Theory) initiative has been proposed for addressing some of the aforementioned problems. Based on the main ideas related to SEMAT, in this paper we propose a first course that introduces students to the main issues about SEMAT. This course is planned to be included in the System and Informatics Engineering Program belonging to the Universidad Nacional de Colombia, Medellin Branch. Also, we discuss the way in which this course addresses the previously diagnosed problems.

Key words: SEMAT, curriculum design, software industry, software academy, software community.

RESUMEN: Aunque desde hace años se habla del diseño curricular de la ingeniería de software, algunos problemas aún subsisten: la brecha existente entre la academia y la industria del software, los problemas para mantener actualizados los cursos, las dificultades para combinar la teoría con la práctica y la carencia de una sonora y ampliamente aceptada base teórica para la ingeniería de software. La iniciativa SEMAT (nombrada así por el acrónimo en inglés de Teoría y Métodos de la Ingeniería de Software) se propuso para solucionar algunos de los problemas en mención. Tomando como base las ideas relacionadas con SEMAT, en este artículo se propone un primer curso que suministra a los estudiantes una introducción a los principales tópicos de SEMAT. Se planea incluir este curso en el programa de Ingeniería de Sistemas e Informática de la Universidad Nacional de Colombia, Sede Medellín. También, se discute la manera en que este curso contribuye a solucionar los problemas que se diagnosticaron previamente.

Palabras clave: SEMAT, diseño curricular, industria del software, academia del software, comunidad del software.

1. INTRODUCTION

Since the early work of Farley [1], several authors have been working on curriculum design of software engineering courses [2–5]. Some authors are focused on complete curricula [1, 2] while others are focused on courses for meeting the industry needs [3–5]. However, all of them recognize the same remaining problems in the software engineering curriculum: (i) the growing gap among the software engineering courses created by academy and industry needs; (ii) the failure experienced by professors in order to keep up-to-date the contents of the courses, since there are fads emerging day-by-day related to software engineering methods and practices; (iii) the actual difficulties for offering practices matching the theories related to this field of knowledge; and (iv) the lack of a theoretical basis for the software engineering, even though there is a strong work on formal methods and conceptualization.

Jacobson et al. [6] proposed the SEMAT (Software Engineering Methods and Theory) initiative as a way to deal with several problems related to software engineering as a discipline by re-founding it. So, the SEMAT community has focused on two major goals: (i) finding a kernel of widely-agreed elements, and, (ii) defining a solid theoretical basis of the software engineering.

The main ideas of the SEMAT community lead us to propose a first course in software engineering methods

and theory to be included in the System and Informatics Engineering program belonging to the Universidad Nacional de Colombia, Medellín Branch. The course syllabus comprises several activities, ranging from lectures—some of them held by international guests—to active strategies such as forums, games, and practical projects. Also, some discussion is promoted in order to show the way in which we can address the aforementioned problems related to software industry and software academy, since they are closely linked to the progress of this discipline.

The reminder of this paper is organized as follows: in Section 2 we present some background related to the software engineering curriculum. The main description of the SEMAT initiative is presented in Section 3. Then, in Section 4 we propose a first course in software engineering methods and theory and we promote some discussion about the way in which some problems are addressed by this course. Finally, we present conclusions and future work.

2. SOFTWARE ENGINEERING CURRICULUM BACKGROUND

Farley [1] has probably the first attempt to define a software engineering undergraduate program. Such a program was designed for providing some core elements of the software engineering, but—as the author specifically recognized—these elements did not provide the adequate training for a software engineer. Mead [2] gives some insight about the further development of the software engineering curriculum. She also points out the strong urge for correlating the software engineering curriculum to the industrial practice.

Ludewig and Reißing [3] discuss the importance of theoretical courses as a way to provide tools to the software engineer for addressing real problems. Amiri et al. [4] make a survey to several managers of the software industry to the extent of determining the appropriateness of the software engineering courses. Moreno et al. [5] make a similar study about the software engineering courses versus the needs of the industry.

Being related to either curriculum as a whole or isolated software engineering courses, the previously reviewed work exhibits a consensus about some difficulties experienced by the software engineering education nowadays. Further information about some common problems is provided as follows:

Industry vs. Academy. The software Industry is demanding a different software engineer than the one promoted by the academic curriculum. In fact, some of the most demanded skills are usually far beyond the software engineering curricula. Some of the software industry managers claim the strong need to re-train the software engineering newcomers when they arrive from universities.

Courses vs. New methods. New software engineering methods are emerging day-by-day, creating new trends to be followed by practitioners. This fact poses a big challenge upon the software engineering course designers: how to keep up-to-date the course contents when several trends are coming from the academic world and they are being adopted by the software industry? How software engineering courses can evolve to follow the new trends in methods and practices?

Theory vs. Practice. Lectures are common ways of teaching software engineering and practice is usually limited to "toy" projects in some software engineering courses. Software industry needs practical, skillful engineers. How can we train practical professionals when the education is mainly theoretical?

Software Engineering Theoretical Basis. Methods and practices are emerging at a fast rate, since many research groups are promoting new ideas for dealing with the software engineering problems. However, we can argue against the novelty of such ideas, because many of them use previously known concepts. Even though some effort is devoted to define bodies of knowledge and glossaries about software engineering, we still lack a sound, widely-agreed theoretical basis for software engineering.

The SEMAT initiative is intended to deal with the aforementioned problems, as described in the following Section.

3. THE SEMAT INITIATIVE

Jacobson et al. [6] propose the SEMAT initiative for refunding the software engineering by defining

a theoretical basis—a small set of widely-agreed elements. Such elements are useful for defining past, present, and future methods, so the practices included on them can be reused. As a result, some practices seem to belong to obsolete methods and they can be incorporated into modern ones, as software engineering evolves. Theoretical basis is gathered into the so-called SEMAT kernel—the Essence of software engineering. The kernel is represented into a language with a limited number of elements (called alphas, see Fig. 1). The alphas can be used for assessing the health and progress of a software endeavor

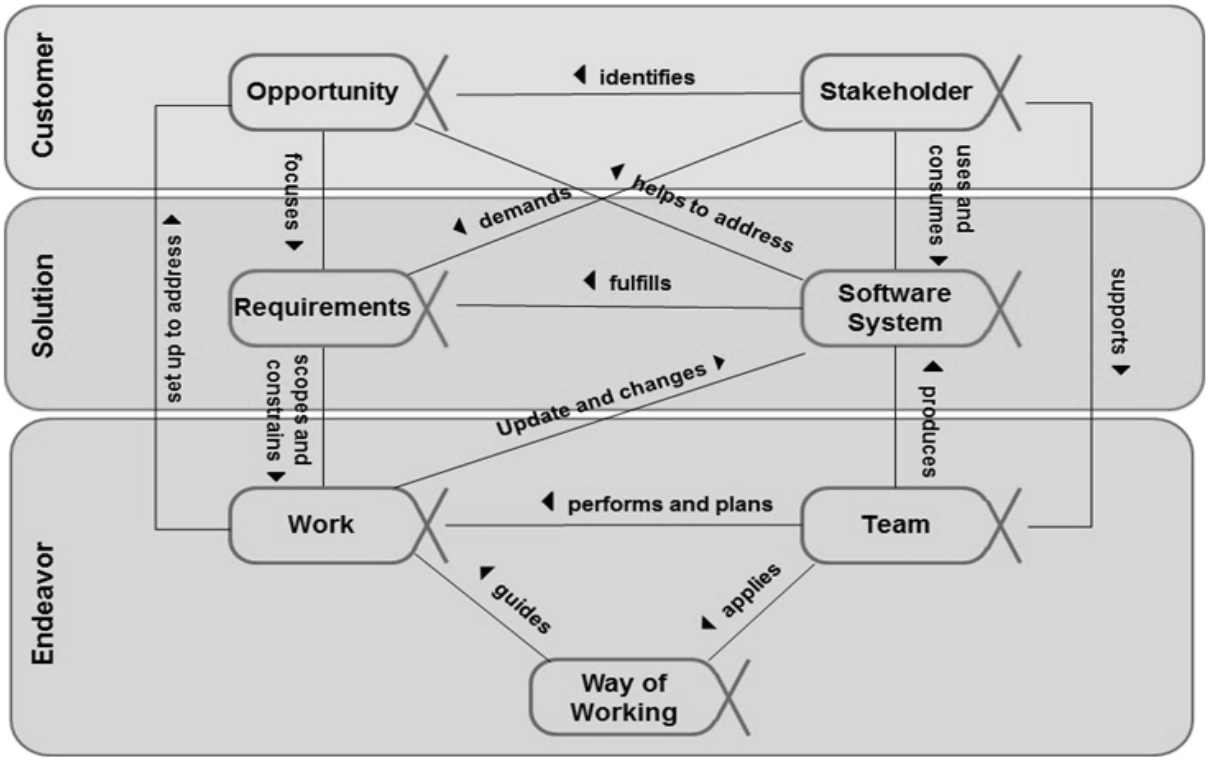


Figure 1. Alphas of the SEMAT kernel [6].

Health and progress can be estimated by using the alpha states. Some cards can be used in the SEMAT kernel for (i) describing alphas and possible states (see an example in Fig. 2), and (ii) linking alpha states to checklists for the sake of assessing their compliance (see an example in Fig. 3). Some other elements complement the theoretical basis of the software engineering represented by the SEMAT kernel. Some of them are: areas of concern, activity spaces, activities, competencies, and work products.

