Ingeniería e Investigación
Journal

Abbreviated Journal Title: **Ing. Investig.**

Editor-in-chief
Sonia C. Mangones, PhD

Associate Editor
Andrés Pavas, PhD, MSc

Technical Editor
Lenin Alexander Bulla Cruz, PhD, MSc

Editorial Assistants
Julian Arcila-Forero, MSc, BSc
Ingri Gisela Camacho, BSc

Frequency
Continuous periodicity (three issues per year)

Cover Layout
Carlos Andres Ortiz Valle

Proofreader
José Daniel Gutiérrez-Mendoza

Layout Artist
David Mauricio Valero

For additional information contact
revii_bog@unal.edu.co
Bogotá - Colombia
2023

Table of Contents

Papers published in *Ingeniería e Investigación* are abstracted/indexed in

- Science Citation Index Expanded
- (SciSearch®), Clarivate Analytics
- Scopus - Elsevier
- Scientific Electronic Library Online - SciELO, Colombia
- Chemical Abstract
- Índice de Revistas Latinoamericanas
- en Ciencias Periódica
- Dialnet
- Sistema Regional de Información en Línea para
- Revistas Científicas de América Latina, el Caribe, España y Portugal - Latindex
- Ebsco Publishing
- DOAJ - Directory of Open Access Journals
- Redib - Red Iberoamericana de Innovación y Conocimiento Científico

The journal *Ingeniería e Investigación* was created in 1981. It is an entity in charge of spreading the teaching, Scientific, and technical research conducted at Universidad Nacional de Colombia's Department of Engineering and other national and international institutions. *Ingeniería e Investigación* deals with original, unedited scientific research and technological developments in the various disciplines related to engineering. *Ingeniería e Investigación* contributes the development of knowledge, generating a global impact on academia, industry, and society at large through an exchange of knowledge and ideas while maintaining a set of serious and recognized quality standards.

Ingeniería e Investigación publications are developed for the academic community who is interested in research and engineering knowledge development. We invite readers to be part of this Journal and participate either as authors, peer reviewers, or subscribers.

For additional information contact:
www.revistas.unal.edu.co/index.php/ingeinv
E-mail: revii_bog@unal.edu.co
Tel: 57(1) 3 16 5000 Ext. 13674

Tabla de Contenido

Research Article / Agricultural Engineering

http://doi.org/10.15446/ing.investig.99102

Near-Infrared Spectroscopy: Assessment of Soil Organic Carbon Stock in a Colombian Oxisol

Espectroscopia de infrarrojo cercano: evaluación del almacenamiento de carbono orgánico del suelo en un oxisol colombiano

Felipe Fernández-Martínez[1], Jesús Hernán Camacho-Tamayo[2], and Yolanda Rubiano-Sanabria[3]

ABSTRACT

Soil organic carbon (SOC) is a property known for its influence on the physical, chemical, and biological characteristics of soils, which are essential when assessing their quality. SOC stock (SOCS) monitoring is a key task in climate change mitigation studies. However, the resources necessary to obtain the information required by these studies tend to be high. The objective of this study was to develop a model for estimating the SOCS of a Colombian oxisol using near-infrared (NIR) diffuse reflectance spectroscopy. In a sampling scheme of 70 points distributed over 248 ha, 313 soil samples were collected in five defined depth intervals of 10 cm each, from 0 to 50 cm. SOC was determined through an elemental analyzer, and bulk density (BD) by means of sampling cylinders. A NIRFlex spectrometer was used to acquire spectral signatures in the NIR range from the processed soil samples, and, together with the data measured in the laboratory, a statistical analysis was performed using partial least squares regression (PLSR) in order to calibrate the spectral models. Based on the residual prediction deviation (RPD), the root mean square error (RMSE), and the coefficient of determination (R^2) of the validation groups, a highly representative model was achieved for the estimation of SOCS (R^2 = 0,93; RMSE = 2,12 tC ha^{-1}; RPD = 3,69), which was also corroborated with geostatistical interpolation surfaces and depth splines. This research showed NIR diffuse reflectance spectroscopy to be a viable technique for SOCS estimation in the study area.

Keywords: soil organic carbon, bulk density, soil spectroscopy, spline, geostatistics

RESUMEN

El carbono orgánico del suelo (COS) es una propiedad conocida por su influencia en las propiedades físicas, químicas y biológicas de los suelos, que son fundamentales para evaluar su calidad. El monitoreo del stock de COS (SCOS) es una labor clave en los estudios de mitigación del cambio climático. Sin embargo, los recursos necesarios para obtener la información requerida en estos estudios suelen ser elevados. El objetivo de este estudio fue desarrollar un modelo para estimar el SCOS de un oxisol colombiano utilizando espectroscopía de reflectancia difusa de infrarrojo cercano (NIR). En un esquema de muestreo de 70 puntos distribuidos en 248 ha, se recolectaron 313 muestras de suelo en cinco intervalos de profundidad definidos de 10 cm cada uno, de 0 a 50 cm. El COS se determinó mediante un analizador elemental, y la densidad aparente (DA) mediante cilindros de muestreo. Se utilizó un espectrómetro NIRFlex para adquirir firmas espectrales en el rango NIR de las muestras de suelo procesadas, y, junto con datos medidos en laboratorio, se realizó un análisis estadístico usando regresión de mínimos cuadrados parciales (RMCP) para calibrar los modelos espectrales. Con base en la desviación de predicción residual (DRP), raíz del error cuadrático medio (RECM) y el coeficiente de determinación (R^2) de los grupos de validación, se logró un modelo de alta representatividad para la estimación de SCOS (R^2 = 0,93; RECM = 2,12 tC ha^{-1}; DRP = 3,69), lo cual también se corroboró con superficies de interpolación geoestadística y splines de profundidad. Esta investigación mostró que la espectroscopia de reflectancia difusa NIR es una técnica viable para la estimación de SOCS en el área de estudio.

Palabras clave: carbono orgánico del suelo, densidad aparente, espectroscopía de suelos, spline, geoestadística

Received: October 19th, 2021
Accepted: March 21th, 2022

Introduction

In recent decades, scientists around the world have been increasingly interested in studying soil organic carbon (SOC) and its relationship with climate change. A potential impact of global warming is the accelerated decomposition of SOC and an increase in the carbon released into the atmosphere (Jia *et al*., 2017). Rising levels of CO_2 concentration in the atmosphere have led a considerable part of the scientific community to contemplate SOC sequestration as an alternative to mitigate climate change, which can be achieved through various management practices based on increasing the SOC stock (SOCS) (Huang *et al*., 2019).

[1] Agricultural engineer, MSc in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Graduate research assistant and PhD candidate in Civil Engineering – Water Resources Engineering, Mississippi State University, United States. Email: ff181@msstate.edu

[2] Agricultural engineer, Universidad Nacional de Colombia, Colombia. MSc in Agricultural Engineering, Universidad de Campinas, Brazil. PhD in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Associate professor, Universidad Nacional de Colombia, Colombia. Email: jhcamachot@unal.edu.co

[3] Agrologist, Universidad Jorge Tadeo Lozano, Colombia. PhD in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Associate professor, Universidad Nacional de Colombia, Colombia. Email: yrubianos@unal.edu.co

Through photosynthesis, plants are able to capture carbon from the atmosphere and integrate it into their own structure (biomass), as well as into the soil through their radicular systems. This phenomenon can also be called *carbon sequestration* (Stockmann *et al*., 2013) and has a special focus on agricultural land due to the fact that around 37% of the planet's surface is destined for agricultural use (Sommer and Bossio, 2014).

In order to monitor the SOCS, it is necessary to acquire information on the soil bulk density (BD) and the SOC and subsequently analyze the changes in these variables over different periods of time according to the established needs. However, the resources needed to conduct these kinds of studies are expensive and time-consuming because of the standard laboratory analyses and procedures involved (Camacho-Tamayo *et al*., 2014; Liu *et al*., 2019). Soils are also characterized by being a heterogeneous material in both spatial and temporal scales, often causing SOCS studies to require a high sampling density (Cambou *et al*., 2016). Therefore, it is necessary to explore SOCS analysis techniques that represent a reduction in the resources required.

Chemometric approaches constitute a potential tool to evaluate soil attributes rapidly and accurately in the laboratory (Ben Dor *et al*., 2015; Davari *et al*., 2021). Diffuse reflectance spectroscopy is a technique developed to accelerate data acquisition regarding materials' properties. Soil spectroscopy is based on the idea that the characteristics of the radiation reflected by a material when exposed to an electromagnetic spectrum depend on the composition of the material itself. This means that studying soil reflectance can provide information about the properties and condition of the soil under study (Davari *et al*., 2021; Viscarra Rossel and Webster, 2012). The features of the reflected spectra of a soil sample in the NIR spectrum have allowed for optimal results in terms of predicting soil properties, especially when dealing with SOC assessments (Liu *et al*., 2019; Nawar and Mouazen, 2019; Nocita *et al*., 2014).

It is important to mention that the acquisition of spectral signatures can be carried out *in situ* or using laboratory equipment. This can affect the information obtained from the reflectance spectra, causing variations in the model performance with regard to the estimated properties (Ahmadi *et al*., 2021; Ben Dor *et al*., 2015). When spectral signatures are taken *in situ*, factors such as the contact probe device of the equipment and its operator, the moisture of the soil sample, and environmental conditions can alter the spectral reflectance of the samples in comparison with those acquired in the laboratory. The latter would then be assumed to have been acquired under controlled environmental conditions, as well as after being dried and sieved for the sake of homogeneity.

The objective of this research was to evaluate the potential of NIR spectroscopy in order to estimate the SOCS of oxisols in the study area. Several models were developed by varying the number of samples used for calibration, followed by the evaluation of the spatial variability of the measured and estimated data by means of depth splines and geostatistical interpolated surfaces.

Materials and methods

Study area and field sampling

This study was carried out at the Carimagua Experimental Station of Agrosavia, located in the municipality of Puerto Gaitán (Meta, Colombia), with geographic coordinates 4° 34′ 48″ N, 71° 21′ 00″ W, an average altitude of 175 masl, an average annual temperature of 28 °C, an average rainfall of 2 240 mm, and a warm-humid climate. The relief of the study area is flat to slightly undulating, with slopes ranging between 0 and 7%. Soils are mostly oxisols whose dominant taxonomic component is typic hapludox. These soils have a low pH (<5), are well drained, and are covered with native savanna. Their predominant use is extensive livestock farming.

The study area has an extension of 248 ha, where a directed sampling scheme was established. Sampling took place in 2018 and consisted of 70 points spaced perpendicularly by 200 m, from which soil samples were extracted in depth intervals of 0-10, 10-20, 20-30, 30-40, and 40-50 cm, for a total of five samples per trunk (Figure 1). The total number of samples collected was 313, out of which 70 corresponded to the first and second depths, 68 to the third, 59 to the fourth, and 46 to the last. The difference in samples collected per depth is attributed to the presence of the water table in the study area. Each of the 313 soil samples consisted of a disturbed sample and an undisturbed cylinder sample.

Laboratory analysis

The disturbed samples were dried at 35 °C and then sifted through a 2 mm mesh. Total soil carbon was determined using an elemental analyzer (TruSpec CN Carbon Nitrogen Determinator, LECO Corp., St. Joseph, MI, USA). Since the soil samples did not show the presence of carbonates in the hydrochloric acid test, the total carbon was considered to be the SOC. The undisturbed soil sample cylinders were oven-dried at 105 °C for 48 h, and then the BD was calculated by dividing the mass of the soil by the volume of the cylinder. The SOCS was determined according to Equation (1):

$$SCOS\left[\frac{t}{ha}\right] = DT\left[dm\right] * SOC\left[\frac{gC}{kg}\right] * \left(1 - Pd * 0{,}01\right) \quad (1)$$

where DT stands for the depth and thickness of the soil layer, and Pd stands for the stoniness factor. In this case, Pd was assumed to be 0 because no significant stoniness was evidenced in the sampling process. Spectral signatures were recorded using a NIRFlex N-500 spectrometer (BÜCHI Labortechnik AG), which recorded measurements in

reflectance units in the NIR range, from 1 000 nm to 2 500 nm, averaging 32 scans for each wavelength. The spectral signatures were taken from the dried and sieved soil samples.

Spectral modeling

To perform spectral modeling, it was first necessary to define the calibration and validation groups of the spectral models. The Kennard-Stone algorithm (Kennard and Stone, 1969) made it possible to establish uniformly distributed groups according to the variance of the spectral signatures. The calibration groups for BD and SOC were defined with 70% of the total signatures, while the validation group was assigned the remaining 30%. For SCOS spectral modeling, calibration groups consisting of 10, 20, 30, 60, 90, 120, 150, 180, 210, 240, and 270 samples were built with the Kennard-Stone algorithm. A subsequent validation of each group was carried out with the samples that did not take part in the calibration model.

In the model calibrations, spectral data were submitted to a series of pretreatments to normalize the responses and smooth out any possible noise that could be present in the signatures. The pretreatments considered were the transformation of raw spectra to absorbance, the Savitsky-Golay derivate, and standard normal variation. Once the combinations of pretreatments were applied, the model was built by means of PLSR, which is a commonly used technique when dealing with high-dimensional data (such as the wavelengths of the NIR spectral range) with highly collinear predictor variables. This technique summarizes the data into a few orthogonal factors, which are linear combinations of the predictor variables, so that the covariance of both the predictor and dependent variables is maximized. These orthogonal factors are then employed to develop a linear model to estimate the soil's attribute of interest (Katuwal *et al.*, 2020).

The criteria to establish the performance of each model were set by the values obtained for the coefficient of determination (R^2), root mean square error (RMSE), and residual prediction deviation (RPD), the latter being the ratio of the standard deviation of the laboratory-measured data to the validation RMSE. The lower the RMSE and the higher the R^2, the better the estimations reached. According to Viscarra Rossel *et al.* (2006), the RPD can rate a model as unrepresentative if it reaches a value lower than 1,4; as regular with values between 1,4 and 1,8; as well-performing with 1,8-2,0; as a very good model with 2,0-2,5; and as an excellent model with a RPD value greater than 2,5. The construction and evaluation of spectral models was carried out with the R software (R Core Team) and the libraries *prospectr* (Stevens and Ramírez-Lopez, 2020) and *pls* (Liland *et al.*, 2021).

Spatial variation analysis

The vertical behavior of the estimated and measured data was verified by means of depth splines. Splines are a set of quadratic functions joined by nodes (depths measured) that represent a smoothed curve (Ponce-Hernandez *et al.*, 1986). Regarding the horizontal plane, by means of geostatistical interpolation surfaces, the likeness of the

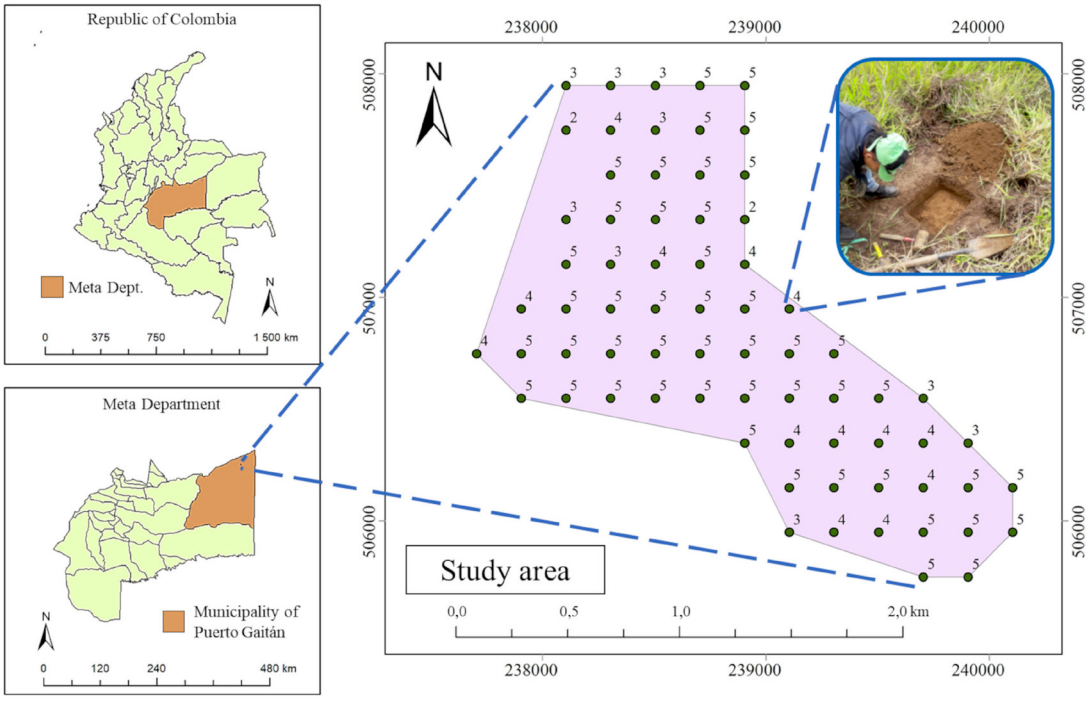


Figure 1. Location of the study area detailing the number of samples obtained for each trunk
Source: Authors

measured and estimated SOCS was observed. To this effect, experimental semivariograms were built and then adjusted to fit the theoretical models. These models have three common parameters: the nugget effect (C_0), the sill (C_0 + C_1), and the range. The adjustment of the theoretical model was based on the lowest value of the sum of the squared residuals, the highest values of R^2 of the fitted model, and the cross-validation coefficient (CVC). Afterwards, spatial interpolation was carried out by ordinary kriging, which is regarded as an optimal estimation method that provides unbiased predictions while minimizing the variance (Oliver and Webster, 2015). In addition, the degree of spatial dependence (DSP) was assessed, which is the relationship between the nugget effect and the sill of the semivariograms (C_1 * (C_0 + C_1)$^{-1}$). According to Cambardella *et al*. (1994), the DSP is classified as weak if it is lower than 25%, as moderate if it is between 25 and 75%, and as strong above 75%. Geostatistical processing was performed with the GS+ software v.9 (Gamma Design Software, LLC, Plainwell, MI, USA), in combination with ArcGIS v.10.8 (ESRI), while depth splines were built with R (R Core Team) and the *ithir* library (Malone, 2016).

Results and discussion

Exploratory analysis and spectral signatures

For the statistical analysis of each variable, an exploratory data study was initially carried out with R, calculating measures of location (mean, median, minimum, and maximum), variability (coefficient of variation, CV), and central tendency (skewness and kurtosis). Values close to zero for skewness and kurtosis indicated that the precept of normality was met. The descriptive statistics for the behavior of BD and SOC across the studied profile are shown in Table 1. It was evidenced that, across the soil profile, the behaviors of SOC and BD are inversely proportional. The BD reached a mean value of 1,33 g cm^{-3} at the first depth, which kept increasing until 1,49 g cm^{-3} for the last interval. For the SOC, the average content in the range of 0-10 cm was 2,61%, which kept decreasing until reaching a value of 0,95% at 40-50 cm. The higher values of SOC at lower depths can be attributed to the presence of vegetation and the natural contribution of surface residues. A similar behavior of these two properties was also reported by other investigations in the oxisols of the Eastern Plains in Colombia (Camacho-Tamayo *et al*., 2014; Ramirez-Lopez *et al*., 2008)

The average spectral signature for all soil samples grouped by soil depth is shown in Figure 2. The behavior shown by the spectral signatures at each sampling depth can also be qualitatively related to some of the soil properties. In the NIR, at wavelength ranges of 1 000-1 350 nm and 1 400-2 200 nm, spectral responses tend to be smoother and have a lower magnitude of reflectance in soils with a higher content of organic material (Poppiel *et al*., 2018). This is because organic matter affects the albedo of soils, which makes them reflect less light at higher levels of organic content (organic carbon) (Camacho-Tamayo *et al*., 2014). The spectral signatures obtained in this study showed a spectral reflectance according to these characteristics, as higher SOC contents were observed at lower sampling depths, which is consistent with the lower magnitudes of reflectance at a greater proximity to the soil surface.

BD and SOC modeling performance

The performance of the most accurate spectral models for BD and SOC is shown in Table 2. These were calibrated with 70% of the total signatures and validated with the remaining 30%. The BD spectral model turned out to be unrepresentative, reaching a RPD of 1,35 with an R^2 equal to 0,41 during its validation. On the other hand, SOC showed excellent performance, reaching RPD and R^2 values of 5,63 and 0,97, respectively.

Table 1. Descriptive statistics for BD and SOC per depth with measured data

Property	Depth [cm]	Mean	Median	Max,	Min,	SD	CV [%]	Skewness	Kurtosis
BD [g cm^{-3}]	0-10	1,33	1,33	1,58	1,14	0,10	7,78	0,18	-0,38
	10-20	1,41	1,42	1,55	1,28	0,06	4,43	-0,01	-0,62
	20-30	1,41	1,41	1,58	1,25	0,09	6,14	-0,06	-0,74
	30-40	1,43	1,46	1,63	1,24	0,10	7,07	-0,11	-0,89
	40-50	1,49	1,50	1,73	1,29	0,10	6,70	-0,09	-0,28
SOC [%]	0-10	2,61	2,63	3,24	1,98	0,26	10,08	-0,01	-0,12
	10-20	1,75	1,72	2,16	1,33	0,18	10,51	0,07	-0,51
	20-30	1,39	1,37	1,75	1,08	0,16	11,33	0,38	-0,42
	30-40	1,12	1,08	1,50	0,85	0,17	14,78	0,57	-0,59
	40-50	0,95	0,98	1,31	0,58	0,16	16,50	-0,06	-0,30

SD: standard deviation; CV: coefficient of variation; Max: maximum; Min: minimum

Source: Authors

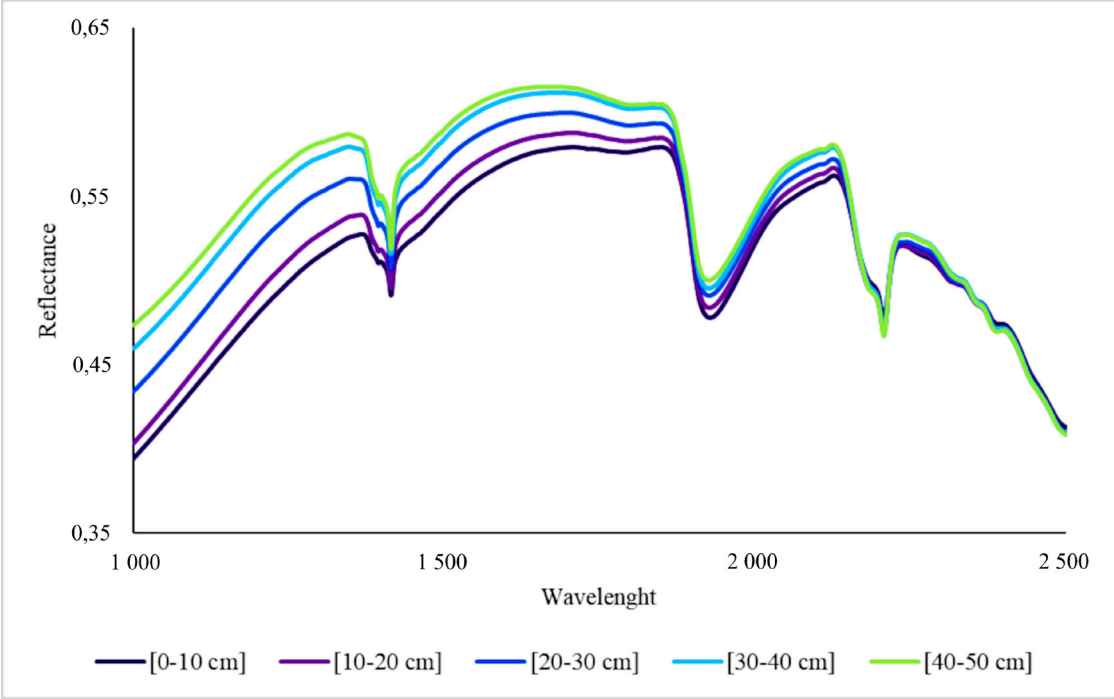


Figure 2. Spectral signatures averaged for each sampling depth
Source: Authors

Table 2. Calibration and validation results for the BD and SOC spectral models

Property	Model	R^2	RMSE	RPD	Range
BD [g cm^{-3}]	Calibration	0,45	0,08	1,36	1,14 - 1,73
	Validation	0,41	0,08	1,35	1,15 - 1,63
SOC [%]	Calibration	0,97	0,11	5,46	0,58 - 3,24
	Validation	0,97	0,11	5,63	0,75 - 3,24

Source: Authors

The underperformance obtained for the spectral model of BD has also been reported in other studies (Katuwal *et al*., 2020; Moreira *et al*., 2009). Particularly, the work by Moreira *et al*. (2009) consisted of 1 184 samples with BD values in the range of 0,45-1,95 g cm^{-3}. They managed to validate models with R^2 between 0,10 and 0,34 by applying different pretreatments to the spectral signatures. Meanwhile, Katuwal *et al*. (2020) achieved slightly higher performances, with R^2 between 0,10 and 0,52 while working with a BD range of 1,02-2,01 g cm^{-3}. BD is a soil property whose estimation through spectral signatures depends on the correlations that it may have with other attributes with recognized associations in the spectral range (Askari *et al*., 2015; Viscarra Rossel and Webster, 2012). For example, Al-Asadi and Mouazen (2014) coupled auxiliary variables such as water content, texture fractions, and organic matter with spectral models of BD in order to improve estimations, going from R^2 values between 0,23 and 0,69 to a 0,70-0,81 range.

The level of performance obtained for the SOC estimations has also been reported in other works, corroborating the potential use of NIR spectroscopy for the prediction of this property (Jia *et al*., 2017; Nawar and Mouazen, 2019; Nocita *et al*., 2014). An example of this can be found in the large-scale study conducted by Liu *et al*. (2019), in which a model composed of 11 213 soil samples with a mean SOC content of 1,84% and magnitudes ranging from 0,02% to 9,96% achieved a validation R^2 equal to 0,96 and a RMSE of 0,29%.

SOCS spectral estimation

The result of the product between BD [kg dm^{-3}] and SOC [gC kg^{-1}] is the volumetric SCOS [gC dm^{-3}] (SCOSv), which was the value estimated from spectral models. The performance of the models calibrated via the established groups ranging from 10 to 300 samples is shown in Table 3. It was evidenced that, by increasing the number of samples used to calibrate the models, the RMSE decreased while the RPD and R^2 increased. This bears out the fact that more robust estimation models can be achieved with more samples. The stabilization of these parameters occurs at around 90 samples for the calibration model, so this number was chosen as optimal for SCOS spectral modeling. However, it is worth highlighting that each of the models defined in Table 3 showed RPD and R^2 values greater than 2,0 and 0,80 in the validation parameters.

According to the RPD, the model made with the calibration group of 90 soil samples gave excellent estimates for SCOSv, with a RMSE of 2,12 gC dm^{-3} and a R^2 of 0,93 (Figure 3). The

calibration range for this model was 8,39-45,55 gC dm^{-3}, while the range was 9,81-41,61 gC dm^{-3} for the validation group. The behavior of the measured and estimated SCOS in the soil profile is shown in Table 4. No large differences between the measured and estimated SOCS were evidenced. The datasets showed similar magnitudes regarding the mean and median, as well as for skewness, kurtosis, and the CV, which was in fact lower at each one of the sampling ranges (estimated *vs.* measured SOCS). The average content regarding the SOCS measured in the whole depth range of 0-50 cm was 109,28 tC ha^{-1}. This value was 110,56 tC ha^{-1} for the estimated data.

Table 3. Results for the calibration and validation of SOCSv spectral modeling

Number of samples	Calibration			Validation		
	R^2	RMSE [gC dm^{-3}]	RPD	R^2	RMSE [gC dm^{-3}]	RPD
10	0,98	1,04	2,87	0,82	3,35	2,34
20	0,92	2,36	1,97	0,84	3,13	2,50
30	0,88	2,74	2,11	0,86	2,95	2,66
60	0,93	2,14	2,95	0,92	2,19	3,58
90	0,93	2,16	3,01	0,93	2,12	3,69
120	0,93	2,17	3,08	0,92	2,15	3,64
150	0,92	2,16	2,97	0,93	2,13	3,68
180	0,93	2,15	3,20	0,93	2,07	3,78
210	0,93	2,09	3,34	0,93	2,04	3,84
240	0,93	2,08	3,42	0,93	2,08	3,77
270	0,93	2,09	3,44	0,93	2,08	3,77

Source: Authors

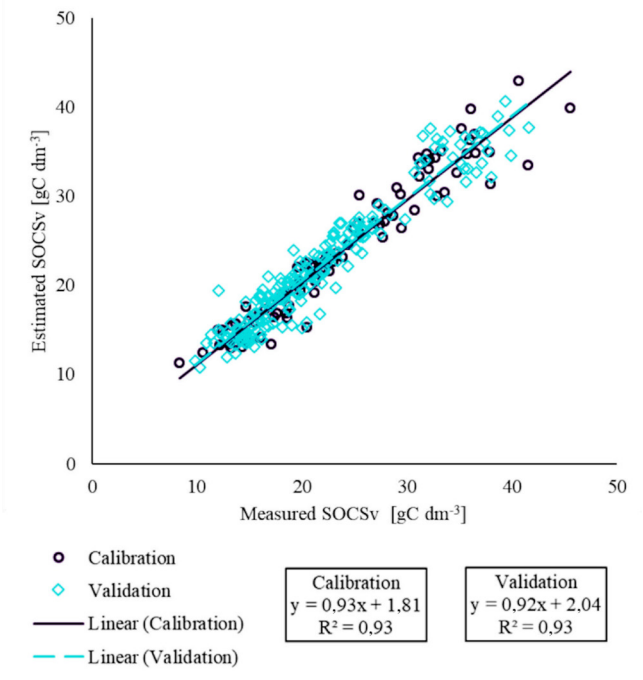


Figure 3. Spectrally estimated *vs.* measured values of SOCSv
Source: Authors

A study carried out by Cambou *et al*. (2016) reported acceptable models regarding the estimation of SCOS from spectral signatures acquired in the field for a haplic luvisol, reaching R^2 and RPD values of 0,70 and 1,80, respectively, for SCOSv values ranging between 7,90 and 27,70 gC dm^{-3}. Part of the flaws in the estimation model of said research were attributed to the fact that the spectral signatures were collected by manually extracted auger soil cores that altered the natural structure of the samples. This, in addition to the fact that conventional laboratory analyses for BD and SOC were not performed on the soil samples from which the spectral signatures were collected (Cambou *et al*., 2016). On the other hand, Allory *et al*. (2019) compared the performance of models calibrated for urban soils using spectral signatures acquired in the field and in the laboratory. With a SCOSv range of 0,30-54,60 gC dm^{-3}, a validation model was obtained with a RPD equal to 2,20 and an R^2 of 0,78 for spectral signatures acquired *in situ*, while, for the laboratory-acquired spectra, the model parameters reported were 3,1 for the RPD and 0,89 for the R^2. The performance improvement of the latter was attributed to the more stable measurement conditions in the laboratory, in comparison with those in the field.

For tropical volcanic soils that included andosols, cambisols, and ferralsols, Allo *et al*. (2020) calibrated models with spectral signatures obtained in the field and in the laboratory, obtaining better results in the validation of the model elaborated from signatures taken *in situ* (R^2 = 0,91 and RPD = 3,29) when compared to the validation of the laboratory spectral model (R^2 = 0,86 and RPD = 2,63). Allo *et al*. (2020) used a mechanical column auger to collect spectra in the field, allowing them to collect 30 signatures uniformly every 10 cm of the sampled soil column to establish an average signature for a single sample, achieving a closer representation of natural conditions. Doing this could have considerably improved the predictive ability of the model with spectra taken *in situ vs.* those acquired in the laboratory.

Given that the magnitudes of SOC and SCOS tend to increase at depths closer to the surface and decrease at lower depths in the soil profile (FAO, 2019; Stockmann *et al*., 2013), the Intergovernmental Panel on Climate Change recommends monitoring these organic C concentrations at least in the depth range of 0-30 cm (IPCC, 2019). This recommendation is also motivated by the fact that the first 30 cm from the soil surface usually experience substantial changes in SOC due to management practices. Hence, the detection of variations in SOC reserves at this depth range could be noticed in shorter periods of time, depending on the activities carried out on them (Huang *et al*., 2019). In this regard, the oxisols of the Eastern Plains in the study area show a cumulative mean value of 78,85 tC ha^{-1} in the first 30 cm of soil depth, which, for estimated spectral data, equals 79,86 tC ha^{-1} and corresponds to approximately 72% of the total SOCS present in the profile considered.

Table 4. Results for the calibration and validation of SOCSv spectral modeling

SCOS [tC ha^{-1}]	Depth [cm]	Mean	Median	Max,	Min,	SD	CV [%]	Skewness	Kurtosis
Measured	0-10	34,64	34,57	45,55	25,44	3,69	10,64	0,18	0,82
	10-20	24,74	25,01	29,85	19,18	2,43	9,83	-0,05	-0,56
	20-30	19,47	19,55	25,81	14,68	2,35	12,05	0,34	-0,07
	30-40	16,11	15,91	21,69	12,05	2,67	16,55	0,42	-0,79
	40-50	14,32	14,64	20,41	8,39	2,43	16,99	-0,10	0,35
Estimated	0-10	34,38	34,31	42,90	27,64	3,08	8,95	0,16	0,12
	10-20	25,32	25,58	29,11	21,56	1,91	7,55	-0,17	-0,85
	20-30	20,16	20,21	25,44	15,76	1,95	9,65	0,22	0,38
	30-40	16,21	15,69	20,89	12,41	2,05	12,65	0,59	-0,38
	40-50	14,49	14,44	18,34	10,84	1,75	12,07	0,28	-0,04

SD: standard deviation; CV: coefficient of variation; Max: maximum; Min: minimum

Source: Authors

SOCS spatial variation

The vertical variation of the soil properties is shown in Figure 4. The splines allowed to estimate the value of soil attributes within each measured depth. Furthermore, the addition of boxplots to these diagrams aided in examining the similarity between measured and estimated data. At first glance, it can be noted that the resulting splines for BD have clear differences, which are attributed to the low representativeness of the spectral model. The interquartile range, as well as the whiskers of the boxplots, showed a considerable reduction for the estimated data in comparison with the measured values. The unrepresentativeness obtained in the spectral modeling for BD can be associated with the alteration of the structure of the soil samples prior to the signature acquisitions, as the soil samples were dried and sieved.

Furthermore, the splines confirmed the excellent performance of spectral models for SOC and SOCSv. Even though a reduction can be perceived in the interquartile ranges of the boxplots of the estimated data when compared to those of the measured data, in general, the ranges of the whiskers of the estimated and measured data are similar. However, it can be pointed out that the loss of performance of the SCOS spectral model in comparison with the SOC one is due to the incorporation of the BD in the estimated variables, since this property could not be estimated directly with the same degree of representativeness.

The results of the geostatistical analysis for the measured and estimated SOCS are shown in Figure 5 and Table 5. The semivariograms were adjusted to spherical and exponential models, which were the same for the measured and estimated data at every depth range except for 10-20 cm. Moreover, for each of the semivariograms in the five sampling depths, the nugget and sill values were lower for the spectral *vs.* the measured data. Therefore, it can be stated that the variance is lower in the estimated data. Lower nugget and sill values for the spectral data have also been reported by other authors (Araújo *et al.*, 2015; Bonett *et al.*, 2016; Camacho-Tamayo *et al.*, 2017; Wetterlind *et al.*, 2008). Among the sampling depths, the greatest variability of the SOCS was evidenced in the semivariograms at the closest range to the soil surface, which is consistent with the natural addition processes taking place in the study area.

The use of spectrally estimated data in the construction of semivariograms did not modify the spatial variation tendency of the SOCS, as was verified in the obtained models and the interpolation surfaces illustrated in Figure 5. DSD shows moderate spatial dependence for each of the sampling depths, which is also consistent for the spectral and measured data. In addition, most R^2 and CVC values were above 0,70.

The geostatistical interpolation surfaces obtained from estimated and measured data showed a high degree of correspondence at each sampling depth. This is a positive indicator for approaching SOCS monitoring through spectral models, given that, once a sufficiently robust model is reached with adequate performance regarding the needs, the volume of information obtained could increase without the need to spend resources on conventional laboratory analyses, which is convenient for addressing soil studies in the context of climate change.

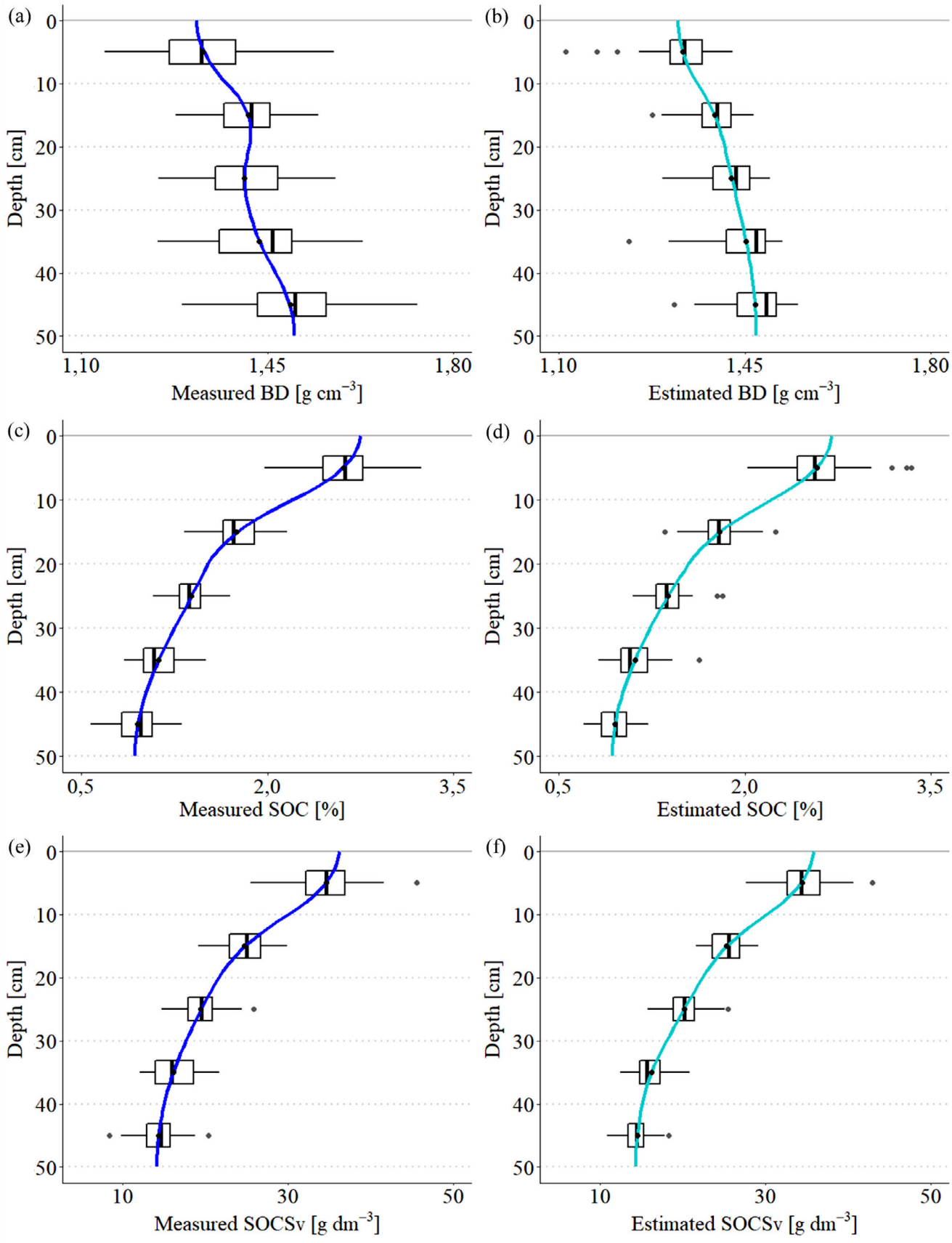


Figure 4. Depth splines for standard laboratory-measured (a, c, e) and spectral estimated data (b, d, f)
Source: Authors

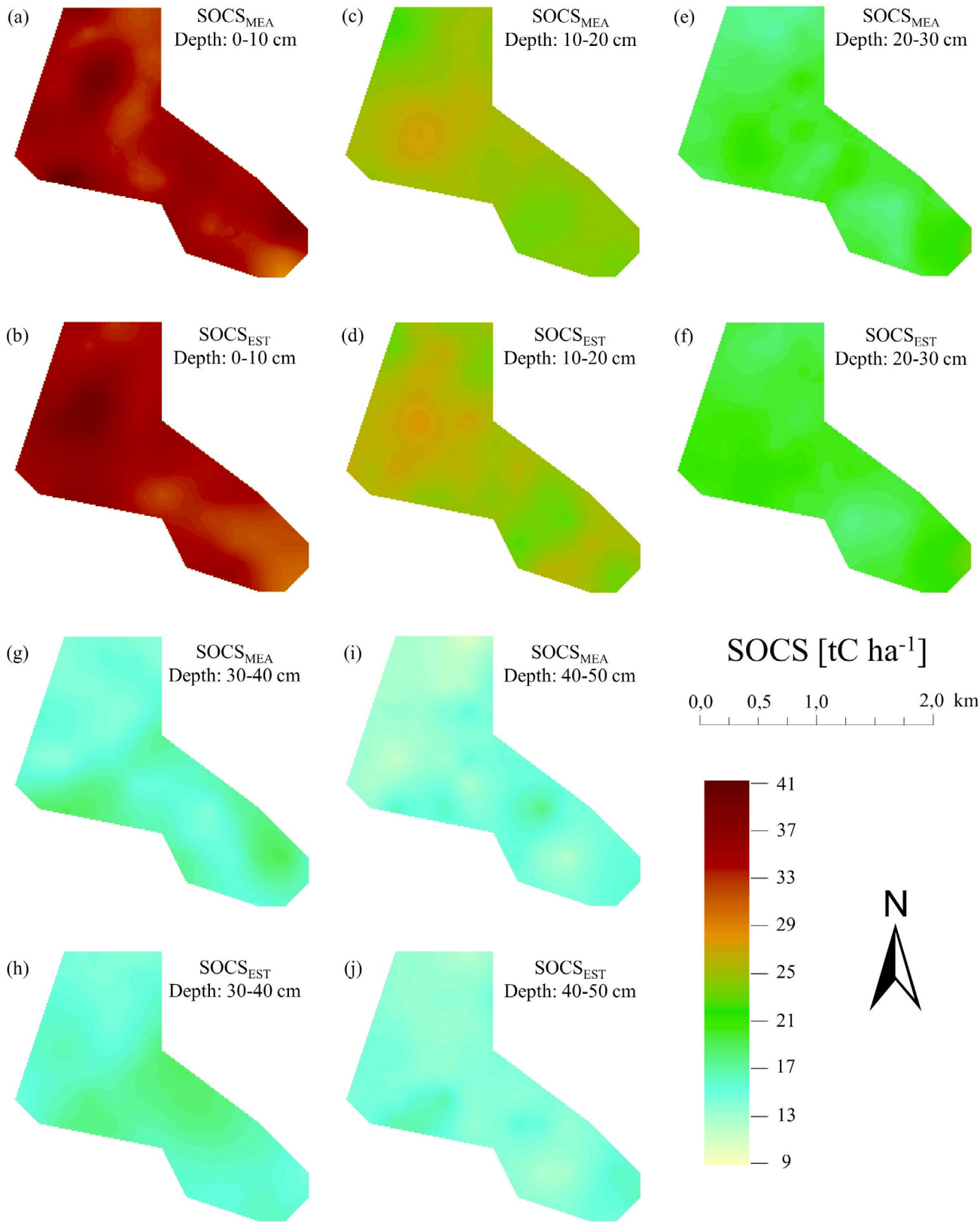


Figure 5. Interpolation surfaces of SOCS from measured (a, c, e, g, i) and estimated (b, d, f, h, j) data
Source: Authors

Table 5. Parameters of the theoretical semivariogram models for SOCS from standard laboratory-measured and spectrally estimated data

SCOS [tC ha^{-1}]	Depth [cm]	Semivariogram model	C_0	C_0+ C_1	Range [m]	DSD [%]	R^2	CVC
Measured	0-10	Spherical	4,50	11,54	424,82	61	0,99	0,73
	10-20	Spherical	2,80	6,00	1 213,49	53	0,83	0,56
	20-30	Exponential	2,29	4,88	558,00	53	0,86	0,67
	30-40	Spherical	2,64	5,75	531,00	54	0,86	0,81
	40-50	Exponential	2,48	6,15	550,00	60	0,74	0,70
Estimated	0 - 10	Spherical	3,83	8,58	627,00	55	0,81	0,79
	10 - 20	Exponential	0,92	3,12	618,00	71	0,91	0,77
	20 - 30	Exponential	1,64	3,48	910,00	53	0,76	0,82
	30 - 40	Spherical	1,76	4,27	1 006,00	59	0,91	0,85
	40 - 50	Exponential	0,68	2,44	630,00	72	0,78	0,98

C_0: nugget; C_0+ C_1: sill; DSD: Degree of Spatial Dependence; CVC: Cross-Validation Coefficient

Source: Authors

Conclusions

NIR spectroscopy allowed estimating, with a high degree of representativeness, the SOCS of oxisol in the study area. The results suggest that laboratory analyses for SOCS could be largely substituted or complemented by an approach involving spectral techniques in the NIR region. This work also showed that data derived from NIR models could be integrated into geostatistical evaluations for Colombian oxisols.

Acknowledgements

The authors gratefully acknowledge the funding received from Universidad Nacional de Colombia via the project with HERMES code No. 37685, which allowed conducting this research.

CRediT author statement

Fernández-Martínez, F., Camacho-Tamayo, J. H., Rubiano-Sanabria, Y.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing (original draft, review, and editing).

Conflicts of interest

The authors declare no conflict of interest.

References

Ahmadi, A., Emami, M., Daccache, A., and He, L. (2021). Soil properties prediction for precision agriculture using visible and near-infrared spectroscopy: A systematic review and meta-analysis. *Agronomy*, *11*(3), 1-14. https://doi.org/10.3390/agronomy11030433

Al-Asadi, R. A., and Mouazen, A. M. (2014). Combining frequency domain reflectometry and visible and near infrared spectroscopy for assessment of soil bulk density. *Soil and Tillage Research*, *135*, 60-70. https://doi.org/10.1016/j.still.2013.09.002

Allo, M., Todoroff, P., Jameux, M., Stern, M., Paulin, L., and Albrecht, A. (2020). Prediction of tropical volcanic soil organic carbon stocks by visible-near- and mid-infrared spectroscopy. *Catena*, *189*, 104452. https://doi.org/10.1016/j.catena.2020.104452

Allory, V., Cambou, A., Moulin, P., Schwartz, C., Cannavo, P., Vidal-Beaudet, L., and Barthès, B. G. (2019). Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions. *Science of the Total Environment*, *686*, 764-773. https://doi.org/10.1016/j.scitotenv.2019.05.192

Araújo, S. R., Söderström, M., Eriksson, J., Isendahl, C., Stenborg, P., and Demattê, J. A. M. (2015). Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy. *Geoderma*, *237*, 308-317. https://doi.org/10.1016/j.geoderma.2014.09.014

Askari, M. S., Cui, J., O'Rourke, S. M., and Holden, N. M. (2015). Evaluation of soil structural quality using VIS-NIR spectra. *Soil and Tillage Research*, *146*, 108-117. https://doi.org/10.1016/j.still.2014.03.006

Ben Dor, E., Ong, C., and Lau, I. C. (2015). Reflectance measurements of soils in the laboratory: Standards and protocols. *Geoderma*, *245-246*, 112-124. https://doi.org/10.1016/j.geoderma.2015.01.002

Bonett, J., Camacho-Tamayo, J. H., and Vélez-Sánchez, J. (2016). Estimating soil properties with mid-infrared spectroscopy. *Revista U.D.C.A Actualidad & Divulgación Científica*, *19*(1), 55-66. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-42262016000100007

Camacho-Tamayo, J. H., Forero-Cabrera, N. M., Ramírez-López, L., and Rubiano-Sanabria, Y. (2017). Evaluación de textura del suelo con espectroscopía de infrarrojo cercano en un oxisol de colombia. *Colombia Forestal*, *20*(1), 5-18. http://dx.doi.org/10.14483/udistrital.jour.colomb.for.2017.1.a01

Camacho-Tamayo, J. H., Rubiano-Sanabria, Y., and Hurtado, M. del P. (2014). Near-infrared (NIR) diffuse reflectance spectroscopy for the prediction of carbon and nitrogen in an Oxisol. *Agronomía Colombiana*, *32*(1), 86-94. http://www.redalyc.org:9081/html/1803/180330697012/

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., and Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Science Society of America Journal*, *58*(5), 1501-1511. https://doi.org/10.2136/sssaj1994.03615995005800050033x

Cambou, A., Cardinael, R., Kouakoua, E., Villeneuve, M., Durand, C., and Barthès, B. G. (2016). Prediction of soil organic carbon stock using visible and near infrared reflectance spectroscopy (VNIRS) in the field. *Geoderma*, *261*, 151-159. https://doi.org/10.1016/j.geoderma.2015.07.007

Davari, M., Karimi, S. A., Bahrami, H. A., Taher Hossaini, S. M., and Fahmideh, S. (2021). Simultaneous prediction of several soil properties related to engineering uses based on laboratory Vis-NIR reflectance spectroscopy. *Catena*, *197*, 1-12. https://doi.org/10.1016/j.catena.2020.104987

FAO (2019). *Measuring and modelling soil carbon stocks and stock changes in livestock production systems Guidelines for assessment* (*Version 1*). FAO. http://www.fao.org/3/ca2934en/CA2934EN.pdf

Huang, J., Hartemink, A. E., and Zhang, Y. (2019). Climate and land-use change effects on soil carbon stocks over 150 years in Wisconsin, USA. *Remote Sensing*, *11*(12), 1504. https://doi.org/10.3390/rs11121504

IPCC (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Volume 4 - Agriculture, forestry and other land use*. The Intergovernmental Panel on Climate Change (IPCC). https://www.ipcc-nggip.iges.or.jp/public/2019rf/vol4.html

Jia, X., Chen, S., Yang, Y., Zhou, L., Yu, W., and Shi, Z. (2017). Organic carbon prediction in soil cores using VNIR and MIR techniques in an alpine landscape. *Scientific Reports*, *7*(1), 2144. https://doi.org/10.1038/s41598-017-02061-z

Katuwal, S., Knadel, M., Norgaard, T., Moldrup, P., Greve, M. H., and de Jonge, L. W. (2020). Predicting the dry bulk density of soils across Denmark: Comparison of single-parameter, multi-parameter, and vis–NIR based models. *Geoderma*, *361*, 1-10. https://doi.org/10.1016/j.geoderma.2019.114080

Liland, K. H., Mevik, B.-H., and Wehrens, R. (2021). *pls: Partial least squares and principal component regression* (R package version 2.8-0). https://cran.r-project.org/package=pls

Liu, S., Shen, H., Chen, S., Zhao, X., Biswas, A., Jia, X., Shi, Z., and Fang, J. (2019). Estimating forest soil organic carbon content using vis-NIR spectroscopy : Implications for large-scale soil carbon spectroscopic assessment. *Geoderma*, *348*, 37-44. https://doi.org/10.1016/j.geoderma.2019.04.003

Malone, B. (2016). *ithir: Functions and algorithms specific to pedometrics. R package version 1.0/r126*. https://rdrr.io/rforge/ithir/

Moreira, C. S., Brunet, D., Verneyre, L., Sá, S. M. O., Galdos, M. V., Cerri, C. C., and Bernoux, M. (2009). Near infrared spectroscopy for soil bulk density assesment. *European Journal of Soil Science*, *60*, 785-791. https://doi.org/10.1111/j.1365-2389.2009.01170.x

Nawar, S., and Mouazen, A. M. (2019). On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. *Soil and Tillage Research*, *190*, 120-127. https://doi.org/10.1016/j.still.2019.03.006

Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., and Montanarella, L. (2014). Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, *68*, 337-347. https://doi.org/10.1016/j.soilbio.2013.10.022

Oliver, M. A., and Webster, R. (2015). *Basics steps in geostatistics: The variogram and kriging*. Springer. https://doi.org/10.1007/978-3-319-15865-5

Ponce-Hernandez, R., Marriott, F. H. C., and Beckett, P. H. T. (1986). An improved method for reconstructing a soil profile from analyses of a small number of samples. *Journal of Soil Science*, *37*(3), 455-467. https://doi.org/10.1111/j.1365-2389.1986.tb00377.x

Poppiel, R. R., Lacerda P., M., Pereira, M., Almeida Junior, D. O., Demattê, J., Romero, D. J., Sato, M., Almeida J, L. R., and Moreira C, L. F. (2018). Surface spectroscopy of oxisols, entisols and inceptisol and relationships with selected soil properties. *Revista Brasilera de Ciencia Do Solo*, *42*, 1-26. https://doi.org/http://dx.doi.org/10.1590/18069657rbcs20160519

Ramirez-Lopez, L., Reina-Sánchez, A., and Camacho-Tamayo, J. H. (2008). Variabilidad espacial de atributos físicos de un Typic Haplustox de los Llanos Orientales de Colombia. *Engenharia Agrícola*, *28*(1), 55-63. https://doi.org/10.1590/S0100-69162008000100006

Sommer, R., & Bossio, D. (2014). Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, *144*, 83-87. https://doi.org/10.1016/j.jenvman.2014.05.017

Stevens, A., and Ramírez-López, L. (2020). *An introduction to the prospectr package* (*R package Vignette R package version 0.2.0*). https://github.com/l-ramirez-lopez/prospectr

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., Courcelles, V. de R. de, Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., … Zimmermann, M. (2013). The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems and Environment*, *164*, 80-99. https://doi.org/10.1016/j.agee.2012.10.001

Viscarra Rossel, R. A., McGlynn, R. N., and McBratney, A. B. (2006). Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy. *Geoderma*, *137*(1-2), 70-82. https://doi.org/10.1016/j.geoderma.2006.07.004

Viscarra Rossel, R. A., and Webster, R. (2012). Predicting soil properties from the Australian soil visible – near infrared spectroscopic database. *European Journal of Soil Science*, *63*, 848-860. https://doi.org/10.1111/j.1365-2389.2012.01495.x

Wetterlind, J., Stenberg, B., and Söderström, M. (2008). The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision Agriculture*, *9*(1-2), 57-69. https://doi.org/10.1007/s11119-007-9051-z

Research Article / Chemical, Food, and Environmental Engineering http://doi.org/10.15446/ing.investig.101497

Leachate Treatment via TiO_2/UV Heterogeneous Photocatalysis: A Multiple Polynomial Regression Model

Tratamiento de lixiviados mediante fotocatálisis heterogénea TiO_2/UV: un modelo de regresión polinomial múltiple

Dorance Becerra-Moreno[1], Fiderman Machuca-Martínez[2], Aymer Y. Maturana[3], Salvador Villamizar[4], Joseph Soto-Verjel[5], and Ángelo Soto-Vergel[6]

ABSTRACT

Advanced oxidation processes such as TiO_2/UV heterogeneous photocatalysis are suitable treatment methods for wastewater with high pollutant loads such as landfill leachates. Optimizing the variables that influence the process is a fundamental aspect. However, in this regard, experimental conditions are limited in terms of resources and time, which is why modeling allows obtaining a general understanding of the phenomenon from a set of experimental data. This work sought to model the photocatalytic process via multivariate polynomial regression, considering variables such as the catalyst concentration, the pH level, and the accumulated energy concerning the percentage of degradation in terms of dissolved organic carbon (DOC). The implemented fitting method resulted in a third-degree polynomial with an R_2 of 0,8652, concluding that the model and its conclusions are valid. Moreover, with greater degrees, the model curve overfitted, even with better R_2. DOC abatement showed a negative correlation with pH and the catalyst dose, while an opposite trend was observed for the accumulated energy. The model predictions allow inferring that, at low catalyst doses and medium and high pH levels, it is possible to find maximum degradations at low cumulative energies.

Keywords: leachates, TiO_2, heterogeneous photocatalysis, model, polynomial regression

RESUMEN

Los procesos de oxidación avanzada como la fotocatálisis heterogénea TiO_2/UV son métodos de tratamiento adecuados para aguas residuales con altas cargas contaminantes como los lixiviados de rellenos sanitarios. La optimización de las variables que influyen en el proceso es un aspecto fundamental. Sin embargo, en este aspecto, las condiciones experimentales son limitadas en términos de recursos y tiempo, por lo que el modelado permite obtener una comprensión general del fenómeno a partir de un conjunto de datos experimentales. Este trabajo pretendió modelar el proceso fotocatalítico mediante regresión polinómica multivariada, teniendo en cuenta variables como la concentración del catalizador, el nivel de pH y la energía acumulada en relación con el porcentaje de degradación en términos de carbón orgánico disuelto (DOC). El método de ajuste implementado dio como resultado un polinomio de grado 3 con un R_2 de 0,8652, concluyendo que el modelo y sus conclusiones son válidos. Además, a mayor grado, la curva del modelo se sobreajustó, incluso con mejor R_2. La remoción de DOC mostró una correlación negativa con el pH y la dosis de catalizador, y se observó una tendencia opuesta para la energía acumulada. Finalmente, las predicciones del modelo permiten inferir que, a dosis bajas del catalizador y niveles medios y altos de pH, es posible encontrar degradaciones máximas con bajas energías acumuladas.

Palabras clave: lixiviado, TiO_2, fotocatálisis heterogénea, modelo, regresión polinómica

Received: March 6th, 2022
Accepted: June 1th, 2023

[1] Master, Universidad Francisco de Paula Santander, Cúcuta, Colombia. Affiliation: Departamento de Ciencias del Medio Ambiente, Facultad de Ciencias Agrarias y del Ambiente, dorancebm@ufps.edu.co

[2] PhD, Universidad del Valle, Cali, Colombia. Affiliation: Escuela de Ingeniería Química, CENM.

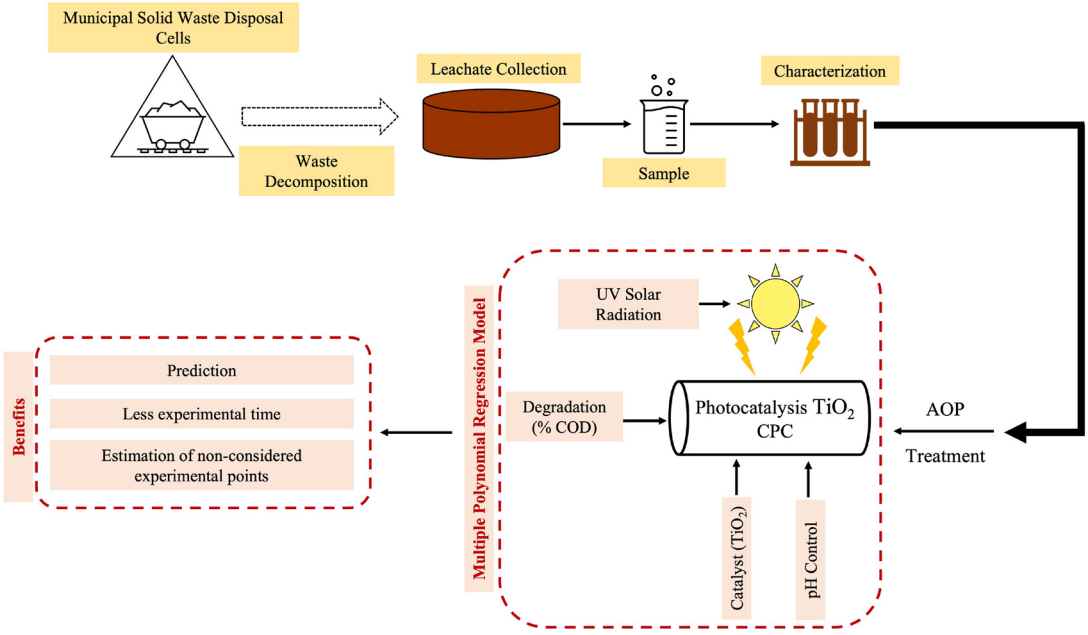
[3] PhD, Universidad del Norte, Barranquilla, Colombia. Affiliation: Departamento de Ingeniería Civil y Ambiental – Instituto de Estudios Hidráulicos y Ambientales IDEHA.

[4] Engineer, Universidad del Norte, Barranquilla, Colombia. Affiliation: Departamento de Ingeniería Civil y Ambiental – Instituto de Estudios Hidráulicos y Ambientales IDEHA.

[5] Master, Universidad Francisco de Paula Santander, Cúcuta, Colombia Affiliation: Departamento de Electricidad y Electrónica.

[6] The authors Joseph Soto-Verjel, and Salvador Villamizar share the affiliation number 4 (Engineer, Universidad del Norte, Barranquilla, Colombia. Affiliation: Departamento de Ingeniería Civil y Ambiental – Instituto de Estudios Hidráulicos y Ambientales IDEHA), while Ángelo Soto-Vergel has affiliation number 5 (Master, Universidad Francisco de Paula Santander, Cúcuta, Colombia Affiliation: Departamento de Electricidad y Electrónica). Also, affiliation number 6 does not correspond to any author.

Graphical abstract



Highlights

- Multivariate polynomial regression can be used for reducing the number of experiments, making predictions, and analyzing values outside of the DoE.
- A third-degree polynomial with an R^2 of 0,8652 was considered to be the best-fitting method in this work.
- Polynomial degrees greater than 3 overfit the model's curves, even if a better R^2 is obtained.
- The catalyst dose and the pH had a negative influence on the percentage of DOC degradation.
- Model predictions allow inferring that, at low catalyst doses and medium and high pH levels, it is possible to find maximum degradations at low cumulative energies.

Introduction

Heterogeneous photocatalysis is an advanced oxidation process (AOP) that uses a semiconductor photocatalyst which, when subjected to a specific radiation, allows for oxidation-reduction processes, enabling the removal of pollutants and providing effluents with lower toxicity (Gout *et al*., 2022). Landfill leachates contain high amounts of heavy metals, solids (suspended and settleable), and organic matter (Vahabian *et al*., 2019). The recalcitrant characteristics of the individual pollutants and the highly variable generation rates are the leading factors that hinder their treatment (Chaturvedi and Kaushal, 2018; Müller *et al*., 2015). It is there that AOPs, by improving the prospects for the treatment of wastewater with high pollutant loads, accomplish their objective (Ruiz-Delgado *et al*., 2020).

In addition, the catalyst is the main factor that directly impacts the efficiency of photocatalytic treatments. During the last few years, several kinds of these semiconductors have appeared, such as metal oxides, nanomaterials, and organic catalysts, among others. According to Al-Mamun *et al*. (2021), TiO_2 is considered one of the best photocatalytic materials due to it is non-toxicity, high photocatalytic efficiency, and photostability. Additionally, TiO_2 can be combined with other compounds such as Ag and MoO_3 (Hasan Khan Neon and Islam, 2019). This material has been used over the years for treating complex contaminants present in pharmaceutical, textile, and industrial wastewater, achieving values of more than 90% regarding the degradation of emerging pollutants (Akter *et al*., 2022).

The complete mineralization of contaminants in any AOP requires a large amount of energy, along with chemicals such as catalysts and oxidants, which increases the

cost of treatments (Thanekar *et al.*, 2018) and limits research due to the availability of resources and time. In wastewater treatment processes, the accurate modeling and optimization of the most suitable conditions is required to obtain the highest possible efficiency (Azadi *et al.*, 2018). Modeling arises, then, as an alternative to knowing the behavior of a phenomenon from actual experimental data, allowing to explain such behavior through a model and to see the prediction in measurement points that were not experimentally considered. This allows obtaining a complete knowledge of the whole phenomenon and, simultaneously, determining whether there are non-measured points or values in which the process could have behaved better (Florea, 2019).

Photocatalytic modeling has been applied from different perspectives, the oldest of which involves reaction kinetics models (Marien *et al.*, 2019; Xu *et al.*, 2020). Most recently, artificial neuronal networks (ANN) have been used to analyze the effects of operating conditions on photocatalyst performance (Ateia *et al.*, 2020) and the efficiency of the photocatalytic process (Jing *et al.*, 2017), as well as to predict temporal variations of leachate COD in photocatalytic treatment processes (Azadi *et al.*, 2018). However, these works have trained ANNs with a minimal data set, thus causing their predictions to have a high percentage of error; according to Ciresan *et al.* (2012), datasets below 50 samples per class show prediction errors of up to 30%. Another recent approach is the response surface methodology (RSM), one of the most common for these applications since it is oriented towards experimental design and the optimization of operating conditions (Ateia *et al.*, 2020; Becerra *et al.*, 2020; Colombo *et al.*, 2013). Moreover, various regression models have been applied to simulate the treatment conditions, determining the significance of variables such as treatment time, catalyst doping percentage, and catalyst dose to the linear model (Nugraha and Fatimah, 2016). Other models have been applied to a lesser extent, as is the case of Support Vector Machines (SVM), which allow optimizing the pollutants removal percentage. This was shown by Ateia *et al.* (2020), who modeled the removal of a pollutant through the least-squares method and optimized it with the Cuckoo algorithm, achieving an R^2 higher than 0,94 for the acid pH conditions and the medium catalyst concentration.

The objective of this research is not to describe the time-dependent relationship between the operating conditions of the system and the degradation rate of organic pollutants (Ateia *et al.*, 2020), as is the case of reaction kinetics models, but to implement a model that allows predicting values of the response variable for which no experimental validations were made. Therefore, this article presents a multiple polynomial regression model, a particular case of multiple linear regression in which the relationship between the independent and the dependent variables is an n-degree polynomial in x. To this effect, based on experimental data of a TiO_2/UV photocatalytic process used for the decontamination of landfill leachates, our work proposes to construct the process model using this polynomial regression, with the catalyst concentration ($mg.L^{-1}$ TiO_2), the pH, and the UV radiation level (accumulated energy in $kJ.L^{-1}$) as independent variables, and the degradation ($mg.L^{-1}$ DOC) as the dependent variable. As a result, an equation is determined which represents the process as a function of the described influential variables, predicting the degradation values in the intermediate of the accumulated energy ranges that could not be measured experimentally. This, in order to determine whether, in each photocatalytic experiment (combination of TiO_2 and pH), there was an optimal UV value different from those calculated by the RSM presented in Becerra *et al.* (2020).

Materials and methods

Our methodology consists of four stages: collection and characterization of the leachate sample, which is the input for heterogeneous photocatalysis, followed by multiple polynomial regression modeling, and finally, results analysis (Figure 1).

The leachate sample was collected from the disposal of ordinary solid waste at a landfill in Norte de Santander (Colombia). The sample was taken directly from the main collection pipe of the different disposal cells. Subsequently, the sample was transported to the Environmental Quality Laboratory of Universidad Francisco de Paula Santander, where it was characterized in terms of dissolved organic carbon (DOC), chemical oxygen demand (COD), total

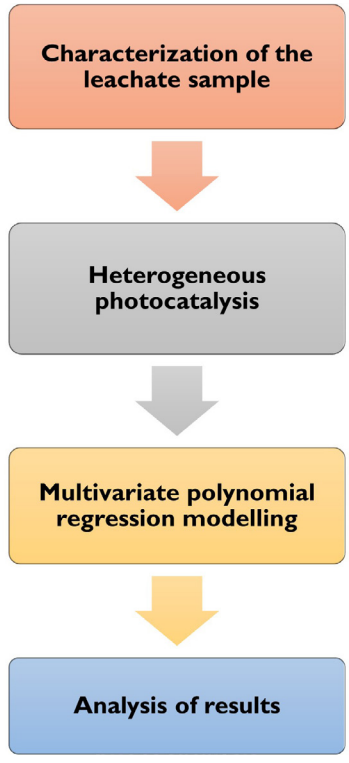


Figure 1. Research methodology
Source: Authors

suspended solids (TSS), volatile suspended solids (VSS), pH, and temperature under the guidelines of the *Standard Methods for the Analysis of Water and Wastewater* (Pawlowski, 1994). As a result of this process, the input concentration of the leachate to the heterogeneous photocatalysis process was obtained.

The heterogeneous photocatalysis, aided by TiO_2 in the presence of UV radiation for the decontamination of leachate from a landfill, was carried out in a laboratory-scale Composite Parabolic Collector (CPC) with a surface area of 0,83 m^2, borosilicate tubes, and aluminum reflective material. The photocatalyst used was commercial TiO_2 P25 (Degussa-Evonik) powder with an anatase:rutile concentration proportion of 70:30, a surface area of 50 m/g, and an average diameter of 20 nm. The material properties of this catalyst were reported by Acosta-Herazo *et al*. (2019), Ohtani *et al*. (2010), and Satuf *et al*. (2005). Our work also employed a storage tank with a capacity of 20 L and a Humboldt pump responsible for recirculating the treatment throughout the system (Figure 2). This operation consists of adding domestic wastewater and leachate (the latter at a concentration of 500 $mg.L^{-1}$ COD) to the storage tank. The pump is turned on to recirculate the mixture for a few moments throughout the CPC. Afterwards, the amount of TiO_2 required for the experiment is added, *i.e.*, 100, 350, or 600 $mg.L^{-1}$. The pH is immediately stabilized at the required level (3, 6, or 9) with a waterproof multiparameter, using solutions of HCl or NaOH (0,1 N). At this point, the zero samples are taken, initiating the procedure described in the section devoted to calculating the accumulated energy. The experiment ends when the 60 $kJ.L^{-1}$ are reached, according to the analysis carried out by Borges *et al*. (2016).

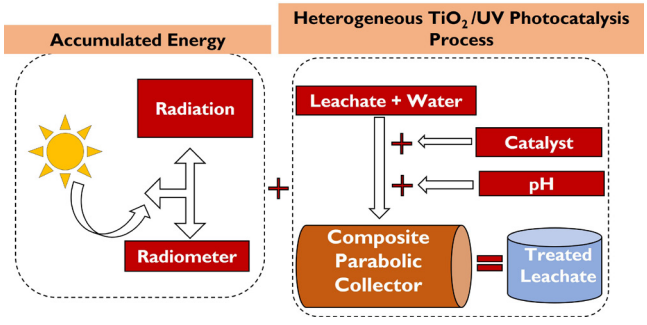


Figure 2. General model for heterogeneous TiO_2/UV photocatalysis
Source: Authors

The StatGraphics Centurion XV statistical tool was used for the experimental design, given its capabilities for randomized testing and data visualization. Furthermore, a 3^2 factorial design with replication at the center point was developed to model the results using a response surface. The summary of the experiments is shown in Table 1. The influential variables measured for the factorial design were the catalyst dose (100, 350, and 600 mg. L^{-1} TiO_2) and the pH (3, 6, and 9). These factors were considered since, when TiO_2 has been used as catalyst in the presence of UV radiation, it generates a strong oxidant radical such as OH^-, in addition to the fact that it is non-toxic and anti-corrosive and it features a high pollutant removal (Al-Mamun *et al*., 2019; Villamizar *et al*., 2022). Fluctuations in the pH level were considered in order to understand which wastewater conditions stimulate the production of hydroxyl radicals (Balarabe and Maity, 2022; Hassan *et al*., 2016; Yasmin *et al*., 2020).

Table 1. Photocatalysis experiments

No.	Catalyst ($mg.L^{-1}$ TiO_2)	pH	UV_{Solar} ($kJ.L^{-1}$)
1	100	3	20
			40
			60
2	100	6	20
			40
			60
3	100	9	20
			40
			60
4	350	3	20
			40
			60
5	350	6	20
			40
			60
6	350	6	20
			40
			60
7	350	9	20
			40
			60
8	600	3	20
			40
			60
9	600	6	20
			40
			60
10	600	9	20
			40
			60

Note: The experiments were conveniently arranged, but StatGraphics Centurion XV randomized the experiments
Source: Authors

Taking into account that photocatalysis is carried out after combining the catalyst and UV radiation to accelerate the chemical transformation of pollutants (Moura and Picão, 2022), the accumulated energy was regarded as a blocking variable in values of 20, 40, and 60 $kJ.L^{-1}$ for each experiment involving the TiO_2/pH combination. Finally, a system was assembled which consisted of a multimeter (Uni-T UT71C)

connected to a pyranometer (SP-110); the latter, when in contact with radiation, sequentially marks a value in millivolts (mV). Collecting precise data depends on the time lapses in which the reading is provided, which, in this case, was every 10 min (600 seconds). The next step was a mathematical conversion from mV to $W.m^{-2}$, where $1\,mV$ represents $5\,W.m^{-2}$. Finally, Equation (1) was used to calculate the energy:

$$Q_n = Q_{(n-1)} + \Delta t_n * I_n * A_f * V_T^{(-1)} \quad (1)$$

Where:

Q_n: total accumulated energy kJ.L^{-1}

Q_{n-1}: previous accumulated energy kJ.L^{-1}

Δt_n: irradiation time (600 seconds)

I_n: average irradiation (W.m^{-2} UV)

A_f: irradiated reactor surface (0,83 m^2)

V_T: volume discussed (20 L)

To calculate the response variable representing the DOC, reported as a percentage (%) of degradation in each TiO_2/pH combination, a sample was taken at time zero of the treatment. This value is then calculated in comparison with the other accumulated energies via Equation (2). This sample collection process is carried out in amber containers with a 60 ml capacity. Once collected, the samples pass through membrane filters and are prepared for DOC measurements in the Teledyne Tekmar Torch Total Organic Carbon Analyzer.

$$\%\ DOC\ Degradation = \frac{mg.L^{-1}DOC(0) - mg.L^{-1}DOC(kJ.L^{-1})}{mg.L^{-1}DOC(0)} \quad (2)$$

Where:

% Degradation DOC **:** percent degradation in each TiO_2/pH combination for each accumulated energy at 20, 40 and 60 kJ.L^{-1}

$mg.L^{-1}DOC(0)$**:** measurement of each DOC value at time zero of treatment in each TiO_2/pH combination

$mg.L^{-1}DOC(kJ.L^{-1})$**:** measurement of each DOC value at each accumulated energy value of 20, 40 and 60 kJ.L^{-1}

The results obtained from these experiments were the input for the multiple polynomial regression model, in order to make predictions about the intermediate of the accumulated energies in each TiO_2/pH combination, *i.e.*, to understand the behavior of the process in the 20-40 and 40-60 kJ.L^{-1} intervals. The accumulated energy was measured every 10 min, so that, in each experiment there was a significant amount of data to calculate the response variable. Multiple polynomial regression was carried out using the MultiPolyRegress function (Cecen, 2021), which implemented Equation (3), a matrix expression of the model. Parameter estimation was performed using the least-squares method.

$$\begin{bmatrix} \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k-1} \\ 1 & x_{21} & \dots & x_{2k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk-1} \\ & & & \end{bmatrix} \begin{bmatrix} \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \\ \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (3)$$

Here,

$\boldsymbol{Y}$ is the vector of observations of the phenomenon – for this case, the percent degradation in terms of DOC.

$\boldsymbol{X}$ is the vector matrix of the independent variables: the catalyst concentration (mg.L^{-1} TiO_2), the pH level, and the solar UV dose applied (kJ.L^{-1}).

β is the vector of the parameters to be estimated.

$\mathcal{E}$ is the vector of model errors.

It is essential to remember that, on many occasions, the aim is to adjust the model in order to obtain very high R^2 values, which measure the relevance of the model. Still, this research did not intend for the model to learn the data by memory, but for it to understand the influence of the independent variables on the response variable. Therefore, the selection regarding the degree of the polynomial that best represents the model was based on the R^2 value, considering that it must be higher than 0,7 if the model is appropriate and pertinent. However, the model must not be so close to 1, which would imply overfitting (Frost, 2020).

The subsequent analysis considered the results obtained by the RSM in comparison with the predictions calculated using the multivariate polynomial regression model. This allowed determining whether, among the values not measured experimentally, there were degradation percentage values higher than those calculated via the response surface. The polynomial regression model analyzes the correlation between the variables, the distribution of the response variable, and the model's fit.

Results and discussion

The leachate characterization is shown in Table 2.

Table 2. Leachate characterization

Parameter	Units	Value
COD	mg O_2*L^{-1}	7 920
DOC	mg C*L^{-1}	2 767
TSS	mg TSS*L^{-1}	19 941
VSS	mg VSS*L^{-1}	8 587
pH	-	7,29

Source: Authors

The COD is used to calculate the inlet concentration of the photocatalytic reactor. From 7 920 mg.L^{-1} COD, an inlet concentration of 500 mg.L^{-1} COD was determined for the process.

The results obtained during the TiO_2/UV heterogeneous photocatalysis are consolidated in Table 3. Each experiment is specified in terms of TiO_2/pH concerning the degradation percentages (DOC) for each measured accumulated energy value. According to what was proposed in Table 1, the latter was set at the values of 20, 40, and 60 kJ. L^{-1}. However, since these values were measured every 10 min, they are close to the fixed values. This is why, in Table 3, this variable appears in decimal form; it depends directly on the solar irradiance on the days of the experiment, some of which were cloudy and others very sunny.

According to these results, it can be inferred that there is an influence of the TiO_2 and pH variables on the percentage of degradation in each combination. This is shown by the fact that the highest degradations are associated with low and medium catalyst doses together with low and medium pH levels. Likewise, the process is favored at low pH – in this case, level 3. For the high catalyst dose (600 mg.L^{-1} TiO_2), the results are not consistent with the pH levels in comparison with the other combinations, given that, in this case, the maximum degradation occurred at maximum pH. Thus, it can be concluded that high doses of this catalyst can unbalance the degradation process in the sense that they do not allow for an efficient absorption of solar radiation. This was confirmed by Yashni *et al*. (2021), who reported the best photocatalytic performance with a pH equal to 3 and concluded that this variable directly impacts the oxidative strength of the generated gap, enhancing the compound's ionization and increasing the removal percentages. Moreover, the concentrations allow stating that, in photocatalytic processes with a low catalyst dosage, hydroxyl radical generation was fostered, improving the pollutant removal efficiency and exhibiting a better light transmittance into the reactor (Rizzo *et al*., 2014). Consequently, the lower efficiencies could be related to the fact that the aim of our studies was to transform toxic waste liquid into one with more biodegradable characteristics and suitable for biological degradation, taking into account that, when AOPs are used as the only treatment, they are usually expensive due to energy and chemical products consumption (Berberidou *et al*., 2017; Oller *et al*., 2011; Paździor *et al*., 2019).

As for the accumulated energies, the highest degradations in all cases occurred at the maximum measurement of 60 kJ.L^{-1}. This can be seen as a limitation of the process since the accumulated energy is related to the duration of the treatment. Therefore, the higher the accumulated energy required, the greater the time needed for the process to reach the desired or required degradations of the evaluated pollutants. It is essential that this variable is regarded as influential in the design of experiments to determine whether there are degradation values in non-maximum accumulated energy ranges, thus contributing not only to optimizing the photocatalytic process, but also to the efficient use of resources in wastewater treatment, an issue that has been controversial in recent years.

In this case, fluctuation was identified between the degradation percentages of each sampled accumulated energy, as is the case of experiment numbers 1, 3, 6, 9, and 10. This fluctuation is generated by the composition of the leachate since, when it is subjected to a chemical process such as photocatalysis, a partial transformation of its contaminant species (*i.e.*, trace metals, recalcitrant compounds, and organic macro- and micro-pollutants) may occur, reaching maximum removal efficiency only when the appropriate UV dose is applied (60 kJ. L^{-1}). This phenomenon was similarly reported by Amigh and Mokhtarani (2022), Çifçi and Meriç (2015), Hassan *et al*. (2017), Nomura *et al*. (2020), and Rocha *et al*. (2011).

Table 3. Photocatalysis results

No.	Catalyst (mg.L^{-1} TiO_2)	pH	UV_{Solar} (kJ.L^{-1})	Degradation (% DOC)
1	100	3	22,66	16,04
			43,86	14,71
			63,46	57,90
2	100	6	20,65	10,11
			40,39	15,34
			62,22	23,89
3	100	9	20,34	1,30
			40,55	-0,62
			60,99	1,43
4	350	3	22,19	25,31
			40,20	27,85
			62,79	29,84
5	350	6	20,35	3,50
			40,91	6,25
			57,27	11,92
6	350	6	22,83	5,13
			39,65	2,97
			60,84	7,18
7	350	9	21,69	5,94
			42,00	8,90
			61,76	5,49
8	600	3	20,68	7,07
			39,38	8,87
			60,67	12,48
9	600	6	21,68	0,72
			40,48	0,26
			58,67	5,02
10	600	9	20,89	21,44
			43,91	2,36
			61,20	22,44

Source: Authors

Figure 3 shows the correlation between the independent variables and the response variable: degradation - ph = -0,47; degradation - UV = 0,34; and degradation - catalyst = -0,21. The highest correlation is found with pH; the higher the pH, the lower the degradation. The same happens with the catalyst, which, in higher quantities, impairs the degradation process. Finally, regarding the UV dose, which maintains a direct proportional relationship, the higher the accumulated energy, the higher the degradation. This was confirmed by Jia *et al*. (2013), who achieved the highest removal percentage in the longest UV-radiation exposure time. Thus, this analysis is consistent and reproducible with the results obtained via the RSM.

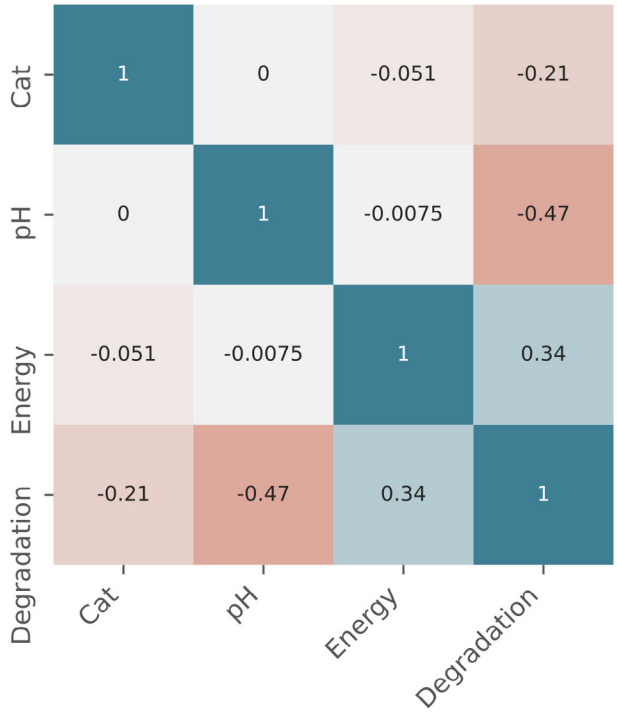


Figure 3. Correlation of variables
Source: Authors

With the dataset in Table 3, a multivariate polynomial regression model was developed in the MATLAB Online R2020b software. The degree of the polynomial was set at n = 3, obtaining an R^2 = 0,8652. For lower degrees, the model did not fit correctly, and, for higher degrees, it showed overfitting. Some researchers state that an R^2 = 0,8 clearly indicates a very good regression model performance and explains a significant amount of the variance in the data (Chicco *et al*., 2021). In our case, other models were tested with a fourth-degree polynomial, obtaining an R^2 of around 0,9. However, we noticed that the model's curve was overfitted to the experimental data. Therefore, the reported R^2 was considered to be more adequate in explaining the variance of our experimental data.

Equation (4) models the heterogeneous photocatalysis in the decontamination of landfill leachate.

$$\begin{aligned} Y = {} & 7,4903X_3 - 0,18517X_3^2 + 0,0082063X_2X_3 - \\ & 0,00032472X_2X_3^2 + 0,039433X_2^2 - 0,01116X_2^2X_3 - \\ & 0,94587X_1 - 0,0030433X_1X_3 - 7,7654e-06X_1X_3^2 - \\ & 0,044475X_1X_2 + 0,00025262X_1X_2X_3 + 0,0022061X_1X_2^2 + \\ & 0,0038557X_1^2 + 2,3428e-06X_1^2X_3 + 2,6147e-05X_1^2X_2 - \\ & 3,9225e-06X_1^3 + 0,013281X_2^3 + 0,0016262 * X_3^3 \end{aligned} \quad (4)$$

Figure 4 reports the accuracy and ecision of the model, and it identifies the type of bias in it. Figure 5 demonstrates that the residuals follow a random pattern, *i.e.*, no trend and no non-constant variance. Figure 6 shows that the residuals fit a normal distribution, *i.e.*, Figures 5 and 6 confirm that the ordinary least-squares assumptions are valid for the model.

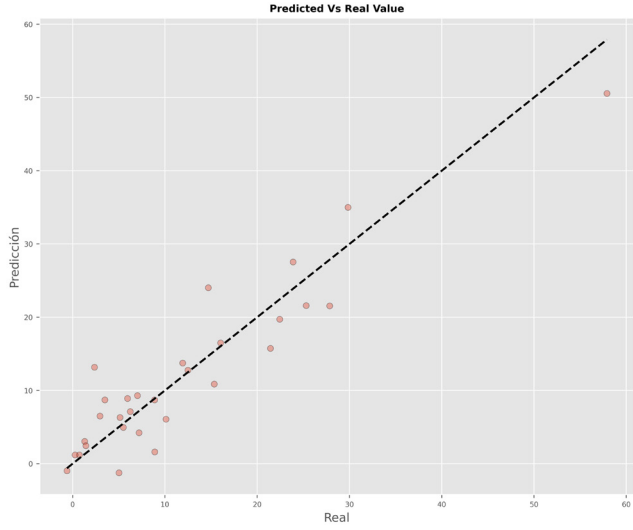


Figure 4. Graph of predicted *vs.* real values
Source: Authors

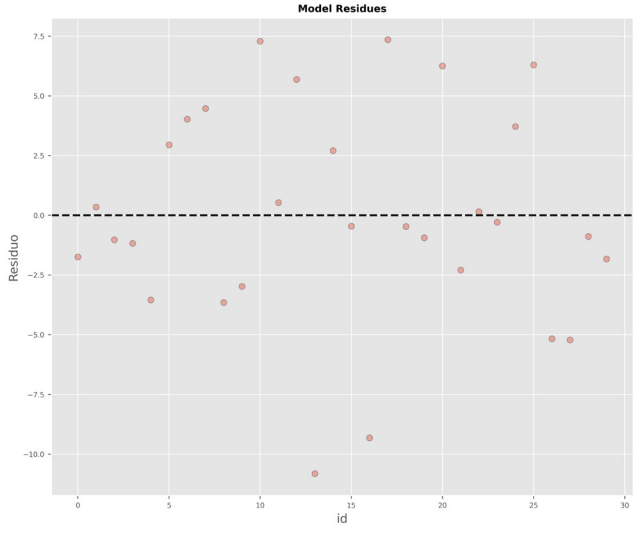


Figure 5. Graph of model residuals
Source: Authors

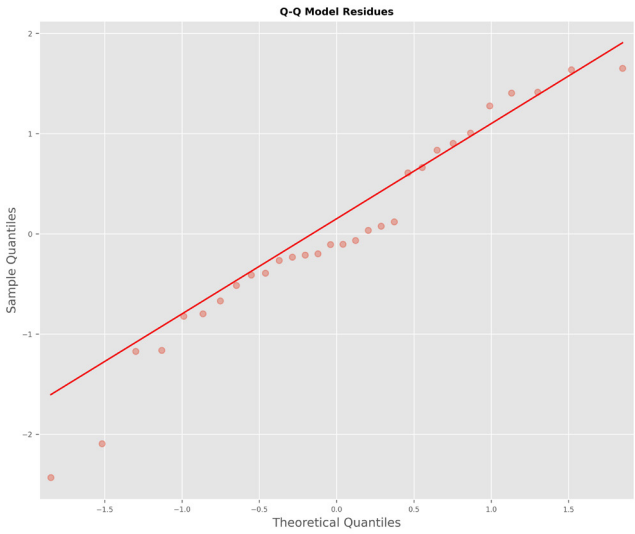


Figure 6. Graph of Q-Q model residues
Source: Authors

Conclusions

This paper modeled the TiO_2/UV heterogeneous photocatalysis process for the decontamination of landfill leachate using experimental data as the input. The fitting method implemented was a multivariate polynomial regression, resulting in a third-degree polynomial with an R^2 of 0,8652. This model was validated using residual analysis, indicating that the errors are independent, follow a normal distribution, and have a constant variance. Therefore, the model and the conclusions derived from it are valid.

The catalyst dose and the pH have a negative influence on the percentage of degradation. This agrees with the RSM analysis, which showed that the maximum degradations occurred at low catalyst levels and acid pH values, so they have a negative correlation, *i.e.*, if either of the two increases, the degradation will decrease, which will affect the treatment process. Based on a similar analysis, it was inferred that the influence of the accumulated energy is positive on the percentage of degradation, and this is coherent with that obtained via the RSM, as the maximum degradations occurred at the highest levels evaluated. This means that, as the UV levels increase, the degradation will also increase, which is the purpose of all wastewater treatments – to obtain the highest possible pollutant removals.

Based on the predictions of the model at intermediate accumulated energy values that were not measured experimentally, it can be concluded that, in the case of low catalyst doses (100 $mg.L^{-1}$) both the polynomial regression model and the RSM agreed that the highest degradation in terms of DOC was achieved at the maximum cumulative energy of 60 $kJ.L^{-1}$ and at a pH = 3. However, at this same low dose, the model predicted higher degradation values at much lower accumulated energies for pH levels of 6 and 9. Thus, for a range between 25 and 30 $kJ.L^{-1}$ the model predicted degradations of 6% (pH=9) and 28% (pH=6), while similar degradation values were obtained in the RSM, but at the maximum accumulated energy. Therefore, it can be inferred that, at low catalyst doses and medium and high pH levels, it is possible to find maximum degradations at low cumulative energies. This phenomenon occurred in a similar way for the other pH and catalyst combinations.

It is important to remember that leachates naturally have a pH close to neutrality (7,29), so evaluating their degradation at a low catalyst dose and their natural pH can provide promising results regarding the inferences made via modeling. This would allow for a decrease in the use of reagents to acidify the pH, thus optimizing not only the influential variables, but also the use of resources.

This work intends to be a partial conclusion concerning leachate treatment via heterogeneous photocatalysis, as well as to contribute to the modeling of this phenomenon given the amount of available data – please note that all the possible combinations that could occur in the treatment were not evaluated.

Acknowledgements

This work was supported by the Universidad Francisco de Paula Santander and Universidad del Valle, to which we would like to express our gratitude.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

Acosta-Herazo, R., Mueses, M. Á., Puma, G. L., and Machuca-Martínez, F. (2019). Impact of photocatalyst optical properties on the efficiency of solar photocatalytic reactors rationalized by the concepts of initial rate of photon absorption (IRPA) dimensionless boundary layer of photon absorption and apparent optical thickness. *Chemical Engineering Journal*, *356*, 839-849. https://doi.org/10.1016/j.cej.2018.09.085

Akter, S., Islam, M. S., Kabir, M. H., Shaikh, M. A. A., and Gafur, M. A. (2022). UV/TiO2 photodegradation of metronidazole, ciprofloxacin and sulfamethoxazole in aqueous solution: An optimization and kinetic study. *Arabian Journal of Chemistry*, *15*(7), 103900. https://doi.org/10.1016/j.arabjc.2022.103900

Al-Mamun, M.R., Kader, S., Islam, M. S., and Khan, M. Z. H. (2019). Photocatalytic activity improvement and application of UV-TiO2 photocatalysis in textile wastewater treatment: A review. *Journal of Environmental Chemical Engineering*, *7*(5), 103248. https://doi.org/10.1016/j.jece.2019.103248

Al-Mamun, M. R., Kader, S., and Islam, M. S. (2021). Solar-TiO2 immobilized photocatalytic reactors performance assessment in the degradation of methyl orange dye in

aqueous solution. *Environmental Nanotechnology, Monitoring & Management*, *16*, 100514. https://doi.org/10.1016/j.enmm.2021.100514

Amigh, P., and Mokhtarani, N. (2022). Leachate post treatment, using Ag-TiO2 nanoparticles immobilized on rotating vanes. *Journal of Water Process Engineering*, *47*, 102842. https://doi.org/10.1016/j.jwpe.2022.102842

Ateia, M., Alalm, M. G., Awfa, D., Johnson, M. S., and Yoshimura, C. (2020). Modeling the degradation and disinfection of water pollutants by photocatalysts and composites: A critical review. *Science of The Total Environment*, *698*, 134197. https://doi.org/10.1016/j.scitotenv.2019.134197

Azadi, S., Karimi-Jashni, A., and Javadpour, S. (2018). Modeling and optimization of photocatalytic treatment of landfill leachate using tungsten-doped TiO2 nano-photocatalysts: Application of artificial neural network and genetic algorithm. *Process Safety and Environmental Protection*, *117*, 267-277. https://doi.org/10.1016/j.psep.2018.03.038

Balarabe, B. Y., and Maity, P. (2022). Visible light-driven complete photocatalytic oxidation of organic dye by plasmonic Au-TiO2 nanocatalyst under batch and continuous flow condition. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *655*, 130247. https://doi.org/10.1016/j.colsurfa.2022.130247

Becerra, D., Soto, J., Villamizar, S., Machuca-Martínez, F., and Ramírez, L. (2020). Alternative for the treatment of leachates generated in a landfill of Norte de Santander – Colombia, by Means of the coupling of a photocatalytic and biological aerobic process. *Topics in Catalysis*, *63*, 1336-1349. https://doi.org/10.1007/s11244-020-01284-1

Berberidou, C., Kitsiou, V., Lambropoulou, D. A., Antoniadis, A., Ntonou, E., Zalidis, G. C., and Poulios, I. (2017). Evaluation of an alternative method for wastewater treatment containing pesticides using solar photocatalytic oxidation and constructed wetlands. *Journal of Environmental Management*, *195*, 133-139. https://doi.org/10.1016/j.jenvman.2016.06.010

Borges, M. E., Sierra, M., Cuevas, E., García, R. D., and Esparza, P. (2016). Photocatalysis with solar energy: Sunlight-responsive photocatalyst based on TiO2 loaded on a natural material for wastewater treatment. *Solar Energy*, *135*, 527-535. https://doi.org/10.1016/j.solener.2016.06.022

Cecen, A. (2021). *Multivariate polynomial regression - File exchange - MATLAB central*. https://github.com/ahmetcecen/MultiPolyRegress-MatlabCentral

Chaturvedi, H., and Kaushal, P. (2018). Comparative study of different Biological Processes for non-segregated Municipal Solid Waste (MSW) leachate treatment. *Environmental Technology and Innovation*, *9*, 134-139. https://doi.org/10.1016/j.eti.2017.11.008

Chicco, D., Warrens, M. J., and Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623

Çifçi, D. I., and Meriç, S. (2015). Optimization of suspended photocatalytic treatment of two biologically treated textile effluents using TiO2 and ZnO catalysts. *Global Nest Journal*, *17*(4), 653-663. https://doi.org/10.30955/GNJ.001715

Ciresan, D. C., Meier, U., and Schmidhuber, J. (2012, June 10-15). *Transfer learning for Latin and Chinese characters with Deep Neural Networks* [Conference presentation]. 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, QLD, Australia. https://doi.org/10.1109/IJCNN.2012.6252544

Colombo, R., Ferreira, T. C. R., Alves, S. A., Carneiro, R. L., and Lanza, M. R. V. (2013). Application of the response surface and desirability design to the Lambda-cyhalothrin degradation using photo-Fenton reaction. *Journal of Environmental Management*, *118*, 32-39. https://doi.org/10.1016/j.jenvman.2012.12.035

Florea, O. A. (2019). *Mathematical modeling of some physical phenomena through dynamical systems*. In C. Flaut, Š. Hošková-Mayerová, and D. Flaut (Eds.), *Models and Theories in Social Systems. Studies in Systems, Decision and Control* (vol. 179, pp. 77-93). https://doi.org/10.1007/978-3-030-00084-4_4

Frost, J. (2020). *Introduction to statistics: An intuitive guide for analyzing data and unlocking discoveries*. Statistics By Jim Publishing.

Gout, E., Monnot, M., Boutin, O., Vanloot, P., Claeys-Bruno, M., and Moulin, P. (2022). Assessment and optimization of wet air oxidation for treatment of landfill leachate concentrated with reverse osmosis. *Process Safety and Environmental Protection*, *162*, 765-774. https://doi.org/10.1016/j.psep.2022.04.046

Hasan Khan Neon, M., and Islam, M. S. (2019). MoO3 and Ag co-synthesized TiO2 as a novel heterogeneous photocatalyst with enhanced visible-light-driven photocatalytic activity for methyl orange dye degradation. *Environmental Nanotechnology, Monitoring & Management*, *12*, 100244. https://doi.org/10.1016/j.enmm.2019.100244

Hassan, M., Wang, X., Wang, F., Wu, D., Hussain, A., and Xie, B. (2017). Coupling ARB-based biological and photochemical (UV/TiO 2 and UV/S 2 O 8 2−) techniques to deal with sanitary landfill leachate. *Waste Management*, *63*, 292-298. https://doi.org/10.1016/j.wasman.2016.09.003

Hassan, M., Zhao, Y., and Xie, B. (2016). Employing TiO2 photocatalysis to deal with landfill leachate: Current status and development. *Chemical Engineering Journal, 285*, 264-275. https://doi.org/10.1016/j.cej.2015.09.093

Jia, C.-Z., Zhu, J.-Q., and Qin, Q.-Y. (2013, January 16-18). *Variation characteristics of different fractions of dissolved organic matter in landfill leachate during UV-TiO2 photocatalytic degradation* [Conference presentation]. 2013 Third International Conference on Intelligent System Design and Engineering Applications, Hong Kong, China. https://doi.org/10.1109/ISDEA.2012.383

Jing, L., Chen, B., Wen, D., Zheng, J., and Zhang, B. (2017). Pilot-scale treatment of atrazine production wastewater by UV/O 3 /ultrasound: Factor effects and system optimization. *Journal of Environmental Management*, *203*, 182-190. https://doi.org/10.1016/j.jenvman.2017.07.027

Marien, C. B. D., Le Pivert, M., Azaïs, A., M'Bra, I. C., Drogui, P., Dirany, A., and Robert, D. (2019). Kinetics and mechanism of Paraquat's degradation: UV-C photolysis *vs.* UV-C photocatalysis with TiO2/SiC foams. *Journal of Hazardous Materials*, *370*, 164-171. https://doi.org/10.1016/j.jhazmat.2018.06.009

Moura, L., and Picão, R. C. (2022). Removal of antimicrobial resistance determinants from wastewater: A risk perspective on conventional and emerging technologies. In H. Sarma, D. C. Domínguez, and W.-Y. Lee (Eds.), *Emerging Contaminants in the Environment* (pp. 603-642). Elsevier. https://doi.org/10.1016/B978-0-323-85160-2.00023-8

Müller, G. T., Giacobbo, A., dos Santos Chiaramonte, E. A., Rodrigues, M. A. S., Meneguzzi, A., and Bernardes, A. M. (2015). The effect of sanitary landfill leachate aging on the biological treatment and assessment of photoelectrooxidation as a pre-treatment process. *Waste Management*, *36*, 177-183. https://doi.org/10.1016/j.wasman.2014.10.024

Nomura, Y., Fukahori, S., and Fujiwara, T. (2020). Removal of 1,4-dioxane from landfill leachate by a rotating advanced oxidation contactor equipped with activated carbon/TiO2 composite sheets. *Journal of Hazardous Materials*, *383*, 121005. https://doi.org/10.1016/j.jhazmat.2019.121005

Nugraha, J., and Fatimah, I. (2016). Modeling of photocatalytic activity of ZnO/AC by using linear probability model, logit and complementary log transformation. *Procedia Engineering*, *148*, 1112-1120. https://doi.org/10.1016/j.proeng.2016.06.613

Ohtani, B., Prieto-Mahaney, O. O., Li, D., and Abe, R. (2010). What is Degussa (Evonik) P25? Crystalline composition analysis, reconstruction from isolated pure particles and photocatalytic activity test. *Journal of Photochemistry and Photobiology A: Chemistry*, *216*(2-3), 179-182. https://doi.org/10.1016/j.jphotochem.2010.07.024

Oller, I., Malato, S., and Sánchez-Pérez, J. A. (2011). *Combination of advanced oxidation processes and biological treatments for wastewater decontamination – A review*. https://doi.org/10.1016/j.scitotenv.2010.08.061

Pawlowski, L. (1994). Standard methods for the examination of water and wastewater, 18th edition: Arnold E. Greenberd, Lenore S. Clesceri, Andrew D. Eaton (Editors) Water Environment Federation, Alexandria, USA, 1992; 1025 pp; US$120 (Hardcover); ISBN 0-87553-207-1. *Science of The Total Environment, 142*(3), 227-228. https://doi.org/10.1016/0048-9697(94)90332-8

Paździor, K., Bilińska, L., and Ledakowicz, S. (2019). A review of the existing and emerging technologies in the combination of AOPs and biological processes in industrial textile wastewater treatment. *Chemical Engineering Journal*, *376*, 120597. https://doi.org/10.1016/j.cej.2018.12.057

Rizzo, L., Della Sala, A., Fiorentino, A., and Li Puma, G. (2014). Disinfection of urban wastewater by solar driven and UV lamp – TiO2 photocatalysis: Effect on a multi drug resistant Escherichia coli strain. *Water Research*, *53*, 145-152. https://doi.org/10.1016/j.watres.2014.01.020

Rocha, E. M. R., Vilar, V. J. P., Fonseca, A., Saraiva, I., and Boaventura, R. A. R. (2011). Landfill leachate treatment by solar-driven AOPs. *Solar Energy*, *85*(1), 46-56. https://doi.org/10.1016/J.SOLENER.2010.11.001

Ruiz-Delgado, A., Plaza-Bolaños, P., Oller, I., Malato, S., and Agüera, A. (2020). Advanced evaluation of landfill leachate treatments by low and high-resolution mass spectrometry focusing on microcontaminant removal. *Journal of Hazardous Materials*, *384*, 121372. https://doi.org/10.1016/j.jhazmat.2019.121372

Satuf, M. L., Brandi, R. J., Cassano, A. E., and Alfano, O. M. (2005). Experimental method to evaluate the optical properties of aqueous titanium dioxide suspensions. *Industrial & Engineering Chemistry Research*, *44*(17), 6643-6649. https://doi.org/10.1021/ie050365y

Thanekar, P., Murugesan, P., and Gogate, P. R. (2018). Improvement in biological oxidation process for the removal of dichlorvos from aqueous solutions using pretreatment based on Hydrodynamic Cavitation. *Journal of Water Process Engineering*, *23*, 20-26. https://doi.org/10.1016/j.jwpe.2018.03.004

Vahabian, M., Hassanzadeh, Y., and Marofi, S. (2019). Assessment of landfill leachate in semi-arid climate and its impact on the groundwater quality case study: Hamedan, Iran. *Environmental Monitoring and Assessment*, *191*, 109. https://doi.org/10.1007/s10661-019-7215-8

Villamizar, S., Maturana Cordoba, A., and Soto, J. (2022). Leachate decontamination through biological processes coupled to advanced oxidation: A review. *Journal of the Air & Waste Management Association*, *72*(12), 1341-1365. https://doi.org/10.1080/10962247.2021.1985012

Xu, B., Ahmed, M. B., Zhou, J. L., and Altaee, A. (2020). Visible and UV photocatalysis of aqueous perfluorooctanoic acid by TiO2 and peroxymonosulfate: Process kinetics and mechanistic insights. *Chemosphere*, *243*, 125366. https://doi.org/10.1016/j.chemosphere.2019.125366

Yashni, G., Al-Gheethi, A., Radin Mohamed, R. M. S., Dai-Viet, N. V., Al-Kahtani, A. A., Al-Sahari, M., Nor Hazhar, N. J., Noman, E., and Alkhadher, S. (2021). Bio-inspired ZnO NPs synthesized from *Citrus sinensis* peels extract for Congo red removal from textile wastewater via photocatalysis: Optimization, mechanisms, techno-economic analysis. *Chemosphere*, *281*, 130661. https://doi.org/10.1016/j.chemosphere.2021.130661

Yasmin, C., Lobna, E., Mouna, M., Kais, D., Mariam, K., Rached, S., Abdelwaheb, C., and Ismail, T. (2020). New trend of Jebel Chakir landfill leachate pre-treatment by photocatalytic TiO2/Ag nanocomposite prior to fermentation using Candida tropicalis strain. *International Biodeterioration & Biodegradation*, *146*, 104829. https://doi.org/10.1016/j.ibiod.2019.104829

Research Article / Chemical, Food, and Environmental Engineering http://doi.org/10.15446/ing.investig.103232

Physicochemical, Rheological, and Thermal Properties of Pot-Honey from the Stingless Bee *Melipona beecheii*

Propiedades fisicoquímicas, reológicas y térmicas de la miel de pote de la abeja nativa sin aguijón *Melipona beecheii*

Larry A. Yah-Rosales[1], Luis A. Chel-Guerrero[2], Julio C. Sacramento-Rivero[3], and Sergio A. Baz-Rodríguez[4]

ABSTRACT

The availability of data on physicochemical properties is crucial to direct efforts towards identifying the quality standards of the Neotropical stingless bee *Melipona beecheii's* pot-honey. In this vein, other properties apart from those typically considered for *Apis mellifera* could also be relevant in characterizing the honey of this stingless bee. The physicochemical, rheological, and thermal properties of pot-honey from *Melipona beecheii* (Yucatán, México) were analyzed. Samples were collected from two annual harvests (2018 and 2019) and from a rural and an urban location. Free acidity, moisture, total reducing sugars, diastase activity, hydroxymethylfurfural content, and electrical conductivity were measured using standard techniques. The rheological and thermal behaviors were determined via Couette rheometry and differential scanning calorimetry, respectively. The physicochemical properties of *Melipona beecheii* pot-honey can be incorporated into a general quality specification for honey of the Neotropical *Melipona genus*, or as the basis for a regional (Mesoamerican) standardization of honey from this particular bee species. The rheological analyses indicated the Newtonian behavior of *Melipona* honey in the full studied range of 10-40 ˚C (7,545-244 cp), showing dynamic viscosities significantly lower than those expected for *Apis mellifera* honey, primarily due to its high water content. Two main endothermic transitions were detected via differential scanning calorimetry: at 96-162 ˚C and at 169-230 ˚C. The *Apis mellifera* honey samples showed the same thermal transitions but differed from *Melipona beecheii* honey in their peak temperatures and enthalpies.

Keywords: *Melipona beecheii* pot-honey, physicochemical characterization, rheology, differential scanning calorimetry

RESUMEN

La disponibilidad de datos de propiedades fisicoquímicas es fundamental para direccionar esfuerzos hacia la identificación de los estándares de calidad de la miel de pote de *Melipona beecheii*. En este contexto, otras propiedades diferentes de aquellas típicamente consideradas para *Apis mellifera* podrían también ser relevantes para caracterizar la miel de esta abeja sin aguijón. Se analizaron las propiedades fisicoquímicas, reológicas y térmicas de la miel de pote de *Melipona beecheii* (Yucatán, México). Se recolectaron muestras de dos temporadas de cosecha anuales (2018 y 2019) y de una locación rural y una urbana. Empleando técnicas estándar, se midieron la acidez libre, la humedad, los azúcares reductores totales, la actividad diastásica, el contenido hidroximetilfurfural y la conductividad eléctrica. Los comportamientos reológico y térmico se determinaron por reometría tipo Couette y calorimetría diferencial de barrido respectivamente. Las propiedades fisicoquímicas de la miel de pote de *Melipona beecheii* se pueden incorporar a una especificación general de calidad para la miel de género Neotropical *Melipona*, o como base para una estandarización regional (Mesoamericana) de la miel de esta especie particular de abeja. Los análisis reológicos indicaron el comportamiento Newtoniano de la miel de *Melipona* para todo el rango estudiado de 10-40 ˚C (7,545-244 cp), mostrando viscosidades dinámicas significativamente menores que aquellas esperadas en la miel de *Apis mellifera*, primordialmente debido a su mayor contenido de agua. Se detectaron dos transiciones endotérmicas relevantes por calorimetría diferencial de barrido: a 96-162 ˚C y a 169-230 ˚C. Las muestras de miel de *Apis mellifera* presentaron las mismas transiciones térmicas, pero discreparon de la miel de *Melipona beecheii* en la temperaturas pico y las entalpías.

Palabras clave: miel de *Melipona beecheii*, caracterización fisicoquímica, reología, calorimetría diferencial de barrido

Received: June 17th, 2022
Accepted: May 19th, 2023

Introduction

Many centuries before the Spanish arrival to the Yucatán Peninsula, the Mayan people systematized the keeping of a stingless bee: *Melipona beecheii*. The pot-honey produced by *Melipona* was the only sweetener used by the Mayans; its honey and cerumen represented a valuable commodity for trade. Their high demand, advanced technification, and the cultural rooting of the Maya for their valuable bees, allowed meliponiculture to survive and thrive throughout the colonial era and the first century of the independent México.

[1] Chemical engineer, Universidad Autónoma de Yucatán, México. Affiliation: MSc student, Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, México. Email: larrytt117@hotmail.com

[2] PhD in Food Science, Instituto Politécnico Nacional, México. Affiliation: Researcher-professor, Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, México. Email: cguerrer@correo.uady.mx

[3] PhD in Chemical Engineering, University of Manchester, United Kingdom. Affiliation: Researcher-professor, Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, México. Email: julio.sacramento@correo.uady.mx

[4] PhD in Chemical Engineering, Universidad Autónoma Metropolitana - Iztapalapa. Affiliation: Researcher-professor, Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, México. Email: sergio.baz@correo.uady.mx

In time, various factors contributed and still contribute to the fall of meliponiculture in Yucatán, mainly deforestation and displacement by the European or Africanized variations of *Apis mellifera* (Villanueva *et al*., 2005).

Despite the relatively small economic relevance of meliponiculture nowadays, it still has a strong cultural and ecological significance. Reassessing their keeping is very important for the preservation of stingless bees and the regional ecosystems' health. One way to do this is promoting the commercialization of *Melipona* honey, which is highly valued in niche markets due to its medicinal and traditional applications and is currently sold as a high value-added product at a market price 30-50 times higher than that of Apis honey. A very important starting point is the characterization of *Melipona* hive products, mainly pot-honey, in order to set quality standards. Needless to say, international standards have only been established for Apis honey (Codex Alimentarius, 2019), whose physicochemical properties are known to significantly differ from those of pot-honeys (Souza *et al*., 2006).

To date, some actions have already been undertaken in South American countries in the form of regulations applicable to the pot-honey of stingless bees in Neotropical regions, namely:

1. For informational purposes only, Annex B of the Colombian Apis honey norm (ICONTEC, 2007) presents ranges of physicochemical properties applicable to species of the genera *Frieseomelitta*, *Melipona*, *Plebeia*, *Scaptotrigona*, *Tetragonisca*, and *Trigona*. The reference values are taken from Souza *et al*. (2006) and are the result of studies conducted in Central and South American countries between 1964 and 2006.

2. Recently, the pot-honey of *Tetragonisca fiebrigi* was included in chapter X of the Argentinian Codex Alimentarius (Código Alimentario Argentino, 2019), currently the only official norm at the national level which regulates the quality standards of pot-honey from a stingless native bee.

3. At the subnational level, an official norm was created by the Brazilian State of Bahia, albeit limited to the *Melipona* genus (ADAB, 2014).

4. Modifications have been suggested to the Venezuelan honey norm (COVENIN, 1984) with specifications for the pot-honey of stingless bees (Vit, 2013). These modifications are based mainly on previous works characterizing regional pot-honeys.

Currently, there is academic interest in compiling data and encouraging the creation of standard quality regulations for the pot-honey stingless bees worldwide (Ávila *et al*., 2018; Braghini *et al*., 2021; Colombo-Pimentel *et al*., 2022; Nordin *et al*., 2018; Souza *et al*., 2021). However, actions 2 and 3 suggest that the most reliable way to propose standards is focusing on specific species or genera. This is reasonable, since bees from the same species or genus have similar foraging behaviors, nectar processing work, morphological hive features, and geographical environments (May-Canché *et al*., 2022).

In the Mesoamerican region, many stingless bee species are harnessed, but few of them stand out as exploitable for economic activity like *Melipona beecheii*. This stingless bee is distributed not only across southwest México and Central America (Mesoamerica) but also across Cuba and Jamaica (May-Itzá *et al*., 2019).

The creation of standards for regulating the quality of pot-honey in México, Central American countries, or the whole Mesoamerican region is desirable because of the cultural, ecological, and economic potential of keeping stingless bees. The pot-honey of *Melipona beecheii* is the single most important stingless bee hive product that is common for the countries in the region. In this sense, it is advisable to direct efforts towards identifying the quality standards of this product, and data availability on physicochemical properties is crucial for this task.

To the best of the authors, the current knowledge for the standardization of pot-honey of *Melipona beecheii* is mainly based on two sources: i) the specifications given by official norms applicable to Apis honey (Codex Alimentarius, 2019; NOM-004-SAG/GAN-2018) and Latin American stingless bee pot-honey (ADAB, 2014; ICONTEC, 2007), and ii) the reported values of physicochemical parameters for honeys of *Melipona becheii* throughout Mesoamerica and Cuba (Alarcón-Sorto and Ibáñez-Salazar, 2008; Álvarez-Suárez *et al*., 2018; Dardón and Enríquez, 2008; Espinoza-Toledo *et al*., 2018; Gutiérrez *et al*., 2008; May-Canché *et al*., 2022; Moo-Huchin *et al*., 2015; Rodríguez-Suazo, 2014; Umaña *et al*., 2021; Vit *et al*., 2006). Both of these sources are reported in Tables 1 and 2, respectively.

Moisture, free acidity, total reducing sugars, sucrose, ash, hydroxymethylfurfural, diastase activity, and electrical conductivity are some of the relevant properties for characterizing honey (Codex Alimentarius, 2019). Beyond these, other physical and chemical features can provide information about the quality and composition of honey with specific biological origins.

The rheological properties of honey are related to its chemical composition. They depend on its botanical and biological origin, moisture, temperature, carbohydrate composition, and granulation. Moreover, they are related to the sensory quality of honey and affect technological operations such as heating, mixing, hydraulic transport, and bottling (Machado de Melo *et al*., 2018). Costa *et al*. (2013) obtained viscosity flow curves for pot-honey of the Brazilian stingless bee *Melipona subnitida* at 25 °C, identifying a Newtonian behavior, with lower viscosity than *Apis mellifera* honey. Other studies report on the viscosity of other South-American pot-honeys, albeit using sensorial analysis (Batista de Sousa *et al*., 2016).

The chemical composition of honeys also defines its thermal properties. In this sense, Differential Scanning Calorimetry (DSC) is a useful technique for characterizing the thermal behavior of foods. Depending on composition and heating and cooling rates, honey undergoes several thermal events. A couple of works have dealt with the calorimetric features of pot-honey samples from South American stingless bees. One of them identified glass transitions and related their features to the water content of *Melipona subnitida* pot-honey (Costa *et al*., 2013). The other one identified thermal transitions at T > 18 °C of pot-honey from *Melipona fuscipes*, *Melipona favosa*, and *Melipona compressipes* (Cardona *et al*., 2018). The authors suggested that DSC is an adequate technique to characterize honey samples for species authentication and detecting adulteration.

To contribute to the knowledge about the physicochemical properties of pot-honey, in this work, the following physicochemical properties were assessed on pot-honey of *Melipona beecheii* from the Yucatán State (México): free acidity, moisture, reducing sugars, ash and hydroxymethylfurfural contents, diastase activity, density, and electrical conductivity. Samples of rural and urban locations of two annual harvests were analyzed. Based on the assessment of existing and the obtained data, a suggestion of ranges of values for standard quality properties was conducted. Moreover, a first report on the flow and thermal behavior of such pot-honey is presented using rheological and calorimetric measurements.

Methodology

The pot-honey samples were collected from two locations of the Yucatán State in México: Mérida (urban) and Tekax (rural) (*factor 1: location*). The nectar sources in the urban location were primarily ornamental trees in thoroughfare and parks, as well as backyard fruit trees and plants to a lesser extent; while, in the rural location, the sources were primarily backyard fruit trees and low deciduous forest and undergrowth. In each location, two multi-flora samples were obtained in April 2018 and April 2019, respectively (*factor 2: harvest*). Each pot-honey sample consisted of 100 ml directly extracted by suction with hose and syringe from two different rational hives by halves. The samples were stored at 25 °C to be analyzed within two months after sampling. It was assumed that no significant changes in the properties of honey occur in such period. The measured variables were physicochemical, rheological, and calorimetric properties (Figure 1).

Physicochemical properties

Free acidity, moisture, total reducing sugars, diastase activity, hydroxymethylfurfural content, and electrical conductivity were measured via well-known standard methods according to the Mexican standard for honey (NOM-004-SAG/GAN-2018). The determination of free acidity (meq/kg) was based on the potentiometric titration of diluted honey, using a solution of NaOH. The moisture content (g/100 g) was determined by measuring the refractive index with an Abbe refractometer (ATAGO NAR-2T, Atago Co.) coupled with a thermostatic bath (25 °C). The total reducing sugars content (TRS, g/100 g) was determined via the modified volumetric method of Lane and Eynon (1923). The diastase activity was obtained with the method of Schade *et al*. (1958). The hydroxymethylfurfural content (HMF, mg/kg) was measured via the spectrophotometric Carrez method. Finally, the electric conductivity (mS/cm) was determined by direct measuring in 20% water diluted honey. Color was not measured in this work; this property is of a sensory nature and dependent on the content of minerals, pigments, and other minor components, and it is not included in conventional grading methods (Moo-

Table 1. Standard physicochemical specifications of honey

Property	*Apis* Honey		Stingless Bee Pot-Honey	
	Codex Standard (Codex Alimentarius, 2019)	**NOM-004-SAG/GAN-2018**	***Melipona* genus** (ADAB, 2014)	**Stingless Bee** (ICONTEC, 2007)
Moisture	≤ 20 %		20-35 %	19.9-41.9 %
Reducing sugars	≥ 60 g/100 g		≥ 60 g/100 g	58.0-75.7 g/100 g
Sucrose	≤ 5 g/100 g		≤ 6 g/100 g	1.1-4.8 g/100 g
Free Acidity	≤ 50 mEq/kg		≤ 50 mEq/kg	5.9-109 mEq/kg
Hydroxymethylfurfural	≤ 40 mg/kg (≤ 80 mg/kg tropical honey)		≤ 10 mg/kg	0.4-78.4 mg/kg
Diastase Activity	≥ 8 Schade (≥ 3 specific cases)		≥ 3 Schade	0.9-23 Schade
Electrical Conductivity	≤ 0.8 mS/cm		Not specified	0.49-8.77 mS/cm
Ash	Not specified		≤ 0.6 g/100 g	0.01-1.18 g/100 g

Source: Authors

Table 2. Reported data regarding physicochemical properties for pot-honey of the Mesoamerican stingless bee *Melipona beecheii*

Location & References	Moisture (%)	TRS (g/100 g)	Acidity (mEq/kg)	HMF (mg/kg)	Diastase activity (Schade)	Electric conductivity (mS/cm)	Sucrose (g/100 g)	Ash (g/100 g)
Costa Rica								
Umaña *et al*. (2021)	23.5±1.3	63.1[1]	23.3±8.2	6.2±5.6	<1	n. r.	0.33±1.32	0.143±0.103
Cuba								
Alvarez-Suarez *et al*. (2018)	28.62±5.61	n. r.	41.52±8.19	9.23±1.32	1.3 ± 0.12	0.58±0.14	n. r.	0.46±0.03
Fonte *et al*. (2013)	24	n. r.	35	n. r.	n. r.	n. r.	n. r.	n. r.
El Salvador								
Alarcon-Sorto & Ibañez-Salazar (2008)	25.05±1.10	75.63±7.95	39.51±19.9	71.55±63.6	Not detected	0.005±0.003	5.34±4.41	0.047±0.04
Guatemala								
Dardón & Enríquez (2008)	17.32±2.64	68.77±3.82	23.23±30.0	0.1±0.007	21.29±32.84	n. r.	3.50±4.14	0.07±0.05
Gutiérrez *et al*. (2008)	27.20±3.14	n. r.	n. r.	0.12±0.08	1.61±0.13	0.15±0.03	3.18±4.55	n. r.
Vit *et al*. (2006)	24.5	59.52[2]	11.83	n. r.	n. r.	n. r.	9.12	0.02
Honduras								
Rodríguez-Suazo (2014)[3]	26.76±1.56 26.53±1.42	n. r.	n. r.	n. r.	n. r.	n. r.	n. r.	n. r.
Mexico (Chiapas)								
Espinoza-Toledo *et al*. (2018)[4]	37.20±0.92 36.93±1.70	68.04±0.01 71.25±0.07	35±0.53 24.79±2.5	9.7±0.79 Not detected	n. r.	0.114±0.013 0.127±0.004	n. r.	n. r.
Grajales-Conesa *et al*. (2018)	25.3±0.17	n. r.	40.33±1.26	2.1±2.17	n. r.	0.507±0.00	n. r.	n. r.
May-Canché *et al*. (2022)	23.41±1.79	64.47±3.79	30.8±11.5				n. r.	n. r.
México (Yucatan)								
Moo-Huchin *et al*. (2015)[5]	23.2±1.0 (21-25.3)	67.7±4.6 (57.1-74.2)	35.0±12.8 (13-71.3)	17.9±11.1 (4-45.5)	n. r.	n. r.	n. r.	0.16±0.12 (0.01-0.6)
Catzín-Ventura *et al*. (2008)	24.1±1.2	n. r.	40.3±10.5	5.7±8.3	n. r.	n. r.	0.1±0.1	n. r.

[1] Reported here as the sum of analyzed reducing sugars
[2] Assuming TRS as glucose plus fructose contents
[3] The data is reported for log and natural hives
[4] The data is separately reported for two locations
[5] Average and data range
n. r. "not reported"

Source: Authors

Huchin *et al*., 2015; Santos *et al*., 2018; Smetanska *et al*., 2021). To determine whether the rheological and calorimetric properties of honey can be related to its composition and/or biological source, samples of Apis honey were acquired from rural beekeepers and analyzed, with their origin being multi-flora (Dzitás, Yucatán) and 'Tzitzilché' (*Gymnopodium floribundum*) dominant (Espita, Yucatán), respectively.

Rheology and calorimetry

Tests were carried out in a rotational viscometer (RVDV-II+Pro, Brookfield) with a small sample adapter (SSA) coupled to a thermostatic bath. A Thermosel Spindle SCA-21 was used as a rotational Couette-type geometry. The readings of shear stress and apparent viscosity at different shear rates allowed constructing flow curves and determining the rheological behavior of the pot-honey samples. Small samples were used. The test temperatures were 10, 20, 30, and 40 °C. Temperature is an additional statistical factor for this case, and the evaluated values correspond to the expected environmental range during harvest and handling. The shear rate was varied by changing the rotation velocity of the spindle, always in the range of % torque recommended by the manufacturer.

The thermal behavior was determined with a differential scanning calorimeter DSC-6 (Perkin-Elmer). The DSC device was calibrated with indium and the data analyzed using the Pyris software. 3-5 µg samples were weighted on aluminum pans (20 µL, max. pressure: 2 bar). Each pan was hermetically sealed and left to reach thermal equilibrium at room temperature. The samples were scanned at temperatures between 15 and 260 °C, at a rate of 5 °C/min. The enthalpy changes (Δh) of the thermal transitions and the associated peak (T_{peak}), initial (T_i), and final (T_f) temperatures were identified from the resulting thermograms.

Data analysis

The data obtained were evaluated with two-way analyses of variance (ANOVA) at a 5% significance level. Duncan tests were applied to determine statistical differences between means. These analyses were performed in Statgraphics Centurion 18.1.12.

The fitting of parameters regarding the viscometric data to a predictive model was made using Matlab R2019a, minimizing the difference between the measurements and the model output. Finally, the thermal parameter values obtained from calorimetric experiments were determined automatically via the TA Universal Analysis 2000 software.

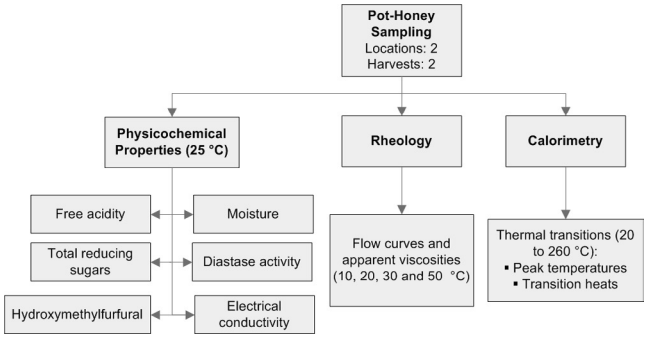


Figure 1. Physicochemical analyses of pot-honey samples.
Source: Authors

Results and discussion

Physicochemical properties

The moisture of the pot-honey samples obtained this work was around 22% in all cases (Table 3), close to the lower limit and within the range suggested as standard for Brazilian species of the genus *Melipona*, but above the maximum value specified for *Apis* honey (Table 1).