In this paper we propose a structure based on the so-called pre-conceptual schemas [7] for representing the SEMAT kernel elements and their relationships. In such schemas, nouns are represented by rectangles and verbs are represented by ovals. Thin arrows are used to connect nouns and verbs in such a way they express phrases, e. g. "method has practice" and "methodologist develops kernel." Thick arrows are used to express cause-and-effect relationships, e. g. "methodologist uses kernel, then methodologist describes practice." Finally, dotted boxes are used to express possible values of the nouns linked to them, e. g. "customer, solution, and endeavor are possible values of area of concern."

Fig. 4 depicts the kernel structure. In this figure, elements shown in blue were added to the schema to the extent of making it more readable and logical.

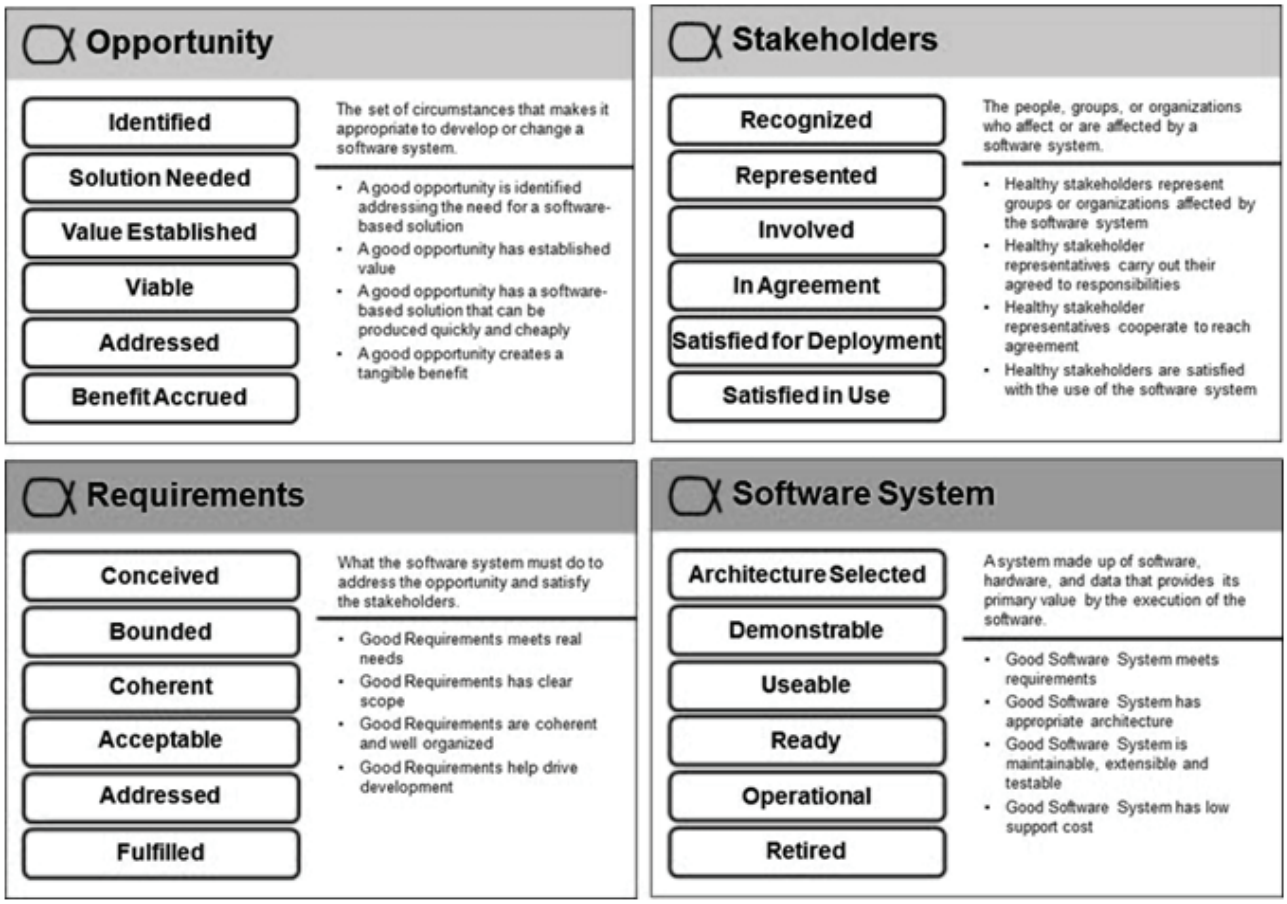


Figure 2. Examples of alphas and their states [6].

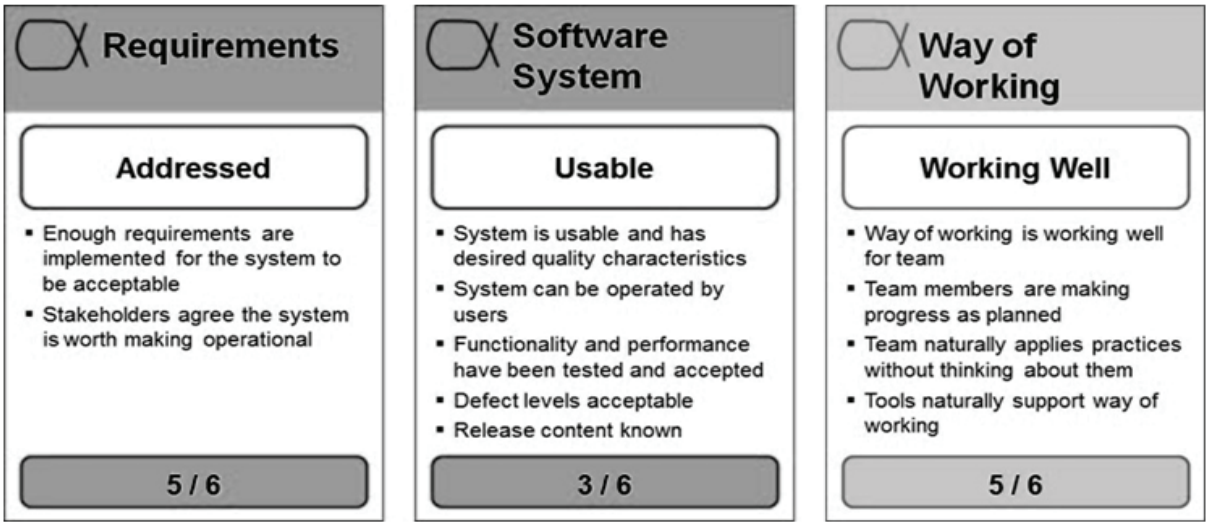


Figure 3. Examples of alpha states and checklists [6].

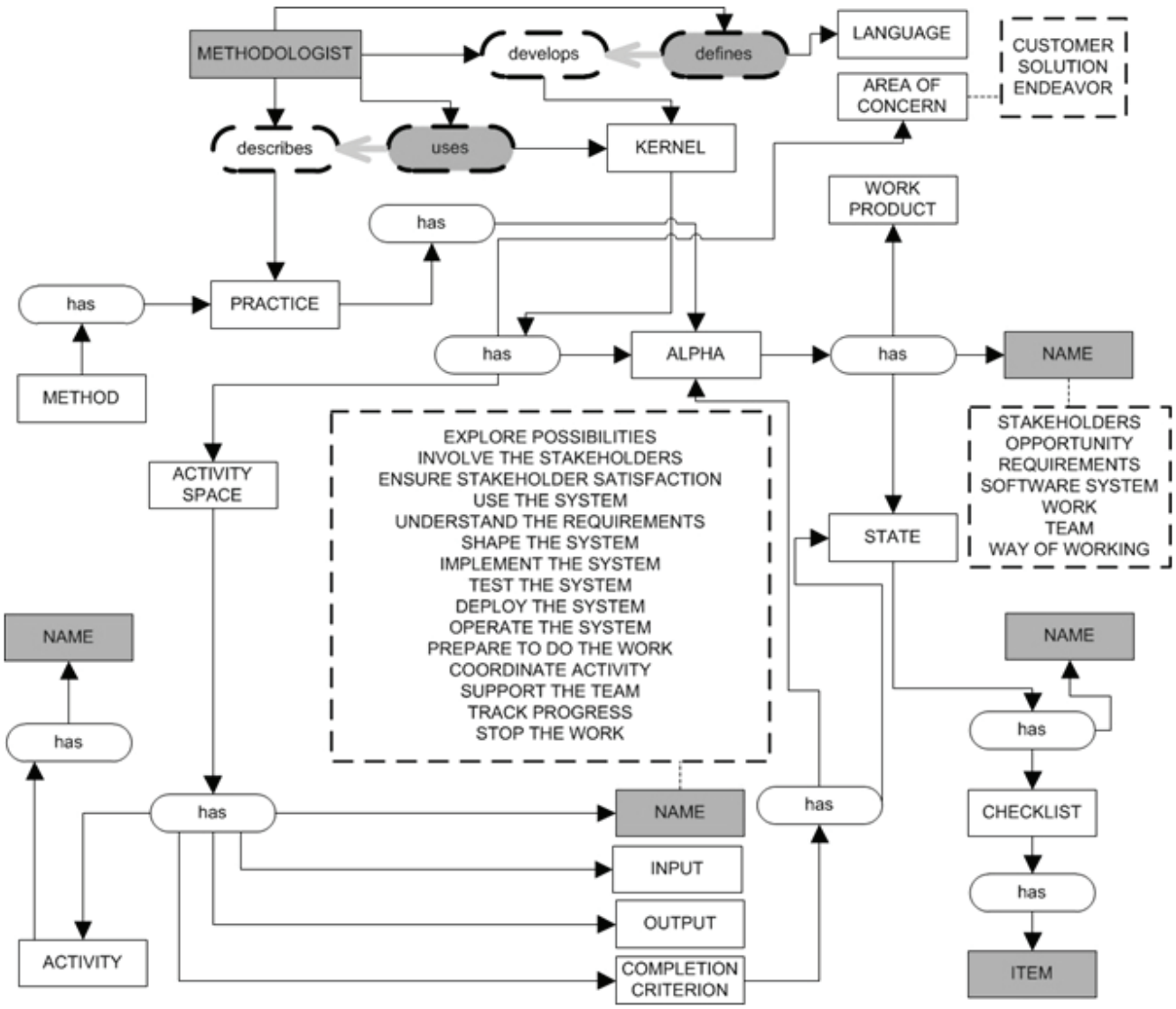


Figure 4. A pre-conceptual schema for representing the SEMAT kernel elements and their relationships. Source: the authors.