Table 3. Measured physicochemical properties of pot-honey samples of *Melipona beecheii* from Yucatán, México

Property	*Melipona* pot-honey		Average
	Urban location	Rural location	
Moisture (%)	22.16^{aa}	22.05^{aa}	22.11
Acidity (mEq/kg)	48.27^{aa}	43.2^{aa}	45.74
Total reducing sugars (%)	50.89^{aa}	50.65^{ab}	50.77
Diastase Activity (Schade)	8^{aa}	11.39^{aa}	9.70
HMF (mg/kg)	Not detected	Not detected	Not detected
Conductivity (mS/cm)	0.38^{aa}	0.34^{aa}	0.36

* Different letters in the same line indicate significant differences at $p < 0.05$; the first one for location and the second one for harvest

Source: Authors

The relatively high moisture (as high as 41,9%) is a distinctive feature of stingless bee pot-honey (Souza *et al.*, 2006). One reason for this relates to the fact that the relative surface area of pots with processed nectar in stingless-bee hives (including free and slightly permeable wall surfaces) is always significantly lower than that of processed nectar in honeycombs of *Apis mellifera*. In particular, the surface area to volume ratio for storage pots of *Melipona beecheii* is roughly 2,5 cm^2/cm^3 (typical 'spherical' pot: 8 cm^3) (Lóriga Peña *et al.*, 2015). On the other hand, honeycomb cells of *Apis mellifera* have a ratio of roughly 12 cm^2/cm^3 (standard cell of hexagonal section: 0,5 cm in diameter and 1 cm deep) (Hailu and Biratu, 2016). The water evaporation of processed nectar occurs at its surface when in contact with intranidal air (primarily) and pot walls (marginally). Thus, the ripening of stingless bee honeys involves a more restrictive dehumidifying process.

The moisture content values in this work fall in the range of those for *Melipona beecheii* honey from other Mesoamerican locations, and they are close to the lower limits (Table 2). This may be due to the fact that sampling was carried in April, a month of the harvest season of the api-botanical cycle in the Yucatán Peninsula, which is characterized by hot and dry environmental conditions (Quezada-Euán, 2018). Under these conditions, the evaporation of water from the pots can reach lower values than those of other seasons and locations.

The free acidity of the analyzed pot-honeys was 45,74 mEq/kg on average. No significant difference was observed between sampling locations or harvesting periods (Table 3). The acidity of honey is due to the presence of organic acids, which come directly from the nectar or are produced by the action of enzymes secreted by bees during ripening and storage (De-Melo *et al.*, 2017). Consequently, it is affected by the floral source and intranidal food processing. Such factors are associated with foraging preferences and the biology of the species, respectively. Thus, it can be hypothesized that this property could be distinctive according to the biological origin and freshness of honey. In this sense, the acidity value reported in this work (Table 2) reinforces the directive given by ADAB (2014) (Table 1) and research observations emphasizing the more acidic nature of stingless bee honeys (Ávila et al., 2018; Colombo-Pimentel et al., 2022; Nordin et al., 2018; Souza et al., 2021).

The total reducing sugars were 50,89 mEq/kg on average. Significant differences were observed only between harvests, not between locations (Table 3). Since the carbohydrate content depends on the botanical source of nectars and climatic conditions (Mohammed, 2020), these variations in the sugar content are expected. It is relevant to point out that the sugar content also can contribute to acidity via enzymatic degradation and fermentation. The obtained values fall within the ranges observed for the species *Melipona* (Braghini *et al.*, 2021) but are relatively low in comparison with the regulations for Brazilian *Melipona* and the reported data for *Melipona beecheii* (Tables 1 and 2).

The obtained values for diastase activity (9.70 Schade) and electric conductivity (0,36 mS/cm) are in the same order of magnitude as the reported data for *Melipona beecheii* (Table 2), and no significant differences were observed between sampling locations or harvests (Table 3). Finally, HMF was not detected, probably because of the relatively short time between sampling and analysis.

Table 4 contains the suggested specifications regarding the standard quality of *Melipona beecheii* honey. They were

elaborated while considering the typical physicochemical parameters used for the standardization of *Apis mellifera* honey, as well as the obtained and available data for *Melipona beecheii* honey. The discordant data regarding moisture (Espinoza-Toledo *et al.*, 2018), total reducing sugars (this work), and sucrose (Vit *et al.*, 2006) must be taken with caution; they were not taken into account for the suggestions in Table 4. The available data on HMF and diastase activity are currently inconclusive; more studies are needed to outline – or confirm – a range. The information for this pot-honey generally agrees with ADAB standards (2014).

Table 4. Suggested standards for *Melipona beecheii* honey from currently available data

Property	Stingless bee pot-honey from *Melipona beecheii*
Moisture	≤ 35 %
Reducing sugars	≥ 60 g/100 g
Sucrose	≤ 6 g/100 g
Free Acidity	≤ 50 mEq/kg
Hydroxymethylfurfural	Inconclusive
Diastase Activity	Inconclusive
Electrical Conductivity	≤ 0.6 mS/cm
Ash	≤ 0.5 g/100 g

Source: Authors

Rheology and calorimetry

A Newtonian behavior was detected in *Melipona beecheii* honey at the four temperatures (Figure 2). No significant difference was observed between harvests, but there was a significant difference between sampling locations (Table 5). Higher values were measured for *Apis mellifera* honey, also exhibiting a Newtonian behavior, in the same order of magnitude previously reported for local samples of *Apis* (Mora-Escobedo *et al.*, 2006). There is a clear contrast between the viscometrical measurements of honeys by species. *Melipona* honey is characterized by a higher fluidity, primarily because of the high water content. However, this property may not be adequate for the direct differentiation of samples, since it is also dependent on composition, basically water and carbohydrate contents, which provide independent information about standard differences (Faustino and Pinheiro, 2021). However, it has been suggested that rheological parameters can be useful to identify adulteration with syrups, as adulterated samples and honey exhibit different behaviors under variable conditions (temperature, concentration, storage time) (Kamboj and Mishra, 2015).

The influence of temperature on honey viscosity was fitted by nonlinear regression to an Arrhenius-type model:

$$\mu = \mu_0 exp\left(-\frac{E_a}{RT}\right)$$

The activation energy (Ea, a barrier to flow) was larger for the honey of *Apis mellifera* than for that of *Melipona beecheii*; the pre-exponential factor (μ_0, viscosity at temperature close to infinity) was two orders of magnitude lower for *Apis mellifera* honey (Table 5).

Regarding the differential scanning calorimetry, three endothermic phenomena were detected in the honey samples (Figure 3 and Table 6):

1. A very weak phenomenon at 18-24 °C, also observed in a previous work with Brazilian *Melipona* honey, could correspond to the vaporization of some volatile compounds (Cardona *et al.*, 2018). None of the thermal parameters showed significant differences between sampling locations or between harvests *Melipona beecheii* honey. Since this first transition is weak and occurs immediately after the initial heat flow from the equipment to the sample pan (or as final part of it), it could also be related to the transient process to thermal equilibrium between the sample and the controlled oven of the equipment.

2. A strong endothermic phenomenon at 96-162 °C may correspond to water vaporization and the rupture of sugar-water and sugar complexes (Cardona *et al.*, 2018). The location (urban and rural) caused statistical differences on Δh, revealing the importance of nectar origin and composition.

3. A transition between 169-230 °C can be attributed to the melting and decomposition of sugars (mono-, di-, tri-, and oligosaccharides) (Cardona *et al.*, 2018). Only the location factor caused statistical differences on the parameters of this third thermal transition (Δh and T*f*).

The samples of honey of *Apis mellifera* also had the same three thermal transitions as that of *Melipona*. To explore whether a differentiation through DSC may be suggested, a statistical comparison through an ANOVA was carried out for the whole set of thermal data with the species as factor. Significant differences between the means of T_{peak} and T_i for transition 1, T_{peak} and T_i for transition 2, and Δh and T_f for transition 3 were detected (Table 6). Despite these observations, the mean values of some of these parameters are close. However, the mean values of T_{peak} for transition 2 and Δh for transition 3 seem to have values different enough to suggest them as a differentiation parameter. In the first case (T_{peak}), the difference could be related to distinctive intermolecular structures, and, in the second case (Δh), to the chemical nature of remaining sugars at such relatively high temperatures.

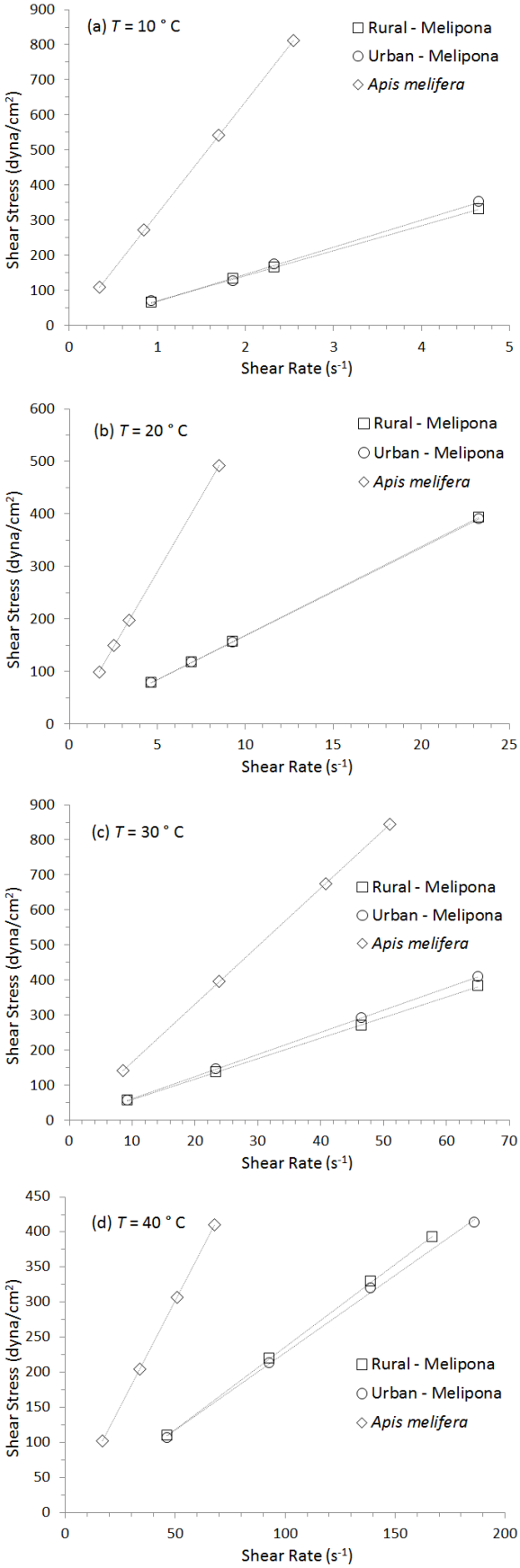


Figure 2. Flow curves for honeys at: a) 10, b) 20, c) 30, and d) 40 °C
Source: Authors

Table 5. Dynamic viscosity of honeys and parameters of Arrhenius-type equation

Temperature (°C)	Dynamic viscosity (cp)			
	Melipona pot-honey*			*Apis melifera* honey
	Urban location	Rural location	Average	
10	7,545^{aa}	6,769^{ba}	7,157	31,995
20	1,826^{aa}	1,686^{ba}	1,756	5,820
30	627^{aa}	5,76^{ba}	602	1,664
40	244^{aa}	2,59^{ba}	251	604
μ_0 (cp):	2.1532×10^{-12}	5.8038×10^{-12}	3.5353×10^{-12}	3.5271×10^{-14}
E_a (cal/mol):	-20,091	-19,538	-19,814	-23,254

* Different letters in the same line indicate significant differences at $p < 0.05$; the first one for location and the second one for harvest

Source: Authors

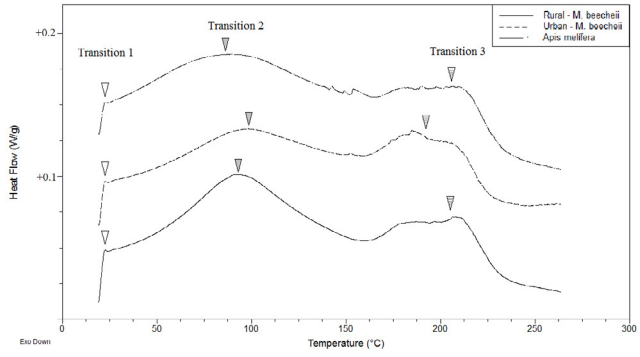


Figure 3. Typical DSC thermograms as observed in this work's samples
Source: Authors

Conclusions

The physicochemical, rheological, and calorimetric properties of urban and rural (Yucatán, México) samples of *Melipona beecheii* honey were characterized in this study. Along with the analysis of the current knowledge about the properties of such pot-honey, the following conclusions arise from this work:

- The standard quality specifications for the physicochemical properties of *Melipona beecheii* honey differ from the specifications for *Apis mellifera*, particularly in moisture content. Moreover, the measured and literature data for hydroxymethylfurfural content and diastase activity seem to be inconclusive, leaving the issue of whether these are relevant properties to characterize this particular product open for debate.

- Since there is great variability in the reported data on the physicochemical properties of stingless bee honey in general, it is advisable to establish standard specifications for honey by stingless be species or genus. In this sense, the measured and literature data on the physicochemical properties of honey from *Melipona beecheii* have proven to be embeddable into a general

Table 6. Enthalpies and transition temperatures of honeys.

Transition	Property	Units	*Melipona* pot-honey*			*Apis* honey
			Urban	Rural	Average	
1	Δh	J/g	3,1 ± 2,1^{aa}	3,4 ± 0,9^{aa}	3,2 ± 1,5	1,7 ± 1,2
	T_{peak}	°C	21,6 ± 1,0^{aa}	20,9 ± 0,8^{aa}	21,2 ± 0,9	22,2 ± 0,1
	T_i	°C	18,8 ± 1,0^{aa}	18,5 ± 1,1^{aa}	18,7 ± 1,0	20,2 ± 0,7
	T_f	°C	23,8 ± 0,7^{aa}	24,2 ± 0,1^{aa}	23,9 ± 0,5	24,0 ± 0,5
2	Δh	J/g	243,6 ± 13,8^{aa}	300,4 ± 27,9^{ba}	267,4 ± 47,4	290,1 ± 46,7
	T_{peak}	°C	129,3 ± 11,6^{aa}	122,0 ± 3,8^{aa}	125,6 ± 8,9	91,4 ± 5,9
	T_i	°C	96,8 ± 12,7^{aa}	100,21 ± 0,4^{aa}	98,5 ± 8,5	38,5 ± 13,1
	T_f	°C	161,2 ± 12,8^{aa}	161,9 ± 7,8^{aa}	161,5 ± 9,8	155,5 ± 10,6
3	Δh	J/g	106,9 ± 17,9^{aa}	77,3 ± 8,1^{ba}	92,1 ± 20,4	158,2 ± 17,4
	T_{peak}	°C	193,8 ± 11,6^{aa}	209,3 ± 4,0^{aa}	201,5 ± 11,5	209,3 ± 2,0
	T_i	°C	169,3 ± 10,7^{aa}	182,3 ± 3,3^{aa}	175,9 ± 9,9	167,5 ± 0,7
	T_f	°C	225,6 ± 1,04^{aa}	230,8 ± 2,1^{bb}	228,2 ± 3,2	235,0 ± 2,8

*Different letters in the same line indicate significant differences at $p < 0,05$; the first one for location and the second one for harvest

Source: Authors

quality specification for honey of the Neotropical *Melipona* genus. They may also be the basis for a regional (Mesoamerican) quality specification norm for honey of this species.

- The honey of *Melipona beecheii* has a Newtonian behavior in the whole range of 10-40 °C (7 545-244 cp), being characterized by a lower dynamic viscosity than honey of Apis mellifera (31 995-604 cp), primarily due to its high water content.

- The thermal decomposition and melting of sugars generate endothermic transitions measurable via differential scanning calorimetry (15-260 °C). It was found that the samples evaluated in this work differ in the peak temperature and the enthalpy for endothermic transitions at 96-162 and 169-230 °C, respectively. These thermal parameters can differentiate the pot-honey of *Melipona beecheii* from that of *Apis mellifera*.

Acknowledgements

This work was supported by the Mexican Council of Science and Technology (Consejo Nacional de Ciencia y Tecnología, CONACyT, México) through grant number 316824.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

ADAB (2014). *Regulamento técnico de identidade e qualidade do mel de abelha social sem ferrão, gênero Melipona* (*PORTARIA ADAB No 207 DE 21-11-2014*). Agência Estadual de Defesa Agropecuária da Bahia.

Alarcón-Sorto, R. C., and Ibáñez-Salazar, L. C. (2008). *Determinación de las características fisicoquímicas de la miel producida por las especies de abejas nativas sin aguijón Melipona beechii* (*jicota*) *y Tetragonisca angustula* (*chumelo*) *de meliponicultores de la zona norte del Departamento de Chalatenango* [Undergraduate thesis, Universidad de El Salvador]. https://ri.ues.edu.sv/id/eprint/2892.

Álvarez-Suárez, J. M., Giampieri, F., Brenciani, A., Mazzoni, L., Gasparrini, M., González-Paramás, A. M., Santos-Buelga, C., Morroni, G., Simoni, S., Forbes-Hernández, T. Y., Afrin, S., Giovanetti, E., and Battino, M. (2018). *Apis mellifera vs. Melipona beecheii* Cuban polifloral honeys: A comparison based on their physicochemical parameters, chemical composition and biological properties. *LWT – Food Science and Technology*, *87*, 272–279. https://doi.org/10.1016/j.lwt.2017.08.079

Ávila, S., Beux, M. R., Ribani, R. H., and Zambiazi, R. C. (2018). Stingless bee honey: Quality parameters, bioactive compounds, health-promotion properties and modification detection strategies. *Trends in Food Science and Technology*, *81*, 37-50. https://doi.org/10.1016/j.tifs.2018.09.002

Batista de Sousa, J. M., Leite de Souza, E., Marques, G., Benassi, M. T., Gullon, B., Pintado, M. M., and Magnani, M. (2016). Sugar profile, physicochemical and sensory aspects of monofloral honeys produced by different stingless bee species in Brazilian semi-arid region. *LWT – Food Science and Technology*, *65*, 645-651. https://doi.org/10.1016/j.lwt.2015.08.058

Braghini, F., Biluca, F. C., Schulz, M., Gonzaga, L. V., Costa, A. C. O., and Fett, R. (2021). Stingless bee honey: A precious but unregulated product – reality and expectations. *Food Reviews International*, *38*(SUP1), 683-712. https://doi.org/10.1080/87559129.2021.1884875

Cardona, Y., Torres, A., Hoffmann, W., and Lamprecht, I. (2018). Differentiation of honey from Melipona species using differential scanning calorimetry. *Food Analytical Methods*, *11*, 1056-1067. https://doi.org/10.1007/s12161-017-1083-z

Catzín-Ventura, G., Delgado, A., and Medina-Medina, L. (2008). Evaluación de la actividad antibacteriana de la miel de Melipona beecheii y Scaptotrigona pectoralis de Yucatán, México. In *Memorias del V Congreso Mesoamericano sobre Abejas sin Aguijón* (pp. 55-62).

Codex Alimentarius (2019). *Codex Alimentarius standard for honey 12-1981* (*Revised 2001, Amended 2019*). FAO/OMS.

Código Alimentario Argentino (2019). *Alimentos azucarados*. *Capitulo X, Articulo 783bis*. https://www.argentina.gob.ar/anmat/codigoalimentario

Colombo-Pimentel, T., Rosset, M., Batista de Sousa, J. M., Gomez de Oliveira, L. I., Meireles-Mafaldo, I., Estevez-Pintado, M. M., Leite de Souza, E., and Magnani, M. (2022). Stingless bee honey: An overview of health benefits and main market challenges. *Journal of Food Biochemistry*, *46*(3), e13883. https://doi.org/10.1111/jfbc.13883

Costa, P. A., Moraes, I. C. F., Bittante, A. M., Sobral, P. J. A., Gomide, C. A., and Carrer, C. C. (2013). Physical properties of honeys produced in the Northeast of Brazil. *International Journal of Food Studies*, 2, 118-125. https://doi.org/10.7455/ijfs/2.1.2013.a9

COVENIN (1984). *Miel de abejas*. Comisión Venezolana de Normas Industriales.

Dardón, M. J., and Enríquez, E. (2008). Caracterización fisicoquímica y antimicrobiana de la miel de nueve especies de abejas sin aguijón (Meloponini) de Guatemala. *Interciencia*, *33*(12), 916-922.

De-Melo, A. A. M., Almeida-Muradian, L. B. De, Sancho, M. T., and Pascual-Maté, A. (2017). Composition and properties of Apis mellifera honey: A review. *Journal of Apicultural Research*, *57*(1), 5-37. https://doi.org/10.1080/00218839.2017.1338444

Espinoza-Toledo, C., Vázquez-Ovando, A., Santos, R. T. D. L., López-García, A., Albores-Flores, V., and Grajales-Conesa, J. (2018). Stingless bee honeys from Soconusco, Chiapas: A complementary approach. *Revista de Biología Tropical*, *66*(4), 1536-1546. https://doi.org/10.15517/rbt.v66i4.32181

Faustino, C., and Pinheiro, L. (2021). Analytical rheology of honey: A state-of-the-art review. *Foods*, *10*(8), 1079. https://doi.org/10.3390/foods10081709

Fonte, L., Díaz, M., Machado, R., Demedio, J., García, A., and Blanco, D. (2013). Caracterización físico-química y organoléptica de miel de *Melipona beecheii* obtenida en sistemas agroforestales. *Pastos y Forrajes*, *36*(3), 345-349.

Grajales-Conesa, J., Vandame, R., Santiesteban-Hernández, A., López-García, A., and Guzmán-Díaz, M. (2018). Propiedades fisicoquímicas y antibacterianas de mieles de abejas sin aguijón del Sur de Chiapas, México. *IBCIENCIAS*, *1*(1), 1-7. http://biociencias.unach.mx/ibciencias/articulos.html.

Gutiérrez, M. G., Enríquez, E., Lusco, L., Rodríguez-Malaver, A., Persano, O.-L., and Vit, P. (2008). Caracterización de mieles de Melipona beecheii y Melipona solani de Guatemala. *Revista de La Facultad de Farmacia*, *50*(1), 2-6.

Hailu, A., and Biratu, K. (2016). Determination of bee space and cell dimensions for Jimma zone honeybee Eco-Races (*Apis melifera*), Southwest Ethiopian. *Journal of Biology, Agriculture and Healthcare*, 6(9), 41-46. https://www.iiste.org/Journals/index.php/JBAH/article/view/30503.

ICONTEC (2007). *Norma Técnica Colombiana para miel de abejas* (*NCT-1273*). Instituto Colombiano de Normas Técnicas y Certificación. https://ecollection.icontec.org/normavw.aspx?ID=590

Kamboj, U., and Mishra, S. (2015). Prediction of adulteration in honey using rheological parameters. *International Journal of Food Properties*, *18*(9), 2056-2063. https://doi.org/10.1080/10942912.2014.962656

Lane, J. H., and Eynon, L. (1923). Determination of reducing sugars by means of Fehling's solution with methylene blue as internal indicator. *Journal of the Society of Chemical Industry*, *42*, 32-36.

Lóriga Peña, W., Álvarez-López, D., Fonte-Carballo, L., and Demedio-Lorenzo, J. (2015). Población inmadura y reservas de alimentos en colonias naturales de *Melipona beecheii* Bennett (Apidae : Meliponini) como factores básicos para su salud. *Revista de Salud Animal*, *37*(1), 47-51. http://scielo.sld.cu/pdf/rsa/v37n1/rsa07115.pdf.

Machado de Melo, A. A., de Almeida-Muradian, L. B., Sancho, M. T., and Pascual-Maté, A. (2018). Composition and properties of *Apis mellifera* honey: A review. *Journal of Apicultural Research*, *57*, 5-37. https://doi.org/10.1080/00218839.2017.1338444 REVIEW

May-Canché, I., Moguel-Ordoñez, Y., Valle-Mora, J., González-Cadenas, R., Toledo-Núñez, B., Arroyo-Rodríguez, L., Piana, L. and Vandame R. (2022). Sensory and physicochemical analysis of honeys of nine stingless bee species of Mexico and Guatemala. *Journal of Food Science and Technology 59*, 4772-4781. https://doi.org/10.1007/s13197-022-05561-7

May-Itzá, W., Peña, W. L., De la Rúa, P., and Quezada-Euán, J. J. G. (2019). A genetic and morphological survey to trace the origin of *Melipona beecheii* (Apidae: Meliponini) from Cuba. *Apidologie*, 50, 859-870. https://doi.org/10.1007/s13592-019-00696-7

Mohammed, M. E. A. (2020). Factors affecting the physicochemical properties and chemical composition of bee's honey. *Food Reviews International*, *38*, 1330-1341. https://doi.org/10.1080/87559129.2020.1810701

Moo-Huchin, V. M., González-Aguilar, G. A., Lira-Maas, J. D., Pérez-Pacheco, E., Estrada-Leon, R., Moo-Huchin, M., and Sauri-Duch, E. (2015). Physicochemical properties of *Melipona beecheii* honey of the Yucatan Peninsula. *Journal of Food Research*, *4*(5), 25-32. https://doi.org/10.5539/jfr./v4n5p25

Mora-Escobedo, R., Moguel-Ordóñez, Y., Jaramillo-Flores, M. E., and Gutiérrez-López, G. F. (2006). The composition, rheological and thermal properties of tajonal (Viguiera dentata) mexican honey. *International Journal of Food Properties*, *9*(2), 299-316. https://doi.org/10.1080/10942910600596159

NOM-004-SAG/GAN-2018 (2018). Norma Oficial Mexicana. Producción de Miel y Especificaciones. *Diario Oficial de la Federación*. *https://www.dof.gob.mx/nota_detalle.php?codigo=5592435&fecha=29/04/2020#gsc.tab=0*.

Nordin, A., Sainik, N. Q. A. V., Chowdhury, S. R., Saim, A. Bin, and Idrus, R. B. H. (2018). Physicochemical properties of stingless bee honey from around the globe: A comprehensive review. *Journal of Food Composition and Analysis*, *73*, 91-102. https://doi.org/10.1016/j.jfca.2018.06.002

Quezada-Euán, J. J. G. (2018). *Stingless bees of Mexico. The biology, management and conservation of an ancient heritage*. Springer. https://doi.org/10.1007/978-3-319-77785-6

Rodríguez-Suazo, G. E. (2014). *Caracterización física, química y microbiológica de la miel de Melipona beecheii* [Undergraduate thesis, Escuela Agrícola Panamericana, Honduras]. https://bdigital.zamorano.edu/items/b8165c8d-dd70-4a1d-a3f9-3932d2a63b50.

Schade, J. E., Marsh, G. L., and Eckert, J. E. (1958). Diastase activity and hydroxy-methyl-furfural in honey and their usefulness in detecting heat alteration. *Journal of Food Science*, *23*(5), 446-463.

Santos, E., Meerhoff, E., García da Rosa, E., Ferreira, J., Raucher, M., Quintana, W., Martínez, A., González, C., and Mancebo, Y. (2018). Color and electrical conductivity of honeys produced by *Apis mellifera* in Uruguay. *INNOTEC, 16*, 51-55. https://doi.org/10.26461/16.08

Smetanska, I., Alharthi, S., and Selim, K. (2021). Physicochemical, antioxidant capacity and color analysis of six honeys from different origin. *Journal of King Saud University – Science, 33*(5), 101447. https://doi.org/10.1016/j.jksus.2021.101447

Souza, B., Roubik, D., Barth, O., Heard, T., Enríquez, E., Carvalho, C., Villas-Bôas, J., Marchini, L., Locatelli, J., Persano-Oddo, L., Almeida-Muradian, L., Bogdanov, S., and Vit, P. (2006). Composition of stingless bee honey: Setting quality standards. *Interciencia*, *31*(12), 867-875.

Souza, E. C., Menezes, C., and Flach, A. (2021). Stingless bee honey (Hymenoptera, Apidae, Meliponini): A review of quality control, chemical profile, and biological potential. *Apidologie*, *52*(1), 113-132. https://doi.org/10.1007/s13592-020-00802-0

Umaña, E., Zamora, G., Aguilar, I., Arias, M. L., Pérez, R., Sánchez, L. A., Solórzano, R., and Herrera, E. (2021). Physicochemical differentiation of stingless bee honeys from Costa Rica. *Journal of Apicultural Research*, *2021*, 1903737. https://doi.org/10.1080/00218839.2021.1903737

Villanueva-G, R., Roubik, D. W., and Colli-Ucán, W. (2005). Extinction of *Melipona beecheii* and traditional beekeeping in the Yucatán peninsula. *Bee World*, *86*(2), 35-41. https://doi.org/10.1080/0005772X.2005.11099651

Vit, P. (2013). Modificaciones comentadas de la norma Miel de Abejas, hacia la norma Miel de Venezuela: inclusión de miel de pote y exclusión de mieles falsas. In P. Vit and D. W. Roubik (Eds.), *Stingless bees process honey and pollen in cerumen pots*. Universidad de Los Andes, Mérida, Venezuela (ch. 17). http://www.saber.ula.ve/handle/123456789/37277.

Vit, P., Enríquez, E., Barth, M. O., Matsuda, A. H., and Almeida-Muradian, L. B. (2006). Necesidad del control de calidad de la miel de abejas sin aguijón. *Revista de Facultad de Medicina, Universidad de Los Andes, 15*(2), 89-95.

Research Article / Chemical, Environmental, and Food Engineering http://doi.org/10.15446/ing.investig.105003

Identification of Eroded and Erosion Risk Areas Using Remote Sensing and GIS in the Quebrada Seca watershed

Identificación de áreas erosionadas y en riesgo de erosión mediante percepción remota y SIG en la microcuenca Quebrada Seca

Cristopher E. Camargo-Roa[1], Carlos E. Pacheco-Angulo[2], Sergio A. Monjardin-Armenta[3], Roberto López-Falcón[4], and Tatiana Gómez-Orgulloso[5]

ABSTRACT

The aim of this research was to identify eroded areas and areas at risk of erosion (EAER) as indicators of soil degradation by water erosion in a semiarid watershed of the Venezuelan Andes in 2017. To this effect, remote sensing techniques and geographic information systems (GIS) were used, focusing on spectral reflectance data from a satellite image, given the absence of continuous pluviographic information and data on soil properties in developing countries. This methodology involved estimating the potential water erosion risk (PWER) and mapping eroded and erosion risk areas (EAER) based on calculating the spectral Euclidean distance to bare soils and a remote sensing technique, which was selected via linear regression. Receiver operating characteristics (ROC) curves were determined to define classification thresholds, which were validated by means of a supervised classification and associated to PWER values. The main results indicate that EAER1 identified more eroded areas with bare soils (229,77 ha) as opposed to EAER2 (195,57 ha). Similarly, it was evident that the first alternative was more successful that the second (sum of the first three principal components). The PWER analysis, in addition to the erosion mapping developed and other data and criteria, such as minimum area size of interest, could help to consider necessary soil conservation measures.

Keywords: spectral Euclidean distance, vegetation indices, principal components analysis, maximum likelihood

RESUMEN

El objetivo de esta investigación fue identificar áreas erosionadas y en riesgo de erosión (AERE) como indicadores de degradación de suelos por erosión hídrica en una cuenca semiárida de los Andes venezolanos en el año 2017. Para ello, se emplearon técnicas de percepción remota y sistemas de información geográfica (SIG), enfocándose en los datos espectrales de reflectancia de una imagen satelital, dada la ausencia de información pluviográfica continua y datos de propiedades del suelo en países en vías de desarrollo. Esta metodología implicó la estimación del riesgo potencial de erosión hídrica (RPEH) y la generación de cartografía de áreas erosionadas y en riesgo (AEER) a partir del cálculo de distancia espectral euclidiana a suelos desnudos y de una técnica de percepción remota seleccionada mediante regresión lineal. Se determinaron curvas ROC (características operativas del receptor) para definir umbrales de clasificación, los cuales fueron validados mediante una clasificación supervisada y asociados a valores de RPEH. Los resultados principales indican que EAER1 identificó más áreas erosionadas con suelos desnudos (229,77 ha) a diferencia de EAER2 (195,57 ha). De igual modo, se evidenció que la primera alternativa tuvo mayores aciertos en contraste con la segunda (sumatoria de los tres primeros componentes principales). El análisis de RPEH, además de las cartografías de erosión desarrolladas y otros datos y criterios como el tamaño del área mínima de interés, podrían ayudar a considerar medidas necesarias en cuanto a conservación de suelos.

Palabras clave: distancia espectral euclidiana, índices de vegetación, análisis de componentes principales, máxima verosimilitud

Received: September 29th, 2022
Accepted: June 23th, 2023

Introduction

Water erosion, a soil degradation process also regarded as an environmental hazard (Mohammed *et al*., 2020; Duguma, 2022), is caused by precipitations falling on vulnerable bare terrain, which, as runoff across the slope, drag the soil along to finally deposit it in low areas or mire and obstruct bodies of water (Ávila and Ávila, 2015; Omuto and Vargas, 2019). This makes it the main soil degradation process, as it quantitatively and qualitatively affects the rootable volume of soils intended for agricultural production (Morales-Pavón *et al*., 2016) and contributes

[1] MSc in Watershed Management, Universidad de Los Andes, Venezuela. Doctoral student, Universidad de Los Andes, Venezuela. Affiliation: Assistant professor, Universidad de Los Andes, Venezuela. Email: ccamargoroa@gmail.com
[2] PhD in Geographic Information Technology, Universidad de Alcalá, España. Affiliation: Full Professor, Universidad de Los Andes, Venezuela. Email: carlosa@ula.ve
[3] PhD in Information Sciences, Universidad Autónoma de Sinaloa, México. Affiliation: Full professor, Universidad Autónoma de Sinaloa, Mexico. Email: sa.monjardin12@info.uas.edu.mx
[4] PhD in Agronomy, University of Georgia, United States. Affiliation: Full professor. Centro Interamericano de Desarrollo e Investigación Ambiental y Territorial, Universidad de Los Andes, Venezuela. Email: rlopez@ula.ve
[5] MSc in Environmental Management, Pontificia Universidad Javeriana, Colombia. Affiliation: Independent researcher, Universidad Distrital Francisco José de Caldas, Colombia. Email: tatiana.orgulloso@gmail.com

to the decline of many other essential ecosystem services (Chaudhary and Kumar, 2018; FAO, 2019).

A way to observe water erosion is through multispectral images captured via remote sensing. From this perspective, eroded soils are characterized by a spectral response similar to that of bare soils, *i.e.*, much more uniform than that of vegetation, which exhibits a flatter reflectivity curve (Chuvieco, 2016), thus indicating the existence of bare soils or parent material outcrops as an effective indicator of areas subjected to erosion (Beguería, 2006). In light of the above, remote sensing techniques and geographical information system (GIS) procedures become necessary. These allow obtaining the spatial and temporal distribution of the diverse factors involved, along with their classifications (Rosales-Rodríguez, 2021). Thereupon, it is also necessary to perform visual and statistical analyses in order to understand and validate the generated cartography, whose purpose is to obtain a more precise and reliable cartographical indicator.

Some remote sensing techniques focus on visual quality by trying to improve the location of the data for analysis, in such a way that the features of interest are more evident (*e.g.*, contrast expansion, color composition, and filtering) (Lillesand *et al*., 2015; Chuvieco, 2016). Other techniques aim to generate continuous variables, such as vegetation indices, which have proven to be efficient at evaluating soil degradation –with erosion among them (Ngandam *et al*., 2016)– by transforming them into Net Primary Productivity (NPP) values (Sartori *et al*., 2018), which has led them to be more frequently employed as quantitative indicators of ecosystem functioning (Orr *et al*., 2017). In the same way, they have been used to monitor vegetation in arid and semiarid lands (Najafi *et al*., 2020), as well as to generate the C-factor for models such as USLE (Universal Soil Loss Equation) (Meinen and Robinson, 2021).

Other techniques have also been created, such as i) soil indices (IS) used to estimate soil degradation types (Ngandam *et al*., 2016; Li and Chen, 2018); ii) principal components analysis (PCA) to discriminate types of bare soils or landslides (Romero *et al*., 2017; Basu *et al*., 2020); and iii) spectral mixture analysis (SMA) to map the C-factor or determine bare soils in order to construct vulnerability indicators (Demaría and Aguado, 2013).

In this sense, all these techniques may be used as steps prior to highlighting erosion and later allowing its semi-automatized classification in order to obtain a risk map (Duguma, 2022). These techniques also aim to overcome the tedious task of visually interpreting satellite images, as well as the considerable amount of time required to carry it out (Leal *et al*., 2018).

By making a special emphasis on erosion risk mapping, which usually indicates the relative probability of it taking place within a certain area as compared to others (Ganasri and Ramesh, 2016; Opeyemi *et al*., 2019), a distinction can be made between *potential*, defined as the maximum possible soil loss in the absence of vegetation cover and conservation practices (*i.e.*, only considering the interaction between the physical factors of the soil: soil erodibility, rain erosivity, and topography), and *actual*, which is determined based on the sum of the land cover/use factor and the previous ones (Plambeck, 2020). The former tends to be substantially higher than the latter (Drzewiecki *et al*., 2014).

Therefore, the importance of remote sensing techniques to map erosion and its risk has become evident. In the same way, these methods have proven to be valuable for generating a cartography of soil degradation in climate change studies. This has been demonstrated in global products, such as soil degradation assessment (GLADA) (Anderson and Johnson, 2016); in regional ones, with erosion risk modeling in Europe (Panagos *et al*., 2015); and in national ones, with Malawi's atlas of soil loss (Omuto and Vargas, 2019). In the same way, some of these techniques have been used to monitor changes in cover or NPP and have been proposed within the analysis of land degradation neutrality (LDN) (Orr *et al*., 2017), whose sustainable development goal (SDG) for 2030 is to tackle desertification, rehabilitate degraded land and soil, and strive towards a world with "neutral soil degradation" (UNGA, 2015).

Identifying areas that have been eroded or are at risk of erosion also allows government agencies (via cross-referencing with population density maps) to initiate rehabilitation and protection activities, to perform territory planning for environmentally sustainable socioeconomic development, and to determine areas that are susceptible to hillslope processes (Ngandam *et al*., 2016; Efiong *et al*., 2021). Additionally, every action taken to address soil degradation may simultaneously contribute to the objectives of the fight against climate change, to the preservation of biological diversity, and to the SDGs (Orr *et al*., 2017).

In light of the aforementioned ideas, the main objective of this study was to map areas that are eroded and at risk of erosion using remote sensing techniques and GIS procedures, given that many models developed for this type of estimation cannot be adequately executed because data are missing to complete their parameters, *e.g.*, soil erodibility, which requires properties such as structure and permeability, among others, which, in practice, are usually scarce or nonexistent for many parts of the world (Ávila and Ávila, 2015; United Nations, 2021). All results were obtained at a watershed located in a semiarid environment, a space where soils often stand out for their susceptibility to hydric erosion (Tsegaye *et al*., 2020).

Materials and methods

Study area: the studied watershed is located between the 72°26′43″ - 72°21′23″O and 7°58′29″ - 7°54′11″N. Its area is 37,31 km^2 (Figure 1), and its altitude oscillates between 253 and 1 622 m.

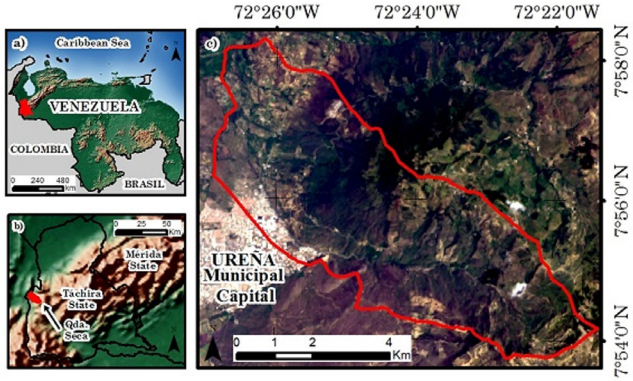


Figure 1. Location of the study area: a) national, b) regional, c) watershed
Source: Authors

Resources: the resources used for generating EAER and PWER are shown in Table 1.

Table 1. Resources

Data input	Denomination	Level of detail	Source
Satellite image	Landsat 8	SR = 30 m	USGS
Satellite image	Digital Globe	SR = 0,5 m	Digital Globe
Digital elevation model	ASTER GDEM v2*	SR = 30 m	USGS
Rain records		Monthly/ annual averages	MARNR - MinTIC
Soil physical properties	Texture and organic mater	Percentage	UNET Bio-environmental Lab
Soils distribution types map	Venezuelan environmental systems	1:250 000	MARNR
Roads	Roads	SR = 0,5 m	OSM

*ASTER Global Digital Elevation Model 2; SR: Spatial Resolution – USGS: US Geological Survey; MARNR: Ministry of the Environment and Renewable Natural Resources; MinTIC: Ministry of Information and Communication Technologies; OSM: Open Street Map.
Source: Authors

General methodology: a systematic scientific literature review was carried out to identify a deterministic PWER (potential water erosion risk) model, remote sensing techniques, and GIS procedures, with the purpose of identifying eroded areas and areas at risk of erosion (EAER). Four main phases were implemented: 1) developing the PWER model, 2) obtaining the EAER, 3) evaluating the degree of accuracy regarding the EAER with a supervised maximum likelihood classification (MLC) and a comparison of PWER and EAER, and 4) analyzing the results (Figure 2).

Landsat 8 OLI/TIRS image processing: a satellite image from October 1, 2017, was downloaded (LC08_L1TP_007055_20171001_20171013_01_T1_sr) with surface reflectance values and L1T corrections (IGAC, 2013), as well as topographic correction performed by Camargo *et al*. (2021).

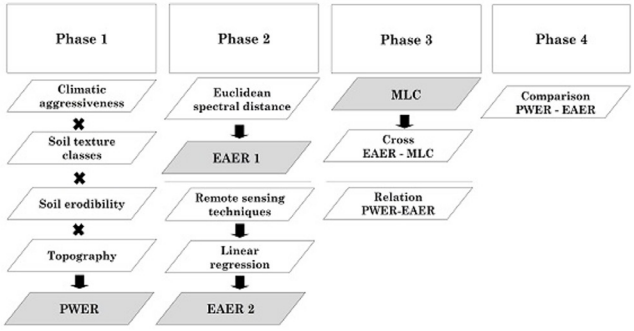


Figure 2. Phases of the methodology
Source: Authors

Digital Globe image: a natural color image from May 7, 2018, was employed. It had a spatial resolution of 0,5 m (SIGIS, 2019) and was supported by Google Earth, which allows its use in non-profit research (Thenkabail, 2016). Said image aided in the selection of 'ground-truth' samples in order to increase risk prediction reliability, based on a qualitative approach of erosion severity classes (Auerswald *et al*., 2018; Fischer *et al*., 2018; Batista *et al*., 2019).

PWER methodology

This methodology is based on two principles. The first one of a fundamental nature, expresses that it is the land that suffers the attack of the forces of climate (climatic aggressiveness) and that it, in turn, offers variable degrees of resistance, which poses a relation that determines the degrees of degradation in a given area. The second one indicates that potential risk assessment is most useful when relatively unstable or non-permanent factors (vegetation and land use) are not included in the calculation (FAO *et al*., 1980). Recent studies that use these principles include, among others, Guerra *et al*. (2020), Allafta and Oop (2021), and Al-Mamari *et al*. (2023). The PWER manifests via Equation (1):

$$PWER = f(C,S,T) \quad (1)$$

PWER is expressed in Mg ha^{-1}year^{-1}; C is the rain erosivity factor; S is the soil factor, estimated by means of texture (*st*) and erodibility (*se*) subfactors; and T is the topography factor (FAO *et al*., 1980; Rosales and García, 2015).

The climate factor (C) was evaluated based on the modification of the Fournier index (1960) (Arnoldus, 1977), which is best correlated with the EI30 value (maximum rain intensity in mm·hr^{-1} with 30 min duration), which has been verified in several parts of the world and is regarded as valid for Venezuela (Pacheco, 2012). The original Fournier index was not employed, as it does not consider that there are areas whose rainfall regime may have more than a monthly precipitation peak (Muñoz *et al*., 2014). Then, interpolated monthly precipitation surfaces were generated using IDW (inverse distance weighting) instead of kriging, as not all assumptions for its use are fulfilled (Hämmerly *et al*., 2019).

The S-factor was estimated based on *st* and *se* features associated with a soil type distribution map (MARNR, 1983) given the scarce availability of detailed soil data, which generally leads to considering reconnaissance studies (Quiñonez and Dal Pozzo, 2008). The texture classes of the first ones were conformed (USDA, 2020) and reclassified into three general categories recognized by FAO-UNESCO (1976), which in turn allowed assigning the valuations necessary for the model as per FAO *et al*. (1980). As for the second ones, the nomogram of the USLE K factor was employed (Foster *et al*., 1981), which is between 0 and 0,09 Mg ha h / ha MJ mm, using organic matter and texture percentages, later associated with erodibility values and classes and valuations as per FAO *et al*. (1980).

Finally, the T-factor was obtained from the ASTER GDEM, in which three slope classes were distinguished: a) flat to mildly undulating (0-8%); b) strongly undulating to hilly (8-30%); and c) strongly eroded to mountainous (>30%) (FAO-UNESCO, 1976), which were rated at (a) 0,35; (b) 3,5, and (c) 11,0 (FAO *et al*., 1980).

EAER methodology

Two mapping alternatives were developed: the first one was based on the spectral Euclidean distance between the reflectivity of each pixel in the satellite image and the bare soils category, and the second one on diverse remote sensing techniques, with which linear regressions were established. ROC (receiver operating characteristics) curves were applied to both products, which allowed defining classification thresholds and associated uncertainties aimed at detecting similar contiguous spectral zones of eroded and erosion risk areas (Beguería, 2006; Alatorre and Beguería, 2009). For thresholds, 100 independent samples (pixels) were selected, which evidenced <10% of vegetation cover, randomly distributed and defined with the help of the Digital Globe image, as the risk of erosion is considered to be high when this value is low (Wang *et al*., 2021).

An ROC curve is a graph that incorporates all sensitivity/specificity pairs resulting from the continuous variation of cutoff points throughout the range of observed results. This offers a global view of diagnostic accuracy by providing significant data on the probability of correctly classifying an individual by means of a determined variable (Ampudia *et al*., 2017). Their equations are:

$$sensitivity = \frac{a}{a+c} \tag{2}$$

$$specificity = \frac{d}{b+d} \tag{3}$$

where *a* are true positives, *b* true negatives, *c* false positives, and *d* false negatives. *Sensitivity* expresses the proportion of correctly predicted positive pixels, and *specificity* represents the proportion of correctly predicted negative pixels. Sensitivity and specificity values of 1 represent the likelihood of omission (type II, or false negative) and commission (type I, or false positive) errors (Alatorre and Beguería, 2009). In order to determine eroded areas, a sensitivity value of 0,9 was fixed, which corresponds to a 10% probability of omission errors. For areas at risk, a value of 0,8 was set (20% probability of omission error).

A supervised maximum likelihood classification (MLC) of the covers was carried out, as it is more precise than an unsupervised one because its classes are previously known (Liang and Wang, 2020). This, to later evaluate accuracy at determining bare soil cover with EAER by means of cross-tabulation. These results were updated with highways (OSM) and validated via global precision (Chuvieco, 2016) and *kappa* statistics (Cohen, 1960).

The remote sensing techniques implemented were vegetation indices based on slope and distance, soil indices (equations in Table 2), and others such as principal components analysis (PCA) (Pearson, 1901), spectral mixture analysis (SMA) (Boardman, 1992) and tasseled cap brightness (*B*) (Kauth and Thomas, 1976), with the coefficients derived by Baig (2014) for Landsat 8 Oli, as shown in Equation (4):

$$B = b2\times 0{,}3029 + b3\times 0{,}2786 + b4\times 0{,}4733 + b5\times 0{,}5599 + b6\times 0{,}508 + b7\times 0{,}1872 \tag{4}$$

Then, a linear regression analysis was carried out in order to determine the degree of dependence present between two variables (Shobha and Rangaswamy, 2018). To this effect, the bare soil spectral Euclidean distance was considered as an independent variable (x) and each technique as a dependent variable (y). Correlation tests were conducted in order to show the degree of the linear relation. Finally, all p-values generated were lower than 0,001, taking into account that values lower than 0,05 had to be considered in the analysis.

The resulting maps show three categories: (i) *eroded areas*, understood as those with no vegetation and denoting active erosion; (ii) *at risk areas*, those with little vegetation and prone to erosion; and (iii) *no erosion*, areas with good vegetation cover that 'apparently' protect against erosion.

Results

PWER

The behavior of the factors (Figure 3) allows defining a C-factor between 78,94 and 106,65. S was defined as fine (0,1) and medium (0,3) in terms of subfactor *st*, and as light (0,5) in terms of subfactor *se* (Table 3). T was between 0,35 and 11. Once the variables were defined, they were multiplied in map algebra in order to obtain PWER and later reclassify it to establish the soil erosion risk classes (Figure 4), thus obtaining 542,07 ha for 'none to light' (14,59%), 1 805,76 ha for 'moderate' (48,60%), and 1 767,37 ha for 'high' (36,80%).

Table 2. Indices

Author	Vegetation index (slope-based)
Jordan, 1969	$DVI = NIR - R$
Pearson and Miller, 1972	$RVI = R / NIR$
Rouse *et al*., 1973	$NDVI = \frac{NIR - R}{NIR + R}$
Baret and Guyot, 1991	$NRVI = \frac{RVI - 1}{RVI + 1}$
Deering *et al*., 1975	$TVI = \sqrt{NDVI} + 0,5$
Thiam, 1997	$TTVI = \sqrt{ABS \mid NDVI + 0,5 \mid}$
Perry and Lautenschlager, 1984	$CTVI = \frac{NDVI + 0,5}{ABS(NDVI + 0,5)} x \sqrt{ABS(NDVI + 0,5)}$
Huete, 1988	$SAVI = \frac{NIR - R}{NIR + R}(1 + L)$
Author	**Vegetation index (distance-based)**
Richardson and Wiegand, 1977. Rewritten by Jackson *et al*.,1983	$PVI_0 = \left(\frac{NIR - a.R - b}{\sqrt{a^2 + 1}} \right)$
Perry and Lautenschlager, 1984	$PVI_1 = \frac{b.NIR - R + a}{\sqrt{b^2 + 1}}$
Walther and Shabaani, 1991	$PVI_2 = \frac{(NIR - a)(R + b)}{\sqrt{a^2 + 1}}$
Qi *et al*., 1994	$PVI_3 = a.NIR - b.R$
Baret and Guyot, 1991	$\mathrm{ATSAVI} = \frac{a.NIR - a.R - b}{(a.NIR + R - a.b + 0,08(1 + a^2))}$
Author	**Soil index**
Nganfdam *et al*., 2016	$NDBSI = \frac{Swir1 - NIR}{Swir1 + NIR} + 0,001$
Celik, 2018	$BSI = \frac{(R + Swir1) - (NIR + B)}{(R + Swir1) + (NIR + B)}$
Li and Chen, 2018	$BI = f(Tasseled\ Cap(Brightness), NDBaI_2)$
FAO-UNESCO, 1976	$NDBaI2 = \frac{Swir1 - Tirs1}{Swir1 + Tirs1}$

Source: Authors

Table 3. Soil subfactors processing

Soil Texture	*st* subfactor		*se* subfactor	
	Texture classes[1] (FAO-UNESCO, 1976)	Valuation textural classes[2] (FAO et al., 1980)	K values[3] (Foster *et al.*, 1981)	Valuation erodibility classes[4] (FAO *et al.*, 1980)
Clay loam	Medium	0,3	0,028	0,5 (light)
Clay	Fine	0,1	0,013	0,5 (light)
Silty clay	Fine	0,1	0,022	0,5 (light)
Sandy loam	Medium	0,3	0,023	0,5 (light)
Sandy clay loam	Medium	0,3	0,013	0,5 (light)
Sandy loam	Medium	0,3	0,019	0,5 (light)

1 Soil texture classes: coarse: <18% clay & >65% sand; medium: <35% clay & < 65% sand 0 < 18% clay & <82% sand; fine: >35% clay
2 *st* Classification: coarse: 0,2 medium: 0,3 fine: 0,1 stony phase: 0,5
3 K-value: light: <0,03; moderate: 0,03-0,06; high: 0,06 >.
4 *se* valuation: light: 0,5; moderate: 1,0; high: 2,0.
Source: Authors

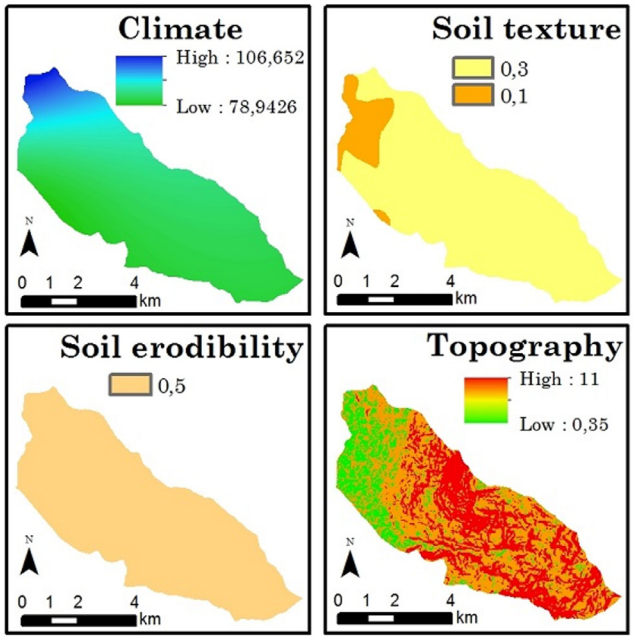


Figure 3. PWER factors and subfactors
Source: Authors

Maximum likelihood classification (*MLC*)

Six categories were established, and their areas were extracted: clouds (13,95 ha), highways (163,98 ha), high vegetation (1 331,10 ha), low vegetation (1 874,07 ha), bare soils (including rocky outcroppings) (277,74 ha), and infrastructure (69,36 ha) (Figure 5). The map had a global accuracy of 81,83% and a *kappa* index of 0,79.

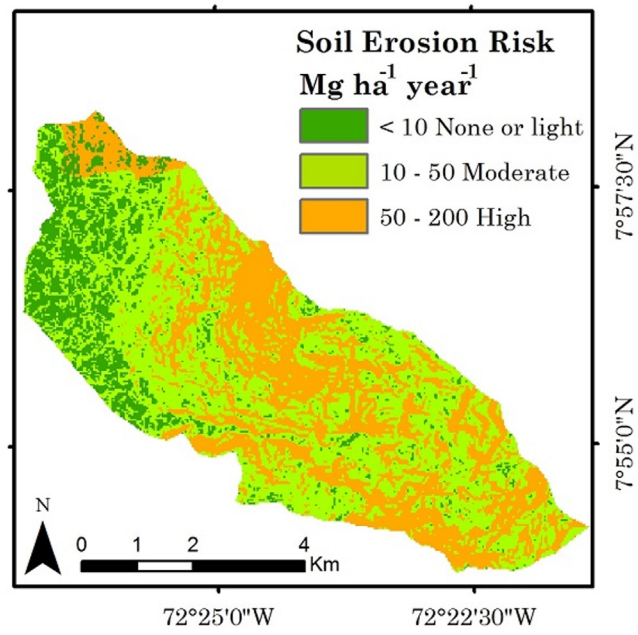


Figure 4. Soil erosion risk map
Source: Authors

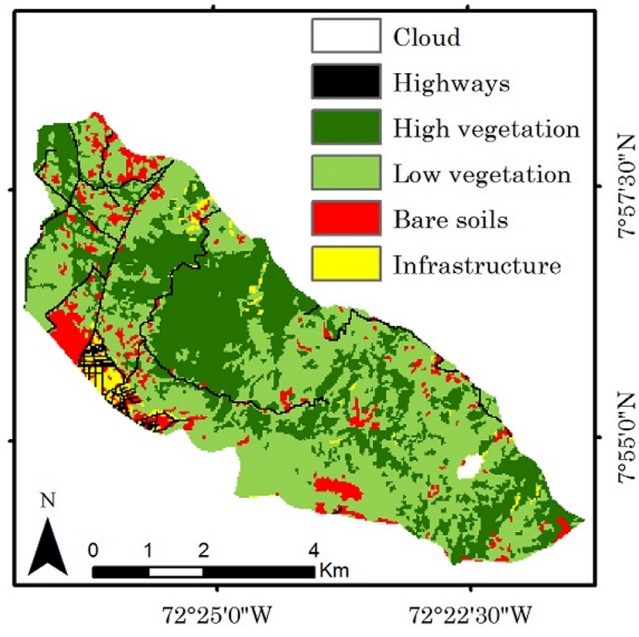


Figure 5. Maximum likelihood classification
Source: Authors

EAER1

The ROC curve showed a high sensitivity, *i.e.*, good capabilities to correctly classify positive pixels (bare soils). The 0,9 (U1) and 0,8 (U2) thresholds indicated spectral distances of 0,114 and 0,087. Both represent a balance between sensitivity and specificity. That is to say, an approach to the most ideal conditions (100%), which allowed generating the map (Figure 6) with the following results: eroded areas: 379,44 ha; erosion risk: 260,91 ha; no erosion: 2 911,95 ha (Figure 6).

Cross-referencing with MLC indicated that the eroded areas coincided with 229,77 ha of bare soils. In the same way, this category intercepted 'high vegetation' in 0,99 ha, 'low vegetation' in 126,99 ha, and 'infrastructure' in 21,69 ha. For its part, the erosion risk class mainly intercepted 'low vegetation' (236,52 ha), 'bare soils' (20,61 ha), 'infrastructure' (3,33 ha), and 'high vegetation' (0,45 ha) (Table 4).

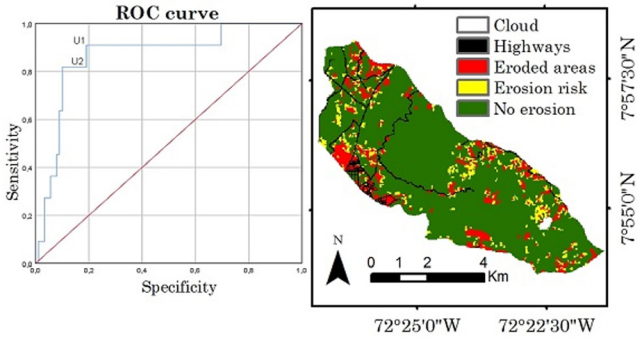


Figure 6. EAER1
Source: Authors

Table 4. Cross tabulation: EAER1-MLC (ha)

		EAER1		
		Eroded areas	Erosion risk	No erosion
MLC	High vegetation	0,99	0,45	1 329,66
	Low vegetation	126,99	236,52	1 510,56
	Bare soils	229,77	20,61	27,36
	Infrastructure	21,69	3,33	44,37

Source: Authors

EAER2

Once the products of the remote sensing techniques were derived, linear regressions were carried out. The analysis showed that the slope value was different from 0 in all cases, *i.e.*, there is a dependence between the variables. Under this criterion, SAVI is the most dependent (2,75), as it has the most inclined line, with a direct (positive) dependence, followed by NDVI (1,83) (Table 5).

By only analyzing the correlation coefficients (*R*), the highest results were obtained by SAVI (0,592), NDVI (0,591), TTVI (0,579), CTVI (0,566), and PCA_1 (0,542). Nevertheless, all of them were surpassed by the PCA sum, which achieved a value of 0,648. As for the R^2, it was found that only the PVI_0 showed a value equal to 0, which means that it does not explain any variation in this index as a function of the independent variable. It is closely followed by PVI_2 and DVI, which may be explained by their low correlations.

Among all the implemented techniques, NDBIa2 is the one that is most determined by the spectral Euclidean distance of the soil, with an R^2 of 0,551. However, it also showed the highest negative correlation (-0,742). Similar situations were reported by BSI, BI, and ATSAVI. On the contrary, the technique with the best R^2 was the PCA sum (0,420), as it showed a positive correlation of 0,648 (the highest one), which makes it the second EAER alternative.

Then, an ROC curve was elaborated based on 100 samples of PCA sum values. The 0,9 (U1) and 0,8 (U2) thresholds indicated values of -0,117 and -0,107 for the PCA sum, respectively, which represent a good balance between sensitivity and specificity, thus allowing to generate the map (Figure 7) with the following results: eroded areas: (261,18 ha); erosion risk: 25,83 ha; no erosion: 3 265,29 ha.

Cross-referencing with MLC indicated that the surface occupied by eroded areas coincided with 195,57 ha of bare soils and intercepted 'high vegetation' in 0,27 ha, 'low vegetation' in 31,95 ha, and 'infrastructure' in 33,39 ha. For its part, 'erosion risk' intercepted 'low vegetation' in 15,84 ha, 'bare soils' in 9,9 ha, and 'infrastructure' in 0,09 ha (Table 6).

Table 5. Linear regressions

Technique	R	R^2	Equation
PCA sum	0,648	0,42	$y = 0{,}44 + 1{,}64X$
SAVI	0,592	0,351	$y = 0{,}6 + 2{,}75X$
NDVI	0,591	0,35	$y = 0{,}40 + 1{,}83X$
TTVI	0,579	0,335	$y = 0{,}95 + 0{,}87X$
CTVI	0,566	0,321	$y = 0{,}95 + 0{,}86X$
PCA_1	0,542	0,294	$y = 0{,}33 + 0{,}83X$
TVI	0,511	0,261	$y = 1{,}15 + 1{,}13X$
PCA_2	0,494	0,244	$y = 0{,}09 + 0{,}71X$
PVI1	0,46	0,211	$y = 0{,}19 + 0{,}33X$
PCA_3	0,294	0,087	$y = 0{,}02 + 0{,}11X$
PVI3	0,277	0,077	$y = 0{,}03 + 0{,}09X$
DVI	0,221	0,049	$y = 0{,}17 + 0{,}27X$
PVI_0	-0,017	0	$y = -0{,}04 - 0{,}02X$
PVI_2	-0,082	0,007	$y = 0{,}02 - 0{,}33X$
RVI	-0,547	0,299	$y = 0{,}41 - 1{,}43X$
Tasseled cap	-0,552	0,306	$y = -0{,}45 - 0{,}85X$
NRVI	-0,595	0,354	$y = -0{,}40 - 1{,}84X$
NDBSI	-0,598	0,358	$y = 0{,}1 - 1{,}64X$
ALME	-0,642	0,413	$y = -0{,}83 - 3{,}55X$
ATSAVI	-0,675	0,456	$y = -1{,}04 - 6{,}03X$
BI	-0,68	0,462	$y = 8{,}05E4 - 8{,}4E4X$
BSI	-0,703	0,494	$y = 0{,}12 - 5{,}72X$
NDBIa2	-0,742	0,551	$y = 4{,}83E4 - 4{,}67E4X$

y= dependent variable; *x*= independent variable
Source: Authors

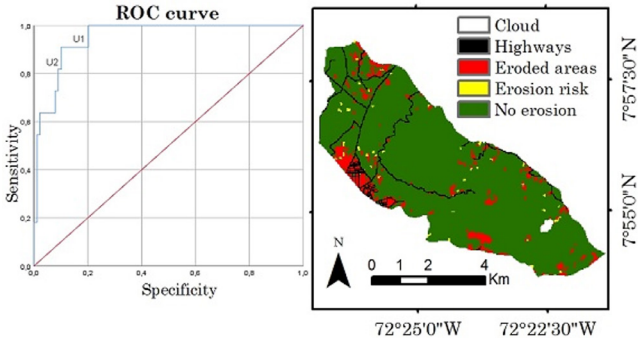


Figure 7. EAER2
Source: Authors

Table 6. Cross-tabulation: EAER2 and MLC in ha

		EAER2		
		Eroded areas	Erosion risk	No erosion
MLC	High vegetation	0,27	0	1 330,83
	Low vegetation	31,95	15,84	1 826,28
	Bare soils	195,57	9,9	72,27
	Infrastructure	33,39	0,09	35,91

Source: Authors

With the purpose of understanding the PWER as an indicator of erosion sensitivity, 14 bare soil pixels (samples) were first analyzed (Figure 8) based on the relationship between the thematic categories defined for the EAER, the classification by FAO *et al*. (1980), and the PWER values (Table 7). It must be highlighted that, given that the samples of bare and erosion-exposed soils were identified based on a higher-resolution image, they should only be classified as eroded or erosion risk areas in the EAER ('no erosion' would be an imprecision generated by the change in spatial resolution or threshold definition). Secondly, the possible soil loss quantities were totalized for eroded and erosion risk areas, with the purpose of evidencing the approach with the most losses.

Table 7. Comparison of bare soil samples with PWER-EAER maps

Sample	Long. / Lat.	EAER1	EAER2	PWER Mg ha^{-1}year^{-1}	FAO *et al.* (1980)
1	-72,442; 7,959	Ea	Ne	1,71	Light
2	-72,436; 7,948	Er	Ne	1,53	Light
3	-72,384; 7,930	Ea	Ne	44,34	Moderate
4	-72,382; 7,924	Ea	Ea	43,94	Moderate
5	-72,421; 7,918	Ne	Ea	132,61	High
6	-72,393; 7,906	Ea	Ea	134,66	High
7	-72,394; 7,907	Ea	Ea	134,64	High
8	-72,359; 7,902	Ea	Ea	42,82	Moderate
9	-72,360; 7,900	Ea	Ea	42,77	Moderate
10	-72,371; 7,898	Ea	Ea	134,46	High
11	-72,368; 7,917	Ea	Ea	43,48	Moderate
12	-72,404; 7,921	Ea	Ea	43,33	Moderate
13	-72,426; 7,939	Ea	Ea	44,20	Moderate
14	-72,434; 7,966	Ea	Ea	52,98	High

Source: Authors

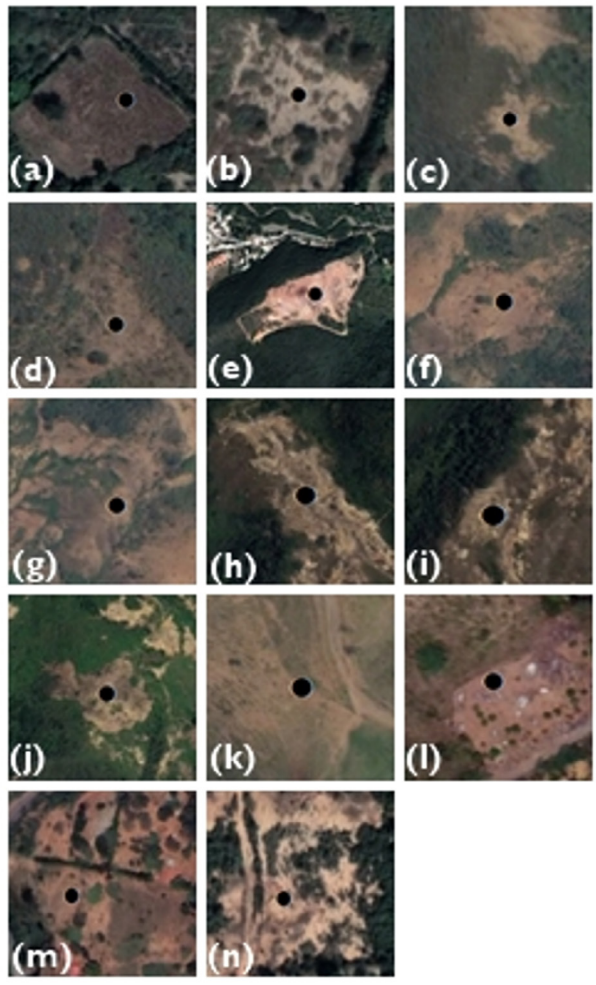


Figure 8. Eroded soil samples
Source: Authors

The results allowed defining a higher success rate in EAER1. However, an imprecision was found in sample 5, which was considered to belong to 'no erosion' (Figure 8e), as it is completely uncovered. This was the only incongruence generated in EAER1, since the correct category should have been 'eroded areas', *i.e.*, spectral distance values lower than 0,087. For its part, EAER2 showed three contradictions when it categorized samples 1, 2, and 3 as 'no erosion', even when they were uncovered (Figures 8a, 8b, and 8c). These pixels should have been lower than the PCA sum value of -0,107 (incongruences refer to values lower than those established by the thresholds in the ROC curves).

Considering the high success rate of the samples in both EAER, the PWER condition of the 14 points could be observed. In the case of sample 5 in EAER1, potential degradation is considered to be high, with 132,61 Mg ha^{-1}year^{-1}. On the other hand, sample 3 in EAER2 showed a loss of 44,34 Mg ha^{-1}year^{-1}, which represents a moderate degradation (the remaining samples can be interpreted in the same way). Finally, the soil losses from eroded and erosion risk areas in the first map were much greater than those of the second (by 162,88%) (Table 8).

Table 8. Comparison EAER1 and EAER2

Class	Mg ha^{-1}year^{-1}	Class	Mg ha^{-1}year^{-1}
Ea 1	219 045,65	**Ea 2**	136 147,80
Er 1	181 670,42	**Er 2**	16 284,01
Total	**400 716,07**	**Total**	**152 431,81**

Source: Authors

Discussion

Even though the PWER methodology dates back to 1980, it is still valid because its application is ideal when there is no adequate information on precipitation, evapotranspiration, lithology, and soils, which is required by methodologies such as USLE or the one proposed by Pacheco (2012) or more recently by Nasir *et al*. (2023). On the contrary, the need for minimal information such as monthly precipitations, free MDE, and reconnaissance-level soil mapping provides this methodology with execution plausibility, simplicity, and scientific robustness. It is also important to note that many of the models currently used require the inclusion of impractical variables (Demirel and Tüzün, 2011), which highlights the need for new approaches in the modeling of erosion processes (Nasir *et al*., 2023).

A case that illustrates the above is the comparison between the potential erosion results obtained in this study and those of Condori-Tintaya *et al*. (2022) and Nasir *et al*. (2023). While the former used soil structure information and calculated a more complex length and slope factor based on equations, the latter employed a multi-criteria decision making (MCDM) approach involving the development of 15 variables and an analytical hierarchy process (AHP).

Access to GIS and satellite images has had great impact on erosion modeling. Models can now be applied with relative ease at a large scale and in a distributed fashion, and they can present results in pixels that allow identifying where they occur and their magnitude, as well as at different temporal and spatial scales (Batista *et al*., 2019). However, it is also recognized that the predictive capability of large-scale erosion models is not the best. Therefore, Alewell *et al*. (2019) have argued that the same should not strive for them to make accurate predictions of soil losses, but instead to explore scenarios and focus on understanding relative differences in erosion rates, which would help to identify areas prone to these degradation processes, which is the objective of this study.

In the same way, very high-resolution images may be used to test erosion models, which has not been widely performed by researchers, with Fischer *et al*. (2018) perhaps being the first to focus completely on its interpretation. They found promising results, such as a high correlation (R^2= 0,91) of visually defined erosion classes with moderate soil losses, thus allowing them to define a semi-quantitative assessment approach, which is much simpler than proposing hypothesis tests (Batista *et al*., 2019).

Other reasons to support the use of this methodology is that it excludes speculations on the validity of the models' predictions, and it allows identifying scenarios leading to great or small soil losses (Auerswald *et al*., 2018). This study did not employ a semi-quantitative approach because the real soil loss values of the samples selected were not calculated independently, so as to be able to determine the R^2, which denotes a predominantly qualitative assessment to address comparative analyses.

The key to proceeding with the identification of EAER lies in the accurate extraction of vegetation cover and land use (Wang *et al*., 2013). In this case, it involved aiming for the highest possible accuracy regarding bare soils, with which the spectral Euclidean distance was obtained to later define thresholds in ROC curves, since the presence of vegetation cover reflects resistance capabilities against erosion or its risk (Wang *et al*., 2021).

Starting with 'eroded areas' and its intersection with bare soils (379,44 ha = 100%), cross-tabulation between MLC and the EAER allowed corroborating that there was a higher intersection in EAER1 (60,55%) in comparison with EAER2, which coincided by 51,54%. This result evinces the former's greater reliability, which is very likely due to the not-so-high correlation and determination coefficients of the PCA sum with respect to the spectral Euclidean distance.

As for the cross-reference between 'eroded areas' and 'low and high vegetation', EAER1 showed a 3,43% coincidence, unlike EAER2, which reported 0,87%. The intersection between 'eroded areas' and 'infrastructure' was lower in EAER1, with 0,58%, in contrast with the EAER2's 0,90%. These results are also explained by the differentiation inability shown by the latter, which is due to the not-so-high correlation and determination of the PCA sum.

These results position EAER1 as a baseline alternative in contrast with the PCA sum. This conclusion is reasonable, as this input was used as an independent variable in linear regressions.

As for the selection of thresholds in ROC curves, given the varying sensitivity/specificity, other thresholds with more or less discrimination power could be selected. Selecting two with a high sensitivity would reciprocally allow achieving a high probability of correctly classifying a pixel whose real situation is defined as positive. The advantage of the ROC curve lies in the fact that it uses all possible cutoff points in the database, through which better thresholds are determined, thus indicating that this test has a very useful and correct classification power (Bernui *et al*., 2022).

Therefore, the proposed method for identifying EAER differs from associating higher loss values (Mg ha^{-1} year^{-1}) yielded by USLE and/or similar models as an indirect measure of diverse erosion risk degrees (*e.g.*, Chaudhary and Kumar, 2018; Mohammed *et al*., 2020), which is supported by the assumption that low values are less vulnerable and high

values refer to a greater erosion sensitivity (Meshesha *et al*., 2012). This also contrasts with the idea of obtaining them by ranking and later zoning, which results from intersecting diverse slope gradients from an MDE or land uses with soil losses (Meshesha *et al*., 2012; Khosrokhani and Pradhan, 2013; Wang *et al*., 2013).

The equivalence between remote sensing techniques obtained by linear regression was manifested by Ngandam *et al*. (2016), who proposed a vegetation index as the independent variable, which placed an exaggerated reliability in said product as a predictor. This, unlike the design implemented in this study, which was defined based on the spectral distance of bare soils.

The comparison between the qualitative and the quantitative results allows establishing a bidirectional complementary analysis relationship. In the first place, EAER helped to identify areas with erosion or highlight those at risk, in which soil losses can be interpreted. Secondly, it allows establishing the inverse relationship, *i.e.*, the observation of possible high and moderate losses identified by the PWER provides information on what happens in said areas from a qualitative perspective.

For a wider understanding and for determining the possible existence of erosion or of areas that might be subjected to it, it is suggested that eroded and erosion risk areas be identified with both mappings. This consideration is supported by the identification of possible substantial differences at potential points of interest.

Finally, based on the study by Batista *et al*. (2019) regarding the fact that the current, different erosion models do not systematically surpass each other, this research agrees that calibration is the only mechanism to improve their performance (*i.e.*, a better conceptual understanding of their operation). Therefore, this study rejects the notion that these can be validated (rather than evaluated), stressing the need to define adjustment tests (or evaluating degrees of reliability) based on multiple data sources, which allow for a wide study of the usefulness and consistency of the developed methodologies. This, considering that the more thorough the tests, the more likely it is that deficient performances are found (critical awareness of the methods).

Conclusions

Using remote sensing images allows conducting research on erosion both qualitatively and quantitatively. However, these must be used with care, as their unthinking use may lead to over- or underestimation. The use of very high-resolution images must also be considered as a mechanism to evaluate model performance.

Resorting to higher resolution imagery or covering larger extensions implies a higher computational cost because it involves larger amounts of data. This could be overcome by using cloud computing platforms such as Google Earth Engine (GEE), which can compute and process large volumes of geospatial data in very short time intervals and have been recently applied to the preparation of the variables required by methods such as USLE and RUSLE (Papaiordanidis *et al*., 2019; Kumar *et al*., 2022), so their application with the models developed in this study is also plausible.

Although technological advances are evident, it should not be forgotten that erosion models are not necessarily true or free of apparent flaws; recognizing them improves the attitude towards evaluating them and changes the way their performances are characterized and communicated, ultimately leading to a better understanding of soil erosion (Batista *et al*., 2019).

This study constitutes a contribution to the lack of precipitation and soil data, which is necessary in parametric methodologies. Therefore, it adopted the premise that this lack should be only resolved with free and available digital inputs that help identify potential erosion hot spot areas and simulate erosion responses to land use and climate change, which makes it a solution that could be associated with variables derived from MDE (*e.g.*, humidity indices) or other kinds of categorical or continuous mapping. Nevertheless, regardless of the methods to be employed, an on-field survey of geo-referenced measurements (when possible) regarding erosion characteristics must not be discarded, with the purpose of broadening the model's evaluation.

Even though it is true that the methods developed for identifying EAER are metric-static and implemented with satellite images from diverse dates, these could help to obtain time series allowing to better understand the dynamics of erosion processes and therefore acquire greater knowledge for soil conservation and ecosystem management. They could also be replicated in other spaces in a semi-automatized fashion, and they could serve as first inputs to define areas where observations should be focused or to map types and degrees of erosion.

Acknowledgements

To the doctoral program in Forest and Environmental Sciences (CEFAP – FCFA – ULA, Mérida, Venezuela), to which this document is linked. To the UNET Bio-Environmental Laboratory for their support. We would like to thank the Journal's Reviewers and Editorial Committee for their contributions and observations.

CRediT author statement

C. E. C-R. conceived the research, did the background research, collected the data, developed the workflow, and performed the assessment. C. E. P-A., R. L-F., S. A. M-A., and T. G-O., supervised the research and provided critical feedback. All authors contributed to writing the manuscript and approved its definitive version for publication.

References

Alatorre, L. C., and Beguería, S. (2009). Identification of eroded areas using remote sensing in a badlands landscape on marls in the central Spanish Pyrenees. *Catena*, *76*(3), 182-190. https://doi.org/10.1016/j.catena.2008.11.005

Anderson, W., and Jhonson, T. (2016). Evaluating global land degradation using ground-based measurements and remote sensing. In E. Nkonya, V. Mirzabaev, A., and J. Von Braun (Eds.), *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development* (85-116). Springer Open. https://doi.org/10.1007/978-3-319-19168-3

Al-Mamari, M., Kantoush, S., Al-Harrasi, T., Al-Maktoumi, A., Abdrabo, K., Saber, M., and Sumi, T. (2023). Assessment of sediment yield and deposition in a dry reservoir using field observations, RUSLE and remote sensing: Wadi Assarin, Oman. J*ournal of Hydrology, 617*, 128982. https://doi.org/10.1016/j.jhydrol.2022.128982

Allafta, H., and Opp, C. (2022). Soil erosion assessment using the RUSLE model, remote sensing, and GIS in the Shatt Al-Arab Basin (Iraq-Iran). *Applied Sciences*, *12*, 7776. https://doi.org/10.3390/app12157776

Alewell, C., Borrelli, P., Meusburger, K., and Panagos, P. (2019). Using the USLE: chances, challenges and limitations of soil erosion modelling. *International Soil and Water Conservation Research*, *7*, 203-225. https://doi.org/10.1016/j.iswcr.2019.05.004

Ampudia, A. Sánchez, G., and Jiménez, F. (2017) Precisión diagnóstica del MMPI-2 con la personalidad delictiva: un análisis con la curva ROC. *Revista de Psicología*, *35*(1), 167-192. http://dx.doi.org/http://doi.org/10.18800/psico.201701.006

Arnoldus, H. (1977). Methodology used to determine the maximum potential average annual soil loss due to sheet and rill erosion in Morocco. Assessing Soil Degradation. *FAO Soils Bulletin, 34*, 39-48.

Auerswald, K., Fischer, F.K., Kistler, M., Treisch, M., Maier, H., and Brandhuber, R. (2018). Behavior of farmers in regard to erosion by water as reflected by their farming practices. *Science of the Total Environment*, *613-614*, 1-9. https://doi.org/10.1016/j.scitotenv.2017.09.003

Ávila, B. D., and Ávila H. F. (2015). Spatial and temporal estimation of the erosivity factor r based on daily rainfall data for the department of Atlántico, Colombia. *Ingeniería e Investigación*, *35*(2), 23-29. http://dx.doi.org/10.15446/ing.investig.v35n2.47773

Baig, M. H. A., Zhang, L., Shuai, T., and Tong, Q. (2014). Derivation of a Tasselled Cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, *5*(5), 423-431. https://doi.org/10.1080/2150704X.2014.915434

Baret, F., and Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment, 35*(2-3), 161-173. https://doi.org/10.1016/0034-4257(91)90009-U

Basu T., Das A., and Pal, S. (2020). Application of geographically weighted principal component analysis and fuzzy approach for unsupervised landslide susceptibility mapping on Gish River Basin, India. *Geocarto International*, *37*(5), 1294-1317. https://doi.org/10.1080/10106049.2020.1778105

Batista, P. V. G., Davies, J., Silva, M.L.N., and Quinton, J. N. (2019). On the evaluation of soil erosion models: Are we doing enough? *Earth-Science Reviews*, *197*, 102898. https://doi.org/10.1016/j.earscirev.2019.102898

Beguería S. (2006). Identifying erosion areas at basin scale using remote sensing data and GIS: A case study in a geologically complex mountain basin in the Spanish Pyrenees. *International Journal of Remote Sensing*, *27*(20), 4585-4598. https://doi.org/10.1080/01431160600735640

Bernui, G., Del Aguila, L., Sanes, M., Prochazka, R., and Bussalleu, A. (2022). Evaluación de un test del aliento con carbono 13 para el diagnóstico de Helicobacter pylori. *Rev. de Gastroenterología del Perú*, *42*(1), 1341. http://dx.doi.org/10.47892/rgp.2022.421.1341

Boardman, J. W. (1992). *Sedimentary facies analysis using imaging spectrometry: a geophysical inverse problem* [Doctoral dissertation, University of Colorado Boulder]. ProQuest.

Camargo, C. E., Pacheco, C. E., and López, R. (2021). Evaluación de métodos de corrección atmosférica y sombreado topográfico en imagen Landsat 8 OLI sobre un área montañosa semiárida. *UD y la Geomática*, *16*, 23-39. https://doi.org/10.14483/23448407.17040

Celik, N. (2018). *Change detection of urban areas in Ankara through Google Earth Engine* [Conference presentation]. 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece. https://doi.org/10.1109/TSP.2018.8441377

Chaudhary, B. and Kumar, S. (2018). Soil erosion estimation and prioritization of Koshalya-Jhajhara watershed in North India. *Indian Journal of Soil Conservation*, *46*(3), 305-311.

Chuvieco, E. (2016). *Fundamentals of satellite remote sensing: An environmental approach* (2nd ed.). CRC Press Taylor and Francis Group. https://doi.org/10.1111/phor.12184

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Condori-Tintaya, F., Pino-Vargas, E., and Tacora-Villegas. (2022). Pérdida de suelos por erosión hídrica en laderas semiáridas de la subcuenca Cairani-Camilaca, Perú. *Idesia*, *40*(2), 7-15. https://dx.doi.org/10.4067/S0718-34292022000200007

Deering, D. W., Rouse, J. W., Iiaas, R. H., and Schell, J. A. (1975). *Measuring forage production of grazing units from landsat MSS data* [Conference presentation]. Proceedings of the Tenth International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, MI, USA.

Demaría, M. R., and Aguado, I. (2013). Dinámica espacio-temporal del porcentaje de suelo desnudo en pastizales semiáridos de Argentina. *GeoFocus, 13*(2), 133-157.

Demirel, T., and Tüzün, S. (2011). *Multi criteria evaluation of the methods for preventing soil erosion using fuzzy ANP: The case of Turkey*. [Conference presentation]. World Congress on Engineering, London, England.

Drzewiecki, W., Wężyk, P., Pierzchalski, M., and Szafrańska, B. (2014). Quantitative and qualitative assessment of soil erosion risk in Małopolska (Poland), supported by an object-based analysis of high-resolution satellite images. *Pure and Applied Geophysics , 171*, 867-895. https://doi.org/10.1007/s00024-013-0669-7

Duguma, T. A. (2022). Soil erosion risk assessment and treatment priority classification: A case study on Guder Watersheds, Abay River Basin, Oromia, Ethiopia. *HELIYON*, *2022*, e10183. https://doi.org/10.1016/j.heliyon.2022.e10183

Efiong, J., Imoke, D., Nwabueze, J., and James, S. (2021). Geospatial modelling of landslide susceptibility in Cross River State of Nigeria. *Scientific African*, *14*, e01032. https://doi.org/10.1016/j.sciaf.2021.e01032

Food and Agriculture Organization of the United Nations (FAO) (Eds.) (2019). *Soil erosion: The greatest challenge to sustainable soil management*. FAO.

Food and Agriculture Organization of the United Nations (FAO), Programa de Naciones Unidas para el Medio Ambiente (PNUMA), United Nations Educational, Scientific, and Cultural Organization (UNESCO) (1980). *Metodología provisional para la evaluación de la degradación de los suelos*. FAO.

Food and Agriculture Organization of the United Nations (FAO), United Nations Educational, Scientific and Cultural Organization (UNESCO) (1976). *Mapa mundial de suelos 1:5.000.0000*. FAO.

Fischer, F. K., Kistler, M., Brandhuber, R., Maier, H., Treisch, M., and Auerswald, K. (2018). Validation of official erosion modelling based on high-resolution radar rain data by aerial photo erosion classification. *Earth Surface Processes and Landforms*, *43*, 187-194. https://doi.org/10.1002/esp.4216

Foster, G. R., Mccool, D. K., Renard, K. G., and Moldenhauer, W.C. (1981). Conversion of the universal soil loss equation to SI metric units. *Journal of Soil and Water Conservation*, *36*(6), 355-359.

Ganasri, B. P., and Ramesh, H. (2016). Assessment of soil erosion by RUSLE model using remote sensing and GIS – A case study of Nethravathi Basin. *Geoscience Frontiers*, *7*, 953-961. http://dx.doi.org/10.1016/j.gsf.2015.10.007

Guerra, C. A., Rosa, I. M. D., Valentini, E., Wolf, F., Filipponi, F., Karger, D. N., Nguyen Xuan, A., Mathieu, J., Lavelle, P., and Eisenhauer (2020). Global vulnerability of soil ecosystems to erosion. *Landscape Ecology,* 35, 823-842. https://doi.org/10.1007/s10980-020-00984-z

Hämmerly, R. C., Paris, R. C., and Paz-González, A. (2019). Assessment of domain areas for precipitation and evapotranspiration on the left bank of the Paraná watershed at Argentine territory. I: Thiessen polygons and kriging. *Cadernos Lab. Xeolóxico de Laxe Coruña*, *41*, 75-97. https://doi.org/10.17979/cadlaxe.2019.41.1.5818

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295-309. https://doi.org/10.1016/0034-4257(88)90106-X

Instituto Geográfico Agustín Codazzi (IGAC) (2013). *Descripción y corrección de productos Landsat 8 LDCM (Landsat Data Continuity Mission), versión 1.0*. Instituto Geográfico Agustín Codazzi.

Jackson, R. D., Slater, P. N., and Pinter, P. 1983. Discrimination of growth and water stress in wheat by varius vegetation indices through clear and a turbid atmospheres. *Remote Sensing of Enviroment 13*, 187-208. https://doi.org/10.1016/0034-4257(83)90039-1

Jordan, C. F. (1969). Deviation of leaf-area index from quality of light on the forest floor. *Ecology*, *50*(4), 663-666. https://doi.org/10.2307/1936256

Kauth, R. J., and Thomas, G. S. (1976). *The Tasseled Cap – A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat* [Conference presentation]. Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, USA.

Khosrokhani, M., and Pradhan, B. (2013). Spatio-temporal assessment of soil erosion at Kuala Lumpur metropolitan city using remote sensing data and GIS. *Geomatics, Natural Hazards and Risk*, *5*(3), 252-270. http://dx.doi.org/10.1080/19475705.2013.794164

Kumar, R., Deshmukh, B., and Kumar, A. (2022). Using Google Earth Engine and GIS for basin scale soil erosion risk assessment: A case study of Chambal river basin, central India. *Journal of Earth System Science*, *131*, 228. https://doi.org/10.1007/s12040-022-01977-z

Leal, J., Pérez, U., and Ortiz, N. E. (2018). Distribución espacial y temporal de deslizamientos (1999 – 2015) en la cuenca del río Combeima, Colombia. *Revista Geográfica Venezolana*, *59*(2), 346 -365.

Li, S., and Chen, X. (2018). New bare-soil index for rapid mapping developing areas using Landsat 8 data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40*(4), 139. https://doi.org/10.5194/isprsarchives-XL-4-139-2014

Liang, S., and Wang, J. (2020). *Advanced remote sensing: Terrestrial information extraction and applications* (2nd ed.). Academic Press. https://doi.org/10.1016/C2017-0-03489-4

Lillesand, T., Kiefer, R. W., and Chipman, J. (2015). *Remote sensing and image interpretation* (7th ed). John Wiley & Sons.

Meinen, B. U., and Robinson, D. T. (2021). From hillslopes to watersheds: Variability in model outcomes with the USLE. *Environmental Modelling and Software*, *146*, 105229. https://doi.org/10.1016/j.envsoft.2021.105229

Meshesha, D. T., Tsunekawa, A., Tsubo, M., and Haregeweyn, N. (2012). Dynamics and hotspots of soil erosion and management scenarios of the Central Rift Valley of Ethiopia. *International Journal of Sediment Research*, *27*(1), 84-99. https://doi.org/10.1016/S1001-6279(12)60018-3

Ministerio del Ambiente y de los Recursos Naturales Renovables (MARNR) (1983). *Sistemas ambientales venezolanos. Regiones naturales: 7A Depresión del Táchira, Proyecto Ven/79/001, Código II-2-7A. Proyecto VEN/79/001*. MARNR.

Ministerio de Tecnologías de la Información y Las Comunicaciones (MINTIC) (2018). *Promedios precipitación y temperatura media. Promedio de los años 1981-2010*. https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/Promedios-Precipitaci-n-y-Temperatura-media-Promed/nsxu-h2dh/data

Mohammed, S., Alsafadi, K., Talukdar, S., Kiwan, S., Hennawi, S., Alshihabi, O., Sharaf, M., and Harsanyie, E. (2020). Estimation of soil erosion risk in southern part of Syria by using RUSLE integrating geo informatics approach. *Remote Sensing Applications: Society and Environment*, *20*, 100375. https://doi.org/10.1016/j.rsase.2020.100375

Morales-Pavón, J., Valdés-Rodríguez, O., Servín-Martínez, A., Hernández-Zárate, J., Tejero-Andrade, J., and Domínguez-Sánchez, G. (2016). *Plantas tropicales para contener suelo y evitar deslizamientos superficiales: estudio de caso Ricinus communis* [Conference presentation]. II Reunión Internacional, Científica y Tecnológica; XXIX Reunión Científica y Tecnológica Forestal y Agropecuaria, Veracruz, México.

Muñoz, J. L., Morante, J., and Miranda, P. (2014). Erosión potencial por reconversión productiva en subcuenca Llay-Llay, Chile. Aplicación de unidades de respuesta a la erosión. *Ciencia y Tecnología*, *7*(2), 35-47. https://doi.org/10.18779/cyt.v7i2.138

Najafi, M., Fakhireh, A., Pahlavan, A., Moradzadeh, M., and Noori, S. (2020). Determining the suitable indices for assessment of cover change in west of Karkheh river using satellite data. *Journal of Applied Sciences and Environmental Studies*, *3*(1), 1-14. https://doi.org/10.48393/IMIST.PRSM/jases-v3i1.18928

Ngandam, A. H., Etouna, J., Nongsi, B. K., Mvogo, F. A., and Noulaquape, F. G. (2016). Assessment of land degradation status and its impact in arid and semi-arid areas by correlating spectral and principal component analysis neo-bands. *International Journal of Advanced Remote Sensing and GIS*, *5*(2), 1539-1560. https://doi.org/10.23953/cloud.ijarsg.77

Nasir, N. S. B., Mustafa, F. B., and Muhammad Yusoff, S. Y. (2023). Spatial prediction of soil erosion risk using knowledge-driven method in Malaysia's Steepland Agriculture Forested Valley. *Environment, Development and Sustainability, 2023,* s10668-023-03251-8. https://doi.org/10.1007/s10668-023-03251-8

Omuto, C. T., and Vargas, R. (2019). *Soil loss atlas of Malawi*. Food & Agriculture Organization.

Opeyemi, O. A., Abidemi, F. H., and Otokiti, V. (2019). Assessing the Impact of Soil Erosion on Residential Areas of Efon-Alaaye Ekiti, Ekiti-State, Nigeria. *International Journal of Environmental Planning and Management*, *5*(1), 23-31.

Orr, B. J., Cowie, A. L., Castillo, V. M., Chasek, P., Crossman, N. D., Erlewein, A., Louwagie, G., Maron, M., Metternicht, G. I., Minelli, S., Tengberg, A. E., Walter, S., and Welton, S. (2017). *Scientific conceptual framework for land degradation neutrality. A report of the science-policy interface*. United Nations Convention to Combat Desertification (UNCCD). https://doi.org/10.1016/j.envsci.2017.10.011

Pacheco, H. A. (2012). El índice de erosión potencial en la vertiente norte del Waraira Repano, estado Vargas, Venezuela. *Cuadernos de Geografía: Revista Colombiana de Geografía*, *21*(2), 85-97. https://doi.org/10.15446/rcdg.v21n2.32215

Panagos, P., Ballabio, C., Borrelli, P., Meusburger, K., Klik, A., Rousseva, S., Tadic, M. P., Michaelides, S., Hrabalikova, M., Olsen, P. Aalto, J., Lakatos, M., Rymszewicz, A., Dumitrescu, A., Begueria, S., and Alewell, C. (2015). Rainfall erosivity in Europe. *Science Total Environment*, *511*, 801-14. https://doi.org/10.1016/j.scitotenv.2015.01.008

Papaiordanidis, S., Gitas, I. Z., and Katagis, T. (2019). Soil erosion prediction using the revised universal soil loss equation (RUSLE) in Google Earth Engine (GEE) cloud-based platform. *Dokuchaev Soil Bulletin*, *100*, 36-52. https://bulletin.esoil.ru/jour/article/view/538

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, *2*(11), 559-572. https://doi.org/10.1080/14786440109462720

Pearson, R. L., and Miller, L. D. (1972). *Remote mapping of standing crop biomass for estimation of the productivity of the short-grass prairie, Pawnee National Grasslands, Colorado* [Conference presentation]. Eighth International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, MI, USA.

Perry, C. R., and Lautenschlager, L. F. (1984). Functional equivalence of spectral vegetation indices. *Remote Sensing of Environment*, *14*(1-3), 169-182. https://doi.org/10.1016/0034-4257(84)90013-0

Plambeck, N. O. (2020). Reassessment of the potential risk of soil erosion by water on agricultural land in Germany: Setting the stage for site-appropriate decision-making in soil and water resources management. *Ecological Indicators*, *118*, 106732. https://doi.org/10.1016/j.ecolind.2020.106732

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., and Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119-126. https://doi.org/10.1016/0034-4257(94)90134-1

Quiñonez, E., and Dal Pozzo, F. (2008). Distribución espacial del riesgo de degradación de los suelos por erosión hídrica en el estado Lara, Venezuela. *Geoenseñanza*, *13*(1), 59-70.

Richardson, A. J., and Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogramnetric Engineering and Remote Sensing*, *43*(12), 1541-1552.

Romero, W., Ramos, R., Vázquez, R., Arrogante, P. and Arroyo, M. (2017). *Detección de deslizamientos de laderas por el método de regresión lineal utilizando imágenes Aster en la zona centro del estado de Guerrero, México* [Conference presentation]. XXV Congreso de la Asociación de Geógrafos Españoles, Madrid, Spain. https://www.age-geografia.es/downloads/Naturaleza_Territorio_y_Ciudad_AGE2017.pdf

Rosales-Rodríguez, C. A. (2021). Hazard maps of shallow landslides associated with infiltration processes in the Sapuyes river basin. *Ingeniería e Investigación*, *41*(1), e84611. https://doi.org/10.15446/ing.investig.v41n1.84611

Rosales, A., and García, P. (2015). La cuenca hidrográfica y su gestión integral. In A. Gabaldón, A. Rosales, E. Buroz, J. Córdova, G. Uzcátegui, and L. Iskandar (Eds.), *Agua en Venezuela: Una Riqueza Escasa* (pp. 867-914). Fundación Polar.

Rouse, J., Haas, R. H., Schell, J. A., and Deering, D. W. (1973). *Monitoring vegetation systems in the Great Plains with ERTS* [Conference presentation]. Third ERTS Symposium, Washington DC, USA.

Sartori, A., Cano, J., Montaner, D., Mattar, C., Moraga, J. Alfaro, W., Soto, G., Morales, L., Quintanilla, O., Andrés, E. Gavilán, C., and Trujillo, G. (2018). *Reporte de neutralidad en la degradación de las tierras (NDT) ante la Convención de las Naciones Unidas de Lucha Contra la Desertificación (CNULD), Estrategia Nacional de Cambio Climático y Recursos Vegetacionales (2017-2025) de Chile*. Unidad de Cambio Climático y Servicios Ambientales (UCCSA), Gerencia de Desarrollo y Fomento Forestal (GEDEFF), Corporación Nacional Forestal (CONAF).

Shobha, G., and Rangaswamy, S. (2018). Machine learning. In C. R. Gudivada and N. Venkat (Eds), *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications* (pp. 197-228). Elsevier. https://doi.org/10.1016/bs.host.2018.07.004

Soluciones Integrales GIS (SIGIS) (2019). *DIGITALGLOBE*. http://www.sigis.com.ve/index.php/imagenessatelitalesar/digitalglobe

Tsegaye, K., Addis, H. K., and Hassen, E. E. (2020). Soil erosion impact assessment using USLE/GIS approaches to identify high erosion risk areas in the lowland agricultural watershed of Blue Nile Basin, Ethiopia. *International Annals of Science*, *8*(1), 120-129. https://doi.org/10.21467/ias.8.1.120-129

Thenkabail, P. S. (2016). *Remotely sensed data characterization, classification, and accuracies*. CRC Press Taylor and Francis Group. https://doi.org/10.1201/b19294

Thiam, A. K. (1997). *Geographic information system and remote sensing methods for assessing and monitoring land degradation in the Shale: The case of Southern Mauritania* [Doctoral dissertation, Darks University].

United Nations (UN) (2021, April 16). *Aplicación de datos del mes: erosión del suelo*. https://www.un-spider.org/es/enlaces-y-recursos/fuentos-de-datos/daotm-erosion-suelo#USLE

United Nations General Assembly (UNGA) (2015). *Transforming our world: The 2030 agenda for sustainable development*. https://sdgs.un.org/2030agenda

United States Department of Agriculture (USDA) (2020). *Soil texture calculator*. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_054167

Walther, D., and Shabaani, S. (1991). *Large scale monitoring of rangelands vegetation using NOAA/AVHRR LAC data: Application to the rainy seasons 1989/90 in northern Kenya*. Ministry of Livestock Development.

Wang, L., Huang, J., Du, Y., Hu, Y., and Han, P. (2013). Dynamic assessment of soil erosion risk using Landsat TM and HJ satellite data in Danjiangkou reservoir area, China. *Remote Sensing*, *5*(8), 3826-3848. https://doi.org/10.3390/rs5083826

Wang, H., Zhao, W., Li, C., and Pereira, P. (2021). Vegetation greening partly offsets the water erosion risk in China from 1999 to 2018. *Geoderma*, *2021*, 115319. https://doi.org/10.1016/j.geoderma.2021.115319

Zhao, H., and Chen, X. (2005). Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *International Geoscience and Remote Sensing Symposium*, *3*, 1666-1668.

Research Article / Systems and Computer Engineering https://doi.org/10.15446/ing.investig.90748

An Automatic Approach for Bone Tumor Detection from Non-Standard CT Images

Un enfoque automático para la detección de tumores óseos a partir de imágenes de CT no estándar

Hatice Catal Reis[1] and Bulent Bayram[2]

ABSTRACT

Image processing techniques are applied in many fields of science. This study aims to detect tumors in the foot and create 3D models via computed tomography (CT), as well as to produce biometric data. 1 039 CT images were obtained from a server. The parameters used were a collimation of 64 detectors, a scanning thickness of 0,5-3 mm, and a pixel size of 512 x 512, with a radiometric resolution of the 16-bit gray levels. Noise reduction, segmentation, and morphological analysis were performed on CT scans to detect bone tumors. In addition, this study used digital image processing techniques to create a virtual three-dimensional (3D) model of bone tumors. The performance of our proposal was evaluated by analyzing the receptor operating characteristics (ROC). According to the results, the sensitivity, specificity, and precision in tumor detection were 0,96, 1, and 0,98%, respectively, with a 0,99% average F-measure. Radiologist reports were used for the sake of comparison. The proposed technique for detecting bone tumors of the foot via CT can help radiologists with its increased precision, sensitivity, specificity, and F-measure. This method could improve the diagnosis of foot and ankle tumors by allowing for the multidirectional quantification of abnormalities.

Keywords: CT, medical image processing, region-growing algorithm, bone tumor, 3D

RESUMEN

Las técnicas de procesamiento de imágenes se aplican en muchos campos de la ciencia. El objetivo de este estudio es detectar tumores en el pie y crear modelos 3D mediante tomografía computarizada (CT), así como producir datos biométricos. 1 039 imágenes de CT se obtuvieron de un servidor. Los parámetros utilizados fueron una colimación 64 detectores, un grosor de escaneo de 0,5-3 mm y un tamaño del píxel de 512 x 512, con una resolución radiométrica de niveles de gris de 16 bits. Se realizó reducción de ruido, segmentación y análisis morfológico a imágenes CT para detectar tumores óseos. Adicionalmente, en este estudio se aplicaron técnicas de procesamiento de imágenes digitales para crear un modelo virtual tridimensional (3D) de tumor óseo. El rendimiento de nuestra propuesta se evaluó con base en el análisis de las características operativas del receptor (ROC). Según los resultados, la sensibilidad, la especificidad y la precisión en la detección de tumores fue de 0,96, 1 y 0,98 % respectivamente, con un F-*measure* promedio de 0,99 %. Se utilizaron reportes de radiología para efectos de comparación. La técnica propuesta para detectar tumores óseos del pie mediane CT puede ayudar a los radiólogos dada su alta precisión, sensibilidad, especificidad y F-*measure*. Este método puede mejorar el diagnóstico de tumores de pie y tobillo al permitir la cuantificación multidireccional de anomalías.

Palabras clave: CT, procesamiento de imágenes médicas, algoritmo de crecimiento de regiones, tumor óseo, 3D

Received: September 25th, 2020
Accepted: February 17th, 2023

Introduction

Photogrammetry involves scientific methods that calculate an object's 3D coordinates using overlays of two consecutive images (Akcay *et al*., 2017). It is frequently used in fields such as forestry, agriculture, environment, mining, planning, and medicine. Medical photogrammetry can precisely present an organ's metric and morphometric measurements and location data without making contact with the patient or the organ (Catal Reis, 2018). Medical photogrammetry using medical data can detect anomalies and cancer. Cancer occurs due to various environmental and genetic factors. The most prevalent types amount to over 100 different known cancers, such as skin cancer (Parker, 2020), lung cancer (Plodkowski *et al*., 2021), prostate cancer (Ghafouri-Fard *et al*., 2020), melanoma (Esim *et al*., 2019), stomach cancer (Zilberman *et al*., 2015), breast cancer (Riis, 2020), bone cancer (Hosseini *et al*., 2020), *etc*. Bone tumors are more widespread in adults than in children and youngsters (Reddy *et al*., 2015). Tumors and tumor-like lesions of the foot and ankle are rare (Toepfer, 2017; Foo and Raby, 2005; Bakotic and Huvos, 2001). Foot tumors have about a 3% ratio among all skeletal tumors (Yan *et al*., 2018). Benign bone tumors are more common than malignant primary bone tumors in the foot (Ozer *et al*., 2017). However, benign bone tumors are generally asymptomatic and go unexplored (Ladd and Roth,

[1] PhD in Geomatics Engineering, Yildiz Technical University, Turkey. Affiliation: Assistant professor, Department of Geomatics Engineering, Gumushane University, Gumushane, Turkey. E-mail: hatice.catal@yahoo.c-om.tr

[2] PhD in Geomatics Engineering, Yildiz Technical University, Turkey. Affiliation: Professor, Department of Geomatics Engineering, Yildiz Technical University, Istanbul, Turkey. E-mail: bulentbayram65@gmail.com

2017). A tumor is the abnormal development of new cells that may occur in any of the body parts (Mistry *et al*., 2016).

The most often faced primary benign bone tumor of the foot is the unicameral bone cyst (UBC), found in second osteochondroma and osteoid osteoma. Calcaneus tumors are especially rare; 32% of them are benign, and 35% of osseus tumor types occur in the foot (Young *et al*., 2013). Calcaneus tumors are the second most common condition involving the foot after meta-tarsals (Kilgore and Parrish, 2005). However, accurately and reliably describing and quantifying foot tumors is a challenging task for orthopedic surgeons and orthopedists. Bone anomalies in the foot and ankle are complex structures and, therefore, difficult to quantify by standard radiographic measurements. The literature is limited to calcaneus lesions. In particular, studies of calcaneus tumors consist of case reports describing a single primary tumor of the calcaneus (Yan *et al*., 2018). The first significant cases related to calcaneus tumors and tumor-like conditions were described in 1973 by Campbell and Leupold. These authors presented a case of a rare elastofibroma (soft tissue tumor) in the calcaneus of a 79-year-old female (Pirak *et al*., 2020). Another case of a giant cell tumor of the calcaneus with a bone cyst was reported by Kamal *et al*. (2016). The case of a 21-year-old male patient with osteosarcoma of the calcaneus was presented by Diémé *et al*. (2015). Past and present research has focused on classifying tumor types, growth directions, and areas, as well as on determining locations (Costelloe and Madewell, 2013). Foo and Raby (2005) presented the spectrum of tumors and tumor-like lesions of the foot and ankle which are not considered standard. Since calcaneal tumor types are rare, delays in diagnosis, incorrect diagnosis, and unnecessary amputations can be observed (Young *et al*., 2013).

Several diagnosis methods have been proposed for tumor detection. However, radiological images are the gold standard and primary step for tumor evaluation (Ladd and Roth, 2017; Bestic *et al*., 2020). Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are commonly used for diagnosis (Hasbek *et al*., 2013; Gemescu *et al*., 2019; Scotto di Carlo *et al*., 2020). High spatial-resolution CT supplies sensitive bone details, thus making it a vital tool in various cases and with specific tumors (Ladd *et al*., 2017; Mehta *et al*., 2017). CT scans are better at evaluating bone tumors for destruction or breakthrough than MR imaging (Hapani *et al*., 2014). When a tumor metastasizes to the bone, severe pain and deterioration occur on it, considerably reducing the possibility of a treatment (Weilbaecher *et al*., 2011). A significant limitation for studying what regulates tumors in the bone is the lack of non-contact models for tumor reconstruction, as well as the limited ability to detect low tumor levels in the bone. In addition, low-resolution and noisy images may prevent the tumor from being easily detected and are time-consuming. There needs to be more tools in addition to these imaging and diagnosis systems in order to obtain accurate and reliable results. A second opinion is essential to meet the requirements of imaging systems and mitigate the possible failures of technicians and physicians. Therefore, photogrammetry and medical image processing are the best techniques for using non-standard CT images, as they produce highly accurate results. Image processing technology converts data into different formats, and it is useful for measuring and evaluating (Aydın and Kurnaz, 2023). Therefore, Computer-Aided Diagnostics (CAD) has been proposed for diagnosis and treatment (Lodwick *et al*., 1963). Today, techniques for 3D bone measurement have become available (Qiang *et al*., 2014), which allow for surface, distance, location, and angular measurements on models generated from CT or MRI images (Gutekunst *et al*., 2013; Eckstein *et al*., 2006).

Many CAD methods have already been employed, including 2D and 3D approaches for tumor measurement (Nishikawa, 2005). However, little data are currently available in the literature with regard to measuring the 3D morphological parameters of the calcaneus (Qiang *et al*., 2014).

This study aims to evaluate the feasibility of medical photogrammetry for the detection of bone tumors, with the purpose of automatically detecting calcaneus tumors in noisy, non-standard CT images, as well as visualizing them in 3D and evaluating their metric information. CT images were used in this study, and the tumor segmentation and 3D model were created using software. This research provides an overview of the evidence for using CAD with calcaneus tumors. This highly sensitive method to detect bone tumors is ultimately summarized regarding its applicability to each foot tumor model. Medical image processing to detect foot tumors in CT can assist radiologists by increasing their accuracy. The proposed techniques to detect small and significant changes in tumors and tumor-like lesions, combined with these novel models, will be instrumental in studying tumors in the bone (benign or metastatic), as well as their stages.

Our study comprises three sections. The next section characterizes the calcaneus and explains the dataset. Then, general information is presented regarding medical image processing (pre-processing, segmentation, 3D modeling, obtaining biometric information, and statistical analysis). Afterwards, the results are presented and discussed while making a set of recommendations. The final section deals with the conclusions of this work.

Materials and methods

In medical imaging, conversion from 2D to 3D or *vice versa* can be done with the help of photogrammetric techniques. The length, area, volume, and geometric information can be accessed by processing these data. In addition, contactless information and metric data can be obtained through metric measurements (Doğan and Yakar, 2018). This study focuses on tumor detection in the calcaneus bone via the medical photogrammetric approach.

The calcaneus bone is the largest tarsal bone and is irregularly shaped, with complex edges and articular surfaces (Qiang *et al.*, 2014). It is a large bone that forms the base of the back of the foot. The calcaneus 'cuboid' connects bones and the talus. The bond between the talus and the calcaneus forms sub-articular joints. This joint is essential for normal foot function (Epstein *et al.*, 2012; Basile *et al.*, 2012).

As for this study, from 2011 to 2014, 20 cases (left and right foot scans) were taken from the server (retrospective study/stock footage research), for a total of 1 039 CT scans. The foot and ankle CT images of 10 cases (male group age: 20,2±9,8; female group age: 47±31) were included given the presence of calcaneal bone tumors and tumor-like-lesions. The imaging procedures were carried out according to the Helsinki Declaration (Goodyear *et al.*, 2007). This study did not require ethical approval or informed consent by national legislation.

CT images were taken from an open-access public hospital database. The medical photogrammetry and digital medical image format used was DICOM (Digital Imaging and Communications in Medicine). Most commercial software applications support the .dcm format, which facilitates reading raw data and sharing them. Therefore, DICOM has become a standard format for medical imaging. The parameters employed were a detector collimation of 64, a scanning thickness of 0,5-3 mm, and pixel sizes of 512 x 512, with a radiometric resolution of 16-bit gray levels (Table 1). Package software was used (Able Software Corp., USA) for 3D modelling and biometric data, as well as MATLAB for pre-processing and statistical analysis. All this, on a computer with an Intel Core i7-2670QM duo processor @2.20 GHz and 6 GB RAM. The axial images (.dcm) were first transferred into a personal computer with the 3D modeling software used in this study. Calibration was performed automatically.

Table 1. Imaging parameters for CT scanning

Patient Name	Gender	Scanning thickness	Number of slices	Age	Tube current (mA) Tube voltage (kV)
1	Male	2 mm	85	19	100-120
2	Female	3 mm	59	13	100-120
3	Female	2 mm	85	53	100-120
4	Female	0,5 mm	313	48	100-120
5	Female	2 mm	101	43	100-120
6	Male	2 mm	81	14	100-120
7	Male	2 mm	88	30	100-120
8	Female	2 mm	81	78	100-120
9	Male	2 mm	76	20	100-120
10	Male	2 mm	70	18	100-200

Source: Authors

Figure 1 shows a diagram of the proposed architecture, which consists of seven different processes.

This study followed these steps: (1) image pre-processing, (2) calcaneus segmentation, (3) tumor segmentation, (4) 3D modeling – visualization, and (5) biometric calculation. These five steps make up the primary process, in comparison with the seven steps of the physician report.

Our study involved the processing of bone CT images, so as to identify the region of bone cancer and perform a 3D evaluation. Axial raw CT images containing healthy and bone tumors are presented in Figures 2 and 3.

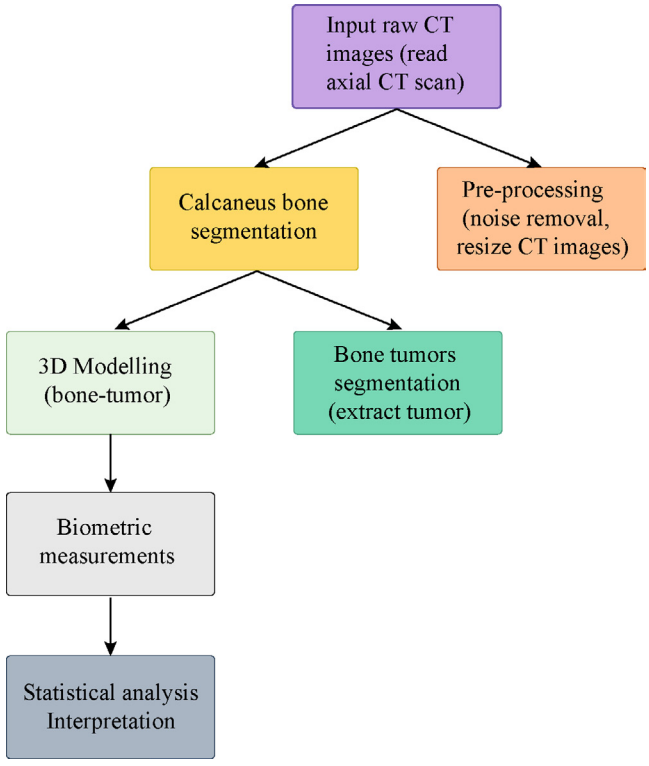


Figure 1. Schematic diagram of the workflow
Source: Authors

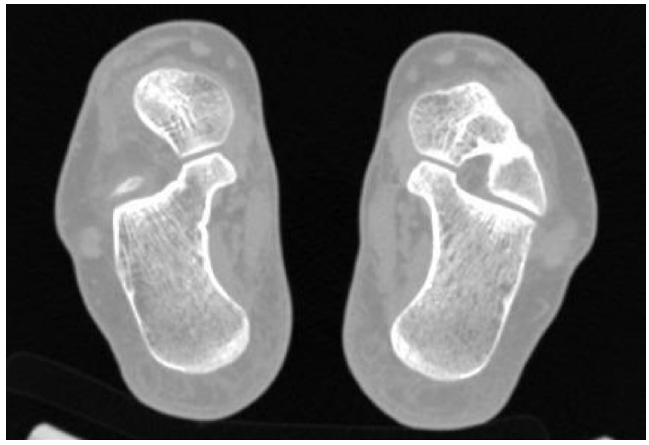


Figure 2. Axial CT image of a healthy person
Source: Authors

The calcaneus CT images employed were noisy images of various sizes. Image pre-processing can improve the accuracy of a 3D model. To this effect, data selection and classification (tumor or non-tumor) were performed, followed by a resizing of the CT images, which were

discarded if there was still noise. The most important goal here was to sort out images of different diseases. Another critical point was using noise removal techniques to enhance the images (the noise outside the main calcaneus bone was removed). Therefore, the images were pre-processed.

Figure 3. (A, B, C) axial CT slices of bone tumors
Source: Authors

In a classification process, the pixels of an image are segmented based on determining features. Segmentation methods and their success rates can vary according to their aim, applied field, and the quality of the image used. In general, segmentation algorithms are dependent on discontinuity and the similarity of grey levels (González and Woods, 2007).

A region-growing algorithm was applied to find the Region of Interest (ROI) of foot bone tumors. This method searches for similarities in the grey levels of an image for segmentation (Figure 4). It is important for the sets created here to be homogeneous, which does not depend on the training process. Although different region-growing algorithm approaches have been developed and proposed, the main general steps of the method can be ordered as thresholding, growing, and split-merge (González and Woods, 2007). This method grows from seeds to adjacent points via thresholding techniques.

Data were collected from the Hounsfield Units of the images to establish a suitable threshold value for the region membership criterion. Tumor segmentation and 3D tumor models were created via commercial software.

Afterwards, a 3D model was created from the segmented images. The performance of this model was evaluated in terms of its accuracy in the segmentation of the calcaneus CT images. Using the region-growing algorithm to extract the correct edge contour increased the performance of the study. Then, biometrical calculations were performed. The steps for image processing consisted of noise reduction, resizing, segmentation, 3D modeling, and morphological analysis.

To summarize this study's workflow, the CT image was read, pre-processed, segmented and edited; a 3D model was generated; and biometric measurements were performed. The tumor location is illustrated in Figure 5.

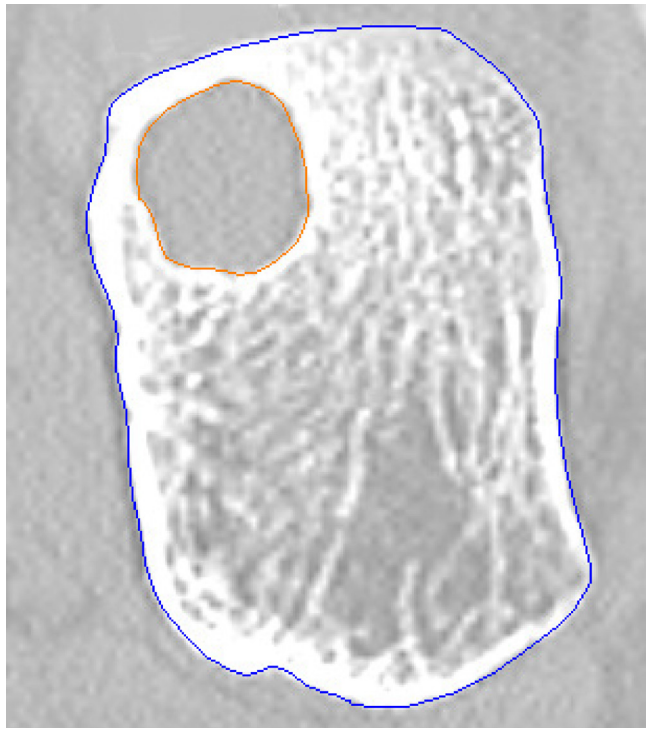


Figure 4. Segmentation result of the calcaneus bone and its tumor using CT
Source: Authors

Figure 5. 3D morphological measurements of the calcaneus and the location of tumors in it (orange denotes a tumor, and blue is the calcaneus body)
Source: Authors

Statistical analysis

The morphological parameters of the 3D measurements were evaluated using statistical analysis. The performance of the study was evaluated with receiver operating characteristic (ROC) analysis. The parameters are the following: true positive (TP), false negative (FN), true negative (TN), and false positive (FP) (Nahm, 2022). The area under the ROC curve (AUC= 0,5) (Nurtanio *et al.*, 2013) can be used as a criterion.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

$$F - Measure = \frac{2TP}{2TP + TN + FB} \quad (4)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

Results and discussion

The bone tumors were extracted, and other bones were removed. The calcaneal tumors' morphological parameters were shown in 3D (Figure 6).

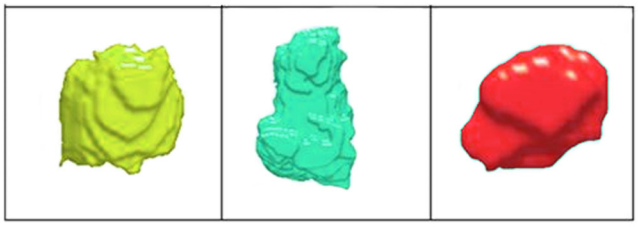


Figure 6. 3D modeling of tumors
Source: Authors

This study focused on automatic calcaneus tumor detection and diagnosis using CT. The results of radiologist reports were compared to those of CAD (Table 2).

Table 2. Comparison between morphological measurements and medical reports of 3D calcaneal tumors

Patients	Our results (mm)	Radiologist reports (mm)
1	20,65 x 16,11	22 (approximately)
2	16,54 x 14,40	15 (approximately)
3	6,25 x 8,99	mm level
4	29,10 x 21,97	cystic structure
5	43,03 x 26,35	40x25
6	57,86 x 45,98	40-50 (approximately)
7	21,50 x 20,47	23x19
8	51,78 x 24,89	cystic structure
9	44,22 x 28,39	cystic structure
10	39,10 x 18,13	cystic structure

Source: Authors

Our method provides precise results, while medical reports provide approximate statements. The experimental results of the diagnostic performance metrics of the 3D process were evaluated with ROC. A threshold value of 0,5 was selected in the AUC plot. An ROC score of 80% and above is a good result. This study proposes an automatic segmentation method for automatic calcaneus tumor segmentation on CT images. Regarding the segmentation process, the average sensitivity, specificity, accuracy, and F-measure were 0,96%, 1,00%, 0,98, and 0,99%, respectively. The 3D method is computationally more complex than the 2D one, although interactive segmentation is highly accurate.

The standard radiological approach may be insufficient due to the irregular shapes of the anomalies that occur in the foot and ankle. Calcaneus tumors are rare (Young *et al.*, 2013) and the second most common tumor site in the foot (Kilgore and Parrish, 2005). Benign bone tumors are generally asymptomatic and go unexplored (Ladd and Roth, 2017). The lack of literature on calcaneal tumors is another compelling factor. Therefore, studies on detection are still insufficient, which constitutes a research gap. Defining and measuring the anomaly in a safe, metric, and reproducible manner can also be challenging for orthopedists (a radiological report is presented in Table 2). It is important to produce fast and reliable results.

This is an efficient technique to determine multiple anomalies in the slices. Benign bone lesions were more common than malignant ones. However, several limitations were noted in this study. Firstly, standardizing CT data could not be generated because they were obtained from a database. It is known that obtaining data under the same conditions and with similar characteristics increases the accuracy and reliability of a study. At the same time, the larger the number of datasets, the more accurate the generalization will be. Secondly, since more patients were needed, no distinction could be made between men and women. However, within the framework of this research, this distinction was observed on the last foot bones because images of healthy individuals were evaluated (Catal Reis, 2010). Thus, the error caused by the amount of data was kept to a minimum. However, computer-aided digital image processing applications have been employed in different fields. This method has some advantages, as it is time-saving, accurate, and economical (Tasdemir and Ozkan, 2018).

Conclusions

The study used image processing, segmentation, clustering techniques, and modeling to detect foot bone tumors from CT images. All this, upon the basis of photogrammetry. This work examined the feasibility of computer-aided diagnosis for detecting foot bone tumors and tumor-like lesions in CT images and from 3D tumor models. The other contribution of this study was the design of an automatic method for 3D morphological measurements of bone tumors based on CT processing techniques. CAD for detecting calcaneus tumors of the foot can aid radiologists given its increased accuracy, sensitivity, specificity, and F-measure. Pre-processing applications can also increase the accuracy of 3D models.

The 3D reconstruction measurements based on image processing techniques were highly reliable and repeatable with regard to the anatomic and morphological measurement of the calcaneus. However, the limitations of our study include its small sample data size, different slice thicknesses,

and its retrospective design. Nevertheless, this technique will help to determine calcaneal tumors and tumor-like lesions. 3D foot bone cancer models can be developed using the approaches and future directions needed to advance these relatively new fields of research. Medical photogrammetry will continue to contribute to the literature as a second eye today and in the future. As future work, we aim to add machine and deep learning algorithms to our work and test them with different approaches.

Ethics approval and consent to participate

Retrospective study/stock footage research.

Funding

None.

Competing interests

The authors declare that they have no competing interests.

CRediT author statement

Hatice Catal Reis: conceptualization, methodology, software, validation, formal analysis, investigation, writing (original draft, writing, review, and editing), data curation. Bulent Bayram: Writing, editing, review, and validation.

References

Aydın, M., and Kurnaz, T. F. (2023). An alternative method for the particle size distribution: Image processing. *Turkish Journal of Engineering, 7*(2), 108-115. https://doi.org/10.31127/tuje.1053462

Akcay, O., Erenoglu, R. C., and Avsar, E. O. (2017). The effect of jpeg compression inclose-range photogrammetry. *International Journal of Engineering and Geosciences*, *2*(1), 35-40. https://doi.org/10.26833/ijeg.287308

Bakotic, B., and Huvos, A. G. (2001). Tumors of the bones of the feet: The clinicopathologic features of 150 cases. *Journal of Foot & Ankle Surgery, 40*(5), 277-286. https://doi.org/10.1016/S1067-2516(01)80063-6

Basile, A. (2012). Subjective results after surgical treatment for displaced intraarticular calcaneal fractures. *Journal of Foot and Ankle Surgery, 51*, 182-186. https://doi.org/10.1053/j.jfas.2011.10.042

Bestic, J. M., Wessell, D. E., Beaman, F. D., Cassidy, R. C., Czuczman, G. J., Demertzis, J. L., Lenchik, L., Motamedi, K., Pierce, J. L., Sharma, A., Sloan, A. E., Than, K., Walker, E. A., Yung, E. Y. K., and Kransdorf, M. J. (2020). ACR appropriateness criteria: Primary bone tumors. *Journal of the American College of Radiology*, *17*(5S),S226-S238. https://doi.org/10.1016/j.jacr.2020.01.038

Campbell, C. J., and Leupold, R. G. (1973). Tumours and tumour-like conditions of the os calcis. *Orthopedic Clinics of North America, 4*,145-156. https://doi.org/10.1016/S0030-5898(20)30510-1

Catal Reis, H. (2010). *Metric analysis of orthopedic changes of ballerina's foot bones by photogrammetric techniques* [Master's thesis, Selcuk University]. http://acikerisimarsiv.selcuk.edu.tr:8080/xmlui/handle/123456789/660.