4. A FIRST COURSE IN SEMAT

SEMAT kernel is intended to address some of the problems exposed in Section 2, but a course related to the kernel is needed for addressing the four problems. In this Section we propose the curricular design of such a course and we discuss the way in which the problems are addressed.

4.1. Curricular Design

General information about the course is the following:

Code: 3009583
Name: Software Engineering Methods and Theory
Program: Systems and Informatics Engineering
Department: Computer and Decision Sciences
Level: Undergraduate
Typology: Free choice
Weekly in-classroom activity: 4 hours
Weekly outside-classroom activity: 5 hours
Weeks per semester: 16
Total activity: 144 hours
Credits: 3

General Objective: Defining and specifying the elements underlying any software engineering method and practice

Specific Objectives: (i) Establishing the cross-cutting elements to explaining software engineering methods and practices in terms of the SEMAT kernel; (ii) generating specific skills for representing any method and its practices in the SEMAT kernel; (iii) specifying any method and its practices in terms of the SEMAT kernel formal language; and (iv) recognizing and assessing tools used for creating representations in the SEMAT kernel

Methodology and Assessment: (i) Lectures; (ii) serialized practical project; (iii) forums; and (iv) experience-based games

The contents of the course are the following:

A. Initial information about software engineering general theories

A.1. Motivation

A.2. General Problems related to software engineering

A.3. Need for a software engineering theory

B. Basic elements of the software engineering kernel

B.1. Alphas and card-based representation

B.2. Activity spaces

B.3. Methods and practices

B.4. Competencies

B.5. Work products

C. Advanced elements of the software engineering kernel

C.1. Patterns

C.2. Resources

C.3. Detail levels

C.4. Competency levels

C.5. Separation of concerns

C.6. Tools for working with the kernel

D. Formal representation of the software engineering kernel

D.1. Introduction

D.2. Kernel meta-model

D.3. Kernel text-based specification

D.4. Object diagrams and executable pre-conceptual schemas

The course syllabus is included in Table 1. Several remarks are needed:

- Each numbered class (1 to 32) is a 2-hour class and has the issue to be covered, the contents related to such issue and the assessment items.

- Almost all the lectures should be held by the course professor, with the exception of class No. 16 and 31, which should be held by an international guest. In this case, we try to take advantage of the excellent relationships we have with the SEMAT Community and invite some international SEMAT researchers to hold a lecture via videoconference or keynote—in such a case the international guest is visiting our University.

- The effort to develop the practical project is intended to be guided in the classroom, with the participation of all the students. Six classes are reserved for this purpose. The topic of the practical project is related to the representation of a method by using the SEMAT kernel.

- The first forum is devoted to practicing the initial skills needed for developing the practical project. The second forum is devoted to activities developed by the students by using active learning: the design of a crossword puzzle with the main concepts of SEMAT and the design of an experience-based game by studying some of the SEMAT concepts.

- Five pre-designed games will be played during the course: the software system alpha game (see Fig. 5), the requirements alpha game (see Fig. 6), the SemCards game (see Fig. 7), the MetricC game (see Fig. 8), and the SEMAT board-crossing game (see Fig. 9). All of these games were previously designed and tested with several groups of heterogeneous people, mostly during Conferences, e. g. the Colombian Computing Conference [8], the Latin American Conference on Informatics [9], and the International Conference on Science and Technology for the Risk Management and the Climate Change Adaptation.

Table 1. SEMAT course syllabus. Source: the authors.

CLASS No.	*ISSUE*	*CONTENTS*	*ASSESSMENT*
1	INTRODUCTION, PRESENTATION, METHODOLOGY	A	
2	INTRODUCTION TO SEMAT TOPICS	A	
3	THINGS TO WORK WITH: ALPHAS. CARD-BASED REPRESENTATION	B	
4	THINGS TO DO: ACTIVITY SPACES AND ACTIVITIES	B	
5	THEMATIC FORUM N° 1: PAPER ABOUT SEMAT		WRITTEN REPORT No. 1
6	GUIDANCE FOR THE PRACTICAL PROJECT		
7	GUIDANCE FOR THE PRACTICAL PROJECT		
8	THE SOFTWARE SYSTEM ALPHA GAME	B	FIRST DELIVERABLE
9	METHODS AND PRACTICES	B	
10	COMPETENCIES	B	
11	DIFFERENCES BETWEEN SEMAT AND OTHER PROPOSALS	A	
12	THEMATIC FORUM N° 2: BEST SEMAT CROSSWORD PUZZLES DESIGNED BY THE STUDENTS		WRITTEN REPORT No. 2
13	BEST SEMAT GAMES DESIGNED BY THE STUDENTS		
14	WORK PRODUCTS, PATTERNS, AND RESOURCES	B, C	
15	SEMCARDS GAME	B	
16	LECTURE HELD BY AN INTERNATIONAL GUEST		
17	GUIDANCE FOR THE PRACTICAL PROJECT		
18	GUIDANCE FOR THE PRACTICAL PROJECT		
19	THE REQUIREMENTS ALPHA GAME	B	SECOND DELIVERABLE
20	SEMAT SPECIFICATION FOUNDATION	D	
21	SEMAT SPECIFICATION FOUNDATION	D	
22	SEMAT SPECIFICATION FOUNDATION	D	
23	DETAIL LEVELS AND COMPETENCY LEVELS	C	
24	GUIDANCE FOR THE PRACTICAL PROJECT		
25	GUIDANCE FOR THE PRACTICAL PROJECT		
26	METRICC GAME	B	
27	SEPARATION OF CONCERNS IN THE SEMAT CONTEXT	B	
28	SEMAT WORK TOOLS	A, B, C, D	
29	GUIDANCE FOR THE PRACTICAL PROJECT		
30	GUIDANCE FOR THE PRACTICAL PROJECT		
31	LECTURE HELD BY AN INTERNATIONAL GUEST		THIRD DELIVERABLE
32	THE BOARD-CROSSING SEMAT GAME	A, B, C, D	

4.2. Discussion

The four problems presented in Section 2 can be addressed by the SEMAT course we propose in this paper. The reasons are the following:

Industry vs. Academy. While the SEMAT initiative is advancing in the world, several discussions between practitioners and academics have been promoted in order to standardize the terminology related to software engineering. The proposed course is a direct result of such discussions and, consequently, it is intended to close the gap between the software engineering curriculum and the industrial training needed by software engineers. Also, two international guests will hold lectures to keep students informed about what is happening in the software industry. Since we can promote such a kind of lectures, we are also working to close the aforementioned gap.

Courses vs. New methods. SEMAT kernel is not related to any particular method or practice. SEMAT kernel is a framework for describing any method or practice. So, there is no need to change the course contents with the emergence of new methods and practices.

Theory vs. Practice. Practical projects and games are strategies for promoting a close relationship between theory and practice. Particularly, some of the games selected for the course are based on real situations with the intention of encouraging decision making in simulated environments.

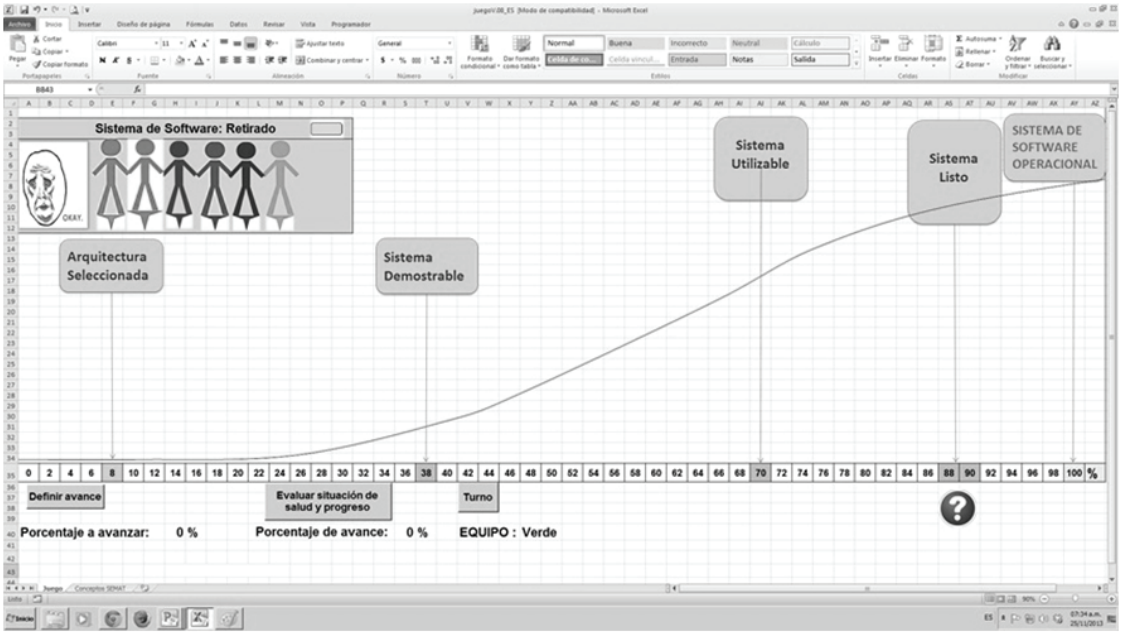


Figure 5. Image from the software system alpha game. Source: the authors.

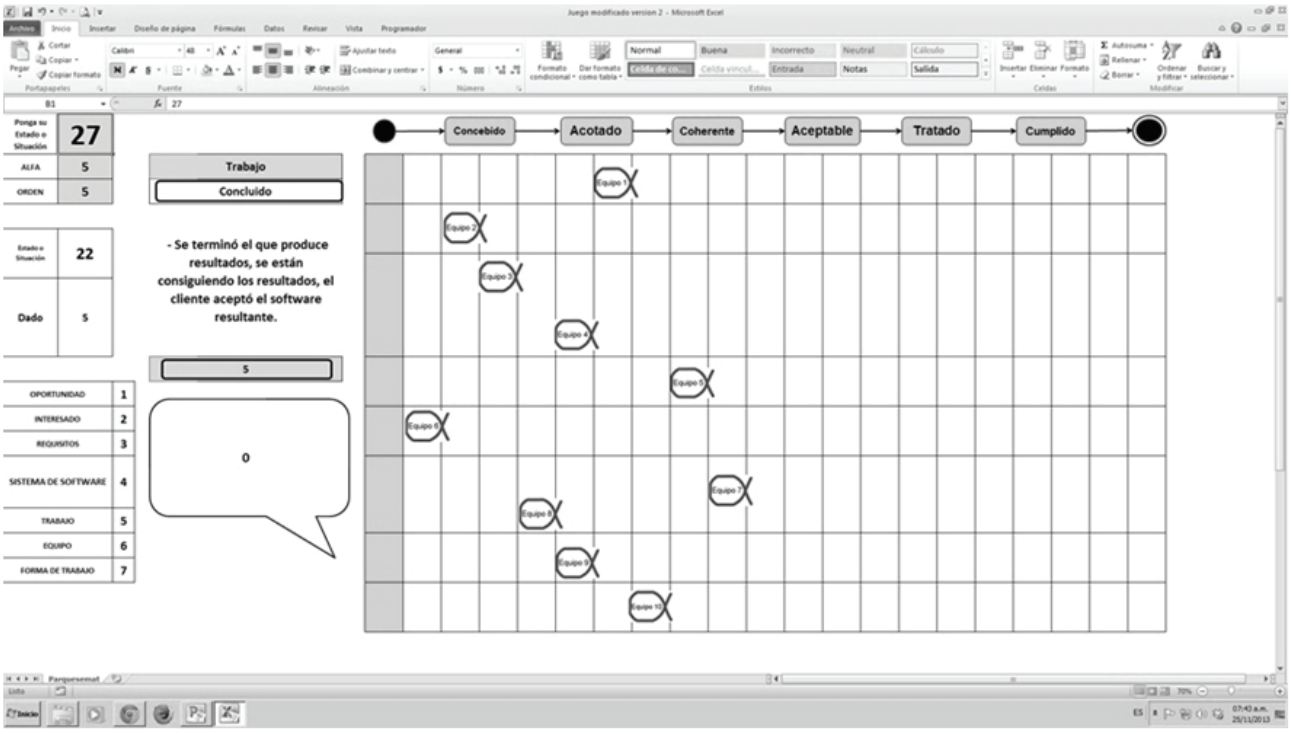


Figure 6. Image from the requirements alpha game. Source: the authors.

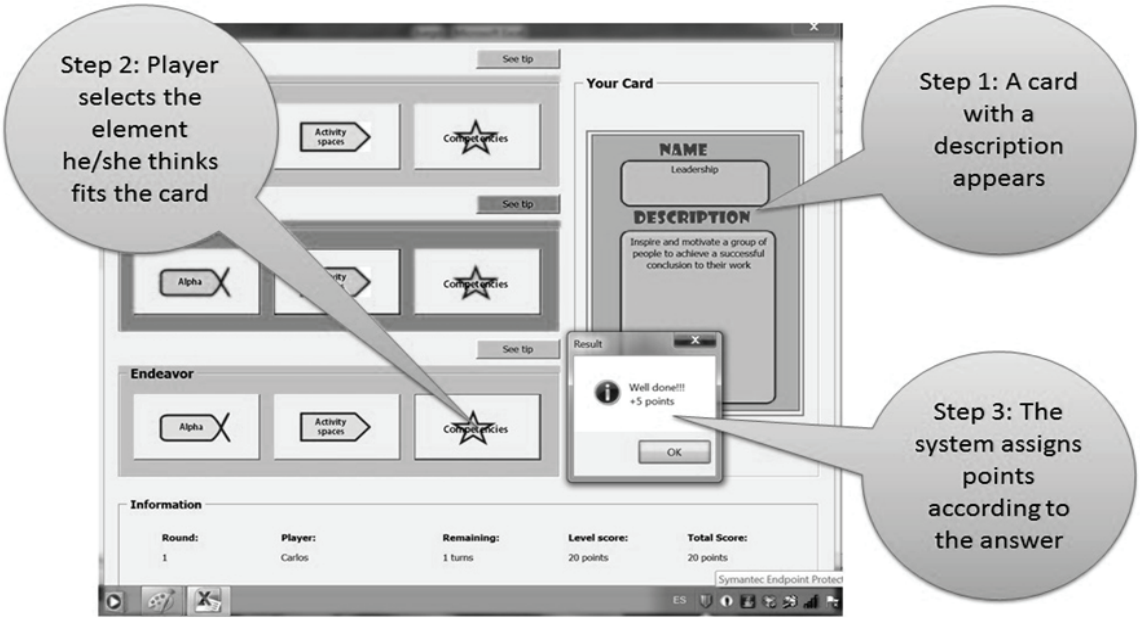


Figure 7. Image from the SemCards game. Source: the authors.

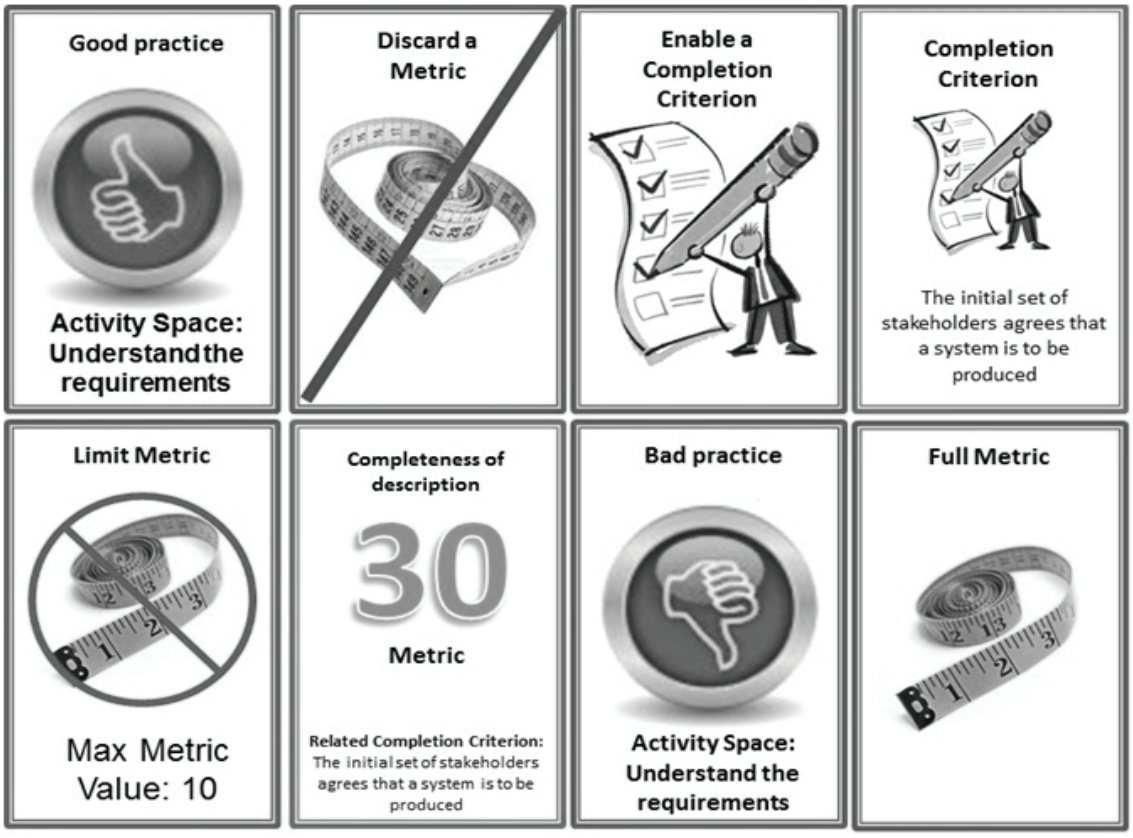


Figure 8. Image from the MetricC game. Source: the authors.