Catal Reis, H. (2018). Detection of foot bone anomaly using medical photogrammetry. *International Journal of Engineering and Geosciences*, *3*(1), 001-005. https://doi.org/10.26833/ijeg.333686

Costelloe, C. M., and Madewell, J. E. (2013). Radiography in the initial diagnosis of primary bone tumors. *American Journal of Roentgenology, 200*, 3-7. https://doi.org/10.2214/AJR.12.8488

Diémé, C., Dembélé, B., Gaye, A.M., Sarr, L., Coundoul, C., Gueye, A.B., Déme, H., Sané, A., and Seye, S. (2015). Osteosarcoma of the calcaneus: A case report. *Médecine et Chirurgie du Pied, 31*, 69-71. https://doi.org/10.1007/s10243-015-0407-1

Doğan, Y., and Yakar, M. (2018). GIS and three-dimensional modeling for cultural heritages. *International Journal of Engineering and Geosciences, 3*(2), 50-55. https://doi.org/10.26833/ijeg.378257

Eckstein, F., Cicuttini, F., Raynauld, J.P., Waterton, J.C., and Peterfy, C. (2006). Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): Morphological assessment. *Osteoarthritis and Cartilage, 14*(1), 46-75. https://doi.org/10.1016/j.joca.2006.02.026

Epstein. N., Chandran. S., and Chou. L. (2012). Current concepts review: Intra-articular fractures of the calcaneus. *Foot & Ankle International, 33*, 79-86. https://doi.org/10.3113/FAI.2012.0079

Esim A. K., Kaya, H., and Alcan, V. (2019). Determination of malignant melanoma by analysis of variation values. *Turkish Journal of Engineering, 3*(3), 120-126. https://doi.org/10.31127/tuje.472328

Foo, L. F., and Raby, N. (2005). Tumors and tumor-like lesions in the foot and ankle. *Clinical Radiology, 60*, 308-332. https://doi.org/10.1016/j.crad.2004.05.010

Gemescu, I. N., Thierfelder, K. M., Rehnitz, C., and Weber, M.-A. (2019). Imaging features of bone tumors conventional radiographs and MR imaging correlation. *Magnetic Resonance Imaging Clinics of North America*, *27*, 753–767. https://doi.org/10.1016/j.mric.2019.07.008

Ghafouri-Fard, S., Shoorei, H., and Taheri, M. (2020). Role of microRNAs in the development, prognosis and therapeutic response of patients with prostate cancer. *Gene, 759*, 144995. https://doi.org/10.1016/j.gene.2020.144995

González, R. C., and Woods, R. E. (2007). *Digital image processing* (2nd ed.). Publishing House of Electronics Industry.

Goodyear, M. D., Krleza-Jeric, K., and Lemmens, T. (2007). The declaration of Helsinki. *British Medical Journal, 335*, 624-625. https://doi.org/10.1136/bmj.39339.610000.BE

Gutekunst, D. J., Liu, L., Ju, T., Prior, F. W., and Sinacore, D. R. (2013). Reliability of clinically relevant 3D foot bone angles from quantitative computed tomography. *Journal of Foot and Ankle Research, 6*, 38. https://doi.org/10.1186/1757-1146-6-38

Hapani, H., Kalola, J., and Hapani, J. (2014). Comparative role of CT scan and MR imaging in primary malignant bone tumors. *IOSR Journal of Dental and Medical Sciences, 13*(11), 29-35. https://doi.org/10.9790/0853-131172935

Hasbek, Z., Salk, I., Yucel, B., Yucel, B., and Babacan, N. A. (2013). Which imaging method to choose for detection of bone metastases? Bone scintigraphy, CT, 18F-FDG PET/CT or MR? *Bozok Medical Journal, 3*(3),44-50. https://dergipark.org.tr/tr/download/article-file/43180

Hosseini, A., Mirzaei, A., Salimi, V., Jamshidi, K., Babaheidarian, P., Fallah, S., Rampisheh, Z., Khademian, N., Abdolbahavi, Z., Bahrabadi, M., Ibrahimi, M., Hosami, F., and Tavakoli-Yaraki, M. (2020). The local and circulating SOX9 as a potential biomarker for the diagnosis of primary bone cancer. *Journal of Bone Oncology, 23*, 100300. https://doi.org/10.1016/j.jbo.2020.100300

Kamal, F., Waryudia, A., Effendia, Z., and Kodrat, E. (2016). Management of aggressive giant cell tumor of calcaneal bone: A case report. *International Journal of Surgery Case Reports, 28*, 176-181. https://doi.org/10.1016/j.ijscr.2016.09.038

Kilgore, W. B., and Parrish, W. M. (2005). Calcaneal tumors and tumor-like conditions. *Foot and Ankle Clinics, 10*(3), 541-565. https://doi.org/10.1016/j.fcl.2005.05.002

Ladd, M. L., Roth, T. D. (2017). Computed tomography and magnetic resonance imaging of bone tumors. *Seminars in Roentgenology, 52*(4), 209-226. https://doi.org/10.1053/j.ro.2017.04.006

Lodwick, G. S., Haun, C. L., Smith, W. E., Keller, R. F., and Robertson, E. D. (1963). Computer diagnosis of primary bone tumors. *Radiology, 80*, 273-275. https://doi.org/10.1148/80.2.273

Mehta, K., McBee, M. P., Mihal, D. C., and England, E. B. (2017). Radiographic analysis of bone tumors: a systematic approach. *Seminars in Roentgenology, 52*(4), 194-208. https://doi.org/10.1053/j.ro.2017.04.002

Mistry, K. D., and Talati, B. J. (2016, Oct. 6-8). *Integrated approach for bone tumor detection from MRI scan imagery* [Conference presentation]. 2016 International Conference on Signal and Information Processing (IConSIP), Vishnupuri, India. https://doi.org/10.1109/ICONSIP.2016.7857471

Nahm, F.S. (2022). Receiver operating characteristic curve: overview and practical use for clinicians. Korean Journal of Anesthesiology, 75 (1), 25-36. https://doi.org/10.4097/kja.21209

Nishikawa, R. M. (2005). *Computer-assisted detection and diagnosis*. Wiley.

Nurtanio, I., Astuti, E. R., Purnama, I. K. E., Hariadi, M., Purnomo, M. H. (2013). Classifying cyst and tumor lesion using support vector machine based on dental panoramic images texture features. *IAENG International Journal of Computer Science, 40*, 1-4. https://www.iaeng.org/IJCS/issues_v40/issue_1/IJCS_40_1_04.pdf

Ozer, D., Aycan, O. E., Er, S. T., Tanritanir, R., Arikan, Y., and Kabukcuoglu, Y. S. (2017). Primary tumor and tumor-like lesions of bones of the foot: Single-center experience of 166 cases. *Journal of Foot & Ankle Surgery, 56*, 1180-1187. https://doi.org/10.1053/j.jfas.2017.05.027

Parker, E. R. (2021). The influence of climate change on skin cancer incidence – A review of the evidence. *International Journal of Women's Dermatology*, *7*(1), 17-27. https://doi.org/10.1016/j.ijwd.2020.07.003

Pirak, J., Brandeisky, J. A., Simon, P., and Khaladj, M. (2020). Elastofibroma in the rear foot: A case report of a rare soft tissue tumor. *Journal of Foot & Ankle Surgery, 59*, 587-589. https://doi.org/10.1053/j.jfas.2019.09.021

Plodkowski, A. J., Arimatei, J., Araujo-Filho, B., Simmers, C. D. A., Girshman, J., Raj, M., Zheng, J., Rimner, A., and Ginsberg, M. S. (2021). Pre-treatment CT imaging in stage IIIA lung cancer: Can we predict local recurrence after definitive chemoradiotherapy? *Clinical Imaging, 69*, 133-138. https://doi.org/10.1016/j.clinimag.2020.07.005

Qiang, M., Chen, Y., Zhang, K., Li, H., and Dai, H. (2014). Measurement of three-dimensional morphological characteristics of the calcaneus using CT image post-processing. *Journal of Foot and Ankle Research, 7*, 19. https://doi.org/10.1186/1757-1146-7-19

Reddy, K. K., Anisha, P. R., and Prasad, L. N. (2015). A novel approach for detecting the bone cancer and its stage based on mean intensity and tumor size. Medicine, *Recent Researches in Applied Computer Science*, 162-171. https://www.semanticscholar.org/paper/A-Novel-Approach-for-Detecting-the-Bone-Cancer-and-Reddy-Prasad/6214771aac055643a74a94a52287cb69dacbd2a8

Riis, M. (2020). Modern surgical treatment of breast cancer. *Annals of Medicine and Surgery, 56*, 95-107. https://doi.org/10.1016/j.amsu.2020.06.016

Scotto di Carlo, F., Whyte, M. P., and Gianfrancesco, F. (2020). The two faces of giant cell tumor of bone. *Cancer Letters, 489*,1-8. https://doi.org/10.1016/j.canlet.2020.05.031

Tasdemir, Ş., and Ozkan, I. A. (2019). ANN approach for estimation of cow weight depending on photogrammetric body dimensions. *International Journal of Engineering and Geosciences, 4*(1), 036-044. https://doi.org/10.26833/ijeg.427531

Toepfer, A. (2017). Tumors of the foot and ankle – A review of the principles of diagnostics and treatment. *Fuß & Sprunggelenk, 15*, 82-96. https://doi.org/10.1016/j.fuspru.2017.03.004

Yan, L., Zong, J., Chu, J., Wang, W., Li, M., Wang, X., Song, M., and Wang, S. (2018). Primary tumours of the calcaneus (Review). *Oncology Letters, 15*, 8901-8914, https://doi.org/10.3892/ol.2018.8487

Young, P. S., Bell, S. W., MacDuff, E. M., and Mahendra, A. (2013). Primary osseous tumors of the hind-foot: Why the delay in diagnosis and should we be concerned. *Clinical Orthopaedics and Related Research, 471*, 871-877. https://doi.org/10.1007/s11999-012-2570-6

Weilbaecher, K. N., Guise, T. A., and McCauley, L. K. (2011). Cancer to bone: A fatal attraction. *Nature Reviews: Cancer, 11*, 411-425. https://doi.org/10.1038/nrc3055

Zilberman, Y., and Sonkusale S. R. (2015). Microfluidic optoelectronic sensor for salivary diagnostics of stomach cancer. *Biosensors and Bioelectronics, 6715*, 465-471. https://doi.org/10.1016/j.bios.2014.09.006

Research Article / Systems and Computer Engineering https://doi.org/10.15446/ing.investig.102248

Towards a Theory of Interoperability of Software Systems

Hacia una teoría de interoperabilidad de los sistemas de software

Diana M. Torres[1], David Chen[2], Mónica K. Villavicencio[2], and Carlos M. Zapata[3]

ABSTRACT

Interoperability is a property of software quality that is related to the cooperation between software systems for exchanging information. However, this concept is not well explained or understood. A theory would be useful to explain interoperability in terms of its essential elements and propositions. Theoretical contributions of interoperability are intended to formalize this concept by using common frameworks, models, and meta-models. However, tentative contributions developed in the past have failed to propose a theory of interoperability due to four reasons: (i) a disunified vocabulary is used, (ii) the essential elements for describing interoperability are not well identified, (iii) only a single level of interoperability is assessed, and (iv) interoperability principles are not well formalized. This paper tentatively proposes an axiomatic theory of interoperability as a complementary approach to the existing knowledge. The proposed theory seeks to better formalize the concepts of interoperability and suggest actions aimed at establishing interoperability. After a brief review of related works and the state of the art, a set of axioms and propositions is presented. This theory is evaluated by a group of experts, and an example is presented to illustrate its use. Conclusions and future works are outlined at the end of the paper.

Keywords: interoperability, axiomatic theory, software system, axiom, proposition, expert consultation

RESUMEN

La interoperabilidad es una propiedad de calidad del software que tiene que ver con la cooperación entre sistemas de software para intercambiar información. Sin embargo, este concepto carece de una explicación o un completo entendimiento. Una teoría permitiría explicar la interoperabilidad en términos de sus elementos esenciales y proposiciones. Las contribuciones teóricas acerca de la interoperabilidad proponen formalizar este concepto utilizando marcos comunes, modelos y metamodelos. No obstante, las contribuciones tentativas desarrolladas en el pasado no logran proponer una teoría. Esto, debido a cuatro razones: (i) usan un vocabulario desunificado, (ii) omiten los elementos esenciales para describir la interoperabilidad, (iii) se enfocan en un nivel particular de interoperabilidad y (iv) presentan una formalización incompleta con respecto a los principios de interoperabilidad. En este artículo se propone una teoría axiomática de interoperabilidad como un enfoque complementario al conocimiento existente. Con la teoría propuesta se pretende formalizar los conceptos de interoperabilidad y sugerir acciones para establecer la interoperabilidad. Con base en una revisión de los trabajos relacionados y el estado del arte, se define un conjunto de axiomas y proposiciones. Un conjunto de expertos valida la teoría, y se expone un ejemplo para ilustrar su uso. Las conclusiones y los trabajos futuros se describen al final del artículo.

Palabras clave: interoperabilidad, teoría axiomática, sistema de software, axioma, proposición, consulta de expertos

Received: April 21st, 2022
Accepted: March 29th, 2023

Introduction

Interoperability is referred to as a software quality property (International Organization for Standardization, 2008), an ability of entities for working together (Liu *et al.*, 2020), and a feature of information systems (Turk *et al.*, 2020), among other definitions. A characteristic of interoperability is the possibility of using exchanged information, *i.e.*, understanding and interpreting the exchanged information without additional effort.

Seven essential elements have been identified and defined for describing interoperability (Torres *et al.*, 2018): software system (source and target), information, language, symbol, context, and interface (Table 1). Uniformity in the terminology of interoperability is a requisite for building a common theory.

A software phenomenon like interoperability can be explained by using a theory. The elements of a theory are constructs (called essential elements) and the relationships between them. Building a theory comprises five steps: (i) defining essential elements, (ii) defining propositions, (iii) providing explanations about the propositions, (iv)

[1] PhD in Engineering – Systems and Informatics, Universidad Nacional de Colombia, Colombia. Affiliation: Universidad Nacional de Colombia, Colombia. Email: dimtorresri@unal.edu.co

[2] PhD in Physics and Engineering, University of Bordeaux, France. Affiliation: Full professor, University of Bordeaux, IMS Laboratory, France. Email: david.chen@ims-bordeaux.fr

[3] PhD in Engineering, Université du Québec in Montreal, Canada. Affiliation: Full professor, ESPOL University, Ecuador. Email: mvillavi@espol.edu.ec

[4] PhD in Engineering – Systems, Universidad Nacional de Colombia, Colombia. Affiliation: Full professor, Universidad Nacional de Colombia, Colombia. Email: cmzapata@unal.edu.co

determining the scope, and (v) testing the theory (Sjøberg *et al.*, 2012). A theory can also be conceived as axiomatic when it introduces new concepts and deduces their properties (Tall, 2004).

Theoretical contributions regarding interoperability (Table 2) attempt to formalize, characterize, and describe it according to different perspectives. To this effect, they use common frameworks defining concepts, practices, and criteria related to interoperability; common models representing interoperability solutions; meta-models including models of problems involving interoperability; and, finally, systematic literature reviews useful for collecting proposals focused on the challenges of interoperability.

Table 1. Essential elements of interoperability

Concept	Definition
Software system (i.e., source and target software system)	*"A system consisting solely of software and possibly the computer equipment on which the software operates".* *"A system made up of software, hardware, and data that provides its primary value by the execution of the software".*
Information	*"The meaning assigned to data by known conventions. The meaning that humans assign to data by means of known conventions that are applied to the data".*
Symbol	*"A representation of something by reason of relationship, association, or convention".*
Language	*"A system consisting of: 1) a well-defined, usually finite, set of characters 2) rules for combining characters with one another to form words or other expressions 3) a specific assignment of meaning to some of the words or expressions, usually for communicating information or data among a group of people, machines, etc.".*
Context	*"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves".*
Interface	*"(i) A common boundary between a considered system and another system, or between parts of a system, through which information is conveyed. (ii) A hardware or software component that connects two or more other components for the purpose of passing information from one to the other".*

Source: Torres Ricaurte *et al.* (2018)

The aforementioned contributions have at least one of the following problems:

1. The authors refer to interoperability by using a disunified terminology. The causes are a lack of consensus about what interoperability is, the lack of a unified vocabulary, and an incomplete characterization of interoperability.
2. The essential elements describing interoperability are not explicitly identified. The causes include the following: a complete view of interoperability is missing; comparing the approaches to interoperability is difficult; and some interoperability solutions which use current approaches are missing.
3. Interoperability principles are left aside because interoperability problems are difficult to identify and the source of the problems is disregarded.
4. A general view of interoperability is not reached because the proposals are focused on a single level of interoperability, *i.e.*, technical, syntactical, semantic, or organizational. A general and complete representation of interoperability is still missing.

This paper presents an axiomatic theory of interoperability based on a set of axioms and propositions. Such an axiomatic theory involves the seven essential elements of this concept. Axioms and propositions are stated in order to explain the role of each essential element during interoperability, as well as the relationships between such elements. Moreover, a discussion about the stated propositions is carried out which involves a group of independent interoperability experts.

An illustrative example is proposed with the aim to demonstrate the applicability of the axioms and propositions in a real situation. The example is taken from a local software development company.

The proposed axiomatic theory seeks to explain the interoperability that occurs between two software systems when information is transmitted by using an interface. Such information (containing symbols) is written in a language and interpreted according to a common context. Additionally, the axiomatic theory is used for describing how the essential elements are related, characterizing related problems based on the essential elements and the rules of interoperability, as well as comparing approaches by using a unified characterization.

Theoretical contributions analysis

To gather theoretical contributions regarding interoperability, a systematic literature review methodology was applied (Software Engineering Group, 2007). This method involved the following phases:

1. *Planning*

The main research question is *What is the set of essential elements for describing interoperability?* Information sources include scientific papers from 2000 to 2022 in databases such as Science Direct, IEEE Xplore, Springer, Scopus, and Google Scholar. The query was defined as follows: (Interoperability OR Interoperate OR Interoperation OR Interoperable) AND (Systems OR Software OR "Software systems") AND (Theory OR Formal OR Axiomatic OR Formalization OR Theoretical OR General).

2) *Execution*

The search was performed once in 2022 and again in 2023 for recent studies. The following selection criteria were applied after reading the abstracts and keywords: a) papers written in English from peer-reviewed conference proceedings, journal articles, book sections, and doctoral dissertations; b) papers with title and keywords related to formalizations regarding the concept, with interoperability as a main topic, and empirical studies; and c) articles including formalizations and models of interoperability in different kinds of software systems.

3) *Results and analysis*

106 (93 + 13) studies were collected for analysis. Three kinds of study were identified: general models, meta-models, and models of interoperability. Given our interest in generality, 38 papers presenting general models and meta-models were analyzed. Table 2 reports the 19 most recent papers (2018-2023).

As a result, it was noted that theoretical contributions regarding interoperability embrace approaches proposing common frameworks, common models, meta-models, and systematic literature reviews.

Common frameworks

Standards are used for reaching a common terminology about some aspects of interoperability. However, such terminology is specific to some issues regarding the concept (*e.g.*, interoperating heterogeneous manufacturing software units) and disunified concerning other proposals.

In other interoperability frameworks, an analysis of some interoperability levels (technical, syntactic, semantic, and organizational) is performed. Interoperability is evaluated within a specific context (*e.g.*, cloud, construction, enterprise, *etc.*). Therefore, a general view of this phenomenon has not been reached. Some proposals discuss the necessary information and conditions for accomplishing systems interoperability. However, such proposals are based on a disunified terminology, and their scope is limited to a given context.

Common models

Some common models aim to implement solutions to specific interoperability problems. Thus, an abstraction of this phenomenon is difficult, and some elements are overlooked. However, the formalizations of interoperability presented in said models are useful for identifying relationships and patterns of interoperability.

Table 2. Findings of the literature review

Reference	Proposal contribution	Interoperability level				Criteria		
		Semantic level	Technical level	Syntactic level	Organizational level	Unified terminology	Identification of some essential elements of interoperability	Statement of principles (rules, propositions, etc.) of interoperability
(Mistry *et al.*, 2022)	Common framework		X		X		X	
(Sana *et al.*, 2021)	Common framework	X	X	X			X	
(Turk, 2020)	Common framework	X						X
(Lafourcade and Lombard-Platet, 2020)	Common framework		X					X
(Liu *et al.*, 2020)	Common framework				X		X	
(Brilhault *et al.*, 2022)	Common model	X						
(Jepsen *et al.*, 2021)	Common model	X	X	X				
(Haile and Altmann, 2018)	Common model	X			X			
(Ribeiro *et al.*, 2018)	Common model		X	X				
(Challita *et al.*, 2018)	Common model	X						
(Horcas *et al.*, 2022)	Meta-model	X		X			X	
(Torres *et al.*, 2022)	Meta-model	X	X	X		X	X	
(Davies *et al.*, 2020)	Meta-model	X					X	
(Delgado *et al.*, 2020)	Meta-model	X					X	X
(Serapio *et al.*, 2019)	Meta-model				X	X		
(Torab-Miandoab *et al.*, 2023)	Systematic literature review	X	X	X				
(Fraga *et al.*, 2020)	Systematic literature review	X			X			
(Burzlaff *et al.*, 2019)	Systematic literature review				X		X	
(Burns *et al.*, 2019)	Systematic literature review		X				X	

Source: Authors

Meta-models

It is possible to identify two types of interoperability meta-models: (i) meta-models focused on data representations by using formalisms (*e.g.*, categorical theory, semantic models, and mathematical foundation, among others) which have partial views of interoperability; for this reason, they are limited for addressing some interoperability problems; (ii) meta-models with a high level of abstraction about interoperability and specific perspectives of the systems (*e.g.*, business processes, software systems, and general systems theory), which are limited by their areas of interest. In both types of meta-models, the principles of interoperability are left aside, and a disunified terminology is addressed.

Systematic literature reviews

There are reviews of current research on interoperability and its future needs. These works are useful for understanding the need to formalize interoperability and for analyzing how some elements of interoperability are used in the descriptions of this concept.

The findings related to the gaps in the previous works are summarized in Table 2.

Methodology

This section proposes an axiomatic theory for explaining what interoperability is and how to achieve it. In this theory, axioms seek to thoroughly describe what interoperability is, while propositions are statements about what is needed for achieving interoperability (Figure 1). Such an axiomatic theory is based on the identified essential elements of interoperability, *i.e.*, source software system, target software system, information, context, language, symbol, interface, and their relationships.

In addition, it is necessary to understand how the seven essential elements relate to each other, what role they play in establishing interoperability (exhibited relationship), and what dependencies and attributes they have. To this effect, the literature was examined by using an adapted methodology comprising three stages: structural and functional analysis, semantic processing, and discourse mapping. As a result, a pre-conceptual schema (PCS) of the essential elements and their structural relationship was obtained (Torres *et al*., 2022).

The relationships observed between the essential elements indicate that interoperability happens at least between two software systems: a source and a target software system. By instantiating the PC of interoperability (Torres *et al*., 2022), it is possible to realize that, even though a two-way relationship is expected (*e.g.*, an answer derived from the exchange), our model allows representing such a situation while changing the roles of the software systems involved. The core element of interoperability is information, which is created by the source software system and used in the target software system. The information is written by using some language and exchanged via an interface. Symbols are part of its content. Each software system has its own context, which describes a set of software system elements used for matching and transforming the exchanged information. Furthermore, a common context is necessary for interpreting the information.

With the results of the mapping and unification process and the interoperability PC about the relationships between essential elements, a new perspective was established, aimed at reaching underlying and intuitive reasoning regarding how the essential elements should be arranged for achieving interoperability. All this, while considering the definitions of axioms as well accepted statements and the propositions as facts accepted by the expert community (Tall, 2004). This work focused on recognizing and analyzing such statements (*i.e.*, well accepted statements and facts) with regard to each essential element (including its mapped concepts), looking at statements like:

"Interoperability is characterized" AND "interoperability means" AND "interoperability requirements are" AND "the means to achieve interoperability" AND "interoperability is provided by means" AND "to ensure interoperability is necessary" AND "an interoperable system meets" AND "to establish interoperability" AND "interoperability depends/ requires".

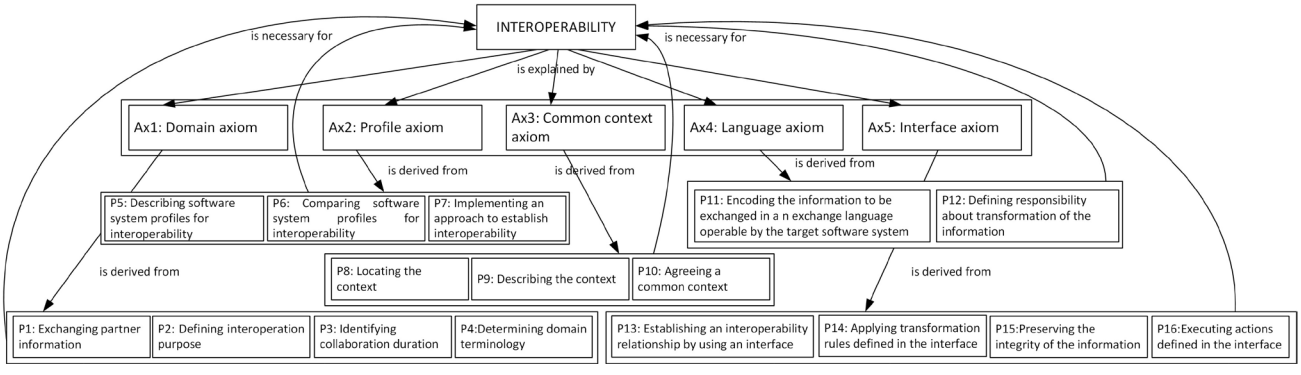


Figure 1. Representation and summary of the axiomatic theory
Source: Authors

As a result, an axiomatic theory of interoperability can be proposed which is supported by the findings of the literature (Figure 2).

Axiomatic theory of interoperability

AXIOM 1. INTEROPERABILITY DOMAIN – Exchanging domain information to prepare interoperability.

> When one needs to establish interoperability between the software systems of two organizations, the domain information of both software systems should first be exchanged for a preliminary mutual understanding of their backgrounds.

For example, if one needs to establish interoperability between two software systems in different domains such as e-Government (encompassing entities as citizens, business, administration, *etc*.) and e-health (encompassing entities as citizens, administration, e-surveillance, *etc*.), this mutual understanding (respect towards partner information, the purpose and duration of interoperability, and the domain terminology) is very important as a starting point.

Proposition 1: Exchanging partner information

> Partners should exchange information to allow their collaboration. This includes organization names, location, economic sector, people in charge of a possible collaboration, and any potential constraints.

Derived axioms and propositions	Reference
Ax1. Domain axiom	(Fraga *et al*., 2020), (Mistry *et al*., 2022)
P1. Exchanging partner information	(Delgado *et al*., 2020), (Liu *et al*., 2020), (Mistry *et al*., 2022), (Torres *et al*., 2022)
P2. Defining the purpose of interoperation	(Delgado *et al*., 2020), (Liu *et al*., 2020), (Haile and Altmann, 2018), (Mistry *et al*., 2022), (Serapio *et al*., 2019), (Torres *et al*., 2022), (Turk, 2020)
P3. Identifying the duration of collaboration	(Liu *et al*., 2020), (Torres *et al*., 2022)
P4. Determining the domain terminology	(Delgado *et al*., 2020), (Fraga *et al*., 2020), (Liu *et al*., 2020)
Ax2. Profile axiom	(Fraga *et al*., 2020), (Haile and Altmann, 2018), (Mistry *et al*., 2022), (Torab-Miandoab *et al*., 2023)
P5. Describing software system profiles for interoperability	(Haile and Altmann, 2018)
P6. Comparing software system profiles for interoperability	(Mistry *et al*., 2022), (Haile and Altmann, 2018)
P7. Implementing an approach to establish interoperability	(Mistry *et al*., 2022), (Turk, 2020)
Ax3. Common context axiom	(Davies *et al*., 2020), (Fraga *et al*., 2020), (Jepsen *et al*., 2020), (Liu *et al*., 2020), (Mistry *et al*., 2022), (Sana *et al*., 2021), (Torres *et al*., 2022), (Turk, 2020)
P8. Locating the context	(Delgado *et al*., 2020), (Liu *et al*., 2020), (Torres *et al*., 2022)
P9. Describing the context	(Brilhault *et al*., 2022), (Delgado *et al*., 2020), (Torab-Miandoab *et al*., 2023), (Torres *et al*., 2022)
P10. Agreeing on a common context	(Brilhault *et al*., 2022), (Davies *et al*., 2020), (Lafourcade and Lombard-Platet, 2020), (Liu *et al*., 2020), (Torres *et al*., 2022)
Ax4. Language axiom	(Burns *et al*., 2019), (Burzlaff *et al*., 2019), (Davies *et al*., 2020), (Fraga *et al*., 2020), (Jepsen *et al*., 2021), (Liu *et al*., 2020), (Sana *et al*., 2021), (Torab-Miandoab *et al*., 2023), (Torres *et al*., 2022)
P11. Encoding the information to be exchanged in an exchange language operable by the target software system	(Horcas *et al*., 2023), (Liu *et al*., 2020), (Torres *et al*., 2022)
P12. Defining responsibility regarding the transformation of the information	(Burzlaff *et al*., 2019), (Davies *et al*., 2020), (Liu *et al*., 2020)
Ax5. Interface axiom	(Brilhault *et al*., 2022), (Haile and Altmann, 2018), (Liu *et al*., 2020), (Sana *et al*., 2021), (Torres *et al*., 2022), (Turk, 2020)
P13. Establishing relationship of interoperability by using an interface	(Brilhault *et al*., 2022), (Challita *et al*., 2018), (Torab-Miandoab *et al*., 2023), (Torres *et al*., 2022)
P14. Applying transformation rules defined in the interface	(Brilhault *et al*., 2022), (Challita *et al*., 2018), (Jepsen *et al*., 2020), (Torres *et al*., 2022)
P15. Preserving the integrity of the information	(Fraga *et al*., 2020), (Horcas *et al*., 2023), (Torres *et al*., 2022), (Turk, 2020)
P16. Executing actions defined in the interface	(Haile and Altmann, 2018), (Jepsen *et al*., 2020), (Torres *et al*., 2022)

Figure 2. Supports of the axiomatic theory
Source: Authors

Proposition 2: Defining the purpose of interoperation

It is important to identify the purpose and what will be exchanged between the two software systems in order to be prepared for establishing interoperability.

Proposition 3: Identifying the duration of collaboration

The duration of the collaboration between two software systems must be defined, as well as the frequency of interoperations.

Proposition 4: Determining the domain terminology

In order to establish semantic interoperability between two software systems in different domains, it is important to identify the differences between their terminologies. This can be done by using an ontology, a glossary, software documentation, *etc.*

Figure 3 presents a template to provide domain information for interoperability purposes.

Organization name	
Area	(e-government, commerce, manufacturing, *etc.*)
Entity name	(name of the organization: company, firm, *etc.*)
Location	(city and country)
Activity	(*e.g.*, software provider, manufacturing company, agriculture fame...)
Purpose	(information to be exchanged: ex: billing document)
System involved	(software or information system involved at both sides)
Duration	(precise the duration: long term (years), short term (weeks) or just one session
Responsibility	(name of the person at both sides responsible for this interoperability)
Constraint	(any particular constraint(s) for this interoperability)

Figure 3. Interoperability domain template
Source: Authors

AXIOM 2. INTEROPERABILITY PROFILE – Exchanging software profile for interoperability

When two software systems seek to interoperate, they should exchange information about their software profiles in order to predict the possibility of interoperation.

Proposition 5: Describing software system profiles for interoperability

Software profiles for interoperability must be first identified and described. These are sets of characteristics describing some particular aspects of a software system relevant for interoperability, such as the coding language, the operating system, the protocol for communication, *etc*.

A template for documenting such characteristics is proposed in Figure 4.

Software system profile	
Name:	(*Name for identifying the software system—e.g., MS Excel, WhatsApp, etc.*)
Version:	(*Number for identifying the current operational software product—e.g., V 1.0.2*)
Programming language:	(*List of software system development languages—e.g., HTML, JavaScript, Python, etc.*)
Communication protocols:	(*List of the communication protocols used by the software system—i.e., http, TCP/IP, etc.*)
Operating system:	(*List of the operating systems supporting the software system—e.g., Windows 10, Mac OS X, etc.*)
Platform:	(*List of devices supporting the software system—e.g., pc, tablet, etc.*)
DBMS:	(*List of DBMSs used by the software system—e.g., PostgreSQL, MySQL, etc.*)
Information exchange format:	(*List of languages used for exchanging information—e.g., XML, Json, Atlas, etc.*)
Communication interfaces available:	(*List of interfaces for exchanging information—e.g., Export/Import Data from*)

Figure 4. Interoperability profile of a software system
Source: Authors

Proposition 6: Comparing software system profiles for interoperability

Exchanging and comparing profiles of two software systems allows knowing whether they are interoperable (according to the differences between their profiles) and consequently searching for an appropriate solution to establish interoperability.

Proposition 7: Implementing an approach to establish interoperability

Incompatibility between two software system profiles leads to failures of interoperability. Therefore, an appropriate approach (*i.e.*, integrated, unified, and federated) (International Organization for Standardization, 2011) should be implemented to establish interoperability.

An integrated approach means using a common format and language in each system. A unified approach implies building a neutral format at the meta level for mapping two systems. A federated approach involves dynamically negotiating and mapping two systems 'on the fly' (*i.e.*, without a pre-defined format/language). The choice depends on the context and the requirements of interoperability.

AXIOM 3. INTEROPERABILITY CONTEXT – Interpreting exchanged information using a common context

Exchanged information can be correctly understood when both software systems have a common context.

Proposition 8: Locating the context

> Before interoperating, both software systems should define and agree on the minimum level of detail regarding each context necessary to understand the information they are going to exchange.

Proposition 9: Describing the context

A prerequisite for establishing interoperability is to describe the context of each software system as a list of software system elements, their attributes, and their rules.

Proposition 10: Agreeing on a common context

Agreeing on a common context consists of identifying relationships between the elements of the source and the target software systems.

A common context should include the relationships stating the direct matching of the elements in both contexts, the necessary rules for transforming source elements into target elements, and rules describing how to interpret particular elements of the source software system.

Without a common context between source and target software systems, the understanding and use of information is uncertain, and it could be misinterpreted. Such lack of agreement on the exchanged information can generate conflicts in the target software system, *i.e.*, business rules violations, loss of consistency in the information, and lack of reliability with regard to the information.

AXIOM 4. INTEROPERABILITY LANGUAGE – Encoding information for interoperability

> An exchange language is used for encoding information to be sent to another software system. The exchange language should be recognized by the two software systems that are going to interoperate.

Proposition 11: Encoding the information to be exchanged in an exchange language operable by the target software system

> Information is written in the source software system by using an exchange language recognized by the target software system. The source software system has one or several exchange languages used for transmitting information.

Proposition 12: Defining responsibility regarding the transformation of the information

> During the interoperation between two heterogeneous software systems, the exchanged information should be transformed into a language/format understandable by the target software system. The responsibility of this transformation should be defined before interoperation, and it may be assigned to the source or to the target software system.

AXIOM 5. INTEROPERABILITY INTERFACE– Establishing interface for interoperability

> An interface is needed for establishing a relationship between the source and the target software systems during interoperation. A relationship between said systems implies linking their elements. When an interface is used, a clear identification of the actions to be executed in the transmission of information is required.

Proposition 13: Establishing relationship of interoperability by using an interface

> An interface is an established agreement between two software systems. This agreement includes specifications, *e.g.*, what language will be used during the exchange, what languages/translators are available for understanding the information, what are the contexts of both software systems, how the common context for the two software systems should be established, what transformation rules are necessary, what conditions are imposed in the elements of each context, and what actions are needed to perform the exchange.

Proposition 14: Applying transformation rules defined in the interface

> Transformation rules are descriptions about the way in which source software system elements are related to target software system elements. Transformations rules seek to describe, in a formal language, the following equivalences: (i) element to element, when an element of source software system is a representation of the same entity in the contexts of the two systems; (ii) attribute to attribute, when attributes of elements representing the same entity or different entities refer to the same property; and (iii) element to attribute, when an attribute is represented as an element in the other context.

Proposition 15: Preserving the integrity of the information

> Conditions in the source software context are constraints regarding the way in which the elements interact within the software system. Conditions are conveyed by using an interface to preserve the integrity of the information. A decision about the need to respect the conditions should be made in the target software system.

Proposition 16: Executing actions defined in the interface

> The interface contains actions seeking to establish interoperability between the source and the target software systems. These actions include (a) establishing communication with the target software system; (b) agreeing on the common context of both software systems; (c) transforming information from a language (chosen by the source software system) to a language recognized by the target software system; and (d) conveying the information to the target software system.

In addition, adequate data must be used for performing these actions (required data), and the results must be delivered as a dataset (provided data).

Results

Eleven experts were invited to review the proposed axiomatic theory and answer a previously approved questionnaire. The participants met the following criteria: experience in software interoperability (*i.e.*, university professors, researchers, consultants, and professionals), work related to interoperability (*i.e.*, interoperability research, software quality, interoperability projects for industry, and formal methods for software engineering), years of experience (two or more), and academic degree (*i.e.*, Master and PhD). The results regarding the level of agreement on a five-point Likert scale (strongly agree, agree, neutral, disagree, and strongly disagree) are presented in Figure 5. Table 3 also summarizes the results concerning some indices used for declaring consensus.

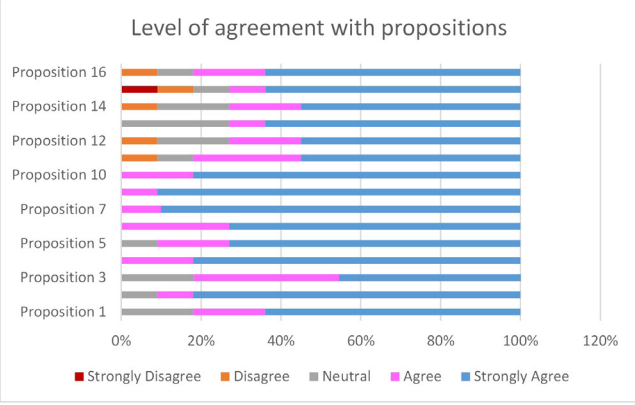


Figure 5. Expert consultation results
Source: Authors

Consensus with all propositions can be declared because there are positive results in at least three of the evaluated indices (Table 2). Nevertheless, the concerns of the experts with propositions 12-15 can be discussed based on the results in the interquartile range (IQR) measure. The expert concerns on proposition 12 are related to the transformation of the information. The experts considered that such a responsibility should be addressed in the interface or in a third system. However, we believe that such a decision should be assigned within an explicit or implicit negotiation between the two systems.

Regarding propositions 13 and 14, some experts referred to the need for a means to exchange information (corresponding to the proposed definition of interface) in the form of APIs, database links, cloud servers, web servers, *etc.* The conception of an interface as a pre-established agreement was also the main concern of other experts. Such an agreement can be unilateral, as in the case of the aforementioned mechanisms (an analysis from a practical point of view could be necessary). In addition, some experts hesitated on the need for transformation rules, as systems sometimes have equivalent profiles and contexts, and interoperability can be directly established.

As for proposition 15, the expert concerns mainly revolve around the time when the preservation of the information should happen (before or after transformation).

Table 3. Evaluated indices for consensus (von der Gracht, 2012)

Variable	LQ	UQ	IQR	Median positive	Mean	4-5%	Standard deviation (SD)	Coefficient of variation (CV)	APMO
P1	*4*	*5*	*1*	*100%*	*4,45*	*82%*	*0,82*	*0,16*	*81,82*
P2	*5*	*5*	*0*	*100%*	*4,73*	*91%*	*0,65*	*0,13*	*90,91*
P3	*4*	*5*	*1*	*100%*	*4,27*	*81%*	*0,79*	*0,20*	*81,82*
P4	*5*	*5*	*0*	*100%*	*4,82*	*100%*	*0,40*	*0,08*	*100,00*
P5	*4*	*5*	*1*	*100%*	*4,64*	*91%*	*0,67*	*0,13*	*90,91*
P6	*4*	*5*	*1*	*100%*	*4,73*	*100%*	*0,47*	*0,09*	*100,00*
P7	*5*	*5*	*0*	*100%*	*4,90*	*100%*	*0,32*	*0,06*	*100,00*
P9	*5*	*5*	*0*	*100%*	*4,91*	*100%*	*0,30*	*0,06*	*100,00*
P10	*5*	*5*	*0*	*100%*	*4,82*	*100%*	*0,40*	*0,08*	*100,00*
P11	*4*	*5*	*1*	*91%*	*4,27*	*82%*	*0,10*	*0,02*	*90,91*
P12	*3*	*5*	*2*	*91%*	*4,18*	*73%*	*1,08*	*0,22*	*81,82*
P13	*3*	*5*	*2*	*100%*	*4,36*	*73%*	*0,92*	*0,18*	*72,73*
P14	*3*	*5*	*2*	*91%*	*4,18*	*73%*	*1,08*	*0,22*	*81,82*
P15	*3*	*5*	*2*	*82%*	*4,09*	*73%*	*1,45*	*0,29*	*90,91*
P16	*4*	*5*	*1*	*91%*	*4,36*	*82%*	*1,03*	*0,21*	*90,91*
Reference values for consensus: LQ>=3; UQ=5; IQR<=1; Median Positive >= 80%; Mean ±5; 4-5% > 50%; SD ± 1,64; CV<=0,5; APMO >=69,7									

Source: Authors

Regarding propositions 13 and 14, some experts referred to the need for a means to exchange information (corresponding to the proposed definition of interface) in the form of APIs, database links, cloud servers, web servers, *etc.* The conception of an interface as a pre-established agreement was also the main concern of other experts. Such an agreement can be unilateral, as in the case of the aforementioned mechanisms (an analysis from a practical point of view could be necessary). In addition, some experts hesitated on the need for transformation rules, as systems sometimes have equivalent profiles and contexts, and interoperability can be directly established.

As for proposition 15, the expert concerns mainly revolve around the time when the preservation of the information should happen (before or after transformation).

Illustrative example

To provide an example, the case of a software development company dedicated to supporting the needs related to the technology and processes of SMEs was considered. This company seeks to develop software based on Microsoft

technologies (*i.e.*, Web, desktop, and mobile applications and service-oriented software). Their main customers are in the judicial, real estate, and CRM sectors.

Specifically, the example corresponds to the software systems of two companies in the commerce and financial sectors. The purpose of interoperability has to do with the target software system's need (the commerce company) for obtaining customer financial information. The source software system (the financial company) only provides such information to authorized partners. The target software system should commit to respecting the regulatory policies regarding information management, as well as the rules pertaining to the permanence of the information, among others.

AXIOM 1. INTEROPERABILITY DOMAIN

According to propositions 1 and 2, the preliminary information of the two organizations is summarized in Figure 6. Based on proposition 3, the duration of the collaboration is limited to the duration of the agreement between the two organizations. The frequency of interoperability is expected to be at least once a day. Additionally, the necessary domain terminology is provided in the form of a regulation document, as well as the API documentation of the source software system, according to proposition 4. This terminology is related to terms such as information holder, information source, and information operator, among others.

Enterprise 1		**Enterprise 2**	
Area	Commerce union	**Area**	Financial
Entity name	Reserved	**Entity name**	Reserved
Location	Medellín, Colombia	**Location**	Colombia
Activity	Commercial activity support	**Activity**	Managing the credit and financial information of individuals
Purpose	Consulting the credit and financial information of potential customers	**Purpose**	Exchanging credit and financial information
System involved	Credit assignment Software	**System involved**	ENTERPRISE 2 Web Application
Duration	Long term	**Duration**	long term
Responsibility	Not provided	**Responsibility**	Not provided
Constraint	---	**Constraint**	No public information available

Figure 6. Interoperability domain, case study
Source: Authors

AXIOM 2. INTEROPERABILITY PROFILE

According to proposition 5, the interoperability profile is filled with the available information, which is shown in Figure 7. However, the main requirements for interoperability are determined in a unilateral agreement using a web service provided by the source software system. Hence, it is not necessary to exchange some information.

Software system profile	**Name**: Credit Assignment Software	**Name**: ENTERPRISE 2 Web Application
Version	--	--
Programming language	Netbeans 8.2 JSP HTML, CSS	C# .Net framework®
Communication protocols	HTTPS	HTTPS
Operating system	Windows Server®	Windows Server®
DBMS	SQL Anywhere® SQL Server®	SQL Server®
Information exchange format	JSON	JSON
Platform	Linux, windows	Windows
Communication interfaces available	Apache Axis 2	REST APIs SOAP

Figure 7: Software interoperability profile, case study
Source: Authors

According to proposition 6, the software system profiles allow identifying a popular standard description for applications intended to consume and deploy web services. This kind of standardization facilitates communication between heterogeneous software systems. Based on the two software system profiles, interoperability between them might be possible. The method employed is more similar to the *unified* approach, according to proposition 7, as there is a neutral format known by both software systems, albeit not jointly agreed.

AXIOM 3. INTEROPERABILITY CONTEXT

According to proposition 8, the necessary details of the context for interoperating both software systems is conditioned by the web service provided, *i.e.*, the source software system only delivers the pre-determined information about the customers, as requested by the target software system, and it should be inferred according to the format of the exchanged JSON file. The source software context (Figure 8) is hidden from the target software system. According to proposition 10, the common context is defined on the target side during the integration of the exchanged information (*i.e.*, for creating and updating information of the customers).

AXIOM 4. INTEROPERABILITY LANGUAGE

According to proposition 11, the exchange language is JSON, which is recognized by both software systems. The source software system is programmed for exchanging a JSON file. The standard structure of this language is commonly recognized and involves context. Transforming the exchanged information is the responsibility of the target software system, according to proposition 12. This responsibility is implicitly defined in the interoperability agreement.

AXIOM 5. INTEROPERABILITY INTERFACE

According to proposition 13, the two organizations have agreed to establish an interface via the web service provided by the source software system. This system seeks to produce an identity certificate and a public key that provides access to partner companies. Settings should be accepted by the target software system in order to achieve communication, exchange, common understanding, and accessibility to the information, according to proposition 16. As stated in proposition 14, the transformations are applied according to the JSON structure, which maps the information in objects of the target software system's model.

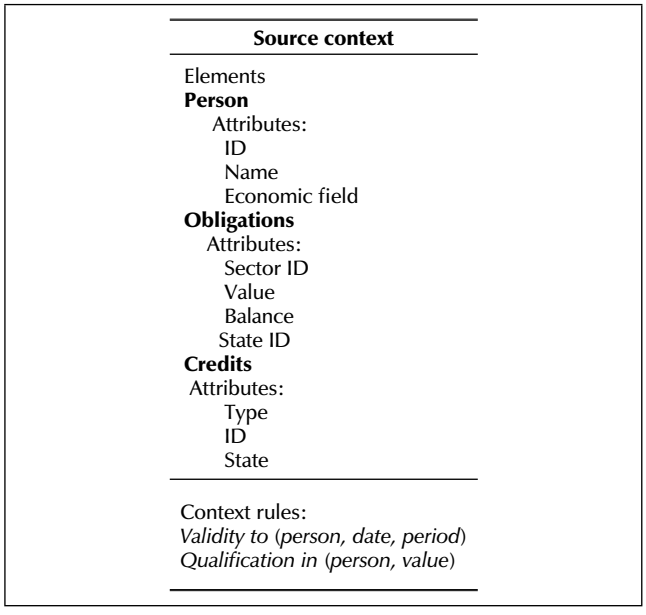


Figure 8. Source software context, case example.
Source: Authors

According to proposition 15, the target software system should respect the conditions contained in the source software system's documentation, given that financial information is sensitive to the precise moment in which the exchange happens, as well as to the interpretation given by the source software system. For example, the expiration period of the report is only seven days. In addition, a positive rating for a company or individual means that all obligations are up to date.

Discussion

The proposed explanation of what software interoperability is and how it works leads to recognizing the interaction between its seven essential elements. The lack of a unified terminology leads to different definitions of interoperability, depending on the approaches and type of interoperability addressed. A definition using the seven essential elements allows characterizing issues with a complete view of all the elements and actors involved. Moreover, a common characterization is useful to compare, discuss, and evaluate different approaches of interoperability.

The above-presented axioms are descriptions of what is needed for achieving interoperability. These descriptions can constitute a roadmap for addressing an interoperability project between organizations. This, with regard to (i) determining the necessities of interoperability (*i.e.*, intention, purpose, duration, and frequency); (ii) collecting information about software systems' interoperability features and profiles as a feasibility study; and (iii) implementing an approach to establish interoperability. The propositions are a reference for understanding the necessary arrangements made between software systems, which should be solved in order to achieve interoperability. The proposed axiomatic theory seeks to analyze and discuss approaches of interoperability and is useful for identifying issues in this regard, as well as their causes.

Conclusions and perspectives

This paper proposes an axiomatic theory (five axioms and sixteen propositions) of interoperability between heterogeneous software systems. This axiomatic theory is based on the seven essential elements of interoperability (and their relationships). An interoperability domain template and software system profiles are also proposed. Additionally, the notions of context and language, regarded as essential elements but lacking in the state of the art, are developed in their corresponding axioms and propositions.

As a result of this research, the proposed axiomatic theory allows explaining how interoperability happens and how it can be achieved. Consultation with some experts allowed validating the pertinence and consistency of the proposal. Finally, the application of the theory in a case study regarding a software development company showed the relevance and feasibility of the approach.

By using this example, it was determined that the relationships between essential elements are appropriately stated for identifying and describing interoperability issues and their solutions. Thus, the compliance of the propositions and the completeness and expressiveness of the axiomatic theory can be verified.

To conclude, the axiomatic theory presented herein allows (i) understanding what interoperability is and how it can be established, (ii) describing interoperability issues, and (iii) recognizing the source of said issues and addressing their solutions.

As future work, our axiomatic theory could be used for characterizing open issues of interoperability and identifying practices and activities intended to achieve it. Moreover, using this theory for creating measures and measurement methods of interoperability can provide the state of the art with added value. Finally, the axiomatic theory could be applied for collecting data and creating methods in different stages of software system development, such as planning,

analysis, implementation, deployment, and maintenance, seeking to predict, detect, and solve interoperability issues. All this, with the aim to advance the maturity of software system interoperability.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (original draft, writing, review, and editing), data curation.

References

Burns, T., Cosgrove, J., and Doyle, F. (2019). A review of interoperability standards for industry 4.0. *Procedia Manufacturing*, *38*(1), 646-653. https://doi.org/10.1016/J.PROMFG.2020.01.083

Brilhault, Q., Yahia, E., and Roucoules, L. (2022). Qualitative analyses of semantic interoperability approaches: toward learning based model transformations. *IFAC-PapersOnLine*, *55*(10), 2348-2353. https://doi.org/10.1016/J.IFACOL.2022.10.059

Burzlaff, F., Wilken, N., Bartelt, C., and Stuckenschmidt, H. (2019). Semantic interoperability methods for smart service systems: A survey. *IEEE Transactions on Engineering Management*, *69*(6), 4052-4066. https://doi.org/10.1109/TEM.2019.2922103

Challita, S., Zalila, F., and Merle, P. (2018). *Specifying semantic Interoperability between heterogeneous cloud resources with the FCLOUDS formal language* [Conference presentation]. IEEE International Conference on Cloud Computing, San Francisco, CA, USA. https://doi.org/10.1109/CLOUD.2018.00053

Davies, J., Welch, J., Milward, D., and Harris, S. (2020). A formal, scalable approach to semantic interoperability. *Science of Computer Programming*, *192*(1), 102426. https://doi.org/10.1016/J.SCICO.2020.102426

Delgado, A., Calegari, D., González, L., Montarnal, A., and Benaben, F. (2020, January 7-10). *Towards a metamodel supporting e-government collaborative business processes management within a service-based interoperability platform* [Conference presentation]. 53rd Hawaii International Conference on System Sciences, Maui, HI, USA. https://doi.org/10.24251/HICSS.2020.246

Fraga, A. L., Vegetti, M., and Leone, H. P. (2020). Ontology-based solutions for interoperability among product lifecycle management systems: A systematic literature review. *Journal of Industrial Information Integration*, *20*(1), 100176. https://doi.org/10.1016/J.JII.2020.100176

Haile, N., and Altmann, J. (2018). Evaluating investments in portability and interoperability between software service platforms. *Future Generation Computer Systems*, *78*(1), 224-241. http://dx.doi.org/10.1016/j.future.2017.04.040

Horcas, J.M., Pinto, M., and Fuentes, L. (2023). A modular metamodel and refactoring rules to achieve software product line interoperability. *Journal of Systems and Software, 197*(1). https://doi.org/10.1016/J.JSS.2022.111579

International Organization for Standardization (2008). *Software engineering – Software product quality requirements and evaluation (SQuaRE) quality model (ISO/IEC 25010:2008)*. ISO.

International Organization for Standardization (2011). *Advanced automation technologies and their applications -- Requirements for establishing manufacturing enterprise process interoperability -- Part 1: Framework for enterprise interoperability (ISO 11354:2011)*. ISO.

Jepsen, S.C., Worm, T., Mork, T. I., and Hviid, J. (2021, Jun 3-3). *Industry 4.0 middleware software architecture interoperability analysis* [Conference presentation]. 3rd International Workshop on Software Engineering Research and Practices for the IoT (SERP4IoT), Madrid, Spain. https://doi.org/10.1109/SERP4IoT52556.2021.00012

Lafourcade, P., and Lombard-Platet, M. (2020). About blockchain interoperability. *Information Processing Letters*, *161*(1), 105976. https://doi.org/10.1016/J.IPL.2020.105976

Liu, L., Li, W., Aljohani, N. R., Lytras, M. D., Hassan, S. U., and Nawaz, R. (2020). A framework to evaluate the interoperability of information systems – Measuring the maturity of the business process alignment. *International Journal of Information Management*, *54*, 102153. https://doi.org/10.1016/J.IJINFOMGT.2020.102153

Mistry, P., Maguire, D., Chikwira, L., and Lindsay, T. (2022). *Interoperability is more than technology*. The King's Fund. https://www.kingsfund.org.uk/sites/default/files/2022-09/Interoperability%20is%20more%20than%20technology%20report.pdf

Ribeiro, E. L. F., Vieira, M. A., Claro, D. B., and Silva, N. (2018, March 19-21). *Transparent interoperability middleware between data and service cloud layers* [Conference presentation]. CLOSER, Funchal, Portugal. https://doi.org/10.5220/0006704101480157

Sana, K., Hassina, N., and Kadda, B. B. (2021, April 9-11). *Towards a reference architecture for interoperable clouds* [Conference presentation]. 2021 8th International Conference on Electrical and Electronics Engineering, Antalya, Turkey. https://doi.org/10.1109/ICEEE52452.2021.9415944

Serapião, G. Guédria, W., and Panetto, H. (2019). An ontology for interoperability assessment: A systemic approach. *Computer in Industry*, *16*(1), 100100. https://doi.org/10.1016/j.jii.2019.07.001

Sjøberg, I. K., Dybå T., Anda B. C. D., and Hannay, J. E. (2008). Building theories in software engineering. In F. Shull, J. Singer, and D. I. K. Sjøberg (Eds.), *Guide to Advanced Empirical Software Engineering* (pp. 312-336). Springer. https://doi.org/10.1007/978-1-84800-044-5_12

Software Engineering Group (2007). *Guidelines for performing systematic literature reviews in software engineering (EBSE-2007-001)*. School of Computer Science and Mathematics, Keele University. https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

Tall, D. (2004). Building theories: The three worlds of mathematics. *For the Learning of Mathematics*, *24*(1), 29-32.

Torab-Miandoab, A., Samad-Soltani, T., Jodati, A., and Rezaei-Hachesu, P. (2023). Interoperability of heterogeneous health information systems: a systematic literature review. *BMC Medical Informatics and Decision Making*, *23*, 18. https://doi.org/10.1186/s12911-023-02115-5

Torres, D. M., Villavicencio, M. K., and Zapata, C. M. (2022). Representing interoperability between software systems by using pre-conceptual schemas. *International Journal on Electrical Engineering and Informatics, 14*(1), 101-127. https://doi.org/10.15676/ijeei.2022.14.1.7

Torres, D. M., Villavicencio, M. K., and Zapata, C. M. (2018, August 29-31). *Towards a terminology unification in software interoperability* [Conference presentation]. 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Prague, Czech Republic. https://doi.org/10.1109/SEAA.2018.00083

Turk, Z. (2020). Interoperability in construction – Mission impossible? *Developments in the Built Environment*, *4*(1), 100018. https://doi.org/10.1016/j.dibe.2020.100018

von der Gracht, H. A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technological Forecasting and Social Change*, *79*(8),1525-1536. https://doi.org/10.1016/j.techfore.2012.04.013

Research Article / Electrical, Electronic and Telecommunications Engineering

http://doi.org/10.15446/ing.investig.99972

Design and Control of a Photovoltaic Distribution System Based on Modular Buck-Boost Converters

Diseño y control de un sistema de distribución fotovoltaico basado en convertidores Buck-Boost modulares

Harrynson Ramírez-Murillo[1], Carlos A. Torres-Pinzón[2], Fabián Salazar-Cáceres[3], Valentina Vera-Saldaña[4], and Carlos J. Mojica-Casallas[5]

ABSTRACT

The main contribution of this research is the design of a series hybrid topology for a photovoltaic distribution system using Buck-Boost converter modules. This design incorporates a maximum power point tracking (MPPT) algorithm based on the perturb and observe method, linear PI controllers, and an energy management algorithm. The controllers' design is validated through simulation using PSIM and SISOTOOL/MATLAB. This work aims to achieve active power sharing in the AC grid through a control loop implemented with a three-phase inverter. The validation of the topology and controller design demonstrates tracking and robustness in four test scenarios for the state variables in microgrids: constant and variable irradiance conditions, auxiliary storage device (ASD) protection and control loops, and power sharing with the AC grid, while considering the DC system dynamics.

Keywords: solar energy, DC-DC power converters, DC-AC power converters, microgrids, energy management systems

RESUMEN

La principal contribucion de este trabajo de investigacion es el diseño de una topologıa hıbrida en serie para un sistema de distribucion fotovoltaico, empleando modulos de convertidores Buck-Boost. Este diseño incorpora un algoritmo de seguimiento del punto de maxima potencia basado en el metodo perturbar y observar, controladores PI lineales y un algoritmo de gestion de energıa. El diseño de los controladores se valida mediante simulacion en PSIM y SISOTOOL/MATLAB. El objetivo de este trabajo es el reparto de potencia activa en la red mediante un lazo de control implementado en un inversor trifasico. La validacion de la topologıa y el diseño de controladores demuestra seguimiento y robustez en cuatro escenarios de prueba para las variables de estado en microrredes: condiciones de irradiancia constantes y variables, lazos de control y proteccion de dispositivos de almacenamiento auxiliar (ASD) y reparto de potencia en la red AC considerando la dinamica de los sistemas DC.

Palabras clave: energıa solar, convertidores de potencia DC-DC, convertidores de potencia AC-DC, microrredes, sistemas de gestion de energıa

Received: January 13th 2021
Accepted: June 31st 2023

Introduction

Non-conventional renewable energy sources (NCRES) are necessary for modern electricity systems, both in the National Interconnected System (NIS) and in non-interconnected zones (NIZ). This type of energy is indispensable for sustainable economic development, the reduction of greenhouse gases (GHG), and the energy security of the system. The latter has great relevance in countries of the Pacific and Caribbean Region, as the El Niño phenomenon and the effects of climate change tend to be increasingly relevant elements for the planning and operation of electrical power systems. Likewise, renewable energies contribute directly to Goal 7 of the Sustainable Development Goals (SDGs), which is related to *Affordable and Clean Energy* (Swain and Karimu, 2020).

Several reviews on AC and DC microgrid operation, applications, modeling, and control strategies have been presented in the literature Shahgholian (2021); Dragičević and Li (2018); Quintero-Molina *et al.* (2020); Gao *et al.* (2019). An advantage of using these networks is that they provide more reliability to the system, since they allow power generation to take place closer to the load centers. In addition, this increases the quality of the energy supplied. Likewise, microgrids can operate in two modes: island and normal (connected to the network). Considering the vulnerability of electrical power systems, the island mode provides safety Luna *et al*. (2020).

DC microgrids have additional benefits, and they include reduced stages in the AC-DC interfaces required for residential loads. Both AC and DC microgrids operate autonomously, controlling parameters within the established limits. However, DC microgrids have a complex control architecture, mainly due to instantaneous power balance. The most important factors in their design are the generation-load power balance and the battery's state of charge (SOC) Hernández *et al.* (2021).

NCRES such as wind and solar photovoltaic (PV) generation could serve as a backup during El Niño events. There is a complementarity between hydroelectric resources and NCRES during drought periods. However, the massive implementation of NCRES constitutes drawbacks for operating electrical power systems according to their

[1] Universidad de La Salle, Bogotá D.C., Colombia. E-mail:haramirez@unisalle.edu.co
[2] Universidad Santo Tomás, Bogotá D.C., Colombia. E-mail:carlostorresp@usta.edu.co
[3] Universidad de La Salle, Bogotá D.C., Colombia. E-mail:jfsalazar@unisalle.edu.co
[4] Electrical engineer, ENEL, Colombia. E-mail: valentina.vera@utp.edu.co
[5] Universidad Santo Tomás, Bogot á D.C., Colombia. E-mail:carlosmojica@usta.edu.co

power and distribution levels, given the randomness of these primary resources. For a massive expansion of these distributed energy resources (DERs), it is essential to develop novel control functionalities that mitigate the large-scale impacts of these technologies Luna *et al.* (2020). In Hernández *et al.* (2021), an islanded microgrid for a rural area was designed and simulated in PSCAD/EMTDC. This design considered a battery to store energy when solar irradiance is low. In this research work, the SOC of the battery design uses the incremental conductance algorithm.