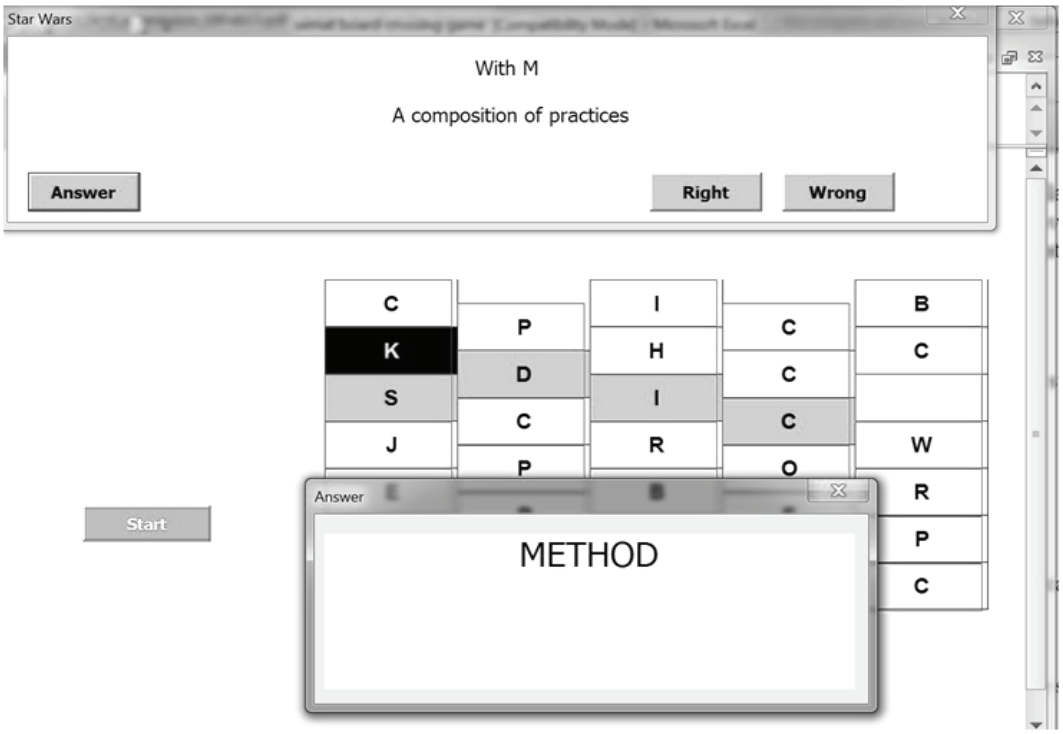


Figure 9. Image from the SEMAT board-crossing game. Source: the authors.

Software Engineering Theoretical Basis. By proposing a first course on SEMAT, we are promoting the dissemination of the SEMAT theoretical basis among students. The difference of this effort versus the previous ones is closely linked to the way in which we will try to teach software engineering. In fact, you can discover many software engineering courses related to specific methods and practices, but this is the first attempt to teach the theoretical basis underlying any methods or practices.

5. CONCLUSIONS AND FUTURE WORK

In this paper we proposed a first course in software engineering methods and theory based on the work we are conducting in the Universidad Nacional de Colombia, Medellin Branch, related to the SEMAT initiative. The course is highly supported by the SEMAT community—for instance in the shape of lectures held by international guests—as a joint effort to keep a close contact among the SEMAT Latin American Chapter and the other regional Chapters. Also, the course thrives on the work we are conducting in games and other teaching strategies, because we are incorporating five SEMAT games created in the Universidad Nacional de Colombia.

We also discussed the way to deal with some common problems of the software engineering teaching, mainly related to the gap between software industry and academy, the continuous emergence of new software engineering methods, the scarce practice of the engineering newcomers and the lack of a software engineering theoretical basis

We propose as future work a complete assessment of the course results in terms of satisfaction surveys to students and professionals linked to the software industry. Also, the creation of new games to be included in the course contents can contribute to improve the SEMAT course.

REFERENCES

[1] Fairley R. Educational Issues in Software Engineering, en 1978 annual conference ACM/CSC-ER (1978, Washington D.C., USA). pp. 58–62.

[2] Mead, N. Software engineering education: How far we've come and how far we have to go The Journal of Systems and Software, 82, pp. 571–575, 2009.

[3] Ludewig, J. and Reißing, R. Teaching what they need instead of teaching what we like—the new software engineering curriculum at the University of Stuttgart, Information and Software Technology, 40, pp. 239–244, 1998.

[4] Amiri A., Banari M., and Yousefnezhad N. An Investigation of Undergraduate Software Engineering Curriculum: Iranian Universities Case Study, en International Conference on Computer Science & Education ICCSE (6th, 2011, Singapore, Singapore). pp. 638–644.

[5] Moreno, A., Sanchez-Segura, M., Medina-Dominguez, F., and Carvajal, L. Balancing software engineering education and industrial needs, The Journal of Systems and Software, 85, pp. 1607–1620, 2012.

[6] Jacobson, I., Ng, P. McMahon, P., Spence, I., and Lidman, S. The essence of software engineering: appying the Semat kernel, 1st edition, New Jersey, Addison Wesley, 2013.

[7] Zapata, C. M., Gelbukh, A., and Arango, F. Pre-conceptual Schema: A Conceptual-Graph-Like Knowledge Representation for Requirements Elicitation, Lecture Notes in Computer Science, 4293, pp. 17–27, 2006.

[8] Zapata C. M., Maturana G., and Castro L. Tutorial sobre la iniciativa SEMAT y el juego MetricC, en Congreso Colombiano de Computación (8th, 2013, Armenia, Colombia).

[9] Zapata C. M. and Montilva J. La esencia de la Ingeniería del Software: el núcleo SEMAT en la práctica, en Conferencia Latinoamericana en Informática (39ª, 2013, Naiguatá, Venezuela).

INSTRUCTIONS FOR OUR CONTRIBUTORS - DYNA JOURNAL

DYNA Journal is a general scientific publication of the technological science area founded in 1933. It is the means of expression of those who work in the area of engineering, science, and technology of the School of Mines of the Universidad Nacional de Colombia, Medellin Branch, and of researchers of the same area who are in Colombia. DYNA mainly presents scientific and technological development research articles, articles with original reflection on a particular problem or topic, and review articles which offer a general perspective of the state of a specific sphere of science and technology. Notes and communications are also included as are discussions on particular points of the theme covered by the journal: technology, science, and engineering. All of the articles published are subject to the peer or referee review by those of recognized suitability on the article's subject.

The presentation of an article should be accompanied by a Presentation Letter, requesting publication in the journal, declaring that the article is approved by all of the authors, that it has not been published or submitted for consideration in another similar publication or journal. If the article has been published in another medium with a different outreach such as seminar memoirs, newspapers, supplements, etc., these should be specified and the possible publication in DYNA should be justified. Finally, the authors should recommend possible evaluators (2 national and 2 international) for the article being presented. It will be necessary to attach copies of permissions for the reproduction of material which was already published.

The submission of articles and the Presentation Letter are handled through the Open Journal System (OJS) of the National University (UN) Virtual Library (Enter through: http://www.revistas.unal.edu.co/index.php/dyna/user).

DYNA publishes the following in each of its editions: an editorial, research articles, essays, brief communications, and diagnostic images. The journal publishes articles from the following areas:

- Organizational Engineering
- Civil Engineering
- Materials Engineering
- Geosciences and the Environment
- Computer Systems
- Chemistry and Petroleum
- Mechatronics
- Engineering and Biotechnology
- Other areas related to Engineering

All that is published in DYNA is considered to be the property of the journal and can only be reproduced completely or partially with the express permission of the director. The articles which are published in DYNA are also published in an electronic version (CD-ROM), on the Internet, as well as in other electronic formats which may become important in the future; the authors of the articles accepting these and other virtual or digital forms of publication. DYNA assumes international norms for the publication of the articles such as ISI, Vancouver, and Harvard, with the specifications which shall be given shortly.

SECTION TYPES

Authors who wish to contribute to a certain section of the journal can send their contributions through the OJS system, by which the author will be notified by e-mail; understanding that this does not imply the acceptation of the paper. The Editorial Committee may reject, suggest changes, or (in certain cases), carry out modifications in the received texts. In the latter case, none of the scientific content shall be altered, but the modifications will be limited only to the literary style. The criteria which will apply to evaluate the acceptation of the articles will be the methodological scientific rigor, innovation, the relevance of the research, expositional conciseness, and the literary quality of the text. The journal will publish one or several of the following sections in each number:

1. The Editorial
These are papers written by someone recommended by the Editorial Committee, or otherwise, by one of its members. Its length should not be more than six pages, double-spaced; the bibliography (if there is one) should not contain more than 10 citations. The editorial should be placed under the consideration of the Editorial Committee for its final approval.