Rajanna (2022) designed a system consisting of a PV system with maximum power point tracking (MPPT), a battery storage system, and an improved three-level neutral point clamped (NPC) inverter. The NPC inverter accurately generates the required AC voltage via a modulation method for a three-level vector, particularly in the face of unbalanced DC voltages. This work evaluated the performance of the NPC inverter in different configurations, considering the SOC characteristics under varying solar irradiance levels.

Srikanth Goud and Sekhar (2023) presented a PV system connected to a DC-DC converter, which was then connected to a DC-link bus and converted to AC using an inverter before being fed into the grid. An optimization technique called *cuckoo search optimization* (*CSO*) *MPPT* maximized the generated power. The system was simulated in MATLAB/Simulink. The results indicate that the CSO approach outperforms the perturb and observe (P&O) and particle swarm optimization (PSO) approaches.

A Buck-Boost configuration with inductors coupled to the current control loop of a hybrid fuel cell system was shown in Ramírez-Murillo (2015). Despite the slow dynamics of proton-exchange membrane fuel cell (PEMFC) systems, this article is relevant because it considers the same scheme as DC-DC modular converters.

Restrepo *et al.* (2011) and Restrepo *et al.* (2012) presented the small-signal model and the experimental closed-loop transfer functions of the selected modular Buck-Boost converter. These works ease the simulation model by providing the gains and bandwidths for its inner current loops.

The main contribution of this research is a series hybrid (SH) topology design based on Buck-Boost converter modules, considering a PV source while implementing an MPPT algorithm based on the P&O method, synthesizing several linear PI controllers for the inner loops of a distributed generation system. These topology and control designs are validated through simulations on PSIM and SISOTOOL/MATLAB. The process is carried out on a 48 V and 768 W regulated DC distribution bus, designing a control loop for power-sharing to the AC grid while employing a three-phase inverter. The simulation model obtained for this microgrid is validated through the most relevant scenarios, testing the tracking and robustness properties of each loop.

This article is organized as follows. The Methodology section shows the DC-DC and DC-AC converters and the selected hybrid topology and the auxiliary storage device (ASD). Moreover, a design criterion is proposed for the energy management algorithm, and a linear state-space model for the converter is defined. In the Results section, a simulation is performed for a set scenarios, testing whether the state variables are suitable operation points. Then, a discussion regarding the results is presented. In the final section, the conclusions and future works of this research are summarized.

Methodology

This section shows the design of the controllers that allow for energy management in the proposed microgrid (Figure 1). It consists of modular Buck-Boost converters and an ASD device, supervised via a master control. The control strategies for the solar array, the voltage regulation in the DC bus, and the control loops of the ASD system were tuned using SISOTOOL/MATLAB, whose design parameters are based on a minimum gain margin of 6 dB and a phase margin in the $[30, 60]°$ range Ramírez-Murillo (2015).

To regulate the DC bus, this study considered the SH topology described in Ramírez-Murillo *et al.* (2014). Unlike that work, this research used an array of solar panels as a primary source, in which an MPPT algorithm was applied in order to ensure maximum power extraction. It is essential to highlight that the development of an on-grid-type microgrid requires the design of DC voltage control loops, in conjunction with controlling the active power injected into the grid via an inverter.

Modular converter

A Buck-Boost converter is a switched-mode converter whose output voltage can be higher, lower, or equal to the input voltage, as it can operate in three different modes: 1) *Buck mode*, where the input voltage is higher than the output voltage ($V_g > V_o$); 2) *Boost mode*, where the input voltage is lower than the output voltage ($V_g < V_o$); and 3) *Buck-Boost mode*, where the input voltage is approximately equal to the output voltage ($V_g \approx V_o$) Ramírez-Murillo (2015). Among the advantages of this type of converter are a high power conversion efficiency and the capability to control input/output currents.

The converter consists of a Boost stage in cascade, connected with a Buck stage with magnetically coupled inductors, with an included damping network. The scheme of a Buck-Boost converter with current sensors is shown in Figure 2. The damping branch consists of a series of resistances R_d and a capacitor C_d connected in parallel to the capacitor C_F Bendib *et al.* (2015).

The averaged model of the converter is represented by a set of differential equations that consider the damping network and the turns ratio $n = 1$, as described in Erickson and Maksimovic (2007), where d_1 and d_2 correspond to the cycles of the switches Q_1 and Q_2. Likewise, the overline in the variables corresponds to the average during a switching period. Likewise, linearizing the set of differential equations, in conjunction with the small-signal state vector given by Equation (1), the dynamic model is shown in Equation (2), where A is the state matrix, and B_1 and B_2 are the input vectors corresponding to $\tilde{d}_1$ and $\tilde{d}_2$. The state matrices are presented in Equation (3) Restrepo *et al.* (2011).

$$\tilde{x} = \begin{bmatrix} \tilde{i}_{Lm} & \tilde{i}_L & \tilde{v}_C & \tilde{v}_{Cd} & \tilde{v}_o \end{bmatrix}^T \tag{1}$$

$$\frac{d\tilde{x}}{dt} = A\tilde{x} + B_1\tilde{d}_1 + B_2\tilde{d}_2 \tag{2}$$

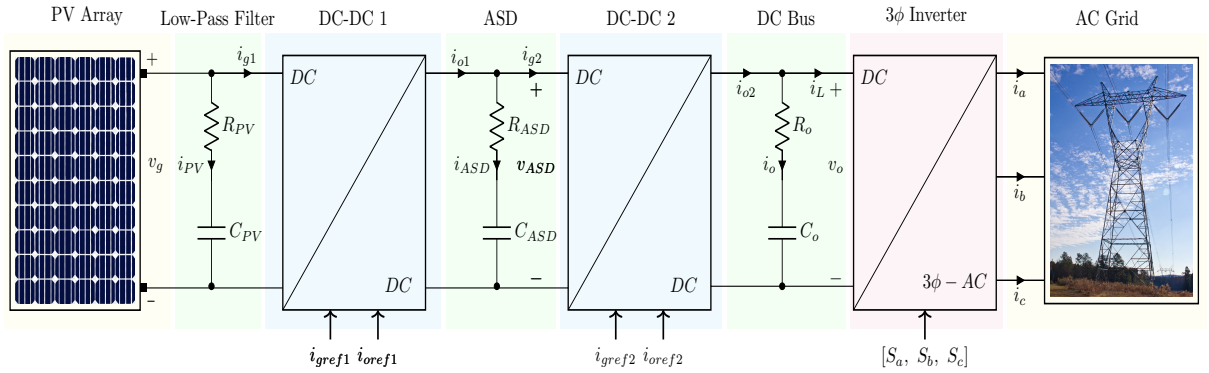


Figure 1: Proposed PV distribution system

Source: Authors

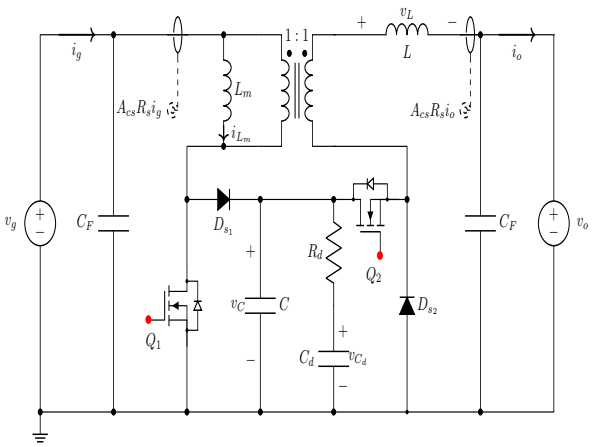


Figure 2: Buck-Boost converter schematic circuit

Source: Authors

$$A = \begin{bmatrix} 0 & 0 & \frac{D_1-1}{L_m} & 0 & 0 \\ 0 & 0 & \frac{D_2+D_1-1}{L} & 0 & \frac{-1}{L} \\ \frac{1-D_1}{C} & \frac{1-D_2-D_1}{C} & \frac{-1}{R_dC} & \frac{1}{R_dC} & 0 \\ 0 & 0 & \frac{1}{R_dC_d} & \frac{-1}{R_dC_d} & 0 \\ 0 & \frac{1}{C_o} & 0 & 0 & \frac{-1}{R_oC_o} \end{bmatrix} \quad (3)$$

$$B_1 = [\frac{-V_g}{L_m(D_1-1)} \quad \frac{-V_g}{L(D_1-1)} \quad \frac{-V_gD_2^2}{CR_o(D_1-1)^2} \quad 0 \quad 0]^T$$

$$B_2 = [0 \quad \frac{-V_g}{L(D_1-1)} \quad \frac{V_gD_2}{CR_o(D_1-1)} \quad 0 \quad 0]^T$$

The transfer functions of the small-signal converter from the current control to the input/output current loops are presented in Figure 3. The corresponding parameters are $\tau_a = 3,3\ \mu s$, $\tau_b = 33\ \mu s$, $K_1 = 590\ (sA)^{-1}$, and $K_2 = 1298\ (sA)^{-1}$ Restrepo *et al.* (2012).

The control variable $\tilde{u}$ shown in Figure 3 refers to the duty cycles of the MOSFETs of the modular converter, as seen in Equation (4).

$$\tilde{u} = \begin{cases} \tilde{d}_1(t) & \text{Boost Mode} \\ \tilde{d}_2(t) & \text{Buck Mode} \end{cases} \quad (4)$$

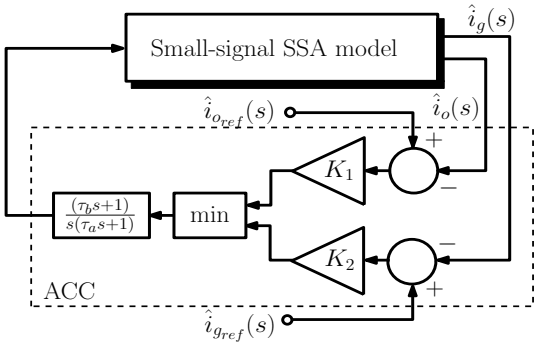


Figure 3: Block diagrams of the average current-mode control (ACC)

Source: Authors

DC bus voltage control loop design

In addition to the small-signal current control loop described in the previous subsection, it is necessary to add a voltage control loop to regulate the converter input/output voltages in each module. The closed-loop transfer function is presented in Equation (5), where ω_{ci} is the natural frequency, whose value is $2\pi 8$ krad/s; and ζ is the damping factor, which corresponds to 0,44 Ramírez-Murillo (2015). These values were selected while aiming for the lowest damping factor and bandwidth values Restrepo *et al.* (2011).

$$H_i(s) = \frac{\tilde{i}_o(s)}{\tilde{i}_{oref}(s)} = \frac{\omega_{ci}^2}{s^2 + 2\zeta\omega_{ci}s + \omega_{ci}^2} \quad (5)$$

The European Cooperation for Space Standardization (ECSS) provides an output impedance mask for the DC bus criterion, which allows for a 1 % variation in the voltage v_o with a 50 % load disturbance i_o. This impedance is presented in Equation (6) Mourra *et al.* (2008).

$$Z_{DCBus} \leq \frac{0,01v_o}{0,5i_o} = \frac{0,02v_o}{i_o} = \frac{(0,02)(48)V}{16A} = 60\ m\Omega \quad (6)$$

The value of C_o is calculated using Equation (7).

$$\frac{1}{w_oC_o} \leq Z_{DCBus}. \quad (7)$$

Here, ω_o corresponds to the bus voltage loop cutoff frequency, where a small ESR value is obtained. A parallel

capacitor array was proposed in order to obtain a capacity of $2,35\ mF$, with a series resistance of $11,8\ m\Omega$. In addition, voltage regulation in the DC bus is performed by the voltage control loop shown in Figure (Figure 4).

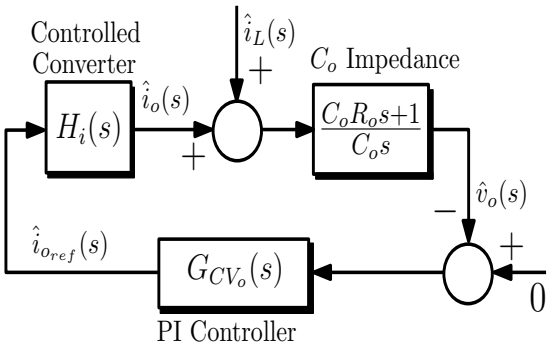


Figure 4: Small-signal model of the DC bus impedance

Source: Authors

In the block diagram of Figure Figure 4, there is an algebraic simplification multiplying the numerator of the impedance C_o by the additional pole of the controller $G_{CV_o}(s)$. The DC bus impedance's closed-loop was adjusted on SISOTOOL with a bandwidth of $1,29\ kHz$, obtaining a proportional gain of $20,894\ kAs/V$ and a time constant τ equal to $890\ \mu s$.

Auxiliary storage device

A simplified lithium-ion battery (LIR18650, $2,6\ Ah$ model) was used in this work EEMB (2010). This model took into account only the nominal region of the discharge capacity, from Q_{top} to Q_{nom} values. Here, the energy management algorithm always works between the voltage ranges of E_{top} and E_{nom}. The main advantage of this model is that it is easy to obtain the parameters from the datasheet provided by the manufacturer. In the linear region, the battery model can be modeled as an RC series branch with a SOC in the [20, 80] % range, where the C-value models the charging and discharging behaviors. The discharge behavior in the nominal region is required in order to size the energy storage. The area under the curve was obtained through Equation (8).

$$\begin{aligned} Energy &= \frac{E_{top}-E_{nom}}{2}(Q_{nom}-Q_{top})Capacity(Ah) \\ Energy &= \frac{3,8V+3,5V}{2}(0,8-0,2)(2,6Ah) = 5,694\ J \end{aligned} \tag{8}$$

The value of the capacitor for one cell was calculated via Equation (9).

$$\begin{aligned} E &= \frac{C(E_{top}^2-E_{nom}^2)}{2} \\ C &= \frac{(5,694J)(2)}{(3,8^2-3,5^2)V} = 5,2\ F \end{aligned} \tag{9}$$

However, to obtain the reference voltage of 48 V, an array of six branches connected in parallel was proposed. Each branch contains 14 cells connected in series. The capacitor value was obtained from Equation (9) while the series resistor value comes from the datasheet provided by the manufacturer EEMB (2010). The equivalent RC series branch for the selected ASD is presented in Equation 10.

$$\begin{aligned} C_{ASD} &= \frac{5,2}{14}(6) = 2,228\ F \\ R_{ASD} &= \frac{(70m\Omega)(14)}{6} = 163\ m\Omega \end{aligned} \tag{10}$$

ASD control loop design

The battery impedances were calculated for the following loops: 1) a nominal or reference SOC (20 %), 2) a minimum SOC (4 %), and 3) a maximum SOC (80 %). This is presented in Equation (11), considering the criteria established by the ECSS.

$$\begin{aligned} Z_{v_{ASDref}} &\leq \frac{0,02v_{ASDref}}{i_{ASD}} = \frac{(0,02)(3,5V)(14)}{16} = 61,25\ m\Omega \\ Z_{v_{ASDmin}} &\leq \frac{0,02v_{ASDmin}}{i_{ASD}} = \frac{(0,02)(3V)(14)}{16} = 52,5\ m\Omega \\ Z_{v_{ASDmax}} &\leq \frac{0,02v_{ASDrmax}}{i_{ASD}} = \frac{(0,02)(3,8V)(14)}{16} = 66,5\ m\Omega \end{aligned} \tag{11}$$

The controller parameters were obtained according to the previously mentioned DC bus impedance mask. The nominal and minimum SOC loops were adjusted while considering a bandwidth of $0,00307\ kHz$, a proportional gain of $11,488\ As/V$, and a time constant τ of $2,9\ s$. The maximum SOC loop was set with a bandwidth of $0,00318\ kHz$, a proportional gain of $6,9632\ As/V$, and a time constant τ of $5\ s$.

Maximum power point tracking loop

The panel filter has the same RC parameters $C_o = C_{PV}$ and $R_o = R_{PV}$ as the DC bus. However, the value of the impedance mask changes, as seen in Equation (12).

$$Z_{v_g} \leq \frac{0,02v_{gmin}}{i_{PV}} = \frac{(0,02)(60V)}{16A} = 37,5\ m\Omega \tag{12}$$

The PI controller parameters were adjusted via SISOTOOL with a $0,903\ kHz$ bandwidth, given a proportional gain of $11,802\ kAs/V$ and a time constant τ of 1,1 ms. On the other hand, the MPPT control loop was tuned considering the P&O algoritm, where, if $\Delta V > 0$, then the voltage is increased by $0,1\ V$. On the other hand, if $\Delta V < 0$, then the voltage is decreased by $0,1\ V$ Mohan *et al.* (2003).

Power sharing algorithm

A three-phase inverter is a key device for interfacing a DC system with an AC grid. The switching of the insulated-gate bipolar transistors (IGBTs) is performed by comparing each of the three current and voltage sine waves, which are obtained after the pulse-width modulation (PWM). These waves have a 120° phase shift, with the same triangular carrier signal and switching frequency of $15\ kHz$. Likewise, the output amplitude is determined by the ratio of the signal carrier to the modulator. The inverter output voltage requires a step-up transformer with a Delta-Star connection and a grounded neutral wire. Thus, the inverter will be able to inject any current level into the grid. On the other hand, the $230\ V$ line's voltage regulation is achieved by modulating the pulse width to obtain the switching signals [S_a, S_b, S_c]. The design criteria are described below.

A series RL first-order low-pass filter is required to eliminate the transmission line current ripple. Its transfer function is

presented in Equation (13). These parameters correspond to R_{abc} and L_{abc}.

$$G_{abc}(s) = \frac{1}{R_{abc}L_{abc}s + 1} = \frac{1}{0,0003183s + 1} \tag{13}$$

The Park transform allows converting the sinusoidal values in time into steady-state values, and it is applied to both the filtered currents and the grid voltages. It is important to consider that the phase of the grid starts at $t = 0\,s$. Likewise, this transform reduces the harmonic content of the output voltage in the inverter.

Active power control is carried out by calculating the reference currents through the network voltages in a steady state. The current references for the PI regulator are obtained via Equation (14) from the active and reactive power reference values, which corresponds to $P_{ref} = (16)(48)\ W = 768\ W$ and $Q_{ref} = 0\ VAR$, respectively, and the voltage references v_d and v_q.

$$\begin{aligned} i_{dref} &= \frac{2}{3}\frac{P_{ref}V_d + Q_{ref}V_q}{V_d^2 + V_q^2} \\ i_{qref} &= \frac{2}{3}\frac{P_{ref}V_q + Q_{ref}V_d}{V_d^2 + V_q^2} \end{aligned} \tag{14}$$

The PI controller was adjusted using the SmartCtrl tool of PSIM. The GM of the PI controller and the BW were set to $45,33°$ and $478,63\ Hz$, respectively.

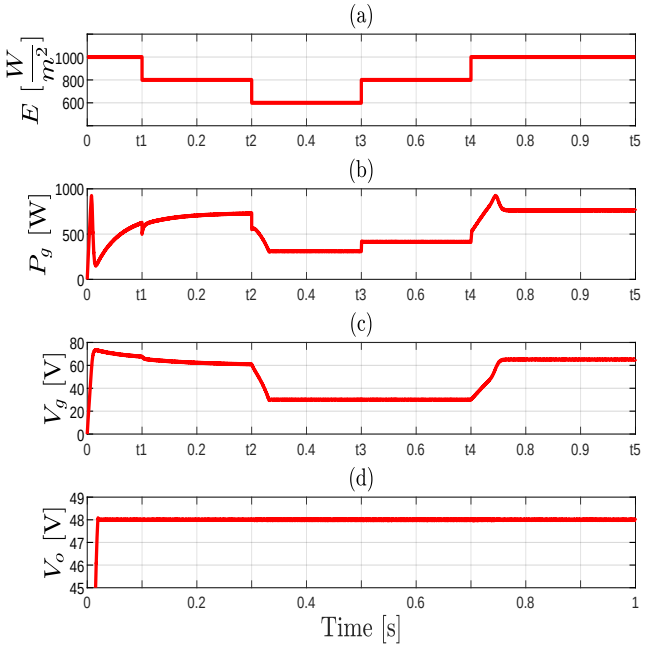


Figure 5: (a) Test irradiance profile, (b) input power, (c) PV array voltage, and (d) DC bus voltage

Source: Authors

Results

Constant load under variations in the irradiance profile

In this subsection, three control loops are designed and validated: first, the DC bus voltage is regulated at $V_{oref} = 48\ V$ using a reference current i_{oref2}; second, the panel minimum voltage is regulated at $V_{gmin} = 30\ V$ through i_{gref1} and the MPPT algorithm; finally, the ASD $V_{ASDref} = 48\ V$ nominal voltage is also controlled by applying the i_{gref1} signal. These control input currents regulate the voltages v_g and v_o in the distribution system shown in Figure 5. In this simulation, a constant commercial load value of $3,3\ \Omega$, $698\ W$ was obtained.

In Figure 5a, an amplitude-varying irradiance profile is depicted, considering values of $[600, 1000]\ W/m^2$. During the interval from 0 to t_1, the PV array fully supplies the power demand, as shown in Figure 5b. However, when the irradiance decreases at $t_1 = 100\ ms$, the array cannot supply all the power required. Therefore, the ASD supplies the power shortage demand. This system continues supplying power from t_1 to t_2, while the power demand increases. Subsequently, in the intervals from t_2 to t_3 and from t_3 to t_4, the PV array power is set to a reference, at 310 and 414 W, which corresponds to irradiance levels of 600 and 800 W/m^2, respectively. Finally, from t_4 to t_5, once the system reaches the maximum irradiance value of 1000 W/m^2, the PV array power stabilizes at a constant power demand. It is observed that, from 0 to t_4, the maximum PV power is supplied. Moreover, from t_4 to t_5, the total power required by the load is delivered by the ASD, and the energy management algorithm prioritizes the ASD nominal voltage control loop at $V_{ASDref} = 48\ V$.

In Figure 5c, once the irradiance decreases, the voltage also decreases, until it stabilizes at $V_{gmin} = 30\ V$, as shown for the time interval from t_2 to t_4, which regulates the PV array voltage at the minimum operating value. On the other hand, from t_4 to t_5, when the irradiance increases to 1000 W/m^2, the panel voltage reaches 65 V in order to provide the total demanded power. In Figure 5d, the DC bus voltage behavior is shown at the reference value of $V_{oref} = 48\ V$ and under a $3,3\ \Omega$ constant load, maintaining this state regardless of variations in demanded power and irradiance.

Impedance-varying load profile under a constant irradiance condition

In this scenario, two control loops are validated: first, the DC bus voltage regulation loop at $V_{oref} = 48\ V$ through I_{oref2}; and second, the ASD nominal voltage regulation at $V_{ASDref} = 48\ V$ using I_{gref1} and the MPPT algorithm. These control loops are tested while considering a constant impedance load of $20\ \Omega$, adding a variable impedance load of $3,9\ \Omega$.

Figure 6a shows that, under a constant irradiance of 1000 $[W/m^2]$, the input power is stabilized at 410 W. Moreover, in Figure 6b, the output voltage V_g is stabilized through the MPPT algorithm, reaching the maximum power point and with no significant increases or decreases in the input voltage. Figure 6c validates the DC bus voltage regulation loop. Finally, in Figure 6d, there are DC bus voltage ripples of $1\ V$ throughout $0,1\ s$. This disturbance in the load represents a demanding scenario, with $76,92\ \%$ of the maximum power, which corresponds to $768\ W$.

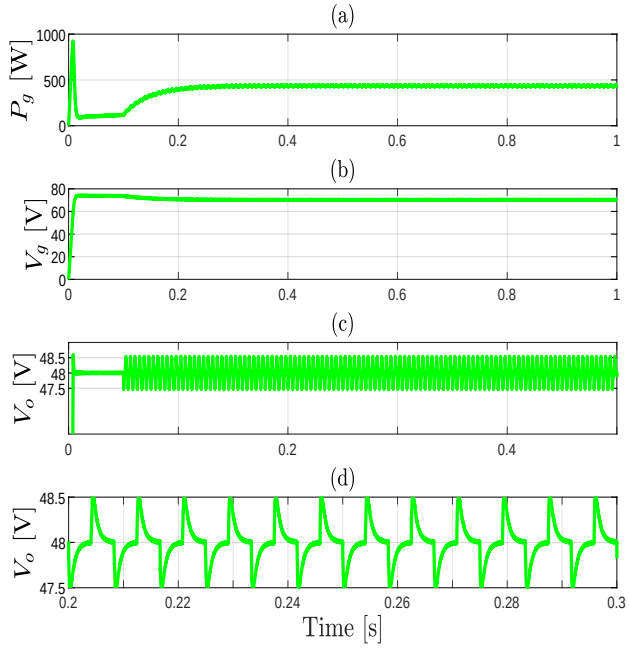


Figure 6: (a) Input power; (b) PV array voltage; (c) DC bus voltage, large signal model; and (d) small-signal variations of the DC bus voltage

Source: Authors

ASD loops scenario test

In this scenario, three control loops are validated: the DC bus voltage, which has been widely considered in the previous subsections; the ASD minimum voltage at $V_{ASDmin} = 42\ V$ using I_{gref2}; and the ASD maximum voltage at $V_{ASDmax} = 53,2\ V$ using I_{oref1}, which implies an MPPT algorithm failure.

Figure 7a shows a saturation of the current reference $I_{gref1} = 16\ A$, with an initial constant load of $3,2\ \Omega$ until $t = 0,1\ s$. Then, from $t = 0,1$ to $t = 1,1\ s$, the maximum ASD voltage V_{ASDmax} exhibits the required performance, while a load shedding is deployed, considering a constant load of $100\ \Omega$. On the other hand, from $t = 1,1\ s$ onward, the load is fully reconnected and shows an adequate behavior in the current output. Finally, Figure 7b shows the DC bus voltage drop V_{oref}, while the ASD minimum reference voltage V_{ASDmin} achieves the reference value in Figure 7c.

Inverter considering the DC system dynamics

In this scenario, the dynamics of the DC system are included using a square signal, whose maximum and minimum values correspond to 768 and $115,2\ W$, respectively. This profile is shown in Figure 8a, whuch includes the active power control loop mentioned in the previous section. Figures 8b and 8c illustrate the three-phase voltages ($V_g = [v_a,\ v_b,\ v_c]^T$) and currents ($I_g = [i_a,\ i_b,\ i_c]^T$) after considering the DC dynamics.

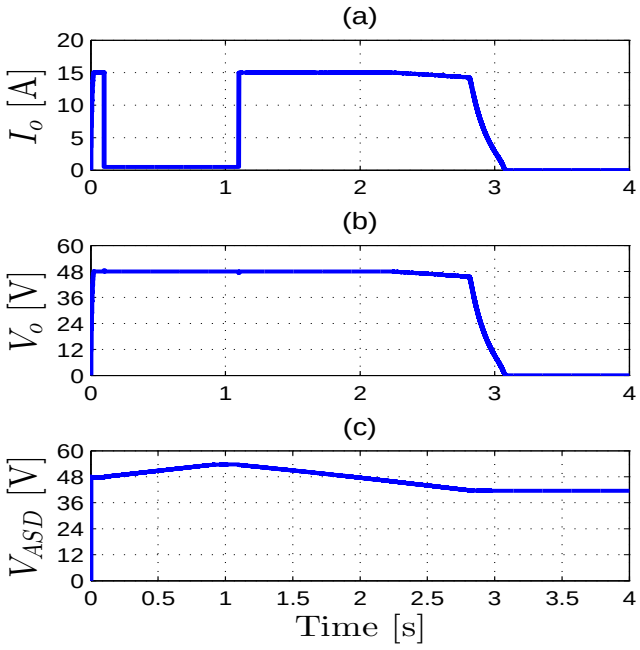


Figure 7: (a) DC bus current, (b) DC bus voltage, and (c) battery voltage

Source: Authors

Discussion

In the previous section, the results obtained for the microgrid were compared with the findings of another study Rajanna (2022), where it was observed that the maximum and minimum SOC of the ASD were not regulated, thus requiring the incorporation of additional protection loops. Furthermore, another recent work Srikanth Goud and Sekhar (2023) highlighted the absence of an ASD. To address these limitations, this study focused on establishing parameters for the protection and operation loops of the ASD. Additionally, an energy management algorithm with an increased number of control loops was designed, using a versatile modular Buck-Boost converter. To regulate the voltages of the buses while considering the power variations, the design of the control loops in the DC stage employed the impedance mask criterion. Moreover, the injection of active power into the AC grid was regulated through the dq model of the inverter. As a remark, the P&O algorithm was integrated into the control loop that regulates the voltage V_g.

Other recent works Zhou *et al.* (2022); Shang *et al.* (2020) explored the optimization of ASDs in microgrids. These studies proposed a novel approach that integrates machine learning techniques with power management algorithms to enhance the efficiency and performance of ASD. Reinforcement learning techniques proved to be effective through simulations and experimental validations, exhibiting significant improvements in energy management and system performance Ali *et al.* (2021).

Conclusions

The main contribution of this research is the design of a series hybrid (SH) topology for a PV distribution system, using two modular Buck-Boost converters.

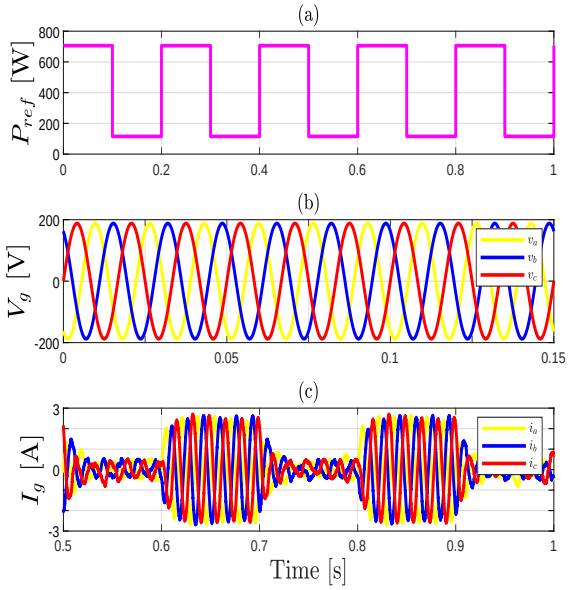


Figure 8: (a) Reference power from DC dynamics, (b) three-phase voltages, and (c) three-phase currents

Source: Authors

These converters allow designing an energy management algorithm, where MPPT, nominal SOC, and maximum ASD SOC control loops are considered for the first module. Likewise, the second stage ensures the minimum SOC and the DC bus voltage reference, as shown in the simulation results. Four test scenarios were proposed to validate the correct operation of the system control loops.

The ASD simplified model drives a suitable transient and shows a steady-state behavior. As a result, when the microgrid is connected to the AC grid (on-grid), the designed energy management scheme allows supplying the demand under irradiance variations. The power injected into the AC grid is controlled using a three-phase inverter, but its waveform is distorted. Therefore, it is recommended that a harmonic compensation device be introduced, such as passive or active filters, line reactors, and hybrid filters, among others.

The impedance mask criterion allows appropriately designing classical controllers for voltage regulation in DC buses, wherein the impedance is modeled. Likewise, the modular converter used in this research facilitates voltage regulation, as it allows for input/output current control with a higher bandwidth. The interaction between these loops does not entail any conflict, given that their control hierarchy goes hand in hand with the lowest energy consumption in the system.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation,writing (original draft, writing, re-view, and editing), data curation.

Acknowledgements

This work was partially funded by the Research and Transfer Vice-Principalship (VRIT) of Universidad de La Salle, through the research project under institutional code IELE221-211, as well as by Universidad Santo Tomás through grant project no. BOG-2023-AI015.

References

Ali, K. H., Sigalo, M., Das, S., Anderlini, E., Tahir, A. A., and Abusara, M. (2021). Reinforcement learning for energy-storage systems in grid-connected microgrids: An investigation of online vs. offline implementation. *Energies*, *14*(18), 5688. https://doi.org/10.3390/en14185688.

Bendib, B., Belmili, H., and Krim, F. (2015). A survey of the most used mppt methods: Conventional and advanced algorithms applied for photovoltaic systems. *Renewable and Sustainable Energy Reviews*, 45:637–648. https://doi.org/10.1016/j.rser.2015.02.009.

Dragičević, T. and Li, Y. (2018). Ac and dc microgrid control. In *Control of Power Electronic Converters and Systems*, pages 167–200. https://doi.org/10.1016/B978--0--12--816136--4.00018--X. Elsevier.

EEMB (2010). *Lithium-ion battery data sheet.* EEMB Co.

Erickson, R. W. and Maksimovic, D. (2007). *Fundamentals of power electronics*. Springer Science & Business Media.

Gao, F., Kang, R., Cao, J., and Yang, T. (2019). Primary and secondary control in dc microgrids: a review. *Journal of Modern Power Systems and Clean Energy*, 7(2):227–242. https://doi.org/10.1007/s40565--018--0466--5.

Hernández, E. D. G., Aldana, N. L. D., and Hernández, A. C. L. (2021). Energy management electronic device for islanded microgrids based on renewable energy sources and battery-based energy storage. *Ingeniería e Investigación*, 41(1):1–5. https://doi.org/10.15446/ing.investig.v41n1.86047.

Luna, A. C., Díaz, N. L., and Narvaez, E. A. (2020). Optimal coordination of active generators in a grid-connected microgrid. *Ingeniería e Investigación*, 40(3):47–54. https://doi.org/10.15446/ing.investig.v40n3.82665.

Mohan, N., Undeland, T. M., and Robbins, W. P. (2003). *Power electronics: converters, applications, and design*. John Wiley & Sons.

Mourra, O., Blancquaert, T., Signorini, C., and Tonicello, F. (2008). European cooperation for space standardization – The new standard for electric and electronic ECSS-E-ST-20-C [Conference presentation]. 8th European Space Power Conference.

Quintero-Molina, V., Blanco, A. M., Romero-L, M., Meyer, J., and Pavas, A. (2020). Power quality in ac islanded microgrids: Technical framework and state of the art review. *Ingeniería e Investigación*, 40(3):29–37. https://doi.org/10.15446/ing.investig.v40n3.89091.

Rajanna, B. V. (2022). Grid connected solar pv system with mppt and battery energy storage system. *International Transactions on Electrical Engineering and Computer Science*, 1(1):8–25, https://iteecs.com/index.php/iteecs/article/view/5.

Ramírez-Murillo, H. (2015). *Diseño, supervisión y control de sistemas híbridos pemfc*. [PhD thesis, Universitat Rovira i Virgili].

Ramírez-Murillo, H., Restrepo, C., Calvente, J., Romero, A., and Giral, R. (2014). Energy management dc system based on current-controlled buck-boost modules. *IEEE Transactions on Smart Grid*, 5(5):2644–2653. https://doi.org/10.1109/TSG.2014.2330198.

Restrepo, C., Calvente, J., Romero, A., Vidal-Idiarte, E., and Giral, R. (2011). Current-mode control of a coupled-inductor buck–boost dc–dc switching converter. *IEEE Transactions on Power Electronics*, 27(5):2536–2549. https://doi.org/10.1109/TPEL.2011.2172226.

Restrepo, C., Konjedic, T., Calvente, J., Milanovic, M., and Giral, R. (2012). Fast transitions between current control loops of the coupled-inductor buck–boost dc–dc switching converter. *IEEE transactions on power electronics*, 28(8):3648–3652. https://doi.org/10.1109/TPEL.2012.2231882.

Shahgholian, G. (2021). A brief review on microgrids: Operation, applications, modeling, and control. *International Transactions on Electrical Energy Systems*, 31:e12885]https://doi.org/10.1002/2050--7038.12885.

Shang, Y., Wu, W., Guo, J., Ma, Z., Sheng, W., Lv, Z., and Fu, C. (2020). Stochastic dispatch of energy storage in microgrids: An augmented reinforcement learning approach. *Applied Energy*, 261:114423. https://doi.org/10.1016/j.apenergy.2019.114423.

Srikanth Goud, B. and Sekhar, G. C. (2023). Cuckoo search optimization mppt technique for grid connected photovoltaic system. *International Transactions on Electrical Engineering and Computer Science*, 2(1):14–19, https://iteecs.com/index.php/iteecs/article/view/3.

Swain, R. B. and Karimu, A. (2020). Renewable electricity and sustainable development goals in the eu. *World Development*, 125:104693. https://doi.org/10.1016/j.worlddev.2019.104693.

Zhou, K., Zhou, K., and Yang, S. (2022). Reinforcement learning-based scheduling strategy for energy storage in microgrid. *Journal of Energy Storage*, 51:104379, https://doi.org/10.1016/j.est.2022.104379.

Research Article / Mechanical Engineering, Mechatronics, and Materials Science http://doi.org/10.15446/ing.investig.93696

A Study on the Performance and Emission Characteristics of a Diesel Engine Operated with Ternary Higher Alcohol Biofuel Blends

Un estudio sobre las características de rendimiento y emisión de un motor diésel operado con mezclas ternarias de biocombustible de alcohol superior

Tayfun Ozgur[1]

ABSTRACT

This study focused experimentally on the evaluation of a higher alcohol biofuel utilization with diesel fuel in a compression ignition engine. To this effect, two different ternary higher alcohol biofuel blends with different percentages by volume were used, i.e., DA5 (85% diesel + 5% butanol +5% pentanol + 5% hexanol) and DA10 (70% diesel + 10% butanol + 10% pentanol + 10% hexanol). The performance and emission results were compared against reference standard diesel fuel. Experiments revealed the brake power (BP) values of the engine to be reduced due to the lower calorific value of the alcohols. In addition, this phenomenon led to increased levels of brake-specific fuel consumption (BSFC), which resulted in a higher fuel consumption to obtain unit power. Regarding emissions, carbon monoxide (CO) was reduced as a result of the additional oxygen content of alcohols, which triggered the more oxidized CO with oxygen. The results regarding unwanted nitrogen oxides (NO_x) emissions were also attractive according to the alcohol utilized. Considerable reductions were obtained with alcohol mixtures when compared to diesel fuel, which can be explained by the reduced combustion temperatures in the cylinder, given the higher latent heat in the evaporation of alcohols.

Keywords: diesel engine, higher alcohol biofuel, performance, emissions

RESUMEN

Este estudio se centró experimentalmente en la evaluación de una mayor utilización de biocombustible de alcohol con combustible diésel en un motor de encendido por compresión. Para ello, se utilizaron dos mezclas ternarias de biocombustible de alcohol superior distintas y con diferentes porcentajes en volumen DA5 (85 % diésel + 5 % butanol + 5 % pentanol + 5 % hexanol) y DA10 (70 % diésel + 10 % butanol + 10 % pentanol + 10 % hexanol). Los resultados de rendimiento y emisiones se compararon con combustible diésel estándar de referencia. Los experimentos revelaron que los valores de la potencia de frenado (BP) del motor se redujeron debido al menor poder calorífico de los alcoholes. Además, este fenómeno llevó a obtener mayores niveles de consumo de combustible específico de frenado (BSFC), lo que resultó en un mayor consumo de combustible para obtener potencia unitaria. Con respecto a las emisiones, el monóxido de carbono (CO) se redujo como resultado del contenido de oxígeno adicional de los alcoholes, lo cual activó el CO más oxidado con oxígeno. Los resultados de la emisión no deseada de óxidos de nitrógeno (NO_x) también fueron atractivos según el alcohol utilizado. Se obtuvieron reducciones considerables con las mezclas de alcoholes en comparación con el combustible diésel, lo que puede deberse a temperaturas de combustión reducidas en el cilindro, dado el mayor calor latente en la evaporación de alcoholes.

Palabras clave: motor diésel, biocombustible con alto contenido de alcohol, rendimiento, emisiones

Received: February 19th, 2021
Accepted: July 03th, 2023

Introduction

Diesel engines are widely used due to their various advantages compared to gasoline engines. The highly efficient, durable, and reliable nature of diesel engines makes them more attractive and preferable (Kumar *et al*., 2020). Miscellaneous sectors such as transportation, agriculture, and power generation need this kind of engine due to its superior features.

Diesel engines have been conventionally powered by fossil-based fuels since their invention, although there have been some attempts for alternative fuel utilization. Unfortunately, fossil fuels have been faced with the scarcity of reality since the late 1970s and early 1980s (Tüccar *et al*., 2014). Therefore, efforts to decrease the dependency on fossil fuels are currently being made by scientists. Alcohols are one of the most impressive alternatives, as they can be directly used as an additive in engines. The production of alcohols

[1] PhD in Mechanical Engineering, Çukurova University, Turkey. Affiliation: Assistant professor, Çukurova University, Automotive Engineering Department, Turkey. E-mail: tozgur@cu.edu.tr

is biologically achieved via fermentation, which is a cost-effective and environmentally friendly method (Yesilyurt, 2020). On the other hand, no engine modification is needed while using alcohols as a primary fuel or fuel additive (Çelebi and Aydın, 2019).

Methanol (CH_3OH) and ethanol (C_2H_5OH) are regarded as lower alcohols, and they have been widely studied in engine experiments by researchers. The extra oxygen content in their structure and their enhanced soot oxidation due to the hydroxyl group (which leads to lower levels of opacity, especially at higher engine loads) can be listed among the prominent advantages of methanol and ethanol (Nour *et al*., 2017, 2019). Despite this common use, these alcohols have some disadvantages, such as their low heating value and high evaporation enthalpy, which causes a weak engine performance, increased carbon monoxide and hydrocarbon emissions, high auto-ignition resistance, long ignition delay periods, high vapor pressure, a low flash point (important for storage), poor lubrication, and weak miscibility with diesel fuel (Nour *et al*., 2019).

Besides lower alcohols, another group that has recently attracted attention is higher alcohols. The term *higher alcohol* indicates series of straight-chain alcohols that contain more than three carbon atoms, such as butanol (C=4), pentanol (C=5), hexanol (C=6), octanol (C=8), dodecanol (C=12), phytol (C=20), *etc.* (No, 2020).

Lately, there has been great interest in higher alcohols due to their from enhanced ignition quality, which can be attributed to the increased carbon chain lengths in their molecular structure (Koivisto *et al*., 2015). Other featured properties of higher alcohols include a high energy density, an increased cetane number, a better miscibility, and a lower hygroscopic tendency (Rajesh Kumar and Saravanan, 2016).

In recent years, alcohols, especially higher ones, have begun to replace conventional fuels, as their usability in internal combustion engines has become increasingly remarkable. Therefore, research has recently focused on this subject.

Pan *et al*. (2020) studied the effects of butanol isomers on diesel engine performance, emissions, and combustion behaviors under varying load conditions. They concluded that cylinder pressure and the heat release rate increased with diesel/butanol blends when compared to diesel fuel. As for emissions, butanol helps to reduce soot emissions, but it causes an increase in nitrogen oxides and carbon monoxide emissions at low loads. At high loads, butanol reduces particulate matter emissions.

Singh *et al*. (2020) used N-butanol as an additive to a eucalyptus biodiesel/diesel mixture in order to observe its effects on engine performance and emissions. They found that the brake power values of all fuel samples were nearly the same. When the butanol concentration was increased in the blends, lower brake specific fuel consumption and higher brake thermal efficiency values were obtained. On the other hand, butanol generated reduced levels of nitrogen oxides and increased levels of carbon dioxide with respect to pure diesel.

Pan *et al*. (2019) experimentally evaluated the spray and combustion characteristics of a compression ignition engine operated with n-pentanol/diesel blends. Results demonstrated that n-pentanol/diesel blends help to obtain better atomization characteristics than diesel fuel. Although the addition of n-pentanol to diesel fuel decreased soot emissions, increments in brake-specific fuel consumption and hydrocarbon and nitrogen oxide emissions were observed.

Radheshyam *et al*. (2020) studied the effects of adding 1-pentanol to diesel, as well as EGR (exhaust gas recirculation) on a common rail direct injection diesel engine. They observed that the engine can be operated with up to 30% 1-pentanol-mixed fuel with no need to modify the engine. Increasing the 1-pentanol percentage in mixtures yields a higher brake-specific fuel consumption and a lower thermal efficiency. Nitrogen oxide emissions decreased in all blends with 1-pentanol, while hydrocarbon and carbon monoxide values increased.

Thomas *et al*. (2020) sought to determine the effects of hexanol and biodiesel produced from waste vegetable oil on the combustion of a reactivity-controlled compression ignition engine. Their experiments revealed that nitrogen oxides and smoke emissions were reduced with biodiesel/hexanol blends when compared to diesel. They finally suggested that biodiesel/hexanol mixtures should replace fossil fuels as an effective solution for reactivity-controlled compression ignition.

Ramesh *et al*. (2019) conducted an experimental study to understand the effects of *Calophyllum inophyllum* biodiesel and hexanol on diesel engine applications. Diesel/biodiesel/n-hexanol blends were prepared. Hexanol in blends caused an increase in brake thermal efficiency and a decrease in specific fuel consumption in comparison with pure biodiesel. While hydrocarbon, smoke, and carbon monoxide emissions were reduced, those of nitrogen oxides and carbon dioxide increased when compared to diesel.

Aneeque *et al*. (2021) studied the effects of alcohols and *Calophyllum inophyllum* biodiesel while employing a response surface optimization methodology. They evaluated the impacts of alcohol (N-octanol, N-butanol) and biodiesel on the performance and emissions of a diesel engine, and they performed a response surface analysis to obtain the optimal engine efficiency values and emission profiles. The researchers finally suggested that hybrid blends of alcohol and biodiesel can be used directly in diesel engines with no need for engine modification.

Khan *et al*. (2020) evaluated the impacts of synthesized waste milk scum oil methyl ester and ethanol blends on diesel engine performance characteristics. At different ratios,

binary blends (20% biodiesel + 80% diesel, 60% biodiesel + 40% diesel, 100% biodiesel) and ternary blends (20% biodiesel + 5% ethanol + 75% diesel, 20% biodiesel + 10% ethanol + 70% diesel, 20% biodiesel + 15% ethanol + 65% diesel) were prepared. The physicochemical properties of the blends were evaluated, and performance tests were conducted on a single-cylinder diesel engine. The authors concluded that the 20% biodiesel + 5% ethanol + 75% diesel ratio yields the best values regarding fuel consumption, the air-fuel ratio, and thermal efficiency.

Although many experimental studies on compression ignition engines operated with lower alcohols (methanol, ethanol) have been published, the combustion characteristics of higher alcohols still require further research. As stated before, higher alcohols have been intensively studied by scientists in recent times. Therefore, this study focused on determining the performance and emission characteristics of a diesel engine fed with diesel fuel and ternary higher alcohol additives (butanol, pentanol, and hexanol).

Methodology

Engine tests were performed at the Automotive Engineering Laboratories of Çukurova University. A naturally aspirated, 4-stroke, 4-cylinder compression ignition engine was employed for conducting performance and emissions tests at full load. The technical specifications of the engine are presented in Table 1.

Table 1. Engine specifications

Brand	Mitsubishi Canter
Model	4D34-2A
Configuration	Inline 4
Type	Direct injection diesel with glow plug
Displacement	3 907 cc
Bore/Stroke	104 mm/115 mm
Power	89 kW @ 3 200 rpm
Torque	295 Nm @ 1 800 rpm
Oil cooler	Water-cooled
Air cleaner	Paper element type
Weight	325 kg

Source: Author

Figure 1 presents a schematic of the engine test rig. The engine test set-up was basically composed of a diesel engine, a hydraulic dynamometer, performance and emissions measurement devices, and a control unit.

Diesel fuel and alcohols (butanol, pentanol, and hexanol) were purchased from a commercial firm in Adana, Turkey. Three different fuels were used in this study. The fuels employed were diesel fuel and diesel plus ternary alcohol mixtures with different ratios. Table 2 summarizes the fuels used in the experiments.

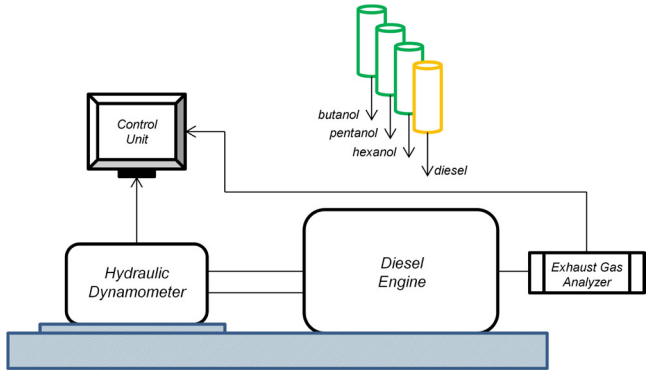


Figure 1. Schematic of the engine test rig
Source: Author

The mixing of diesel and alcohols was achieved with an ultrasonic processor (Sonic Vibra-Cell VC 750 model) in order to homogenize and stabilize the blends before using them in the engine.

Table 2. Fuels and abbreviations

	Diesel (% vol.)	Butanol C_4H_9OH (% vol.)	Pentanol $C_5H_{11}OH$ (% vol.)	Hexanol $C_6H_{13}OH$ (% vol.)
D	100	0	0	0
DA5	85	5	5	5
DA10	70	10	10	10

Source: Author

An IKA-Werke C2000 Calorimeter, a Zeltex ZX 440 NIR petroleum analyzer, a Kyoto Electronics DA-130 density meter, a Tanaka AKV-202 automatic kinematics viscosity meter, a Tanaka MPC-102L pour point analyzer, and a Tanaka APM-7 flash point analyzer were used to determine the heating value, cetane number, density, viscosity, and flash point, respectively. Table 3 shows the physicochemical fuel properties of the D, DA5, and DA10 fuels. The measured properties were also compared to EN 590 standards.

Table 3. Fuel properties

Property	D	Butanol*	Pentanol*	Hexanol*	DA5	DA10	EN 590
Lower Heating Value (MJ/kg)	43,3	33,09	34,65	39,1	41,9	40,1	-
Cetane Number	57,1	17	18,2-20	23	52,2	46,7	Min. 51
Density at 15 °C (kg/m^3)	835	809,7	814,8	821,8	828	823	820 - 845
Viscosity at 40 °C (mm^2/s)	2,45	2,22	2,89	5,32	2,68	2,55	2 - 4.5
Flash Point (°C)	64	35-37	49	59	63	61	Min. 55
Latent Heat of Evaporation (kJ/kg)	270-375	581,4	308,05	486	-	-	

Source: Author, adapted from Rajesh Kumar and Saravanan (2016)

The evaluation of fuel physicochemical properties is a crucial priority in interpreting the combustion, performance, and emission characteristics of an engine.

Results

Brake power (*BP*)

Figure 2 depicts the variations in BP with the engine speed. Diesel fuel showed the best performance. In comparison with reference diesel fuel, the average decrements in BP were 13,06 and 24,25% for DA5 and DA10, respectively. Increased alcohol concentrations in the mixtures caused further decrements in BP levels. As seen in Table 3, the lower calorific values of alcohols and blending them with diesel fuel led to a slight decrement in the heating values of the mixtures. The reduced BP values could be explained by the lower energy content of the mixtures (Appavu *et al.*, 2020; Karabektas and Hosoz, 2009; Sundar Raj and Saravanan, 2011).

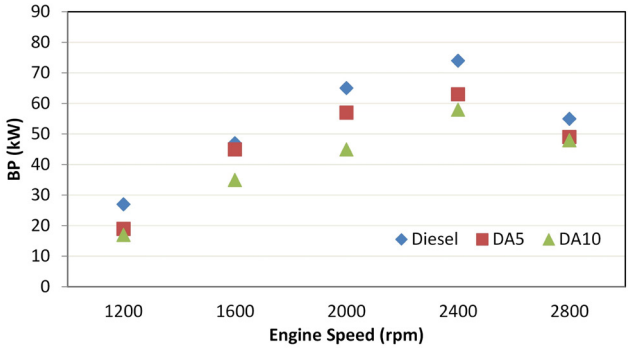


Figure 2. BP *vs.* engine speed
Source: Author

Brake-specific fuel consumption (*BSFC*)

The consumed mass of fuel per unit time can be expressed as the BSFC in order to obtain the unit power. This important parameter is directly related to the engine's fuel economy. Figure 3 shows that there is an increment in BSFC with the diesel-alcohol blending procedure when compared to diesel fuel. On average, 7,98 and 14,57% increases were obtained for DA5 and DA10 in comparison with diesel. One of the possible causes for the increased BSFC might be the lower calorific values of alcohols. Another reason could be the higher latent heat of evaporation, which contributes to extracting more heat from the combustion chamber. This cooling effect in the cylinder can reduce the combustion efficiency (Karabektas and Hosoz, 2009; Sundar Raj and Saravanan, 2011; Yilmaz *et al.*, 2017).

Carbon monoxide (*CO*)

CO appears in engines as an incomplete combustion product, and it is one of the most hazardous and toxic gases released after combustion. As seen in Figure 4, alcohol utilization in the mixtures enhanced the CO emissions characteristics of the tested engine. Compared to diesel, on average, CO decreased by 8,07 and 12,87% for DA5 and DA10, respectively. The oxygen content in the molecular structure of alcohols may trigger a further oxidization of CO molecules. Therefore, a more complete combustion can be achieved, which leads to reduced CO levels (Devarajan *et al.*, 2017; Kattela *et al.*, 2019; Ramesh *et al.*, 2019).

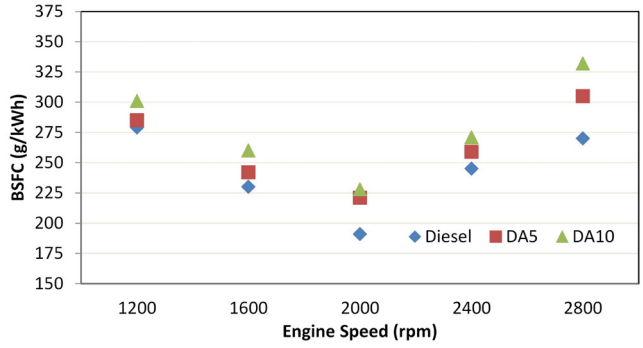


Figure 3. BSFC *vs.* engine speed
Source: Author

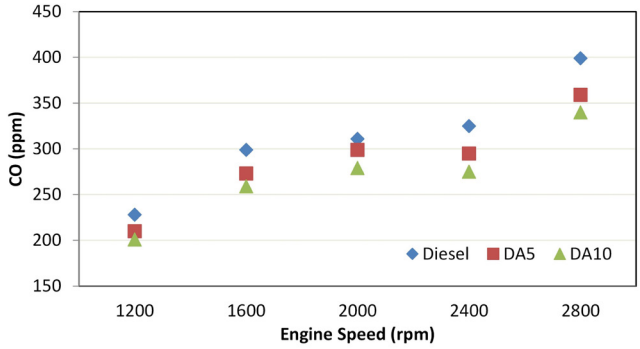


Figure 4. CO *vs.* engine speed
Source: Author

Nitrogen oxides (NO_x)

The burning of fuel occurs in the presence of oxygen. Air contains not only oxygen, but also nitrogen and a small amount of other gases. Moreover, approximately 79% of the air consists of nitrogen. At elevated temperatures in the combustion chamber (~ above 1 600 °C), nitrogen reacts with oxygen to form nitrogen oxides. Therefore, temperature is an important factor in NO_x formation. Figure 5 shows a decrement in NO_x emissions with diesel-alcohol blends when compared to diesel fuel. On average, 8,73 and 14,92% drops were obtained for DA5 and DA10, respectively. These drops may be associated with the high latent heat of evaporation values of the alcohols in the blends. A higher latent heat of evaporation means a higher heat removal from the combustion chamber, which has a cooling effect. This phenomenon has been validated by many researchers in their studies (De Poures *et al.*, 2017; Jeyakumar and Narayanasamy, 2019; Joy *et al.*, 2020).

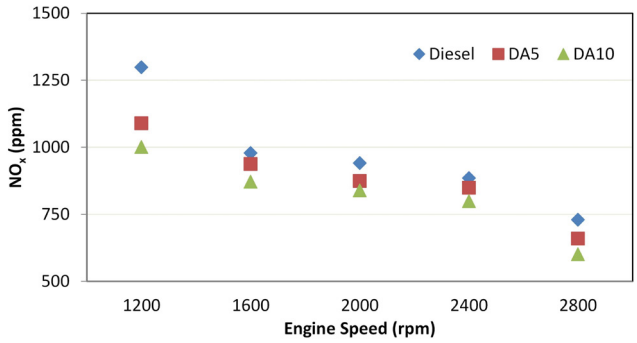


Figure 5. NOx *vs.* engine speed
Source: Author

Conclusions

In this study, the performance and emission characteristics of a diesel engine operated with ternary higher alcohol biofuel blends were analyzed.

From this experimental evaluation, the following conclusions can be drawn:

- The fuel properties of mixtures seem to be suitable for direct use in engines.
- Blending alcohol with fossil fuels contributes to reducing the dependence on conventional fuels, which are becoming scarce.
- Although there was a decreasing trend in performance parameters with alcohol blends, the emission values were improved. If less dependence on conventional fuels is desired and environmental concerns are a priority, using these fuels is recommended.
- The increased alcohol ratio in the blends led to a weaker engine performance, albeit with reduced levels of CO and NO_x emissions.
- The worst BP (due to reduced energy contents) and BSFC values (due to both the lower energy content and the higher latent heat of evaporation of alcohols) were obtained with a 24,25% reduction and a 14,57% increment for DA10 compared to diesel fuel. In addition, the best emission reductions were obtained with DA10 fuel, *i.e.*, 12,87 and 14,92% decrements for CO and NO_x, respectively. A decrease in CO emissions may be obtained by supplying more oxygen to the cylinder with alcohols. On the other hand, the lower NO_x levels can be explained by the higher latent heat of evaporation of alcohol, which lowers the cylinder's internal temperature.

Acknowledgements

This work was supported by the Çukurova University Scientific Research Project Coordination (grant number FBA-2019-12154).

Conflicts of interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Credit author statement

Tayfun Ozgur: conceptualization, investigation, visualization, writing-original draft, and writing-review and editing.

References

Aneeque, M., Alshahrani, S., Kareemullah, M., Afzal, A., Saleel, C. A., Soudagar, M. E. M., Hossain, N., Subbiah, R., and Ahmed, M. H. (2021). The combined effect of alcohols and calophyllum inophyllum biodiesel using response surface methodology optimization. *Sustainability*, *13*(13), 7345. https://doi.org/10.3390/su13137345

Appavu, P., Madhavan, V. R., Venu, H., and Jayaraman, J. (2020). A novel alternative fuel mixture (diesel–biodiesel–pentanol) for the existing unmodified direct injection diesel engine: performance and emission characteristics. *Transactions of the Canadian Society for Mechanical Engineering*, *44*(1), 1-9. https://doi.org/10.1139/tcsme-2019-0049

Çelebi, Y., and Aydın, H. (2019). An overview on the light alcohol fuels in diesel engines. *Fuel*, *236*, 890-911. https://doi.org/10.1016/j.fuel.2018.08.138

De Poures, M. V., Sathiyagnanam, A. P., Rana, D., Rajesh Kumar, B., and Saravanan, S. (2017). 1-Hexanol as a sustainable biofuel in DI diesel engines and its effect on combustion and emissions under the influence of injection timing and exhaust gas recirculation (EGR). *Applied Thermal Engineering*, *113*, 1505-1513. https://doi.org/10.1016/j.applthermaleng.2016.11.164

Devarajan, Y., Munuswamy, D. B., and Nagappan, B. K. (2017). Emissions analysis on diesel engine fuelled with cashew nut shell biodiesel and pentanol blends. *Environmental Science and Pollution Research*, *24*(14), 13136-13141. https://doi.org/10.1007/s11356-017-8915-7

Jeyakumar, N., and Narayanasamy, B. (2019). Effect of pentanol-biodiesel blends on performance and emission characteristics of the diesel engine. *International Journal of Ambient Energy*, *42*(8), 900-906. https://doi.org/10.1080/01430750.2019.1568910

Joy, N., Balan, K. N., Nagappan, B., and Justin Abraham baby, S. (2020). Emission analysis of diesel and butanol blends in research diesel engine. *Petroleum Science and Technology*, *38*(4), 289-296. https://doi.org/10.1080/10916466.2019.1702680

Karabektas, M., and Hosoz, M. (2009). Performance and emission characteristics of a diesel engine using isobutanol-diesel fuel blends. *Renewable Energy*, *34*(6), 1554-1559. https://doi.org/10.1016/j.renene.2008.11.003

Kattela, S. P., Vysyaraju, R. K. R., Surapaneni, S. R., and Ganji, P. R. (2019). Effect of n-butanol/diesel blends and piston bowl geometry on combustion and emission characteristics of CI engine. *Environmental Science and Pollution Research*, *26*(2), 1661-1674. https://doi.org/10.1007/s11356-018-3704-5

Khan, H., Kareemullah, M., Ravi, H. C., Rehman, K. F., Kumar, R. H., Afzal, A., Soudagar, M. E. M., and Fayaz, H. (2020). Combined effect of synthesized waste milk scum oil methyl ester and ethanol fuel blend on the diesel engine characteristics. *Journal of The Institution of Engineers (India): Series C*, *101*(6), 947-962. https://doi.org/10.1007/s40032-020-00605-3

Koivisto, E., Ladommatos, N., and Gold, M. (2015). Systematic study of the effect of the hydroxyl functional group in alcohol molecules on compression ignition and exhaust gas emissions. *Fuel*, *153*, 650-663. https://doi.org/10.1016/j.fuel.2015.03.042

Kumar, V., Singh, A. P., and Agarwal, A. K. (2020). Gaseous emissions (regulated and unregulated) and particulate characteristics of a medium-duty CRDI transportation diesel engine fueled with diesel-alcohol blends. *Fuel*, *278*, 118269. https://doi.org/10.1016/j.fuel.2020.118269

No, S.-Y. (2020). Utilization of pentanol as biofuels in compression ignition engines. *Frontiers in Mechanical Engineering*, *6*, 15. https://doi.org/10.3389/fmech.2020.00015

Nour, M., Attia, A. M. A., and Nada, S. A. (2019). Combustion, performance and emission analysis of diesel engine fuelled by higher alcohols (butanol, octanol and heptanol)/diesel blends. *Energy Conversion and Management*, *185*, 313-329. https://doi.org/10.1016/j.enconman.2019.01.105

Nour, M., Kosaka, H., Sato, S., Bady, M., Abdel-Rahman, A. K., and Uchida, K. (2017). Effect of ethanol/water blends addition on diesel fuel combustion in RCM and DI diesel engine. *Energy Conversion and Management*, *149*, 228-243. https://doi.org/10.1016/j.enconman.2017.07.026

Pan, M., Huang, R., Liao, J., Jia, C., Zhou, X., Huang, H., and Huang, X. (2019). Experimental study of the spray, combustion, and emission performance of a diesel engine with high n-pentanol blending ratios. *Energy Conversion and Management*, *194*, 1-10. https://doi.org/10.1016/j.enconman.2019.04.054

Pan, M., Tong, C., Qian, W., Lu, F., Yin, J., and Huang, H. (2020). The effect of butanol isomers on diesel engine performance, emission and combustion characteristics under different load conditions. *Fuel*, *277*, 118188. https://doi.org/10.1016/j.fuel.2020.118188

Radheshyam, Santhosh, K., and Kumar, G. N. (2020). Effect of 1-pentanol addition and EGR on the combustion, performance and emission characteristic of a CRDI diesel engine. *Renewable Energy*, *145*, 925-936. https://doi.org/10.1016/j.renene.2019.06.043

Rajesh Kumar, B., and Saravanan, S. (2016). Use of higher alcohol biofuels in diesel engines: A review. *Renewable and Sustainable Energy Reviews*, *60*, 84-115. https://doi.org/10.1016/j.rser.2016.01.085

Ramesh, A., Ashok, B., Nanthagopal, K., Ramesh Pathy, M., Tambare, A., Mali, P., Phuke, P., Patil, S., and Subbarao, R. (2019). Influence of hexanol as additive with Calophyllum Inophyllum biodiesel for CI engine applications. *Fuel*, *249*, 472-485. https://doi.org/10.1016/j.fuel.2019.03.072

Singh, R., Singh, S., and Kumar, M. (2020). Impact of n-butanol as an additive with eucalyptus biodiesel-diesel blends on the performance and emission parameters of the diesel engine. *Fuel*, *277*, 118178. https://doi.org/10.1016/j.fuel.2020.118178

Sundar Raj, C., and Saravanan, G. (2011). Influence of hexanol-diesel blends on constant speed diesel engine. *Thermal Science*, *15*(4), 1215-1222. https://doi.org/10.2298/TSCI101001089S

Thomas, J. J., Sabu, V. R., Nagarajan, G., Kumar, S., and Basrin, G. (2020). Influence of waste vegetable oil biodiesel and hexanol on a reactivity controlled compression ignition engine combustion and emissions. *Energy*, *206*, 118199. https://doi.org/10.1016/j.energy.2020.118199

Tüccar, G., Tosun, E., Özgür, T., and Aydin, K. (2014). Diesel engine emissions and performance from blends of citrus sinensis biodiesel and diesel fuel. *Fuel*, *132*, 7-11. https://doi.org/10.1016/j.fuel.2014.04.065

Yesilyurt, M. K. (2020). A detailed investigation on the performance, combustion, and exhaust emission characteristics of a diesel engine running on the blend of diesel fuel, biodiesel and 1-heptanol (C7 alcohol) as a next-generation higher alcohol. *Fuel*, *275*, 117893. https://doi.org/10.1016/j.fuel.2020.117893

Yilmaz, N., Atmanli, A., and Trujillo, M. (2017). Influence of 1-pentanol additive on the performance of a diesel engine fueled with waste oil methyl ester and diesel fuel. *Fuel*, *207*, 461-469. https://doi.org/10.1016/j.fuel.2017.06.093

Research Article/ Mechanical, Mechatronics, and Materials Engineering https://doi.org/10.15446/ing.investig.95685

Upgrade and Modification of a Machine for Micro-Abrasion Wear Testing in Simulated Biological Environments with Oscillatory Motion

Actualización y modificación de una máquina para ensayos de micro abrasión-desgaste en entornos biológicos simulados con movimientos oscilatorios

Diego F. Prieto[1], José L. Caballero[2], Willian A. Aperador[3], and Juan H. Martínez[4]

ABSTRACT

Aiming to evaluate the useful life of biomaterials used in joint prostheses, this study performed different wear tests in stainless steel 316L, a biomaterial used in hip joint replacements. The tests were carried out in a dry medium, with the help of an equipment that was improved regarding some of its characteristics and allows conducting wear tests via the contact of two bodies, one of them being the biomaterial under study and the other one a sphere of a harder material. For the evaluation, a device was developed to change the rotation of the sphere, varying the angle it traveled and the frequency with which it did it. Once the improvements were made to the aforementioned equipment, tests were conducted which involved obtaining wear tracks in order to observe the surface morphology through scanning electron microscopy (SEM) and to measure the length and the width of the tracks, with which the biomaterial wear coefficient was obtained for each case studied. In these tests, the wear coefficient showed variations with respect to the sphere's angle of travel.

Keywords: wear, biomaterial, scanning electron microscopy, reciprocating motion

RESUMEN

Con el objetivo de evaluar la vida útil de los biomateriales utilizados en prótesis articulares, este estudio llevó a cabo ensayos de desgaste en acero inoxidable 316L, un biomaterial usado en prótesis articulares de cadera. Los ensayos se realizaron en seco, con la ayuda de un equipo que se mejoró en algunas características y que permite realizar ensayos de desgaste mediante el contacto de dos cuerpos, siendo uno de estos el biomaterial a estudiar y el otro una esfera de un material más duro. Para la evaluación se desarrolló un dispositivo para cambiar el giro de la esfera, variando el ángulo que recorría y la frecuencia con la que lo hace. Una vez realizadas las mejoras al equipo mencionado, se realizaron ensayos donde se obtuvieron huellas de desgaste para observar la morfología superficial mediante microscopía electrónica de barrido (SEM) y medir el largo y el ancho de las huellas, con los cuales se obtuvo el coeficiente de desgaste del biomaterial para cada caso estudiado. En estas pruebas, el coeficiente de desgaste presentó variaciones con respecto al ángulo de recorrido de la esfera.

Palabras clave: desgaste, biomaterial, microscopía electrónica de barrido, movimiento reciprocante

Received: May 10th, 2021
Accepted: January 15th, 2023

Introduction

Since ancient times, humans have sought to improve their quality of life, a field that has been widely developed thanks to many advances. This, in addition, has made it possible to increase people's life expectancy, allowing them to live long and healthy lives, which represents a challenge for field of healthcare, given that, as one becomes older, diseases tend to emerge which are complex to treat (Gómez and Sabeh, n.d.).

The United Nations Population fund (UNFPA) submitted a report detailing that, by 2050, one in five people will be over 60 old, which also means that there will be more people over 60 than under 15 years old. In addition, it is expected

[1] Mechatronics engineer, Universidad Militar Nueva Granada, Colombia. Master in Mechatronics Engineering, Universidad Militar Nueva Granada, Colombia. Affiliation: Consultant, Management Solutions, Colombia. Email: diegofelipe-prietomorales@gmail.com

[2] Mechatronics engineer, Universidad Militar Nueva Granada, Colombia. Master in Mechatronics Engineering, Universidad Militar Nueva Granada, Colombia. Affiliation: Professor, Universidad Militar Nueva Granada, Colombia. Email: jose.caballero@unimilitar.edu.co

[3] Physicist, Universidad Pedagógica y Tecnológica de Colombia, Colombia. PhD in Engineering, Universidad del Valle, Colombia. Affiliation: Professor, Universidad Militar Nueva Granada, Colombia. Email: william.aperador@unimilitar.edu.co

[4] Mechanical engineer, Universidad Central, Colombia. Master in Mechanical Processes, Universidad Nacional de Colombia. Affiliation: Professor, Universidad Militar Nueva Granada, Colombia. Email: juan.martinezp@unimilitar.edu.co

that the number of inhabitants who are around 80 years of age will multiply by five. Massive longevity has individual implications related to maintaining the quality of life, as well as new social and economic challenges that, if left unsolved, will cause people to lack the optimal conditions for living (BBC, 2012).

These sociological factors have driven a breakthrough in biomaterials, as well as research in this field. Furthermore, technological and medical advances allowed improving surgical techniques, which resulted in a greater use of prostheses, implants, medical systems, and devices that must usually be in contact with body tissues (BBC, n.d., 2012).

Biomaterials must have certain characteristics in order for them to be used in patients. One of the most critical is the compatibility of a biomaterial with the patient and the assurance of a long half-life, a factor that is currently being increased, given that, by prolonging the useful life of the biomaterial, future interventions on the patient are avoided, which are inherently risky (Vallet, 2004).

In light of the above, some factors involved in the extension of the useful life of a biomaterial when in external or internal contact with the body have been highlighted by the literature, *i.e.*, the reduction of friction between the interacting materials – where wear must be minimized – and the implementation of methods to prevent and control the corrosion of materials, such as the deposition of coatings and a correct selection of materials (Caicedo *et al.*, 2013; Rasor, 2009).

Several tests are being carried out to find new biomaterials capable of efficiently withstanding the wear produced by the body's internal fluids and by the friction of the body itself. One of these is micro-abrasion wear testing, which has different variations and involves several systems, aiming to determine the characteristics of the analyzed material.

The purpose of this paper is to present an improvement plan for micro-abrasion wear testing equipment. Different mechanisms were built, adapted, and coupled with the system, and an invertible turning mechanism was incorporated, seeking to recreate typical movement behaviors in joint replacements, specifically the hip.

Calculating micro-abrasive wear

Tribology is a science that studies the friction, wear, and lubrication phenomena of surfaces with relative motion. It is a science that interrelates different fields of study such as chemistry, physics, and mechanics, among others (Losada-Prieto *et al.*, 2001). Tribological tests employ different standards according to the type of test to be performed and the mechanism used to carry it out.

This research analyzes the results obtained from a system used for micro-abrasion tests (Figure 1), with a variant called *fixed-ball*.

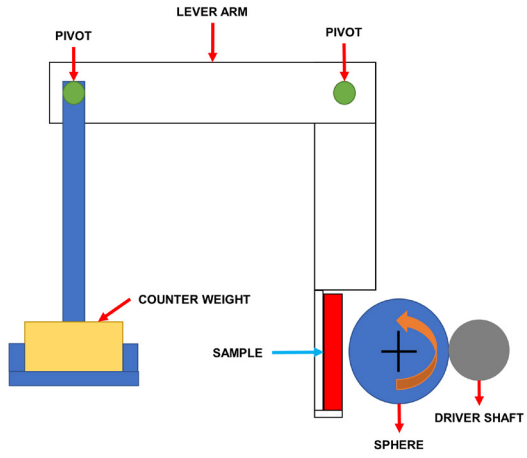


Figure 1. Schematic diagram of the fixed-ball test
Source: Authors

In this test, the sphere rotates in just one direction, and it is clamped against two coaxial transmission shafts that allow it to be easily removed and replaced. The square specimen is pressed by the rotating sphere from the side by means of loads placed on the test stand (Gant and Gee, 2011). The contact between the sphere and the specimen creates friction, which subsequently generates a crater-shaped wear track (Staia *et al.*, 1998).

Normally, micro-abrasion tests, as the name suggests, are carried out with an abrasive fluid, but this type of system can also be used to carry out lubricated or dry sliding tests.

Conventional systems have a limitation when it comes to testing, since the sphere rotating with the sample cannot perform reversals. Thus, its operation clearly involves only one direction at a certain speed. The equipment developed in this research, however, can perform tests with rotation inversion and a certain angle of travel at a desired frequency.

This article presents different mathematical models that allow relating the dimensions of the tracks obtained during testing with a rate of volume loss and a constant wear rate. Therefore, after each test performed, parallel and perpendicular measurements of each track were taken (Figure 2), and an average of these measurements was calculated, since, for each study case, two tests were performed in order to obtain reliable diameter values.

To analyze the measurements of wear tests, some authors state that the wear volume of the sample is given by the following equation, taking into account that $b \ll R$ (Gee *et al.*, 2002; Sampaio *et al.*, 2016; Caballero *et al.*, 2016):

$$V = \pi \frac{b^2}{64R} \quad (1)$$

where R is the radius of the sphere and b is the diameter of the crater generated during the test. Once the wear volume

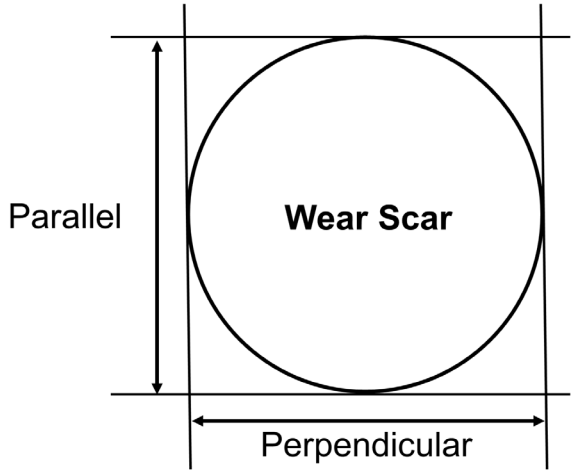


Figure 2. Measurement of wear track
Source: Authors

has been determined, the Archard relation is used, which relates the wear volume to the normal load N and the sliding distance S:

$$V = K_c SN \tag{2}$$

Equating (1) and (2) describes the wear coefficient as:

$$K = \frac{\pi b^4}{64R}\frac{1}{SN} \tag{3}$$

To calculate the sliding distance S for a continuous rotation of the system, the following relation is used:

$$S = \#cycles * 2 * \pi * R \tag{4}$$

When tests are performed with reciprocating rotation, with an angle of travel β expressed in degrees, the sliding distance S is determined as follows:

$$S = \frac{\#cycles * \beta * 4 * \pi * R}{360} \tag{5}$$

Materials and methods

To develop the improvement plan, some aspects were taken from the roadmap for engineering design (Figure 3), which considers different elements that allow carrying out each proposed task.

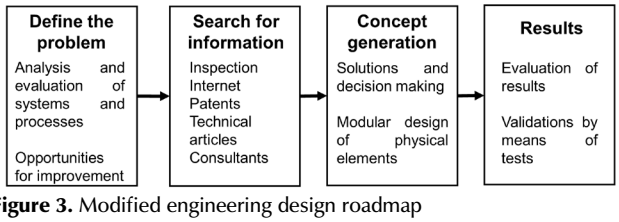


Figure 3. Modified engineering design roadmap
Source: Authors

The redesign sought to improve the existing system. In this process, there can be different tasks aimed to improve the operation of certain components, thus allowing the general assembly to work properly. It is worth noting that the redesign is achieved without any change in the operating principle or the original concept of the evaluated system (Dieter and Schmidt, 2009).

For the project, some parts assemblies were redesigned, and one system was replaced, which improved the functionality of the equipment.

There is no single formula for a viable design or redesign; several authors have different methods, although one of the simplest and most accepted, which is taken as a basis for other models, is the one given by Morris Asimow, who believed that a process can be framed in different cyclical design operations resulting in a product (Dieter and Schmidt, 2009).

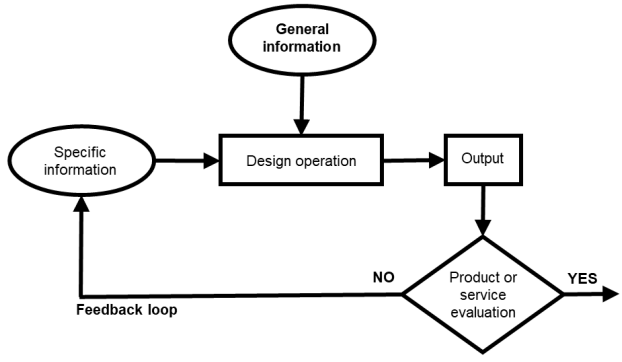


Figure 4. Basic module of the design process
Source: Authors

Based on Figure 4, it can be inferred that design is a sequential process consisting of multiple design operations, where one moves to the next step after evaluating whether what has been done satisfies the required specifications.

Design can be approached as a problem to be solved. The methodology selected for solving the problems involved in the improvement plan designed in this research consist of the following steps (Dieter and Schmidt, 2009):

- Defining the problem
- Collecting information
- Generating solutions and decision-making
- Results (testing)

Experimentation

Defining the problem

To develop the improvement plan, the starting point was a system (Figure 5) for conducting wear and micro-abrasion tests which has a square profile structure with a lever arm,

where there are counterweights that, when moved along the axis that supports them, apply a certain load to the sample under analysis. A dynamometer is used to determine the load exerted by the dead weight and the counterweight. Moreover, there is a sphere that rotates against the sample for a certain number of cycles and at certain velocity, producing wear and abrasion on the sample.

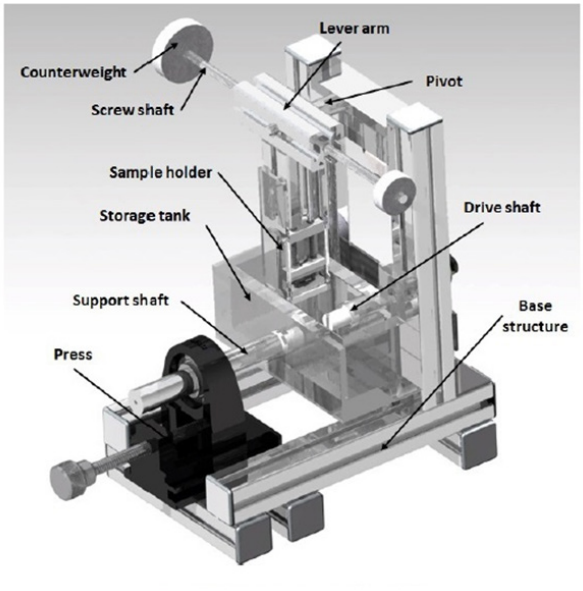


Figure 5. Micro-abrasion testing equipment
Source: Caballero *et al*. (2016)

Based on an analysis of the system's operation, the most important requirements for any user are the compatibility of the samples, the modularity of the system, and the accuracy of the tests. These factors are fundamental in equipment for materials testing, which is why the possible points for improvement are framed in the requirements mentioned Table 1 (Caballero *et al*., 2016).

In addition, it was observed that, in most studies involving fixed-ball micro-abrasion tests, multiple parameters are varied, such as, the sample and sphere material, the applied loads, whether the test is performed dry or with abrasive particles, among others, without considering that the movement of the sphere could be oscillatory and have a certain angle of travel, which, depending on the application, could be modified at will (Stachowiak *et al*., 2006; Cozza *et al*., 2009; Ardila *et al*., 2020).

The materials to be used in a biological environment, specifically inside the human body (*e.g.*, prostheses), can have oscillatory movements, as is the case of hip replacements, which participate in human motion (Figure 6).

Due to the above, this study proposes that the most significant change to the system will be allowing to invert the rotation of the sphere in contact with the analyzed sample. It is stressed that the system will be able to vary the rotation of the sphere to any desired value.

Collecting information

An inspection of the equipment revealed four points for improvement and the possibility of changing one of its systems, with the aim to improve the equipment's performance.

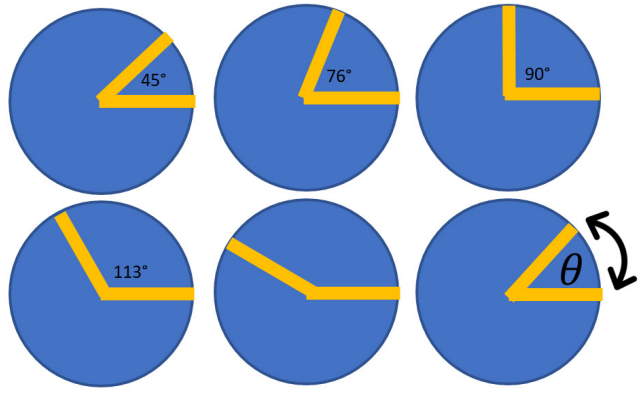


Figure 6. Turning angles of the hip during human motion
Source: Authors

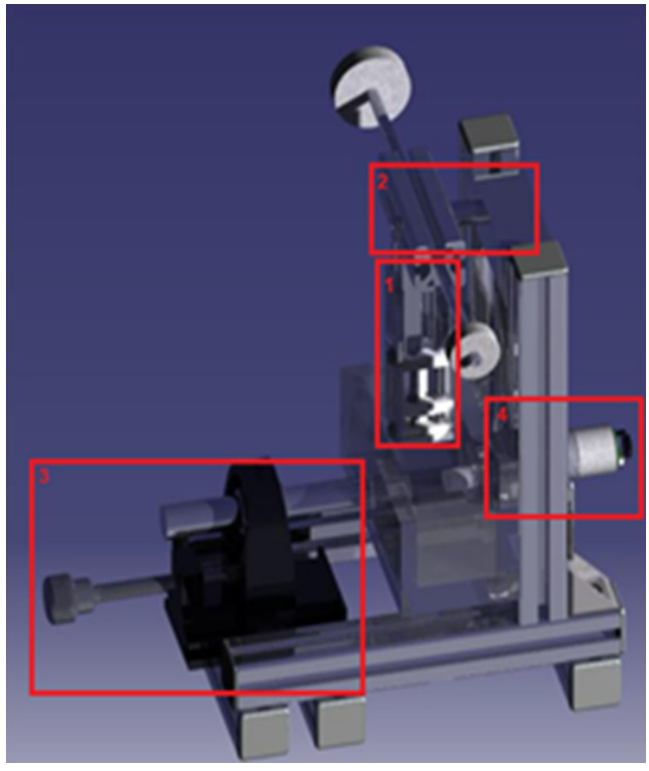


Figure 7. Current micro-abrasion testing system: 1) specimen clamping, 2) arm clamping, 3) sphere clamping, 4) motor system
Source: Adapted from Caballero *et al*. (2016)

The above-presented elements and assemblies (Figure 7) were subjected to different modifications, which allowed for some variations in the tests performed.

Below is a table with the improvement points, the characteristics to be improved, and the proposed solutions.

Table 1. Improvement points for the micro-abrasion testing system

Subsystem	Feature to improve	Solution implemented
Specimen clamping	Although the system has an outstanding grip, it limits the shape of the samples that can be analyzed. Additionally, in some tests, it was observed that the sample showed a slight movement, which could generate variations in the test results.	A circular cut was made in the parts that exert pressure on the sample, which is now a circular sample holder, which will allow the clamping to be more stable. A cut was made on the lateral plates, which avoids their contact with the sphere clamping system, thus avoiding misalignments.
Sample connector arm	The assembly is stable and allows the support to hold the specimen at the end of the square profile that composes it, although it only allows for one degree of freedom (vertical), since the bearing is housed inside the support and does not allow the pivot to move horizontally.	The element that holds the shaft connecting the sample arm to the support was replaced by a bearing, allowing for an additional degree of freedom (horizontal).
Sphere clamping	This subsystem consists of a support shaft on which the sphere rests. The support shaft rests on a bearing that is positioned at the top of the press. The limitation of the clamping system is that the press is oversized for the function that it must perform, and it takes up a large amount of space in the assembly of the system. Additionally, there is no certainty as to how much force the press exerts on the sphere and therefore on the motor shaft (this is adjusted manually).	To hold the sphere, a system with a plunger was designed, which will be responsible for exerting a certain pressure on the sphere, so that it is not offset during the tests.
Motor, drive shaft, and bushing	The current system has a DC motor, which has a magnetic encoder coupled to its shaft. This motor is controlled by a DAQ data acquisition card. The motor used in the current system, together with its respective encoder, are difficult to manipulate, since they require an acquisition system that can process the pulses of the encoder, which, being magnetic, have high frequencies and do not feature rotation inversion.	The DC motor was replaced by a stepper motor (Nema 17), which does not require a sensor, since each of its steps reports exactly how many degrees the motor shaft moves. A bushing was designed to allow for a solid coupling between the motor and the drive shafts. The drive shaft was designed with a smaller diameter in the part where it comes into contact with the groove of the sphere, preventing it from colliding with the sample.

Source: Authors

Generating solutions and decision-making

Based on Table 1, the following images show the development of the improved subsystems.

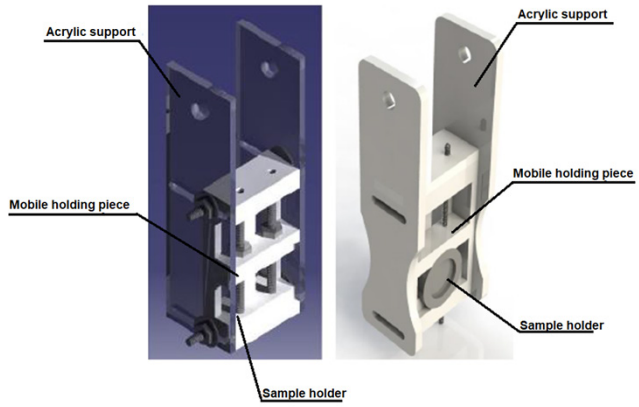


Figure 8. Fastening system: a) current b) redesigned
Source: Authors

Figure 8 shows that the sample holder and the mobile holding part were made with EMPAC. In the redesign, they were manufactured by means of 3D printing.

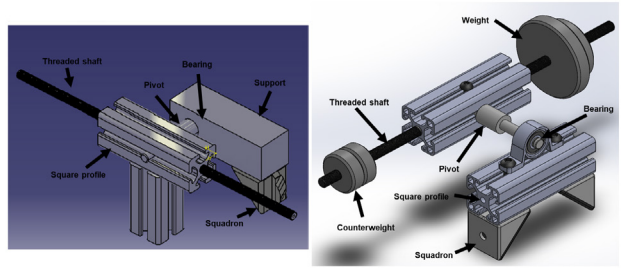


Figure 9. Sample connector arm clamping system: a) current b) redesigned
Source: Authors

Figure 9 shows how the rigid support that contained the bearing that in turn supported the pivot connected to the counterweight system was replaced by a square profile tube on which a bearing rests, thus allowing the system to move horizontally as needed.

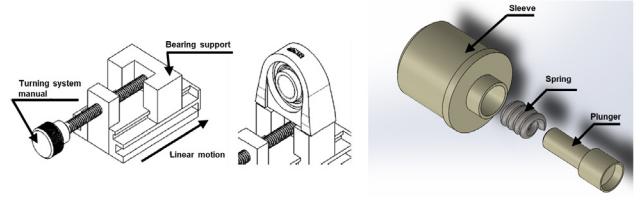


Figure 10. Sphere clamping system: a) current b) redesigned
Source: Authors

For the calculation of the spring (Figure 10), it is necessary to keep the sphere pressed against the shaft that connects to the motor. To find the spring constant to be used, the force exerted vertically by the sphere must first be calculated.

$$\textbf{\textit{Sphere mass}} = 0,11074\ \textbf{\textit{kilograms}}$$

$$\mathbf{F = m * g} \tag{6}$$

$$\mathbf{F} = 0,11074\mathbf{Kg} * 9,8\,{}^{\mathbf{m}}\!/_{\mathbf{S}^2} \tag{7}$$

$$\mathbf{F} = 1,086\,\mathbf{Newton} \tag{8}$$

Considering a 10% safety factor, the vertical force exerted by the sphere is:

$$\mathbf{F_{sphere}} = 1,10 * \mathbf{F} \tag{9}$$

$$\mathbf{F_{sphere}} = 1,10 * 1,086\mathbf{N} \tag{10}$$

$$\mathbf{F_{sphere}} = 1,19\,\mathbf{Newton} \tag{11}$$

The spring recoil is approximately 14 mm, which is equal to the Δx that the spring must have when it compresses the sphere. From the above, by Hooke's law,

$$\mathbf{F = K * \Delta x} \tag{12}$$

From the above, the value of the constant K of the spring is cleared.

$$\mathbf{K} = \frac{\mathbf{F}}{\mathbf{\Delta x}} \tag{13}$$

Replacing the values of the sphere force and the distance differential Δx expressed in meters yields:

$$\mathbf{K} = \frac{1,19\,\mathbf{Newton}}{0,014\,\mathbf{m}} \tag{14}$$

$$\mathbf{K} = 85,35\,{}^{\mathbf{N}}\!/_{\mathbf{m}} \tag{15}$$

The spring constant K to be used in the plunger system to hold the sphere must be $K = 85,35\,{}^{N}\!/_{m}$

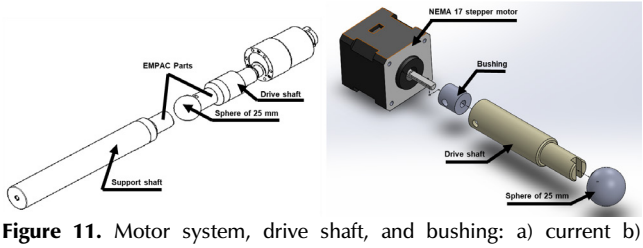


Figure 11. Motor system, drive shaft, and bushing: a) current b) redesigned
Source: Authors

The implementation of the stepper motor (Figure 11) was the most representative change made to the system, which allows controlling the position of the sphere, its velocity, and the direction of its rotation. In this case, the system was tested while considering hip joint replacements as reference.

The Nema 17 stepper motor was configured to travel 1,8° per step, which allows for a relation to be established, so that the sphere can travel at a certain angle. The rotation velocity of the motor is varied by pulse width modulation (PWM), which makes the motor versatile for the proposed implementation.

Results (testing)

Testing was performed with equipment that allows assessing wear. This equipment has a base where there is a lever arm with counterweights, which, from their distance to the center, apply a certain load to the sample under analysis. To determine the load exerted by the dead weight and the counterweight, a dynamometer is used. Moreover, there is a sphere, which rotates against the sample for a number of cycles and at a certain velocity, producing wear and abrasion on the sample.

The sphere is coupled between two coaxial shafts, one of which is supported on a bearing and coupled to a bipolar DC stepper motor, which ensures the velocities and angles of rotation of the sphere during the tests, while the other one has a plunger system to exert a force between the sphere and the other shaft in order to keep the sphere in a single position. There is a control module in charge of the whole system, from which parameters such as the test time, the angle of rotation, and the velocity of the sphere during the test can be varied (Figure 12).

We initially planned to carry out dry tests, with the aim to observe the general behavior of the system and analyze the traces left on the samples subjected to the test.

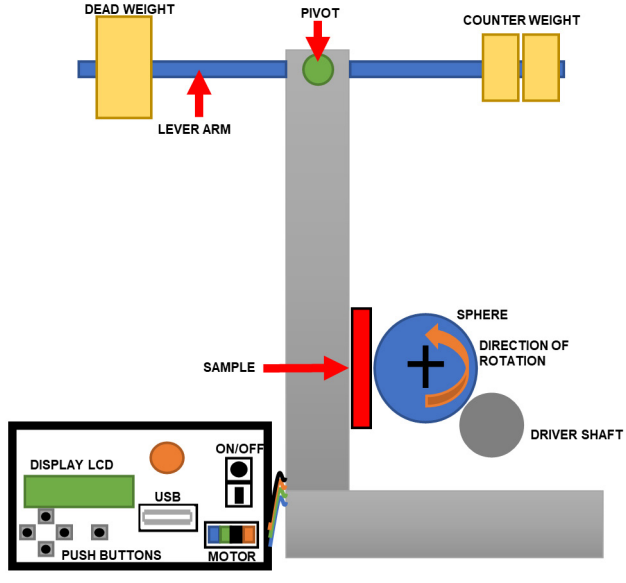


Figure 12. General scheme of the micro-abrasion tests
Source: Authors

It is also important to highlight the use of suitable materials in joint replacements, specifically those used in hip prostheses, where biocompatible steels are used for the femoral head, as well as ultra-high molecular weight polyethylene (UHMWPE) for the acetabular cup (Figure 13) (Diomidis *et al*., 2012).

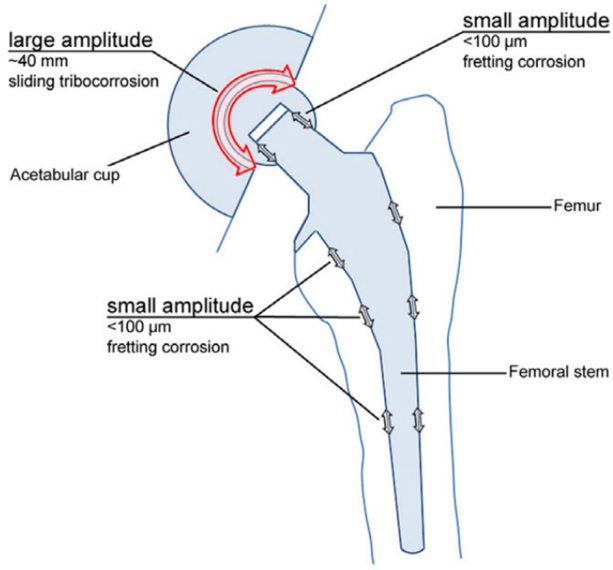


Figure 13. Hip replacement diagram
Source: Diomidis *et al*. (2012)

The samples used in the tests were previously polished until a low surface roughness was obtained. This was done using silicon carbide (SiC) abrasive paper, increasing the size from 100 to 1 200 and finally polishing the specimens with a 0,05-micron alumina (Al2O3) suspension on a rotating disk. Once the samples were polished, they were cleaned with ultrasound equipment, immersing them completely in ketone. Then, the samples were dried, and assembly was carried out in the system. The parameters of the test were adjusted as mentioned below:

- ✓ The applied load was adjusted by moving the dead weight and the counterweight while measuring the force with a dynamometer. For this case, a constant load of 2 N was set by securing the masses with nuts.
- ✓ The sphere was positioned on the drive shaft, and it was later adjusted by moving the support shaft.
- ✓ The specimen was clamped to the system with the help of a 3D-printed clamping piece.
- ✓ The lever arm was lowered so that the specimen gently touched the sphere.
- ✓ The user control module was used to set the time of the test, the speed of the sphere, and a normal or a reciprocating rotation.
- ✓ Finally, the equipment was started and, when the test was complete, the specimen was carefully removed.

At the end of the test, the sphere and the sample were cleaned again, and the wear tracks were analyzed by means of scanning electron microscopy (SEM).

For the tests, two sphere rotation speeds were determined, one of these being the normal walking speed of a person, which is about 1 Hz. Each leg makes 1 cycle in 1 second, so, in 1 minute, there are 60 cycles, which means that the sphere must rotate at a speed of 60 Rev/min to recreate this frequency. The second speed was estimated to be twice the first one (120 Rev/min) (Lavernia and Alcerro, 2008; OrthoInfo, n.d.).

For joint replacements, there is a normal load of approximately 70 or 80 Newtons that falls between the femoral area and the tibial or acetabular component (Cañizo *et al*., 2010). In all tests, a constant load of 2 Newtons was applied to maintain this parameter controlled, while varying other parameters such as the angles of travel of the sphere and the velocity. These values were chosen to closely resemble those observed in a real environment.

In one-way and reciprocating rotation testing, a test time of 30 minutes, a sample material of L 316 steel, and a sphere material of AISI 52100 steel in contact with the sample are used (Gee *et al*., 2002).

During testing with different angles of travel, it has been found that the hip can reach a flexion angle of 113° and an extension angle of 28°, which means that one of the angles to be examined is 113. The other angle is the sum of the flexion and extension angles, which is 141° (Boone and Azen, 1979). In hip joint replacements, one must consider the pressure that exists on the articular surface, which is transmitted by the acetabulum towards the femoral head, which means that the forces present in the hip act on a spherical area that is delimited by two planes at 76° (Figure 14) (Cañizo *et al*., 2010). Thus, 76° should also be included in the tests. Finally, the intermediate angles should also be evaluated, *i.e.*, 45 and 90°.

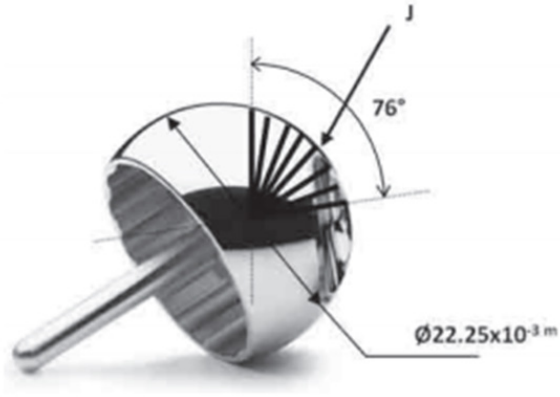


Figure 14. Load acting on the spherical bone
Source: Cañizo *et al*. (2010)

Table 2 shows the testing parameters.

Within each test, two tests are performed, from which two wear constants are obtained and averaged in order to obtain a wear constant for each case evaluated.

Table 2. Wear test specifications

Sample material	Steel L 316
Sphere material	Steel AISI 52100
Normal load	2 Newtons
Velocity 1	60 RPM
Velocity 2	120 RPM
Continuous rotation	360°
Angle 1	76°
Angle 2	141°
Angle 3	113°
Angle 4	45°
Angle 5	90°
Test time	30 minutes

Source: Authors

Results and discussion

Stereoscopy and SEM analysis

Figure 15 shows one of the images of the wear grooves taken by a stereoscope, where the lines generated by the micro-abrasive wear mechanism for dry testing can be seen. In addition, it was observed that, once the test is finished, the specimen shows material adherence by the sphere, which is why, before analyzing them by means of SEM, the sample was cleaned with ultrasound equipment in order to appreciate the morphology of the sample.

Figure 15. Track dry wear test with stereoscope
Source: Authors

Once the samples had been examined with the stereoscope, the morphology of the sample was examined via SEM, as shown in Figure 16.

In Figures 16a and 16b, the borders of the craters generated by the tests exhibit a regular circular limit, which characterizes a centered movement of the sphere during testing (Uehara *et al.*, 2019). Additionally, it can be seen that there is wear in the studied cases, which is generally classified as slight or severe. The distinction is based on the marks left or produced by the wear. Slight wear follows a uniform pattern of defined and shallow lines, while severe wear shows a greater depth in the wear tracks, and amorphous zones may appear (Pérez-Oviedo and Torre-Nieto, 2015).

Figure 17 details a footprint with patterns that meet the description of slight wear. The wear reported by the samples analyzed in all cases is slight, without the presence of abrasive wear, where grooves are generally denoted along the entire length of the crater, with possible drilling points (Cozza, 2013).

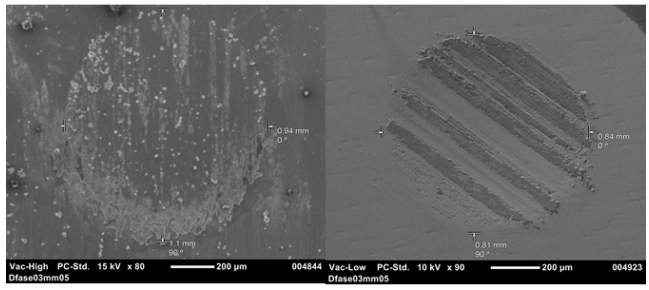


Figure 16. Wear test traces: (a) single direction of rotation at 60 RPM, dry; (b) reciprocating rotation at 120 RPM and 90°, dry
Source: Authors

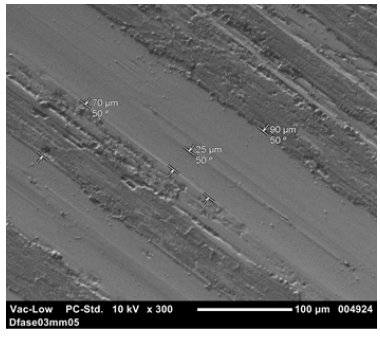


Figure 17. Track dry wear test with reciprocating rotation at 120 RPM and at 90°
Source: Authors

Wear rate estimation

The measurements carried out in each track correspond to the diameter, measured both vertically and horizontally. The values obtained are then averaged. Afterwards, the wear volume of each test (Equation (1)) is calculated while considering the radius of the sphere, which is $R_{sphere} = 0,0125m$.

Finally, the wear coefficient for each test must be calculated (Equation (3)), where the total sliding distance of each test and the average of the previously measured track must be used.

By performing two tests for each experiment, two wear coefficients are obtained, which are averaged to obtain a single value. Table 3 shows the wear constants K_c for the tests performed. The units of the wear coefficient are $K_c = \left[m^2 / N \right]$.

Table 3. Wear coefficient K_c as a function of the velocities and the angles traveled by the sphere during testing

	Velocities	
Angles (Grades)	**60 RPM (1x10^{-8})**	**120 RPM (1x10^{-8})**
45	5,58	2,7
76	4,12	2,91
90	4,94	2,07
113	1,19	2,67
141	4,13	1,81
Continuous rotation (360)	1,23	0,62

Source: Authors

From the data in Table 3, a graph is made with the data shown.

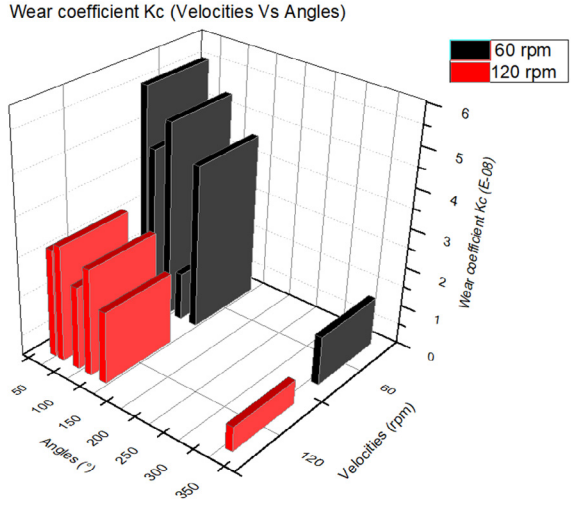


Figure 18. Graph of the wear coefficient Kc (velocities *vs.* angles)
Source: Authors

In Figure 18, a trend is observed among the tests; in five of the six cases, there is a higher wear coefficient at low revolutions (60 RPM).

Conclusions

With the upgrade made to the equipment's subsystems for wear and micro-abrasion tests, the system became variable and open to changes; with the implementation of the reciprocating travel angles, different tests can be performed which allow for a more realistic observation of the behavior of biomaterials, resembling the travel angles of a certain joint, which opens the possibility of studying these biomaterials from more specific approaches.

The modifications to the equipment made it possible to carry out wear testing with variations in the different angles of travel and with the inversion of rotation, obtaining more defined tracks and allowing for an easy analysis. This, in comparison with the tracks obtained before the upgrade.

As for the tests performed, it is evident that, in most cases, there was an adherence of the sphere material (AISI 52100 steel) to the sample, which implies a slight wear.

Finally, with the tests performed, there is a tendency indicating that, as the contact angle of the sphere decreases for reciprocating tests, the wear coefficient increases.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

Ardila, M. A. N., Costa, H. L., and de Mello, J. D. B. (2020). Influence of the ball material on friction and wear in microabrasion tests. *Wear*, *450-451*, 203266. https://doi.org/10.1016/j.wear.2020.203266

BBC (n.d.). BBC Mundo—Noticias—Somos 7.000 millones, ¿cuáles son los desafíos? *BBC News*. https://www.bbc.com/mundo/noticias/2011/10/111026_poblacion_informe_am

BBC (2012). ¿Cuántos habitantes tendrá el planeta en 2050? ¿Y en 2300? *BBC News*. https://www.bbc.com/mundo/noticias/2012/10/121014_poblacion_futuro_habitantes_dp

Boone, D. C., and Azen, S. P. (1979). Normal range of motion of joints in male subjects. *The Journal of Bone and Joint Surgery, American Volume*, *61*(5), 756-759. https://doi.org/10.2106/00004623-197961050-00017

Caballero, J. L., Sierra Melo, N., and Aperador Chaparro, W. (2016). Design and construction of a machine for micro-abrasion-corrosion testing in simulated biological environments. *Tecciencia*, *11*(20), 5-18. https://doi.org/10.18180/tecciencia.2016.20.2

Caicedo, J. C., Aperador, W., and Aguilar, Y. (2013). Tribological performance evidence on ternary and quaternary nitride coatings applied for industrial steel. *Revista Mexicana de Física*, *59*(4), 364-373.

Cañizo, R. G. R., García, L. A. G., Torres, M. V., Cruz, E. A. M., and Pineda, J. M. S. (2010). Análisis experimental del desgaste entre UHMWPE y acero inoxidable 316l empleados en la manufactura de prótesis coxofemorales. *Revista Colombiana de Biotecnología*, *12*(2), 2.

Cozza, R. C. (2013). A study on friction coefficient and wear coefficient of coated systems submitted to micro-scale abrasion tests. *Surface and Coatings Technology*, *215*, 224-233. https://doi.org/10.1016/j.surfcoat.2012.06.088

Cozza, R. C., Tanaka, D. K., and Souza, R. M. (2009). Friction coefficient and abrasive wear modes in ball-cratering tests conducted at constant normal force and constant pressure – Preliminary results. *Wear*, *267*(1), 61-70. https://doi.org/10.1016/j.wear.2009.01.055

Dieter, G., and Schmidt, L. (2009). *Engineering design*. McGraw-Hill Education.

Diomidis, N., Mischler, S., More, N. S., and Roy, M. (2012). Tribo-electrochemical characterization of metallic bioma-

terials for total joint replacement. *Acta Biomaterialia*, *8*(2), 852-859. https://doi.org/10.1016/j.actbio.2011.09.034

Gant, A. J., and Gee, M. G. (2011). A review of micro-scale abrasion testing. *Journal of Physics D: Applied Physics*, *44*(7), 073001. https://doi.org/10.1088/0022-3727/44/7/073001

Gee, M. G., Gant, A., Hutchings, I., Bethke, R., Schiffmann, K., Acker, K. V., Poulat, S., and Gachon, Y. (2002). *Ball cratering or micro-abrasion wear testing of coatings*. National Physical Laboratory. https://eprintspublications.npl.co.uk/2545/1/mgpg57.pdf

Gómez, M., and Sabeh, N. E. (n.d.). *Calidad de vida: evolución del concepto y su influencia en la investigación y la práctica*. http://campus.usal.es/~inico/investigacion/invesinico/calidad.htm

Lavernia, D. C. J., and Alcerro, D. J. C. (2008). Artroplastia total de rodilla. *Actualidades de Posgrado para el Médico General, 13*(7), 6-11.

Losada-Prieto, C., Om-tapanes, N., and Rodríguez, J. (2001). Tribología y lubricación en ensayo de banco. *CIENCIA Ergo-Sum*, *8*(2), 184-190.

OrthoInfo (n.d.) *Reemplazo total de rodilla*. https://www.orthoinfo.org/es/treatment/reemplazo-total-de-rodilla-total-knee-replacement/

Pérez-Oviedo, P., and Torre-Nieto, J. (2015). Guía de diseño para calcular el desgaste abrasivo entre componentes no lubricados. *Ciencia, Ingeniería y Desarrollo Tec Lerdo, 1*(1), 20-29. http://revistacid.itslerdo.edu.mx/coninci2015/Revista%20CID%20Tec%202015.pdf

Rasor, J. S. (2009, March 29). The science of friction. *Verdict Medical Devices*. https://www.medicaldevice-network.com/features/feature52481/

Sampaio, M., Buciumeanu, M., Henriques, B., Silva, F. S., Souza, J. C. M., and Gomes, J. R. (2016). Comparison between PEEK and Ti6Al4V concerning micro-scale abrasion wear on dental applications. *Journal of the Mechanical Behavior of Biomedical Materials, 60*, 212-219. https://doi.org/10.1016/j.jmbbm.2015.12.038

Stachowiak, G. B., Stachowiak, G. W., and Brandt, J. M. (2006). Ball-cratering abrasion tests with large abrasive particles. *Tribology International*, *39*(1), 1-11. https://doi.org/10.1016/j.triboint.2004.10.010

Staia, M. H., Enriquez, c. E., Puchi, E. S., Lewis, D. B., and Jeandin, M. (1998). Application of ball cratering method to study abrasive wear. *Surface Engineering*, *14*(1), 49-54. https://doi.org/10.1179/sur.1998.14.1.49

Uehara, P. N., Iegami, C. M., Tamaki, R., Ballester, R. Y., de Souza, R. M., and Laganá, D. C. (2019). Analysis of behavior of the wear coefficient in different layers of acrylic resin teeth. *The Journal of Prosthetic Dentistry*, *121*(6), P967.E1-967.E6. https://doi.org/10.1016/j.prosdent.2019.02.022

Vallet, R. M. (2004). *Biomateriales para sustitución y reparación de tejidos*. https://www.aecientificos.es/biomateriales-para-sustitucion-y-reparacion-de-tejidos/

Research Article / Mechanical Engineering, Mechatronics, and Materials Science http://doi.org/10.15446/ing.investig.98916

Design and Manufacturing of a Low-Cost Prosthetic Foot

Diseño y fabricación de una prótesis de pie de bajo costo

Saad M. Ali[1] and Shurooq S. Mahmood[2]

ABSTRACT

Below-knee prosthetics are used to restore the functional activity and appearance of persons with lower limb amputation. This work attempted to design and manufacture a low-cost, novel, comfortable, lightweight, durable, and flexible smart below-knee foot prosthesis prototype. This prosthesis foot was designed according to the natural leg measurement of an adult male patient. The foot is composed of rigid PVC layers interspersed with elastic strips of PTFE, and the axis of the ankle joint is flexible and consists of metal layers and a composite of polymeric damping strips with different mechanical properties, making it flexible and allowing it to absorb shocks and store and release energy. The design, modeling, and simulation of the manufactured prosthetic foot were performed via the ANSYS 18.0 software and the finite element method (FEM), where a large number of parallel and oblique planes and sketches were created. This work included four adult patients weighing 50, 75, 90, and 120 kg with different walking cycles. The results show that the highest equivalent von Mises stress and total deformations for the prosthetic limb occur at the beginning of the walking step, while the highest equivalent elastic strains and strain energy release rates are observed at the end of the walking step, regardless of the weight. This prototype can satisfactorily perform the biomechanical functions of a natural human foot, and it can be produced in attractive sizes, models, and shapes to suit different levels of below-knee amputations for different ages and weights, especially for patients with limited income.

Keywords: below- knee prosthesis, lower limb amputation, flexible multi-layered foot, ANSYS 18.0 software simulation, RSM, walking cycle stages, von Mises stress

RESUMEN

Las prótesis debajo de la rodilla se utilizan para restaurar la actividad funcional y la apariencia de personas con amputación de miembros inferiores. Este trabajo intentó diseñar y fabricar un prototipo de prótesis de pie inteligente debajo de la rodilla, novedoso, cómodo, liviano, duradero, flexible y de bajo costo. Esta prótesis de pie fue diseñada según la medida natural de la pierna de un paciente masculino adulto. El pie está compuesto por capas rígidas de PVC intercaladas con tiras elásticas de PTFE, y el eje de la articulación del tobillo es flexible y está formado por capas metálicas y un compuesto de tiras amortiguadoras poliméricas con diferentes propiedades mecánicas, que la hacen flexible y le permiten absorber impactos y almacenar y liberar energía. El diseño, modelado y simulación de la prótesis de pie fabricada se realizó mediante el software ANSYS 18.0 y el método de elementos finitos (FEM), donde se crearon una gran cantidad de planos y bocetos paralelos y oblicuos. Este trabajo incluyó a cuatro pacientes adultos que pesaban 50, 75, 90 y 120 kg con diferentes ciclos de caminata. Los resultados muestran que la tensión de von Mises equivalente más alta y las deformaciones totales para la extremidad protésica ocurren al comienzo del paso de caminar, mientras que las deformaciones elásticas equivalentes y las tasas de liberación de energía de deformación más altas se observan al final del paso de caminar, independientemente del peso. Este prototipo puede realizar satisfactoriamente las funciones biomecánicas de un pie humano natural y puede producirse en tamaños, modelos y formas atractivos para adaptarse a diferentes niveles de amputaciones por debajo de la rodilla para diferentes edades y pesos, especialmente para pacientes con ingresos limitados.

Palabras clave: prótesis por debajo de la rodilla, amputación de miembro inferior, pie flexible de varias capas, simulación en software ANSYS 18.0, RSM, etapas del ciclo de marcha, estrés de von Mises

Received: October 8th, 2021
Accepted: May 23th, 2023

Introduction

In normal environments, people walk about 6 500 steps per day at a preferred walking speed of 1,3 m/s (Michael *et al.*, 2016). Amputation is often a concern for a person, their family, and society. Lower limb amputations lead people to the loss of mobility, a deteriorated the quality of life, decreased functional performance, and difficulties in maintaining their normal daily activities (Awad *et al.*, 2016). Compared to healthy individuals, the majority of these people tire faster, walk slower, and are less sedentary (Alejandro *et al.*, 2020).

[1] B. in Production Engineering and Metallurgy, University of Technology, Iraq. MSc in Industrial Engineering, University of Baghdad, Iraq. PhD in Applied Mechanical Engineering, University of Technology, Iraq. Affiliation: Scientific assistant at the Biomedical Engineering Department, University of Technology, Iraq. Email: saad.m.ali@uotechnology.edu.iq

[2] B. from the College of Science, University of Kerbela. MSc and PhD from the College of Science for Women, University of Baghdad, Iraq. Affiliation: Full-time PhD Lecturer, Department of Physics, College of Education, Al Iraqia University, Iraq. Email: dr.shurooq1988@mauc.edu.iq

The incidence of lower limb amputations is increasing over the years in developed countries, and its main causes are traffic accidents and diabetes (Pirouzi *et al*., 2014). More than half of lower limb amputees are over the age of 65 years due to the increase in vascular disorders as the patients age (Michael *et al*., 2016). A prosthesis restores a large proportion of the normal functions, lifestyle, and stability for a person suffering from lower limb loss (Saif *et al*., 2018; Tavangarian and Proano 2019; Stephen *et al*., 2019).

The process of integrating artificial limbs with the human body poses a great challenge (Safari, 2020). This challenge lies in designing artificial limbs whose movement resembles the natural movement of the human body. The lower prosthesis pattern must be designed to perform the user's daily pre-amputation activities such as simple walking or specified running, as well as to suit their financial capacity and cosmetic requirements (Pearlman *et al*., 2008; Me *et al*., 2012; Pirouzi *et al*., 2014).

A below-knee prosthesis typically consists of ten components, namely a socket, a metal pylon, an ankle joint component, screws, bolts, nuts, padding/suspension materials, and a foot component (Tochukwu *et al*., 2018; Zagoya-López *et al*., 2021). There have been great developments and progress in producing large numbers of prosthetic limbs, which have enabled many amputees to return to their daily activities, but these prosthetic limbs are very expensive, reaching up to 60 thousand dollars. This affects individuals who have scarce economic resources, especially in developing countries, so there is an urgent and continuing need to develop affordable prostheses (Sánchez *et al*., 2012).

In recent years, there has been a marked global interest in improving the mobility of people with lower limb amputations. Despite the great developments and new technologies, commercial lower limb prosthetics are still passively active and requires an external actuator. Many movement functions (such as walking across slopes and climbing stairs) require significant force in the knee and ankle joints in order for the productive apparatus (Au *et al*., 2009; Colombo *et al*., 2011; Ashmi, *et al*., 2014) and the mechanical properties to remain constant with the difference in walking speed and surface topography. These prosthetics typically consist of flexible springs or carbon-composite paper springs that can store and release energy to help the individual move forward during the standing and walking stages (Stephen *et al*., 2019; Dhokiaa *et al*., 2017). These prostheses are assembled using ready-made components and a custom-made socket to attach them to the limb residuum (Albert and Ali, 2012; Awad *et al*., 2016). To provide comfort to lower limb amputees, smart prosthetics, whose development requires a lot of effort (Xie *et al*., 2020), are now used in the field of biological rehabilitation.

Since most amputees in developing countries are in the low-income categories, it is necessary to develop low-cost prosthetics to support the rehabilitation of those who depend on the public healthcare system (Carlos *et al*., 2015). This project aims to design and manufacture a low-cost, lightweight, durable, comfortable, and smart below-knee prototype prosthetic limb that is easy to wear and remove and cosmetically pleasing, has a good mechanical performance, and easily fulfills reasonable maintenance requirements, using suitable materials so that the low-income wearers can afford to buy them.

Materials and methodology

There are many differences between the mechanical behavior of conventional prosthetics and that of the human ankle-foot complex (Stephen *et al*., 2019). The human foot is made up of complex groups of joints and muscles. It allows the human foot to be stable on any kind of uneven surface (Thilina *et al*., 2017). Most prosthetics below the knee perform poorly because the amputation cannot control the ankle joint (Sander and Dick, 2009). A great degree of walking comfort for amputees can be achieved by increasing the angle of the prosthetic ankle. The prosthesis should be designed to specifically suit the amputated stump and the individual needs of the patient (Hsu *et al*., 2018). It should also be made of light metal alloys and polymer composites with a significantly improved mechanical performance to reduce the stress on the amputated limb and facilitate walking (Au *et al*., 2009; Jenan and Saad 2020; Mohammed *et al*., 2020; Ameer *et al*., 2021). The design goals of this work include restoring a normal and effective gait by developing movable and fixed components, with the aim to manufacture a prosthesis with functional features that correspond to the unique gait dynamics of the natural limb (Tochukwu, *et al*., 2017; Fahad *et al*., 2018; Barrios-Muriel *et al*., 2020).

The desired prosthesis mass should be 2,5% of the total body mass, equal to the percent mass of the missing biological limb. Manufacturers must make choices about their priorities regarding these factors (Pearlman *et al*., 2008; Colombo and Rizzi 2016). Materials selection plays an important role in meeting the requirements of the prosthesis's parts in order to make it effectively functional, and the cost of the materials selected has to make sense (*i.e.*, the device prosthesis must be economical and affordable for low-income amputees, for instance) in mass production, as materials do contribute a lot to the total manufacturing costs of each part. The cost of an artificial limb is recurrent because artificial limbs are usually replaced every 3-4 years due to wear and tear (Pearlman *et al*., 2008; Tavangarian and Proano, 2019).

ANSYS design

In this work, the design and simulation of the below-knee prosthetic limb were performed while using the mechanical design ANSYS 18.0 software (Jenan and Saad, 2020; Allawy and Abdulghafour, 2020). A large number of parallel and oblique planes and sketches were used to complete the design (Figures 1 to 5).

Figures 1 shows the 3D prosthetic foot design with the natural leg measurement of an adult male patient. It includes all the required details of multi-layered feet, a flexible ankle-joint pivot, and damping strips to reduce shock when standing and walking on different terrains to perform daily activities (Jenan and Saad, 2019).

Figure 2 shows the design drawings of the prosthetic foot with the addition of the ankle link, the shank link, and the pyramid adapter.

Figure 3 shows the design process of mechanically linking the laminated foot parts with the ankle pivot, as well as the parts of the ankle link with the pyramid adapter, using steel screws and nuts. Figure 4 shows the rotating process of the front edges of the prosthetic laminated foot parts.

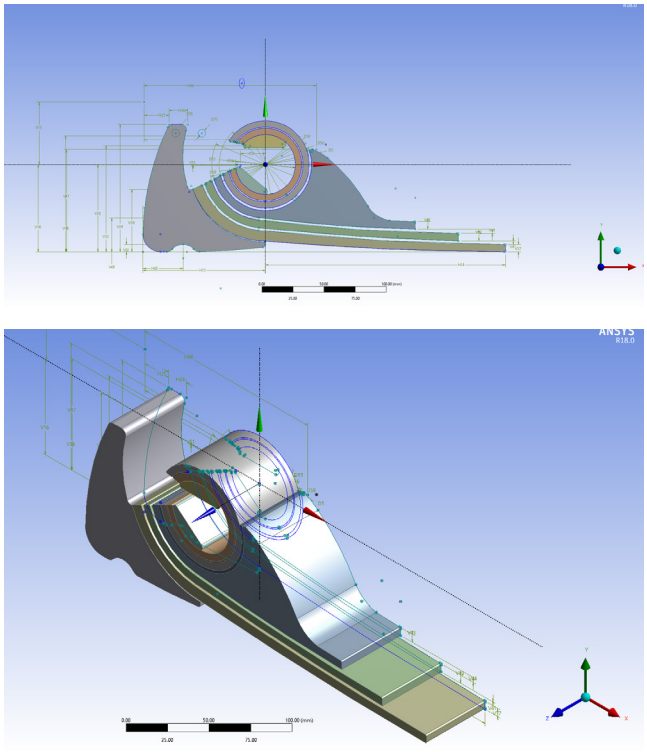


Figure 1. The 3D prosthetic foot was designed in ANSYS with all the required details: multi-layered feet, a flexible ankle-joint pivot, and damping
Source: Authors

Figure 4 also shows the complete isometric design of the below-knee prosthesis after adding the shank (pylon) and its connectors. Its length depends on the height of the amputation from the ground minus the designed foot height, or between the socket and the foot.

Selected materials

An important consideration in the design and fabrication of a limb prosthesis is the type of material used for its construction. The structural materials affect the strength and weight of the overall prosthesis (Pearlman *et al*., 2008).