2. Letters to the Editor
Critical, analytical, or interpretational positions on documents published in the journal, which, by judgment of the Editorial Committee, make an important contribution to the discussion of the subject of the field, on behalf of the scientific community of reference.

3. Articles of Scientific and Technological Research
Those papers which exposit upon the outcomes of research in engineering (within previously-defined areas) will be considered as articles of scientific and technological research. These should have the following generally used structure: an introduction, the methodology, outcomes, and conclusions. They should not have more than 30 bibliographical citations.

4. Reflection Articles
These present outcomes of research from an analytical, interpretive, or critical perspective of the author, on specific topics in the area of engineering and within the previously defined areas, using original sources. They should not go beyond 10 bibliographical citations.

5. Review Articles
Documents which are a result of research in which the outcomes of published or non-published research projects are analyzed, systematized, and integrated in the field of engineering, with the objective of giving an account of the advances and the developmental trends of the previously cited topics. These articles present a careful bibliographic review of around 50 references.

6. Brief Communication and Summaries
This section allows for us to publish brief articles with greater speed. It allows authors to present their observations, initial outcomes of current research, book reviews, and even to make comments on works that have already been published in the journal. The text should not include more than 10 bibliographical citations.

7. Diagnostic Image (Case Studies)
In this section singular cases are presented which the image is decisive. For publication, one or a maximum of two images, or diagnostic cases, may be submitted. These may be from any branch of engineering and should be accompanied by a short commentary with relation to the case. The length of the commentary may not be less than 15 written lines, single spaced. This page should be sent together with another which has the information about the author or authors which shall be described further on in the presentation of papers.

8. Translations
Translations of classic or current texts, the transcriptions of historical documents, or those of particular interest in the field of the journal's publication are presented.

9. Event Notices
This shall include brief information which announces events related to engineering. It has no cost.

PRESENTATION OF PAPERS

Together with the Commitment Letter, the article shall be presented through the Open Journal System – OJS, of the UN Virtual Library (on the webpage, go to Envío de Artículos).

1. The authors shall send their articles through the Open Journal System (OJS), entering according to the instructions which are found in Envío de Artículos and, additionally will accompany it with the Presentation Letter, which they may write in accordance with the model presented in: Modelo de Carta de Presentación (in Web site: dyna.unalmed.edu.co).

2. The article should be unpublished and may not be submitted simultaneously to the scrutiny of other journals or periodical publications.

3. The maximum length of the articles will be 10 pages, single-spaced, in Spanish or English, letter-sized presentation. Include in the pages the abstract in Spanish (RESUMEN) and in English (ABSTRACT), both of a maximum of 150 words, and the "Keywords" also in Spanish (PALABRAS CLAVES) and in English (KEYWORDS). The articles may NOT be class notes and when they are translations or use material protected by intellectual property rights they should be submitted with the due author authorizations.

4. Every article shall be submitted to the evaluation of two referees designated by the Editorial Committee. The exceptions to this requirement shall be decided upon by the same Committee.

5. Regarding graphics, tables, and figures, these should be processed as objects (in order to ease the editing). Authors may use the program that they wish. These should be sent in an image format (JPG, GIF, BMP, etc.), in black and white or grey tones, avoiding polychromes.

6. The figures that are not presented in an electronic medium should be presented as photography or as vellum and Indian ink.

7. If the text includes photography, its presentation is recommended in black and white, matte, with good contrast. Photographs should be placed on pages of the same size as the text. In the rest of the aspects, the treatment given is similar to that of figures.

8. The presentation of the article will be according to the articles of this number. Footnotes are not admitted due to the typesetting requirements of the journal.

9. Under the title of the article should be included the name of the author (or authors), their professional identification or institutional affiliation, their postal and electronic (e-mail) addresses.

10. The references should go at the end of the paper in the following format:

Journal articles:
Alien, J. S., Samuelson, R. and Newberger, A. Chaos in a model of forced quasi-geostrophic flow over topography: An application of Melinkov's method, J. Fluid Mech., No. 226, pp. 511-547, 1991.

Books:
Baker, G. L. and Golub, J., Chaotic Dynamics: An introduction, Cambridge University Press, Cambridge, 1990.

Chapter of a book:
Lewis, P., Ray, B. and Stevens, J.G. Modeling time series by using Multivariate Adaptive Regression Splines (MARS), In: Time Series Prediction: Forecasting the Future and Understanding the Past (Eds. A. S. Weigend and N. A. Gershenfeid), SFI Studies in the Science of Complexity, Proc. Vol. XV, Addison-Wesley, pp. 297-318, 1994.

Conference proceedings:
Álzate, N., Botero, T. and Correa, D. Title of the Paper. Memorias XIX Congreso Latinoamericano de Ponencias Científicas. Córdoba, Argentina, Tomo II, pp. 219-228, Octubre 2000.

Organization or government report:
U.S. EPA. Status of Pesticides in Registration and Special Review. EPA 738-R-94-008. Washington, DC: U.S. Environmental Protection Agency, 1994.

Thesis:
Jacobs J., Regulation of Life History Strategies within Individuals in Predictable and Unpredictable Environments [PhD Thesis]. Seattle, WA: University of Washington, 1996.

Internet references:
NOAA-CIRES. Climate Diagnostics Center. Advancing Understanding and Predictions of Climate Variability. Available: http://www.cdc.noaa.gov [cited August 8, 1998].

Non-published data:
Last name, N. Non-published data.

Personal communication:
Last name, N. Personal communication.

OTHER CONSIDERATIONS

The author shall always keep a copy of the paper. The Editorial Committee will decide on the publication of each manuscript and will suggest possible modifications when the Committee considers it to be appropriate, after it has been submitted to the evaluation of the selected peers. This procedure will be as brief as possible.

The Journal's Editorial Committee will acknowledge receipt of the originals and will inform the author on their acceptation through communication by e-mail, through the Open Journal System.

The editorial staff of the journal will take the greatest care in avoiding errors in the final transcription of the articles to publish, however, it does not respond for the errors that may appear. If the authors detect important errors, they should recur as soon as possible to the editorial staff so that a list erratas may be prepared for the next number of the journal.

The editorial staff reserves the right to carry out small adjustments on the article titles, as well as minor rectifications in the editing, in translations, resúmenes and abstracts; although the general policy to be followed shall be to consult with the authors about these issues.

NORMAS DE PUBLICACIÓN - REVISTA DYNA
Instrucciones a los Autores

La Revista DYNA es una publicación científica general del área de las ciencias tecnológicas, fundada en 1933, es el medio de expresión de los trabajos en el área de ingeniería, ciencias y tecnología de la Facultad de Minas de la Universidad Nacional de Colombia, sede Medellín y de los investigadores del área en Colombia. DYNA presenta principalmente artículos de investigación científica y de desarrollo tecnológico, artículos de reflexión originales sobre un problema o tópico particular y artículos de revisión que brindan una perspectiva general del estado de un dominio específico de la ciencia y la tecnología. También tienen cabida notas y comunicaciones, así como discusiones sobre puntos particulares de la temática cubierta por la revista: tecnología, ciencias e ingeniería. Todos los artículos publicados se someten a revisión de pares o árbitros, de reconocida idoneidad en el tema.

La presentación de un artículo debe ir acompañada de una Carta de Presentación, solicitando su publicación en la revista, declarando que el artículo es aprobado por todos los autores, que no ha sido publicado o sometido a consideración en otra publicación o revista similar. Deben especificar si el artículo ha sido publicado en otro medio de diferente alcance divulgativo como memorias de seminarios, periódicos, separatas, etc., y justificar su posible publicación en DYNA. Finalmente recomendar posibles evaluadores (2 nacionales y 2 internacionales) para el artículo presentado; de ser necesario adjuntar copias de los permisos para reproducir material ya publicado.

El envío de los artículos y de la Carta de Presentación se hace a través del Sistema Open Jornal System (OJS) de la Biblioteca Virtual UN (Ingresar a través de: http://www.revistas.unal.edu.co/index.php/dyna/login).

DYNA publica en sus ediciones: Editorial, Artículos de Investigación, Ensayos, Comunicaciones Breves e Imágenes Diagnósticas. La revista publica artículos sobre las siguientes áreas:

- Ingeniería de la Organización
- Ingeniería Civil
- Ingeniería de Materiales
- Geociencias y Medio Ambiente
- Sistemas
- Química y Petróleos
- Mecatrónica
- Ingeniería y Biotecnología
- Otros relacionados con Ingeniería

Todo lo publicado se considerará propiedad de DYNA y solo podrá ser reproducido total o parcialmente con el permiso expreso del director. Los artículos que sean publicados en la revista DYNA podrán ser también publicados en la versión electrónica de la revista, en Internet, así como en otros formatos electrónicos (CD-ROM y otros) que surjan en el futuro, aceptando los autores de los artículos éstas y otras formas de publicación virtual o digital. DYNA asume para la publicación de los artículos las normas internacionales como ISI, Vancouver y Harvard con las especificaciones que se detallan mas adelante.