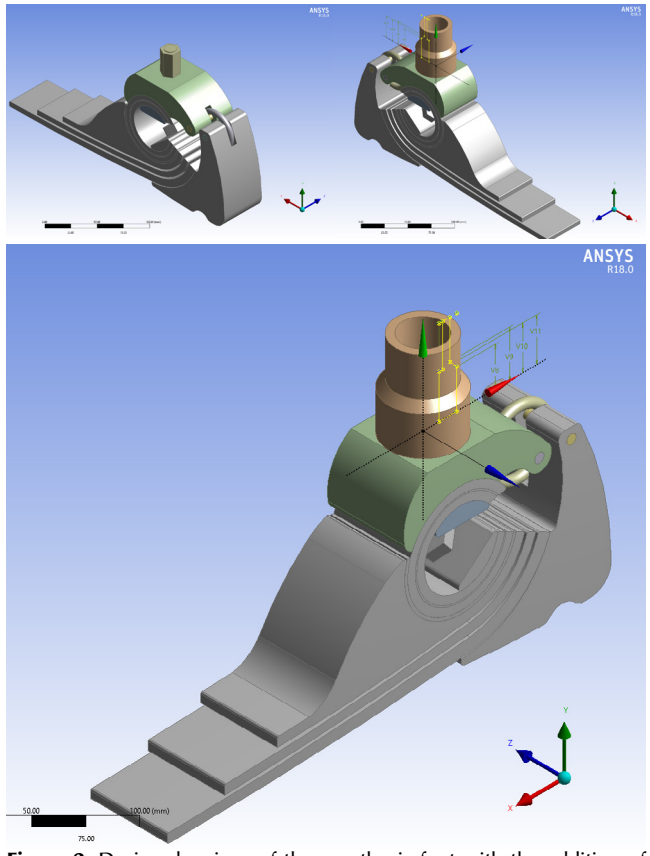


Figure 2. Design drawings of the prosthesis foot with the addition of ankle pivot, the shank link, and the pyramid adapter
Source: Authors

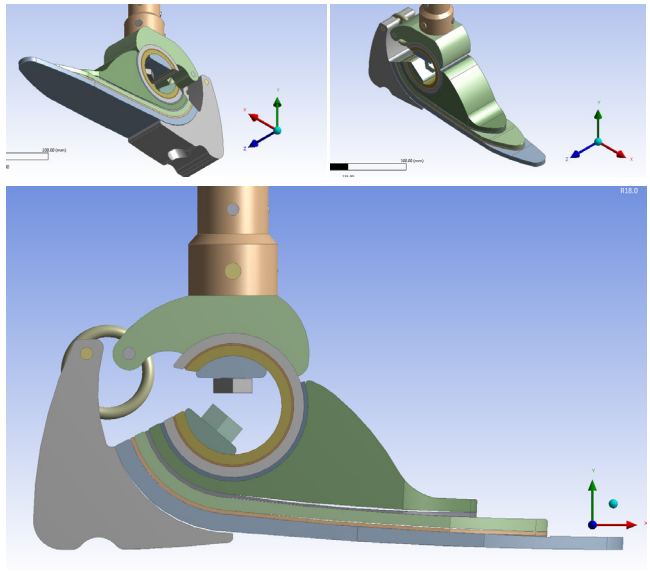


Figure 3. Design process regarding the mechanical linking of the laminated foot parts with the ankle pivot and the parts of the ankle link with the pyramid adapter using steel screws and nuts
Source: Authors

Several available raw materials were purchased for this work, including lightweight aluminum 6061 alloy tubes (tensile strength, yield strength, and ductility of up to 320 MPa, 280 MPa, and 10-12%, respectively) (Guanxia *et al*.,

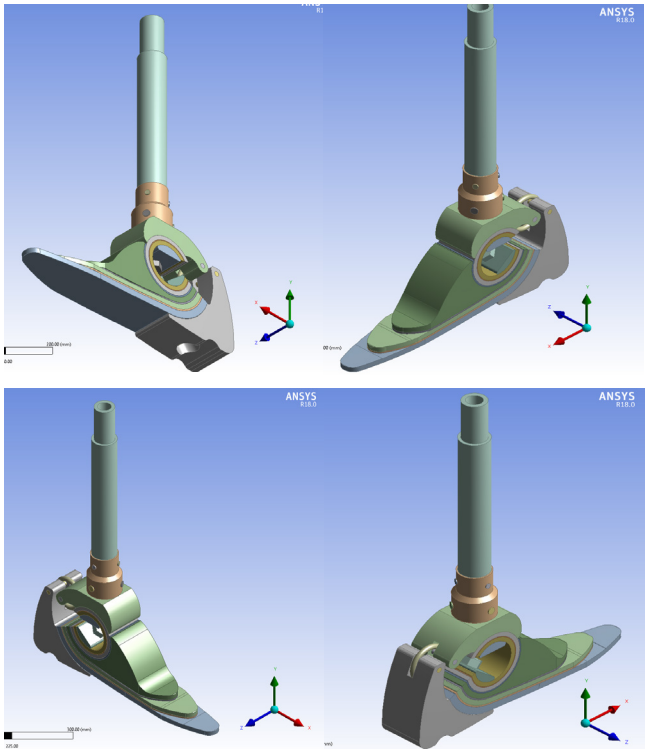


Figure 4. Complete isometric design of the below-knee prosthetic limb after adding the shank (pylon) and its connectors
Source: Authors

2020); a polytetrafluoroethylene (PTFE) ribbon for shock absorption and to increase the flexibility of the lower limb prosthesis, with a tensile strength of 161 MPa, a strain at break of 21,8%, an elastic modulus of 1 380 MPa, and a shear modulus of 27 MPa (Yingying *et al.*, 2020); and two thick black polyvinyl chloride (PVC) plates with a tensile strength of 13,79 MPa, a strength at yield of 3,01 MPa, and a tensile strain at break of 230,18% (Zhang *et al.*, 2021). The raw materials are shown in Figure 5. Table 1 provides a list of the materials required to manufacture the below-knee prosthetic limb.

Aluminum alloy is regarded as a lightweight alternative to steel. It is not as strong but, depending on the particular application, it is often strong enough to meet the design criteria and pass the necessary testing procedures. Thus, the pylon, the shank link, and the ankle joint were made of aluminum alloy, taking advantage of their light weight (Pearlman *et al.*, 2008).

PTFE is a fluorocarbon solid synthetic fluoropolymer of tetrafluoroethylene. It is nonreactive, partly because of the strength of carbon-fluorine bonds, and so it is often used in containers and pipework for reactive and corrosive chemicals. When used as a lubricant, PTFE reduces the friction, wear, and energy consumption of machinery. To manufacture the designed artificial limb, PTFE was used as a shock absorption material while walking or running, as well as to increase the flexibility of the prosthesis.

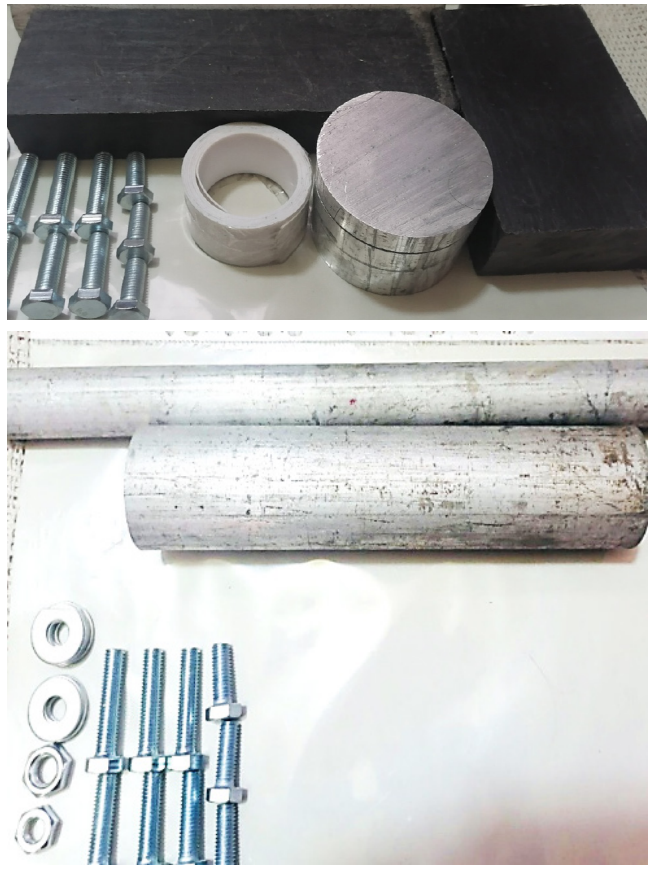


Figure 5. Shaft raw materials including the lightweight aluminum alloy tubes and rods, the polytetrafluoroethylene (PTFE) ribbon, and the thick black polyvinyl chloride (PVC) plate
Source: Authors

Table 1. Materials required to manufacture the below-knee prosthetic limb

Item	Part name	Part material	Quantity	Dimensions
1	Tube	Aluminum	1	D=72 mm, d=62 mm; L =50 mm
2	Tube	Aluminum	1	D=60 mm, d=50 mm; L =50 mm
3	Teflon sheet	PTFE	1	L= 2 m; W=40mm; T=1mm
4	Net (square)	Steel	1	M12 mm, L=27 mm
5	Belt	Steel	1	M12 mm; L =70 mm
6	Belt and nut	Steel	1	M12 mm; L =120 mm
7	Set screw	Steel	6	M10 mm; L =10 mm
8	Thick plate	PVC	1	L= 260 mm; W=120mm; T=50mm
9	Thick plate	PVC	1	L= 160 mm; W=120mm; T=50mm
10	Tube	Aluminum	1	D=30 mm, d=18 mm; L =250 mm
11	Tube	Aluminum	1	D=40 mm, d=17 mm; L =40 mm

Source: Authors

The PCV plastic polymer was used to manufacture the laminated foot, with the aim to create a lightweight and strong product. The advantage of plastic laminates is that the prosthetist has a great deal of control over the strength, stiffness, and thickness of the finished product (Michael *et al.*, 2016).

Manufacturing

The processes related to manufacturing the below-knee prosthetic limb were performed while using various metal cutting machines for turning, milling, sawing, and drilling, as well as for thread tapping and welding. This, in addition to manual work such as filing and polishing, *etc.* (Figure 6).

Figure 7 shows the ANSYS-designed drawings of the prosthesis parts, which were pasted on the PVC plates and manufactured using the band saw machine. Afterwards, these parts were smoothed and polished with an electric polishing machine. Then, the laminated foot parts, the ankle pivot, the parts of the ankle link, and the pyramid adapter were initially assembled using transparent adhesive tape and then drilled with the steel bolts and nuts (Figure 8). The final assembly of the smart prosthetic leg prototype is shown in Figure 9.

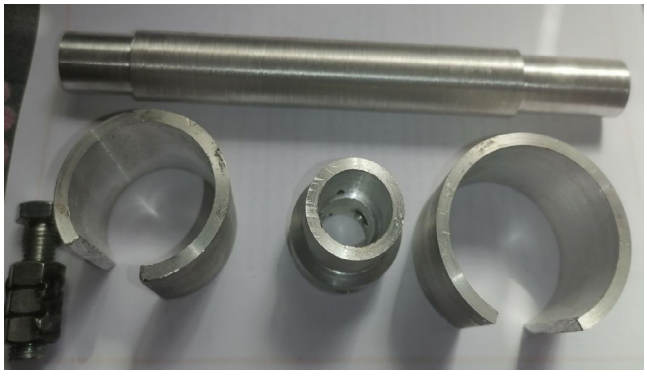


Figure 6. The fabricated prosthesis limb parts were processed in the turning machine
Source: Authors

Figure 7. Cutting the parts of the artificial foot using a band saw machine
Source: Authors

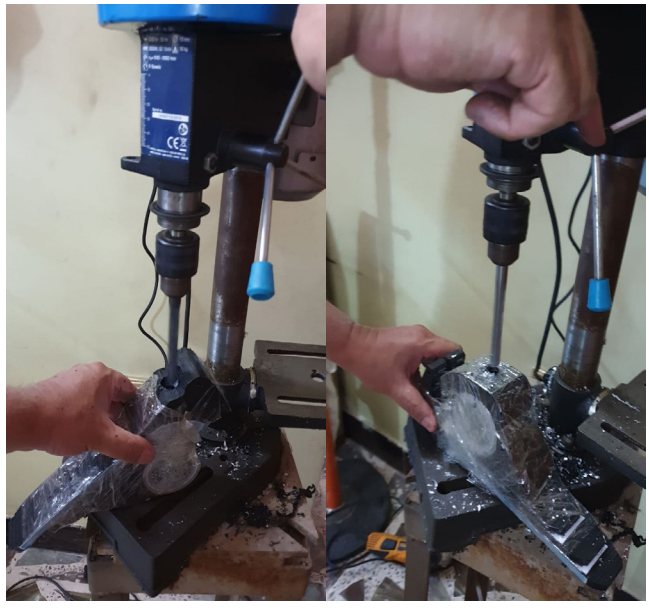


Figure 8. The laminated foot parts, the ankle pivot, the ankle link, and the pyramid adapter were assembled using the drilling machine and the steel bolts and nuts
Source: Authors

Figure 9. Final assembly of the smart prosthetic leg prototype
Source: Authors

Results and discussion

The prosthetic foot was designed for a 55-year-old male who had been amputated below his right knee as a result of a landmine explosion. The patient was wearing a large, hard prosthetic foot, which impeded his daily movement. His condition was experimentally studied, and the required measurements were taken. The foot was manufactured and fitted to the aforementioned patient, and his gait cycle was analyzed. The patient now has a better daily life.

The walking cycle is usually defined in human biomechanics as beginning with a heel strike on the earth and ending with the next heel strike of the same foot. It can be divided into a standing phase (60%) and a swing phase (40%), where the feet are high off the ground. The standing stage begins when the heel touches the ground and ends when the fingers touch (Au *et al.*, 2009).

ANSYS modeling and simulation

In order to model and simulate the proposed prosthesis design, ANSYS 18.0 and the response surface methodology (RSM) were employed (Ahmed *et al.*, 2015). To study the mechanical properties and the distribution of loads and stresses on all parts of the prosthetic limb, the finite element method (FEM) was used. The prototype was first meshed to 55 658 nodes and 24 597 elements (Figure 10a). The static structural model was implemented using three types of boundary conditions, *i.e.*, the fixation areas and those of the applied loads. The first type was used for when the patient is standing, where the prosthetic foot is supported on the whole lower surface of the foot (on the heel and the welt) (Figure 10b). In the second case, when the patient is at the beginning stage of the walking cycle, the foot is supported only on the sole and instep of the foot (Figure 10c). In the third case, when the patient is at the end stage of the cycle, the foot is supported only on the sole and instep end of the foot (on the heel) (Figure 10d).

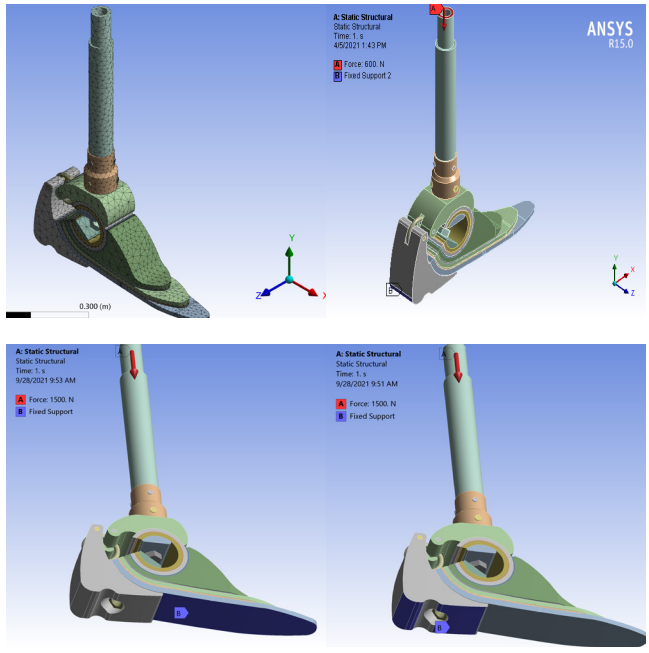


Figure 10. Static structural model with the three types of boundary conditions, *i.e.* the areas of fixation and the area of the applied load: (a) meshed process; (b) patient standing; (b) beginning of the walking cycle; (c) end of the walking cycle
Source: Authors

Mechanical properties

The mechanical properties of our proposed prosthetic limb were determined in four adult patients weighing 50, 75, 90, and 120 kg, using a safety factor of 1,2. The patients' total weights were applied to the prosthesis, *i.e.*, 600, 900, 1 200, and 1 500 N. The results are shown in Table 2.

Equivalent von Mises stress

Figures 11 and 12 show the relationship between the equivalent von Mises stress for the prosthetic limb, the patients' weights, and the walking cycle stages. These simulations were performed using ANSYS 18.0, the RSM, and the statistical Expert Systems 11.0 software.

These Figures show that the highest equivalent von Mises stresses occur at the beginning of the walking cycle. The highest value was 105,17 MPa for the highest patient weight of 1 500 N. This value is higher than those at the end of the walking step and the standing stage by 19,74% and 54,32, respectively.

The increase in the stress value is caused by the fact that, during the initial contact with the ground, the heel of the foot collides with the ground, and the patient's body weight is applied to the small area of the heel in contact with the ground, given the unstable dynamic state of the body at this stage of the walking cycle. Meanwhile, during the standing phase, the shocks are absorbed, and the body achieves its stability and balance.

Table 2. Mechanical properties of the prosthetic limb for different patient weights and walking cycle stages
Source: Authors

Patient weight (N)	Patient condition	Equivalent von Mises stress (MPa)	Total deformation (m) X1^{0}-5	Equivalent elastic strain (m/m)	Strain energy (J) X1^{0}-5
600	Standing stage	34,11	6,04	0,00066	2,28
	Beginning of the walking cycle	42,07	10,86	0,00055	1,99
	End of the walking cycle	35,07	5,15	0,00089	2,28
900	Standing stage	51,16	9,05	0,00099	5,12
	Beginning of the walking cycle	63,10	16,29	0,00082	4,49
	End of the walking cycle	52,61	7,73	0,00100	5,71
1200	Standing stage	68,21	20,72	0,00132	8,09
	Beginning of the walking cycle	84,13	21,72	0,00110	7,97
	End of the walking cycle	70,14	17,76	0,00133	10,15
1500	Standing stage	68,21	12,07	0,00132	14,22
	Beginning of the walking cycle	105,17	27,14	0,00137	12,46
	End of the walking cycle	87,68	12,88	0,00222	15,86

Total deformations

Figures 13 and 14 show the total deformations of the prosthetic leg as a result of applying different loads (*i.e.*, the different weights of the patients who were selected) at different walking cycle stages. The Figures show that the highest total deformations occur at the beginning stage of the walking cycle, reaching the highest value (0,271 mm) with the highest patient weight (1 500 N).

This value is higher than those at the end stage of the walking cycle (by 110,71%) and the standing stage (by 124,86 %). These high deformation rates are explained by the collision state, which immediately produces high dynamic load and strain energy as soon as the foot heel touches the ground.

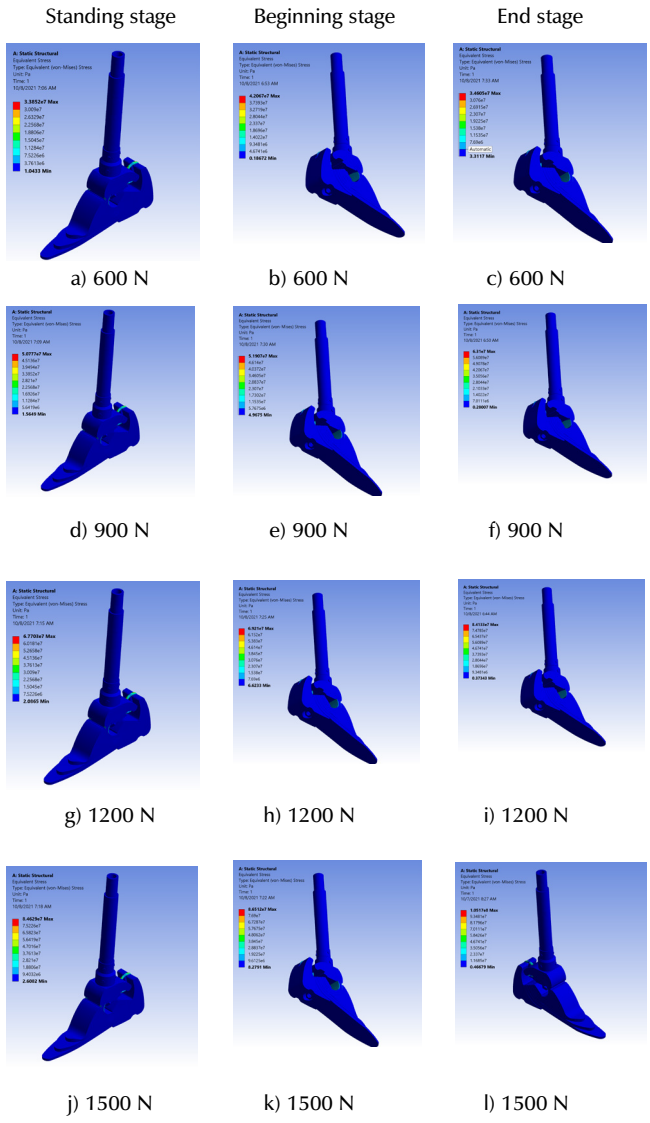


Figure 11. Equivalent von Mises stress for the prosthetic limb, considering the patients' weights and the walking cycle stages
Source: Authors

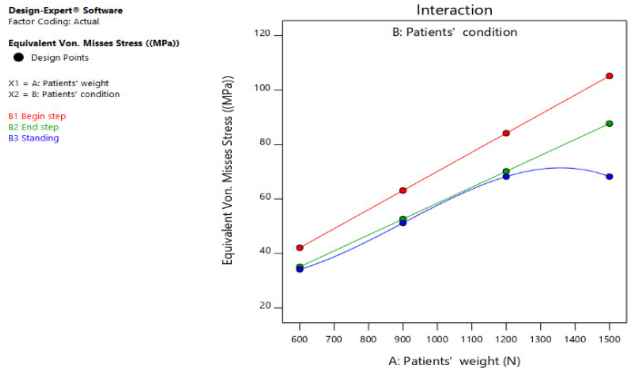


Figure 12. Relationship between the equivalent von Mises stresses of the prosthetic limb, the patients' weights, and the walking cycle stages
Source: Authors

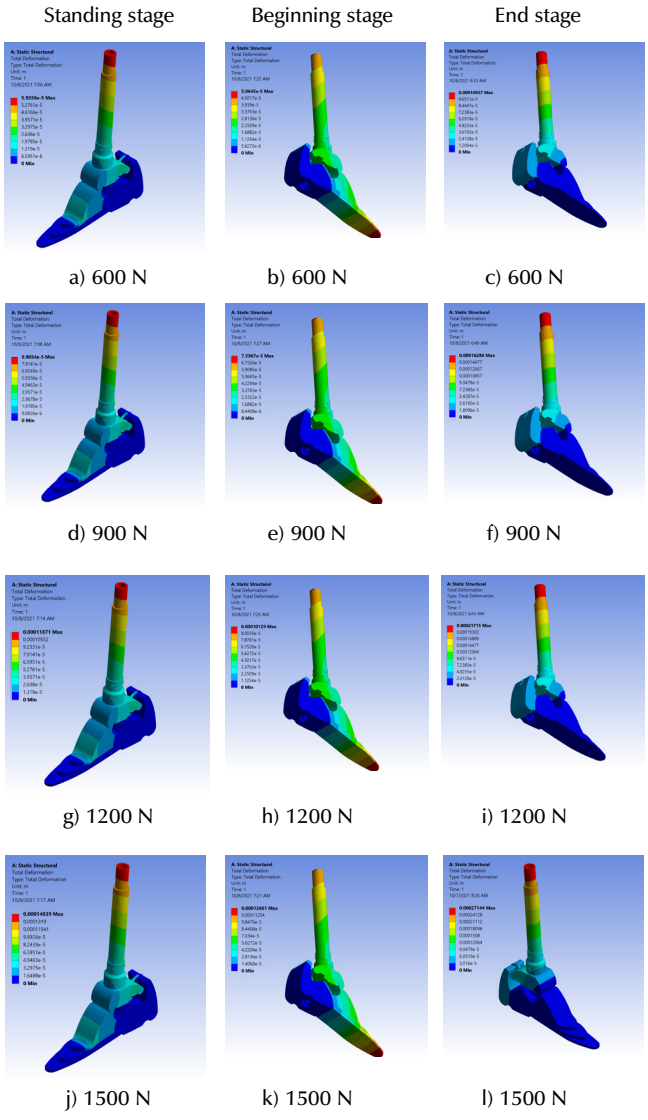


Figure 13. Total deformations in the prosthetic leg for different patient weights and walking cycle stages
Source: Authors

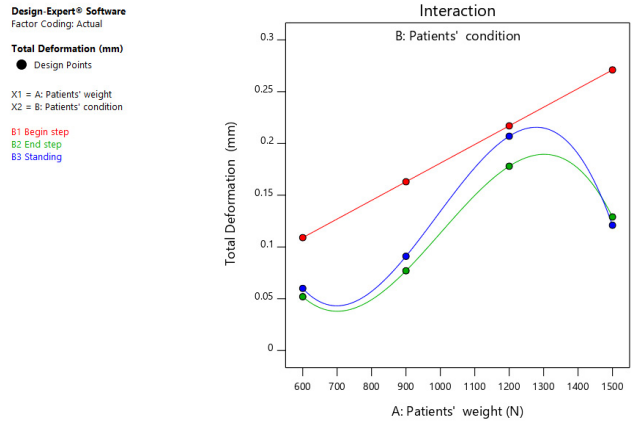


Figure 14. Relationship between the total deformations of the prosthetic limb, the patients' weights, and the walking cycle stages
Source: Authors

Equivalent elastic strains

Figures 15 and 16 show the changing values of the equivalent elastic strain rates of the prosthesis, given the changes in the static and dynamic loads that occur during the walking and standing stages of the four patients analyzed. These Figures show that the highest equivalent elastic strains occur at the end of the walking step, reaching 0,00222 m/m with the highest patient weight (1 500 N). This value is higher than those at the beginning of the walking step (by 62,04%) and the standing position (by 68,18%). This is due to the fact that, at the end of the step, the required force increases during the last stage of contact with the ground. This, in order to help push the body forward, which entails an increase in the equivalent elastic strain rate.

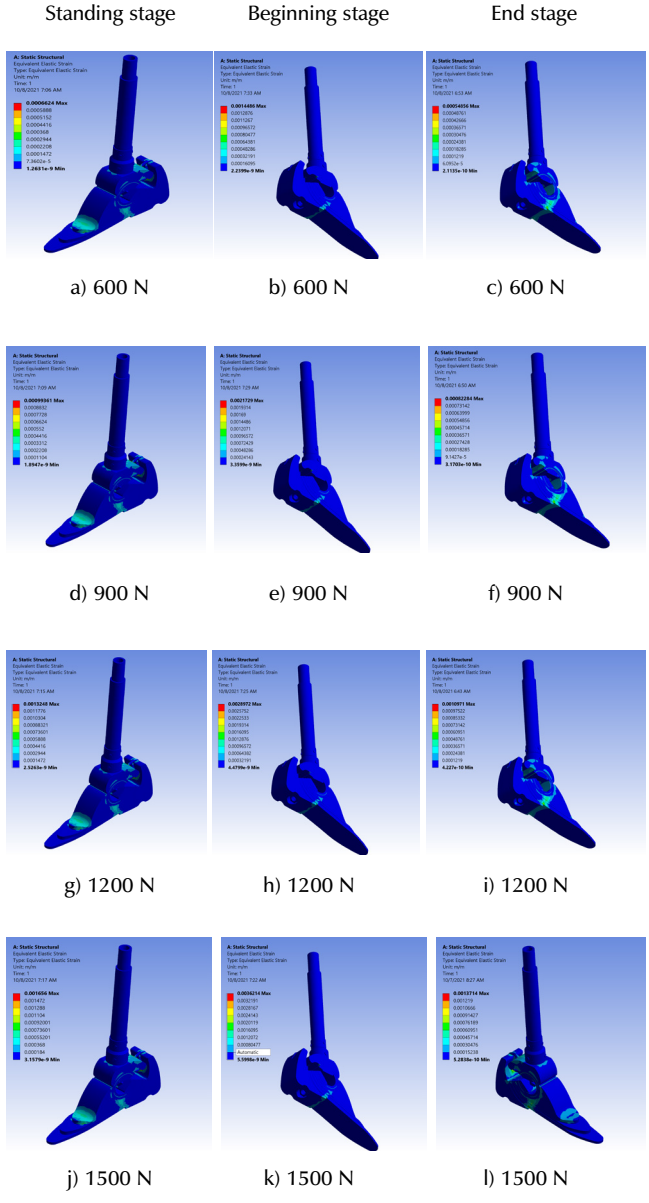


Figure 15. Equivalent elastic strain rates of the prosthetic limb according to the changes in the patient weight and the walking cycle stages
Source: Authors

Strain energy

Figures 17 and 18 show the rates for the strain energy absorbed, stored, and released during the prosthesis's daily activities according to the weight and stages of the patients' walking cycle. These Figures show that the highest strain energy rate occurs at the end of the walking stage and reaches the highest value (0,159 mJ) with the highest patient weight (1 500 N). This value is higher than that at the standing stage (by 11,67%) and the beginning of the walking stage (by 27,29%).

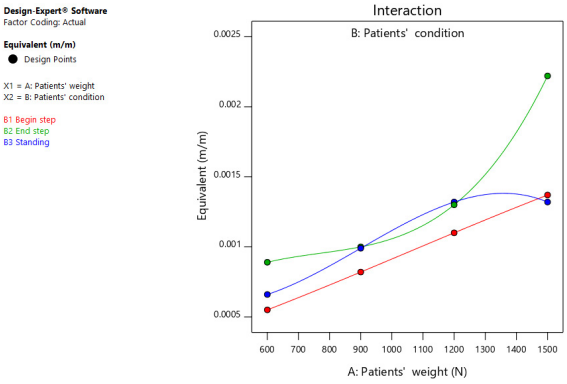


Figure 16. Relationship between the equivalent elastic strains of the prosthesis, the patients' weights, and the walking cycle stages
Source: Authors

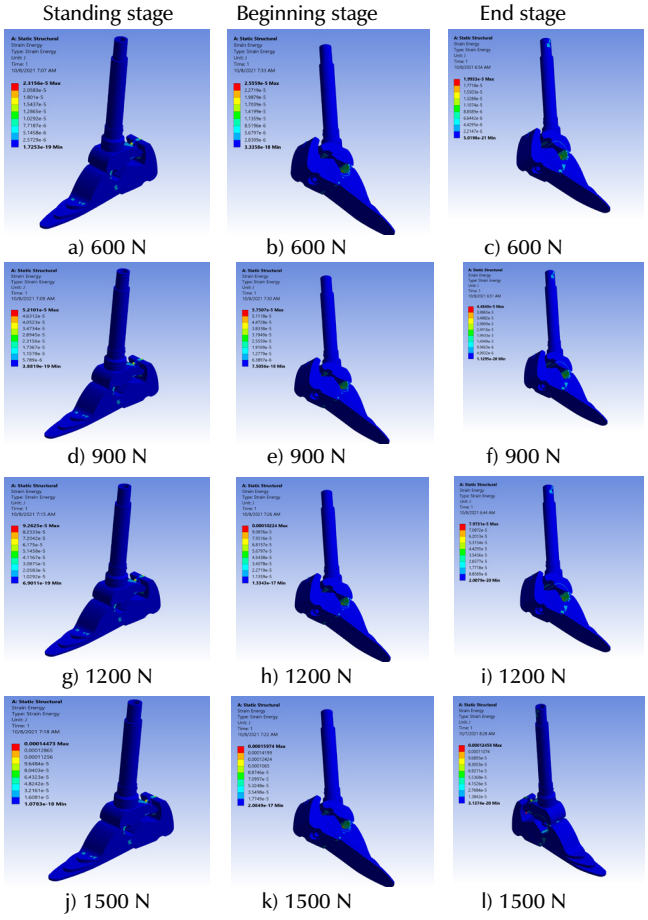


Figure 17. Strain energy rates of the prosthetic limb according to the changes in the patient weight and the walking cycle stages
Source: Authors

The main reason for this change in the strain energy rates is that, at the beginning of the walking step, and given the increase in the dynamic and static loads, the strain energy storing rate increases. During the standing phase, the strain energy expenditure rates decrease because the body quickly returns to static load and the shocks are absorbed in order for it to achieve balance and stability. At the beginning of the walking stage, at the last stage of contact with the ground, the strain energy expenditure increases to help push the patient's body forward.

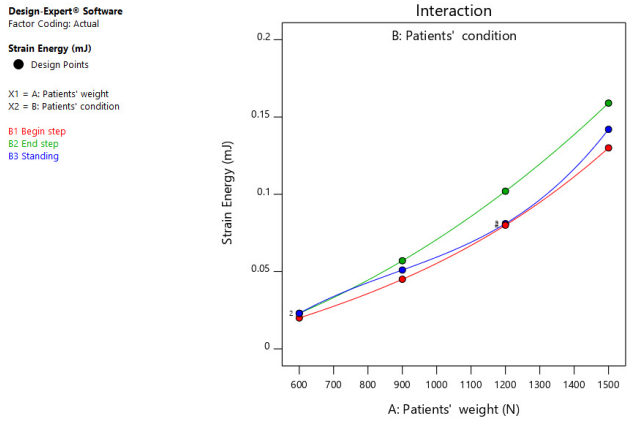


Figure 18. Relationship between the strain energy rate, the patients' weights, and the walking cycle stages
Source: Authors

Conclusions

The main conclusions of this work can be summarized as follows:

1- The results show that the prosthetic limb's highest equivalent von Mises stress and total deformations occur at the beginning of the walking step and with the heaviest patient (1 500 N), while the highest equivalent elastic strains and strain energy rates occur at the end of the walking step with the same patient.

2- The highest equivalent von Mises stress reached 105,17 MPa, which is higher than the stress values at the end of the walking step (by 19,74%) and the standing position (by 54,32%).

3- The highest total deformations reached 0,271 m. This value is higher than those at the end stage of the walking cycle (by 110,71%) and the standing stage (by 124,86%).

4- The highest equivalent elastic strain reached 0,00222 m/m. This value is higher than those at the beginning of the walking step (by 62,04%) and the standing stage (by 68,18%).

5- The highest strain energy rate reached 0,159 mJ. This value is higher than those at the standing stage (by 11,67%) and the beginning of the walking stage (by 27,29 %).

6-Further work is required to design and develop other low-cost models to suit all ages, weights, and amputation levels with high sensitivity and specificity, in order to aid in restoring the overall functions of patients of various professions, adding a high aesthetic value and ensuring comfort.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

Ahmed, N. A., Samir, A. A., and Saad, M. A. (2015). Study the effect of the graphite powder mixing electrical discharge machining on creation of surface residual stresses for AISI D2 die steel using design of experiments. *Engineering and Technology Journal, 33A*(6), 1399-1415. https://doi.org/10.30684/etj.33.6a.10

Albert, E. Y., and Ali, A. S., (2012). The design, development and construction of an adjustable lower extremity. *IOSR Journal of Engineering* (*IOSRJEN*), *2*(10), 30-42. https://doi.org/10.9790/3021-021033042

Alejandro, F. A., Luke, M. M., Jean-François, D., Ann, M. S., Levi, J. H., and Elliott J. R. (2020). Design and clinical implementation of an open-source bionic leg. *Nature Biomedical Engineering, 4*, 941-953. https://doi.org/10.1038/s41551-020-00619-3

Allawy, N. I., and Abdulghafour, A. B., (2020). Integration of CAD/CAE/RP environment for developing new product in medical field. *Engineering and Technology Journal, 38A*(9), 1276-1282. https://doi.org/10.30684/etj.v38i9A.982

Ameer, A. K., Esraa, A. A., Ahmed K. M., Kadhim K. R., and Muhannad, A., (2021). Manufacturing and analyzing of a new prosthetic shank with adapters by 3D printer. *Journal of Mechanical Engineering Research and Developments, 44*(3), 383-391. https://jmerd.net/03-2021-383-391/

Ashmi, M., Jayaraj, S., and Sivanandan, K. S., (2014). *Development of a robust microcontroller based output feedback control for assistive limb* [Conference presentation]. 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia. https://doi.org/10.1109/IECBES.2014.7047484

Au, S. K., Weber, J., and Herr, H., (2009). Powered ankle-foot prosthesis improves walking metabolic economy. *IEEE Transactions on Robotics, 25*, 51-66. Https://doi.org/10.1109/TRO.2008.2008747.

Awad, M. I., Abouhossein, A., Dehghani-Sanij, A. A., Richardson, R., Moser, D., Zahedi, S., and Bradley, D. (2016). Towards a smart semi-active prosthetic leg: Preliminary assessment and testing. *IFAC-PapersOnLine, 49*, 170-176. https://doi.org/10.1016/j.ifacol.2016.10.539

Barrios-Muriel, J., Romero-Sánchez, F., Javier, F. A., and David Rodríguez, D. S., (2020). *Advances in orthotic and prosthetic manufacturing: A technology review*. *Materials, 13*(295),1-15. https://doi.org/10.3390/ma13020295

Colombo, G., Gabbiadini, S., Rizzi, C., and Regazzoni, D., (2011). *Design procedure and rules to configure lower limb prosthesis* [Conference presentation]. ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Washington DC, USA. https://doi.org/10.1115/DETC2011-47651

Colombo, G. F., and Rizzi, C. (2016). Automatic below-knee prosthesis socket design: A preliminary approach. In V. Duffy (Ed.), *Digital Human Modeling: Applications in Health, Safety, Ergonomics and Risk Management,* (*pp.* 75-81). Springer. https://doi.org/10.1007/978-3-319-40247-5_8

Carlos, W. D., Aurélio, M. V., and Jean-Jacques, B., (2015). Conception, design and development of a low-cost intelligent prosthesis for one-sided transfemoral amputees. *Research on Biomedical Engineering31*(1), 62-69. https://doi.org/10.1590/2446-4740.0647

Dhokiaa, V., Bilzonb, J., Seminatib, E., David Canepa, D. T., and Youngc, M., Mitchella, W., (2017). The design and manufacture of a prototype personalized liner for lower limb amputees. *Procedia CIRP, 60*, 476-481. https://doi.org/10.1016/j.procir.2017.02.049

Fahad, M. K., Jumaa, S. C., and Ayad, M. T., (2018). *Design* and manufacturing knee joint for smart transfemoral prosthetic. *IOP Conference Series: Materials Science and Engineering, 454*, 1-7. https://doi.org/10.1088/1757-899X/454/1/012078

Guanxia X., Gu Z., Shipeng L., Hu-Tian L., Xinghui G., and Lei Z. (2020). Study on microstructure and mechanical properties of Al-Mg-Si-Cu aluminium alloy with high ductility. *MATEC Web of Conferences, 326*,1-6. https://doi.org/10.1051/matecconf/202032603003

Hsu, C. H., Ou, C. H., Hong, W. L., and Gao, Y. H., (2018). Comfort level discussion for prosthetic sockets with different fabricating processing conditions. *BioMed Eng OnLine, 17*(Suppl 2),145, 75-91. https://doi.org/10.1186/s12938-018-0577-2

Jenan, S. K., and Saad, M. A., (2020). 3D model of bone scaffolds based on the mechanical behaviour for a hybrid nano bio-composites. *Journal of Mechanical Engineering, 17*(2), 45-67. https://doi.org/10.24191/jmeche.v17i2.15300

Jenan, S. K., and Saad, M. A., (2020). Ansys modeling for bone reconstruction by using hybrid nano bio composite. *Journal of Solid Mechanics, 12*(4), 774-790. https://doi.org/10.22034/JSM.2020.1864464.1411

Jenan, S. K., and Saad, M. A., (2019). Modeling and simulation for mechanical behavior of modified biocomposite for scaffold application. *Ingeniería e Investigación, 39*(1), 63-75. https://doi.org/10.15446/ing.investig.v39n1.73638

Me, R. C., Ibrahima, R., and Tahi, P. M., (2012). Natural based biocomposite material for prosthetic socket fabrication. *ALAM CIPTA, International Journal of Sustainable Tropical Design Research and Practice, 5*, 27-34. https://core.ac.uk/download/pdf/153813659.pdf

Michael, W., Martin, G., Oliver, C., Stephan, R., B., Beckerle, P. (2016). Active lower limb prosthetics: a systematic review of design issues and solutions. *Biomedical Engineering Online, 15*(S3, 140), 7-19. https://doi.org/10.1186/s12938-016-0284-9

Mohammed, A. A., Oleiwi, J. K., and Al-Hassani, E. S., (2020). Influence of nanoceramics on some properties of polye-

theretherketone-based biocomposites. *Engineering and Technology Journal, 38*A(08), 1126-1136. https://doi.org/10.30684/etj.v38i8A.703

Pearlman, J., Cooper, R. A., Krizack, M., and Noon, J., (2008). Lower-limb prostheses and wheelchairs in low-income countries (an overview). *IEEE Engineering in Medicine and Biology Magazine, 27*, 12-22. https://doi.org/10.1109/EMB.2007.907372.

Pirouzi, Gh., Abu Osman, N. A., Eshraghi, A., Ali, S., Gholizadeh, H., and Wan Abas, W.A. (2014). Review of the socket design and interface pressure measurement for transtibial prosthesis. *The Scientific World Journal, 2014*, 849073. https://doi.org/10.1155/2014/849073.

Safari, R., (2020). Lower limb prosthetic interfaces: Clinical and technological advancement and potential future direction. *Prosthetics and Orthotics International, 44*(6), 384-401. https://doi.org/10.1177/0309364620969226.

Saif, M. A., Ayad, M. T., Ayad, M., Mohsin, A. A., and Muhannad, A., (2018). Manufacturing and analysis of ankle disarticulation prosthetic socket (systems). *International Journal of Mechanical Engineering and Technology* (*IJMET*), *9*(7), 560-569. https://doi.org/10.34218/IJMET.9.7.2018.059

Sánchez, J., Hernández, R. J., and Torres, J. E., (2012). The mechanical design of a transfemoral prosthesis using computational tools and design methodology. *Ingeniería e Investigación, 32*(3), 14-18. https://doi.org/10.15446/ing.investig.v32n3.35934

Sander, L. M., and Dick, I. P., (2009). Design, fabrication, and preliminary results of a novel below-knee prosthesis for snowboarding: A case report. *Prosthetics and Orthotics International, 33*(3), 272-283. https://doi.org/10.1080/03093640903089576

Stephen, H. R., Nishant, P., Aasish, Y., Saurabh, G., Mudit, G., and Kailash, M., (2019). Current trends & challenges in prosthetic product development: Literature review. *International Journal of Science and Research* (*IJSR*), *8*(6), 1554-1563. https://www.ijsr.net/get_count.php?paper_id=ART20198727

Tavangarian, F., and Proano, C. (2019). The need to fabricate lower limb prosthetic devices by additive manufacturing. *Biomedical Journal of Scientific & Technical Research, 15*, 11662-11673. https://doi.org/10.26717/BJSTR.2019.15.002772

Thilina, H. W., Thilina, D. L., and Gopura, R. C., (2017). Adaptive foot in lower-limb prostheses. *Journal of Robotics, 2017,* 9618375. https://doi.org/10.1155/2017/9618375

Tochukwu, U. N., Chidi, O. G., Jervas E., Alice, I., and Wilson, O., (2017). Design and fabrication of improved below-knee prosthesi better mimic natural limb. *Current Trends in Biomedical Engineering & Biosciences, 10*(2), 25-28 https://doi.org/10.19080/CTBEB.2017.10.555781

Tochukwu, N. U., Jervas, E., Chioma, K. A., William, I. O., and James, I. N., (2018). Inclusion of ankle joint in the design and fabrication of below-knee prosthesis. *International Journal of Health and Rehabilitation Sciences, 7*(3),131-135. https://doi.org/10.5455/ijhrs.0000000155

Xie, H., Li, Z., and Li, F., (2020). Bionics design of artificial leg and experimental modeling research of pneumatic artificial muscles. *Journal of Robotics, 2020*, 3481056. https://doi.org/10.1155/2020/3481056

Yingying Z., Qilin Z., Chuanzhi Z., and Ying Z. (2010). Mechanical properties of PTFE coated fabrics. *Journal of Reinforced Plastics and Composites, 29*(24) 3624-3630. https://doi.org/10.1177/0731684410378542

Zhang, X., Ma, Z., Wu, Y., and Liu, J. (2021). Response of mechanical properties of polyvinyl chloride geomembrane to ambient temperature in axial tension. *Applied Sciences, 11*(22), 10864. https://doi.org/10.3390/app112210864

Zagoya-López, J., Zúñiga-Avilés, L. A., Vilchis-González, A. H., and Ávila-Vilchis, J. A., (2021). Foot/ankle prostheses design approach based on scientometric and patentometric analyses. *Applied Sciences, 11*(12), 5591. https://doi.org/10.3390/app11125591

Research Article / Civil and Sanitary Engineering https://doi.org/10.15446/ing.investig.101082

Study on Closely Spaced Asymmetric Footings Embedded in a Reinforced Soil Medium

Estudio de zapatas asimétricas poco espaciadas incrustadas en un medio de suelo reforzado

Anupkumar G. Ekbote[1] and Lohitkumar Nainegali[2]

ABSTRACT

In practice, footings are rarely laid on the surface or at ground level; usually, they are embedded in the soil medium. Most studies focus on surface footings. This research examines the behavior of two interfering asymmetric footings while considering their widths to be dissimilar and the effect of embedment depth to enhance the ultimate bearing capacity and limit the settlement within the working range. This was evaluated through the finite element method of the ABAQUS software. The soil was assumed to have a Mohr-Coulomb failure, and the asymmetry corresponded to the footing widths. The results are presented in terms of interference factors, *i.e.*, the ultimate bearing capacity (UBC) and the settlement, which are defined as the UBC/settlement ratio of the left/right footing in the presence of the other one placed on reinforced soil. This, in comparison with an identical isolated footing on unreinforced soil. Interference is more significant in small footings than in large ones. Due to behavioral variations, the bearing capacity and settlement are different. This effect increases with an increase in the width of large footings, and the interference factors decrease with an increase in the embedment depth of the footings. When the right footing width is twice that of the other and considering one layer of reinforcement and soil friction angles of 30° and 40°, the per-cent increments in the bearing capacity of interfering left footings, for a spacing of 0,5 times the left footing width, are 104 and 148%, respectively.

Keywords: interference, asymmetric footings, ultimate bearing capacity, settlement

RESUMEN

Las zapatas, en la práctica, rara vez se colocan en la superficie o al nivel del suelo; por lo general, están incrustadas en el medio del suelo. La mayoría de los estudios se centran en zapatas superficiales. Esta investigación examina el comportamiento de dos zapatas asimétricas que interfieren, considerando que sus anchos son diferentes y que el efecto de la profundidad de empotramiento mejora la capacidad portante última y limita el asentamiento dentro del rango de trabajo. Esto se evaluó a través del método de elementos finitos del *software* ABAQUS. Se asume la falla de Mohr-Coulomb en el suelo, y la asimetría corresponde al ancho de las zapatas. Los resultados se presentan en términos de factores de interferencia, *i.e.*, la capacidad portante última (UBC) y el asentamiento, que se definen como la relación de UBC/asentamiento de la zapata izquierda/derecha en presencia de la otra, colocada sobre suelo reforzado. Esto, en comparación con una zapata aislada idéntica sobre suelo no reforzado. La interferencia es más significativa en las zapatas pequeñas que en las grandes. Debido a las variaciones de comportamiento, la capacidad de carga y el asentamiento son distintos. Este efecto aumenta con el aumento del ancho las zapatas grandes, y los factores de interferencia disminuyen con el aumento de la profundidad de empotramiento de las zapatas. Cuando el ancho de la zapata derecha es el doble de la otra, y considerando una capa de refuerzo y ángulos de fricción del suelo de = 30° y 40°, el incremento porcentual en la capacidad portante de las zapatas izquierdas que interfieren, para un espaciamiento de 0,5 veces el ancho de la zapata izquierda, son 104 y 148 % respectivamente.

Palabras clave: interferencia, zapatas asimétricas, capacidad portante última, asentamiento

Received: February 16th, 2022
Accepted: May 8th, 2023

Introduction

Foundations, referred to as *sub-structures*, are an integral part of an engineered structural system. They safely transfer and distribute the load from the superstructure to the underlying soil/rock strata, so that neither the soil/rock fails in shear nor the foundation itself. The foundation/sub-structure constitutes an embedded part of the structure and acts as an intermediary between the superstructure and the ground on which the foundations are laid. The bearing capacity and the settlement characteristics of the underlying ground strata significantly govern the design of the foundation system and the structure's serviceability. The shear failure and or excessive settlement of the ground strata results in a distortion of the superstructure. Therefore, great attention must be paid to all the affecting factors.

[1] PhD, Indian Institute of Technology, Dhanbad. Affiliation: Assistant professor, Department of Civil Engineering, BMS Institute of Technology and Management, Yelahanka - 560064, Bengaluru, Karnataka, India. Email: anupge@gmail.com

[2] Corresponding Author, PhD, Indian Institute of Technology, Kanpur. Affiliation: Associate professor, Department of Civil Engineering, Indian Institute of Technology (ISM), Dhanbad, Dhanbad - 826004, Jharkhand, India. Email: lohitkumar@iitism.ac.in

Numerous studies have been conducted with the aim to predict the response of shallow foundations, and many pioneers (Prandtl, 1920; Terzaghi, 1943; Skempton, 1951; Meyerhof, 1963; Hansen, 1970; Vesic, 1773) have contributed to this theory. Thus, the expressions for estimating the bearing capacity and settlement aspects of shallow foundations resting on the soil and the rock strata have been derived. Notably, this theory and these equations are exceptionally valid for an isolated foundation system where no foundation exists in an isolated state. In several situations, individual or groups of foundations are persistently built in close proximity. Some of these situations include deficient construction space, restrictions in the property line, structural design requirements, and building architecture. Recently, with the increased expansion of infrastructures in conjunction with a rapid urbanization, and in order to tackle the problem of construction space, structures and sub-structures have been forced to come ever close to each other. In a way, foundations have been laid with closer spacing. Under such situations, the bearing capacity, settlement, failure mechanisms, and rotational characteristics are invariably different from those of isolated foundations. This may be due to the influence of one foundation on the other, as they are so closely laid. The stress isobars or the failure zones of neighboring foundations may combine or interfere with each other, leading to the phenomenon known as *foundation interference/interaction*. The overlapping of individual stress isobars or failure zones combines to form a single stress isobar or failure zone that extends both laterally and vertically, affecting large soil masses. This phenomenon alters the behavior of closely placed foundations when compared to the fundamental behavior of isolated ones, which also includes those that are placed far enough from each other so that may be considered to be isolated.

Many researchers have studied the problem by performing small-scale laboratory or field tests, as well as by conducting theoretical or numerical analyses on an unreinforced soil medium. These studies have implemented different theoretical or numerical techniques, such as a method of stress characteristics (Graham *et al*., 1984; Kumar and Ghosh, 2007a); upper bound limit analysis (Kumar and Kouzer, 2008; Kouzer and Kumar, 2008, 2010; Kumar and Ghosh, 2007b; Yang *et al*., 2017; Biswas and Ghosh, 2018; Keawsawasvong *et al*., 2021; Yodsomjai *et al*., 2021); lower bound limit analysis (Kumar and Bhattacharya, 2010, 2013; Shiau *et al*., 2021; Keawsawasvong and Boonchai Ukritchon, 2022); the finite difference method (Ghazavi and Lavasan, 2008; Ghosh and Sharma, 2010; Mabrouki *et al*., 2010; Lavasan and Ghazavi, 2012b; Javid *et al*., 2015; Lavasan *et al*., 2017); the finite element method (Lee *et al*., 2008; Lee and Eun, 2009; Kumar and Bhoi, 2010; Nainegali *et al*., 2013, 2018, 2019; Noorzad and Manavirad, 2014; Zidan and Mohamed, 2019; Sekhar *et al*., 2020; Shokoohi *et al*., 2019; Fuentes *et al*., 2019; Ekbote and Nainegali., 2019a, b; Sekhar *et al*., 2020; Alzabeebee, 2020, 2022; Ekbote *et al*., 2022); the analytical method (Kumar and Saran, 2003b, 2004; Ghosh *et al*., 2017); and the probabilistic approach (Griffiths *et al*., 2006).

In this vein, the ground improvement technique is a potentially cost-effective method, as it provides soil reinforcement for shallow foundations, constituting an alternative to conventional deep foundations wherever applicable. Soil reinforcement is carried out in several layers below the foundations, using geotextile, geogrid, geocell, or geocomposite materials. The use of reinforcement in the soil creates a composite material, increasing the bearing capacity and reducing the settlement of the foundations. Several numerical and experimental studies have been reported with regard to the effects of interference on reinforced soil beds. Moreover, several experimental studies on reinforced soil media have also been performed (Khing *et al*., 1992; Al-Ashou *et al*., 1994; Kumar and Saran, 2003a; Ghosh and Kumar, 2009; Lavasan and Ghazavi, 2012a; Naderi and Hataf, 2014; Gupta *et al*., 2018; Saha Roy and Deb 2018; Dehkordi *et al*., 2019; Paikaray *et al*., 2020; Swain and Ghosh, 2015). In this vein, an extensive state of the art review regarding the effect of interference on the behavior of shallow footings was carried out by Ghazavi and Dehkordi (2020).

Most of the studies reported in the literature deal with foundations identical in geometry and loading conditions (symmetrical planning). However, in many situations, structures located close to each other may have different configurations regarding their foundations; the shape of a foundation, as well as its size, may differ, and the super-structural load may be different. This gives rise to asymmetrical interference in the footings.

The subject of interference considering asymmetrical footings is rarely mentioned in the literature. Nainegali *et al*. (2013) highlighted the settlement of asymmetrical interfering strip footings laid on the surface of a linearly elastic non-homogeneous soil medium of finite and semi-infinite extent. Furthermore, they extended the study (Nainegali *et al*., 2018) to examine the interference effect on the bearing capacity and settlement of asymmetrical strip footings on the surface of a non-linearly elastic dense and loose sandy soil medium. In both studies, asymmetry affects the size of the footings, and the surcharge effect is not considered. Ghosh *et al*. (2017) examined the interference of two closely spaced asymmetric footings using the Pasternak model. They considered the hyperbolic stress-strain relation and performed linear and nonlinear elastic analyses. Ekbote and Nainegali (2021) studied two closely spaced asymmetric footings while considering different widths embedded in a homogenous, isotropic soil medium. The study focused on the effect of the width and embedment depth of the two footings, and its results align with earlier reports in conjunction with asymmetry. However, the embedment depth of the footings was also found to have a significant effect.

Most studies have been performed on surface footings, but, from a practical point of view, it is not viable to lay the foundation on the ground level (surface). Therefore, the reported literature does not give a clear picture of the interference effect on embedded footings.

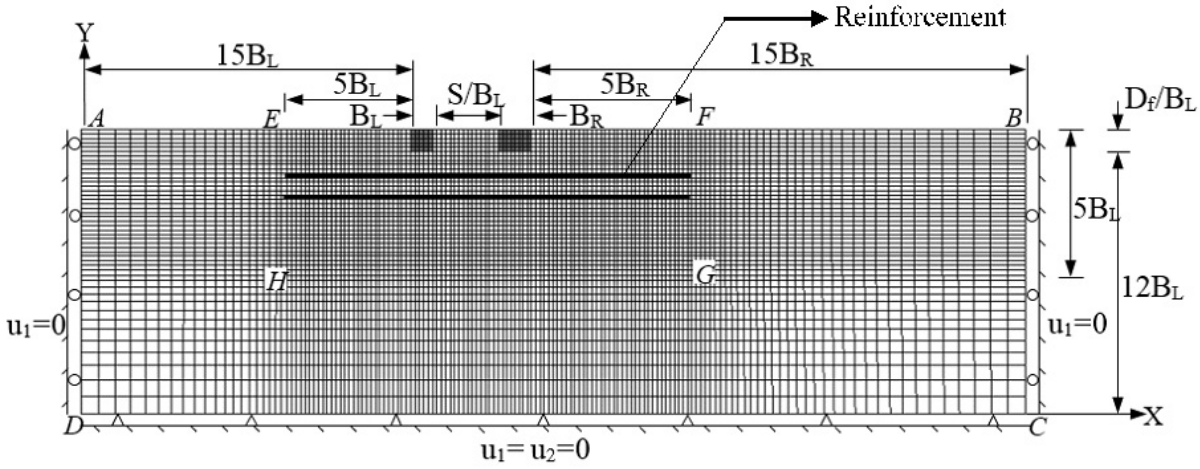


Figure 1. Problem domain along with associated boundary conditions
Source: Authors

For most structures, the foundations are laid below the ground surface (*i.e.*, embedded in nature). The overburden/surcharge from embedment amplifies the bearing capacity and reduces the settlement since there is an enhanced shearing zone. The literature on the effect of interference when footings are in a state of embedment in the reinforced soil medium is virtually null, and there is no study on the subject for embedded shallow footings. Therefore, this phenomenon must be analyzed. Moreover, bearing capacity and settlement are the primary criteria for shallow footing design, so, in order to overcome inadequacies and fill the research gap, this study focuses on these aspects of embedded shallow foundations resting on a reinforced soil medium while considering the embedment depth (same level).

Stating the problem

Figure 1 represents the problem domain, discretized using a structured finite element mesh and associated with appropriate boundary conditions. The problem considers two asymmetrical interfering strip footings with widths of B_L (left footing) and B_R (right footing), embedded at a depth D_f below the ground level, and laid with a clear spacing S. The footings are rigid, rough, and subjected to a load intensity q per-unit length of uniform load intensity. The foundation soil medium is assumed to be a homogeneous, isotropic, and semi-infinite cohesionless medium reinforced with a single and a double layer of geosynthetic material. The first layer (N = 1) of reinforcement is laid at a depth u below the base of the footings, and the consecutive layer (N = 2) at a depth h from the first reinforcement layer. The objective of this research is to perform a parametric study, varying contributing parameters such as the embedment depth of the footings (D_f), the width of the right footing (B_R), the soil friction angle (ϕ), the number of reinforcement layers, and the spacing between the footings (S). This works aims to explore the influence of the interference phenomenon on the behavior of asymmetrical footings (*i.e.*, ultimate bearing capacity, settlement, load-settlement response, and displacement pattern). Note that the right footing width is varied while the left footing width is kept constant. The range of varied parameters is shown in Table 1.

Table 1. Range of the parameters varied in the analyses

Parameters	Range considered for the analysis
B_R/B_L	1,5, and 2,0 (B_L = 1,0 m)
S/B_L	0,0; 0,25; 0,50; 0,75; 1,0; 1,5; 2,0; 3,0; 4,0; 5,0
D_f/B_L	0,25; 0,50; 0,75; 1,0
ϕ	25°; 30°; 35°; 40°
N	1; 2

Source: Authors

Methodology

The numerical analysis of the problem defined in the previous section employs the finite element method via the commercially available ABAQUS V6.13 software. The studied problem is regarded as a plane strain one since the footing length relative to the width is much higher. The interfering footings are considered asymmetrical with respect to the width of the footings, which are embedded at the same depth and loaded with the same magnitude. The width of the left footing is kept constant, and the width of the right interfering footing B_R is varied (Table 1). The Mohr-Coulomb failure criterion is used given its practical significance. It offers the best numerical approximation within the broad confining pressure range, as well as a trustworthy outcome for various nonlinear studies of the ground. Moreover, this criterion is used due to its mathematical simplicity and clear physical significance of the material parameters (Erickson and Drescher, 2002; Labuz and Zang, 2012). Accordingly, the soil medium is modeled as a linearly elastic, perfectly plastic material in compliance with the aforementioned

criterion (Ghazavi and Lavasan, 2008; Mabrouki *et al*., 2010; Shokoohi *et al*., 2018; Zidan and Mohamed, 2019; Sekhar *et al*., 2020). The footings embedded at the required depth are defined as a rigid and rough, linear elastic material (Vivek, 2011; Noorzad and Manavirad, 2014). The thickness of the footings equals their embedment depth. The reinforcement (geosynthetic) material behaves axially and as a linearly elastic material (Basudhar *et al*., 2007; Nazzal *et al*., 2010). The mechanical properties considered for the soil, the footings, and the reinforcement are specified in Table 2.

The foundation soil domain is discretized with structured meshing using two-dimensional plane strain linear continuum elements (CPE4R). A structured discretization meshing scheme is applied, which involves adopting fine mesh in the vicinity of high-stress concentration (around the interfering footings), as well as subsequent coarser meshing in the region of low-stress concentration, such as far-end boundaries. The dimensions are set after a sensitivity analysis regarding domain and mesh size and considering the accuracy of the results. The domain size considered in the analysis is presented in Figure 1. For the element size, elements with a uniform mesh of size 0,2 m are considered within a domain up to five times the width of the footing ($5B_L$) on either side (horizontal direction) and five times the width of the footing ($5B_L$) from the base (vertical direction). Thereupon, the elements vary from 0,2 to 0,8 m at the far-end boundaries using the single bias technique of the ABAQUS meshing module. A detailed mesh convergence study and a domain sensitivity analysis are presented in the recent article published by Ekbote and Nainegali (2021).

Table 2. Mechanical properties of the soil, footings, and reinforcement considered in the analysis

Parameters	Soil[$]	Footing*	Reinforcement (Geosynthetic)[#]
Young's Modulus (E), MPa	32	$25x10^3$	426
Poisson's ratio, υ	0,3	0,2	0,25
Cohesion (c), kPa	2	-	-
Soil friction angle (ϕ)	25° - 40°	-	-
Unit weight, kg/m³	1 600	2 500	946

Source: $Mabrouki *et al*. (2010); *Vivek (2011); #Nazzal *et al*. (2010)

The geosynthetic reinforcement material is modeled as linearly elastic and discretized using two-dimensional quadratic truss elements (T2D3) that resist axial forces and do not resist any bending action. The reinforcement configuration is provided by optimizing the bearing capacity in order to obtain the maximum value. This is also done in Ekbote and Nainegali (2019a). Thus, the depth of the first layer of reinforcement u is 0,35 times B_L from the bottom of the footing, the distance between the first and the second layer h is 0,4 times B_L, and the length of reinforcement layer b is 5,0 times B_L. The interface between the soil and the footing, as well as between the soil and the reinforcement, is modeled using the interaction feature of ABAQUS (surface-to-surface, master-slave contact). Using this feature, there is a master and a slave surface. Defining such surfaces automatically generates a group of contact elements. The interaction simulation consists of normal-to-the-surface and tangential-to-the-surface components. The interface in the normal direction is assumed as hard contact with no separation allowed between the surfaces. The shear interactions are defined using Coulomb's friction law, by virtue of which a tangential behavior is defined between the contact surfaces a coefficient of friction μ, which is equal to the tangent of soil friction angle ϕ ($\mu = \tan \phi$). Appropriate boundary conditions essential for the solution approximation are established along the horizontal and vertical boundaries of the foundation soil domain. The suitable boundary conditions at the extreme boundaries are assigned in line with Potts and Zdravkovic (1999). Roller supports are considered along the far-end boundaries AD and BC in order to restrict the horizontal displacement ($u_1 = 0$) and allow any possible vertical displacement (u_2) at the far ends. Otherwise, the boundary DC is fixed to restrict horizontal and vertical displacements ($u_1 = u_2 = 0$).

Validation

This model was validated with the results of Nainegali *et al*. (2013, 2018). Nainegali *et al*. (2018) conducted a finite element analysis to study the settlement behavior of adjacent footings while considering the conditions of symmetry and asymmetry with respect to footing size and the applied loading with different combinations. Rough, rigid strip footings were placed on the surface of either a linearly elastic finite or an infinite non-homogeneous soil bed. Nainegali *et al*. (2013) presented their results in terms of the interference factor, which is the ratio of the average settlement of left/right footing in the presence of the right/left footing to that of the settlement of isolated footings. Nainegali *et al*. (2018) also performed a finite element analysis of two closely spaced strip footings on the surface of dense, loose, cohesionless soil beds. The footings were considered either symmetrical or asymmetrical based on their size. The soil was modeled following the non-linearly elastic soil, obeying the Duncan and Chang hyperbolic constitutive relationship. The results were presented in terms of interference factors such as the ratio of bearing pressure/settlement of interfering footings to that of isolated footings.