TIPOS DE SECCIONES

Los autores que deseen colaborar en algunas de las secciones de la revista pueden enviar sus aportes (a través del sistema OJS), que será notificado vía e-mail al autor; entendiéndose que ello no implica la aceptación del trabajo. El Comité Editorial podrá rechazar, sugerir cambios o llegado el caso, realizar modificaciones en los textos recibidos; en este último caso, no se alterará el contenido científico, limitándose únicamente al estilo literario. Los criterios que se aplicarán para valorar la aceptación de los artículos serán el rigor científico metodológico, novedad, trascendencia del trabajo, concisión expositiva, así como la calidad literaria del texto. La revista publicará en cada número una o varias de las siguientes secciones:

1. Editorial

Trabajos escritos por encargo del Comité Editorial, o bien, redactados por alguno de sus miembros. Su extensión no debe ser superior a 6 páginas, escritos a doble espacio; la bibliografía si la hay, no debe contener una cifra superior a 10 citas. Esta editorial debe ser puesta a consideración del Comité Editorial para su aprobación final.

2. Cartas al Editor

Posiciones críticas, analíticas o interpretativas sobre documentos publicados en la revista, que a juicio del Comité Editorial constituyen un aporte importante a la discusión del tema ambiental, por parte de la comunidad científica de referencia.

3. Artículos de Investigación Científica y Tecnológica

Tendrán tal consideración aquellos trabajos, que versen sobre resultados de investigaciones en ingeniería, dentro de las áreas definidas con anterioridad. Deben tener la siguiente estructura generalmente utilizada: Introducción, Metodología, Resultados y Conclusiones.

4. Artículos de Reflexión

Son documentos que presentan resultados de investigación desde una perspectiva analítica, interpretativa o crítica del autor, sobre temas específicos en un área de ingeniería y dentro de las áreas definidas anteriormente, recurriendo a las fuentes originales. No debe superar las 10 citas bibliográficas.

5. Artículos de Revisión

Documentos resultados de una investigación donde se analizan, sistematizan e integran los resultados de investigaciones publicadas o no publicadas, sobre el campo de ingenierías, con el fin de dar cuenta de los avances y las tendencias de desarrollo de los temas citados anteriormente. Se caracterizan por presentar una cuidadosa revisión bibliográfica alrededor de 50 referencias.

6. Comunicación Breve y Reseñas

Esta sección permite publicar artículos breves, con mayor rapidez. Ello facilita que los autores presenten observaciones, resultados iniciales de investigaciones en curso, revisiones de libros e incluso realizar comentarios a trabajos ya publicados en la revista. El texto no debe superar las 10 citas bibliográficas.

7. Imagen diagnóstica. (Reporte de Casos)

En esta sección de la revista se presentarán casos singulares en los que la imagen sea resolutoria. Para ello, se aportarán una, o como máximo dos imágenes, o casos diagnósticos, de cualquier rama de la ingeniería y acompañada de un breve comentario en relación con el caso. La extensión del comentario no será menor a 15 líneas escritas a espacio sencillo. Esta página se remitirá junto a otra, con la información del autor o autores y que se detalla mas adelante en presentación de los trabajos.

8. Traducciones

Se presentan traducciones de textos clásicos, o de actualidad, o transcripciones de documentos históricos o de interés particular en el dominio de publicación de la revista.

9. Noticias de Eventos

Incluirá breve información que anunciará eventos relacionados con la ingeniería. No tiene ningún costo.

PRESENTACIÓN DE LOS TRABAJOS

Acompañado de la Carta de Compromiso, el artículo se presentará a través del Sistema Open Journal System – OJS de la Biblioteca Virtual de la UN (ir a Envío de Artículos),

Los autores remitirán sus artículos a través del Sistema Open Journal System (OJS), ingresando según las instrucciones que encuentran en Envío de Artículos y adicionalmente acompañado de la Carta de Presentación, que pueden redactar de acuerdo con el modelo presentado en: Modelo de Carta de Presentación (en sitio Web: dyna.unalmed.edu.co).

El artículo debe ser inédito y no puede someterse simultáneamente al examen de otras revistas o publicaciones periódicas. La extensión máxima de los artículos será de 10 páginas a espacio sencillo, en castellano o en inglés, en papel tamaño carta, incluyendo en ellas el Resumen (máximo de 150 palabras) en castellano (RESUMEN), e inglés (ABSTRACT) y las "Palabras Clave", también en castellano e inglés (Keywords). Los artículos NO pueden ser notas de clase y cuando se trate de una traducción o del

uso de material protegido por "Derechos de propiedad intelectual" deberá contar con las debidas autorizaciones de los autores.

Todo artículo se someterá a la evaluación de dos jurados, designados por el Comité Editorial. Las excepciones a este requisito serán decididas por tal Comité.

Respecto a gráficas, tablas y figuras, éstas deberán procesarse como "objetos" (para facilitar la diagramación de la revista); para ello puede utilizarse el programa que se desee. Se sugiere enviarlas en blanco y negro o tonos de gris, evitando las policromías.

Las figuras que no estén en medio electrónico, se presentan en fotografía o papel albanene y tinta china.

Si el texto incluye fotografías, se recomienda su presentación en blanco y negro, mate, con buen contraste. Las fotografías se deben montar en hojas del mismo tamaño del texto. En los demás aspectos el tratamiento es similar al de las figuras.

La presentación del artículo estará de acuerdo con los artículos de este número. No se admiten "pie de página" por razones de diagramación de la revista.

Al pie del título del artículo debe incluirse el nombre del autor (o autores), su identificación profesional o filiación institucional, su dirección postal y electrónica (e-mail).

Las referencias deberán ir al final con el siguiente formato:

Artículos de revistas:
Alien, J. S., Samuelson, R. and Newberger , A. Chaos in a model offorced quasi-geostrophic flow over topography: an application of Melinkov's method, J. Fluid Mech., 226, 511-547, 1991.

Libros:
Baker, G. L. and Golub , J., Chaotíc Dynamics: An introduction, Cambridge University Press, Cambridge, 1990.

Capítulo de libro:
Lewis, P., Ray, B. and Stevens, J.G. Modeling time series by using Multivariate Adaptive Regression Spiines (MARS), En: Time Seríes Prediction: Forecasting the Futuro and Understanding the Past (Eds. A.S.Weigend y N. A. Gershenfeid), SFI Studies in the Science of Complexity, Proc. Vol. XV, Addison-Wesley, 297-318, 1994.

Memorias de congresos:
Álzate, N., Botero, T. y Correa, D. Título de la Ponencia. Memorias XIX Congreso Latinoamericano de Ponencias Científicas. Córdoba, Argentina, Tomo II, 219-228, Octubre 2000.

Reporte de un organismo o Gobierno:
U.S. EPA. Status of Pesticides in Registration and Special Review. EPA 738-R-94-008. Washington, DC:Ü.S. Environmental Protection Agency, 1994

Tesis:
Jacobs J. Regulación of Life History Strategies within Individuáis in Predictable and Unpredictable Environments [PhD Thesis]. Seattie, WA: University of Washington, 1996.

Referencias de Internet:
NOAA-CIRES Climate Diagnostics Center. Advancing Understanding and Predictions of Climate Variability. Available: http://www.cdc.noaa.gov [citado 8 de Agosto de 1998].

Datos no publicados:
Apellido, N. Datos no publicados

Comunicación Personal:
Apellido, N. Comunicación personal

OTRAS CONSIDERACIONES

El autor siempre conservará una copia del trabajo. Cada manuscrito luego de sometido al proceso de evaluación por pares seleccionados por el Comité Editorial, se decidirá o no su publicación y sugerirá posibles modificaciones cuando lo crea oportuno. Todo este trámite será lo más breve posible.

El Comité Editorial de la Revista acusará recibo de los originales e informará al autor sobre su aceptación, mediante comunicaciones vía correo electrónico.

La redacción de la revista pondrá el máximo cuidado en evitar errores en la trascripción definitiva de los artículos a publicar, no obstante no se responsabiliza de los errores que puedan aparecer. Si los autores detectaran errores importantes deberán dirigirse cuanto antes a la redacción para realizar una fe de erratas en el siguiente número de la revista.

La redacción se reserva el derecho de realizar pequeñas adecuaciones en los títulos de los artículos, así como rectificaciones menores en la redacción, en las traducciones, Resúmenes y Abstracts; aunque la política general que se seguirá será la de consultar a los autores sobre estas cuestiones.

Simultaneous localization of a monocular camera and mapping of the environment in real time
Andrés Díaz, Lina Paz, Eduardo Caicedo & Pedro Piniés

Microwave heating and separation of water-in-oil emulsion from mexican crude oil
Adrian Vazquez V., Arturo López M., Luis J. Andrade C. & Ariana M. Vázquez A.

A stability and sensitivity analysis of parametric functions in a sedimentation model
Carlos D. Acosta, Raimund Bürger & Carlos E. Mejia

E-maintenance platform: A business process modelling approach
Mónica López-Campos, Salvatore Cannella & Manfredi Bruccoleri

Long term performance of existing Portland cement concrete pavement sections – Case study
Shad M. Sargand, Carlos A. Vega-Posada & Luis G. Arboleda-Monsalve

Analysis of the moisture susceptibility of hot bituminous mixes based on the comparison of two laboratory test methods
Fernando Moreno-Navarro, Gema García-Travé, María Del Carmen Rubio-Gámez & Maria José Martinez-Echevarria

Patellofemoral model of the knee joint under non-standard squatting
Gusztáv Fekete, Béla Málnási Csizmadia, Magd Abdel Wahab, Patrick De Baets, Libardo V. Vanegas-Useche & István Bíró

Identifying dead features and their causes in product line models: An ontological approach
Gloria Lucia Giraldo, Luisa Rincón-Pérez & Raúl Mazo

Anaerobic digestion of fique bagasse: An energy alternative
Humberto Escalante H., Carolina Guzmán L. & Liliana Castro M.

Electric vehicle emulation based on inertial flywheel and a DC machine
Iván Alcalá, Abraham Claudio, Gerardo Guerrero, Jesús Aguayo Alquicira & Víctor Hugo Olivares

Barrier property determination and lifetime prediction by electrochemical impedance spectroscopy of a high performance organic coating
Jorge Andrés Calderón-Gutiérrez & Franky Esteban Bedoya-Lora

Effect of the addition of phosphotungstic acid on the thermal stability of zirconium oxide
Juan Manuel Hernández-Enríquez, Ricardo García-Alamilla, Guillermo Sandoval-Robles, José Aarón Melo-Banda & Luz Arcelia García-Serrano

Landmine detection technologies to face the demining problem in Antioquia
Lorena Cardona, Jovani Jiménez & Nelson Vanegas

Electronic, structural and ferroelectric properties of the Ba$_2$ZrTiO$_6$ double Perovskite
David A. Landínez-Téllez, Críspulo Enrique Deluque-Toro & Jairo Roa-Rojas

Effects of flotation variables on feldspathic sand concentration
Alejandro Argüelles-Díaz, Javier Taboada-Castro, Fernando García-Bastante & María Araújo-Fernández

The content of manufacturing strategy: A case study in colombian industries
Jorge Andrés Vivares-Vergara, William Ariel Sarache-Castro & Julia Clemencia Naranjo-Valencia

Computer-aided diagnosis of brain tumors using Image enhancement and fuzzy logic
Jean Marie Vianney Kinani, Alberto J. Rosales-Silva, Francisco J. Gallegos-Funes & Alfonso Arellano

Free and constrained sintering of 3-mol % yttria stabilised zirconia
Luz Stella Arias-Maya

Iron ore sintering. Part 2. Quality indices and productivity
Javier Mochón, Alejandro Cores, Íñigo Ruiz-Bustinza, Luis Felipe Verdeja, José Ignacio Robla & Fernando García-Carcedo

Advance of foundation techniques in Brazil since colonial times
Renato P. Cunha & Paulo J. R. Albuquerque

Analysis of CSP-1 under infallible and fallible inspection systems
Rodrigo A. Barbosa-Correa & Gina M. Galindo-Pacheco

Two-dimensional computational modeling of the electrokinetic remediation of a copper-contaminated soil. Part I: Model validation
Virgilio Rubio-Nieblas, Manuel Pérez-Tello, Richard A. Jacobs, Ronaldo Herrera-Urbina & Sergio A. Moreno-Zazueta

Two-dimensional computational modeling of the electrokinetic remediation of a copper-contaminated soil Part II: Sensitivity analysis for a triangular soil field
Virgilio Rubio-Nieblas, Manuel Pérez-Tello, Richard A. Jacobs & Ronaldo Herrera-Urbina

Evaluation of stress cracking on geomembranes after accelerated tests
Fernando Luis Lavoie, Benedito De Souza-Bueno & Paulo César Lodi

Smart meters adoption: Recent advances and future trends
Natalia Castaño-Jaramillo, Carlos Jaime Franco-Cardona & Juan David Velásquez-Henao

A first course in software engineering methods and theory
Carlos Zapata & Ivar Jacobson

Localización de una cámara monocular y mapeo simultáneo del entorno en tiempo real
Andrés Díaz, Lina Paz, Eduardo Caicedo & Pedro Piniés

Calentamiento por microondas y la separación de emulsión de agua en aceite del petróleo crudo mexicano
Adrian Vazquez V., Arturo Lopez M., Luis J. Andrade C. & Ariana M. Vazquez A.

Un análisis de estabilidad y sensibilidad de las funciones definidas por parámetros en un modelo de sedimentación
Carlos D. Acosta, Raimund Bürger &Carlos E. Mejia

Modelado de una plataforma de e-mantenimiento, CMMS y CBM
Mónica López-Campos, Salvatore Cannella & Manfredi Bruccoleri

Desempeño a largo plazo de secciones de pavimentos existentes de concreto con cemento tipo Portland – Estudio de caso
Shad M. Sargand, Carlos A. Vega-Posada & Luis G. Arboleda-Monsalve

Análisis de la susceptibilidad al agua en mezclas bituminosas en caliente mediante el estudio comparativo de dos métodos de ensayo de laboratorio
Fernando Moreno-Navarro, Gema García-Travé, María Del Carmen Rubio-Gámez & Maria José Martinez-Echevarria

Modelo patelofemoral de la articulación de la rodilla bajo sentadilla no estándar
Gusztáv Fekete, Béla Málnási Csizmadia, Magd Abdel Wahab, Patrick De Baets, Libardo V. Vanegas-Useche & István Bíró

Identificando características muertas y sus causas en modelos de líneas de productos: Un enfoque ontológico
Gloria Lucia Giraldo, Luisa Rincón-Pérez & Raúl Mazo

Digestión anaerobia del bagazo de fique: Una alternativa energética
Humberto Escalante H., Carolina Guzmán L. & Liliana Castro M.

Emulación de un vehículo eléctrico basada en volante de inercia y máquina de CD
Iván Alcalá, Abraham Claudio, Gerardo Guerrero, Jesús Aguayo-Alquicira & Víctor Hugo-Olivares

Determinación de propiedades barrera y de predicción de tiempo de vida por espectroscopía de impedancia electroquímica de un recubrimiento orgánico de alto rendimiento
Jorge Andrés Calderón-Gutiérrez & Franky Esteban Bedoya-Lora

Efecto de la adición de ácido fosfotúngstico en la estabilidad térmica del óxido de circonio
Juan Manuel Hernández-Enríquez, Ricardo García-Alamilla, Guillermo Sandoval-Robles, José Aarón Melo-Banda & Luz Arcelia García-Serrano

Tecnologías para la detección de minas frente al problema de desminado en Antioquia
Lorena Cardona, Jovani Jiménez & Nelson Vanegas

Propiedades electrónicas, estructurales y ferroeléctricas de la doble Perovskita Ba$_2$ZrTiO$_6$
David A. Landínez-Téllez, Críspulo Enrique Deluque-Toro & Jairo Roa-Rojas

Efectos de las variables de flotación en la concentración de arenas feldespáticas
Alejandro Argüelles-Díaz, Javier Taboada-Castro, Fernando García-Bastante & María Araújo-Fernández

El contenido de la estrategia de manufactura: Un estudio de caso en las industrias colombianas
Jorge Andrés Vivares-Vergara, William Ariel Sarache-Castro & Julia Clemencia Naranjo-Valencia

Diagnóstico asistido por computadora de tumores cerebrales utilizando mejora de la imagen y lógica difusa
Jean Marie Vianney-Kinani, Alberto J. Rosales-Silva, Francisco J. Gallegos-Funes & Alfonso Arellano

Sinterizado libre y con restricción de zirconia parcialmente estabilizada con itria 3-mol %
Luz Stella Arias-Maya

Sinterización de minerales de hierro. Parte 2. Índices de calidad y productividad
Javier Mochón, Alejandro Cores, Íñigo Ruiz-Bustinza, Luis Felipe Verdeja, José Ignacio Robla & Fernando García-Carcedo

Progreso de técnicas de cimentación en Brasil desde el tiempo colonial
Renato P. Cunha & Paulo J. R. Albuquerque

Análisis de planes de muestreo CSP-1 considerando sistemas de inspección con y sin error de muestreo
Rodrigo A. Barbosa-Correa & Gina M. Galindo-Pacheco

Modelación computacional en dos dimensiones de la remediación electrocinética de un suelo contaminado con cobre. Parte I: Validación del modelo
Virgilio Rubio-Nieblas, Manuel Pérez-Tello, Richard A. Jacobs, Ronaldo Herrera-Urbina & Sergio A. Moreno-Zazueta

Modelación computacional en dos dimensiones de la remediación electrocinética de un suelo contaminado con cobre. Parte II: Análisis de sensibilidad para un campo de suelo triangular
Virgilio Rubio-Nieblas, Manuel Pérez-Tello, Richard A. Jacobs & Ronaldo Herrera-Urbina

Evaluación de fisuración bajo tensión de geomembranas pos ensayos acelerados
Fernando Luis Lavoie, Benedito De Souza-Bueno & Paulo César-Lodi

Adopción de medidores inteligentes: Avances recientes y tendencias futuras
Natalia Castaño-Jaramillo, Carlos Jaime Franco-Cardona & Juan David Velásquez-Henao

Un curso inicial sobre teoria y métodos de la ingeniería de software
Carlos Zapata & Ivar Jacobson

DYNA

Publication admitted to the
Sistema Nacional de Indexación y Homologación de
Revistas Especializadas CT+I - PUBLINDEX, Category A1

80 DYNA años

"Si ésta revista despierta en algunos curiosidad por la investigación científica, nuestro trabajo quedará cumplido"
Joaquín Vallejo, 1933. Fundador

UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS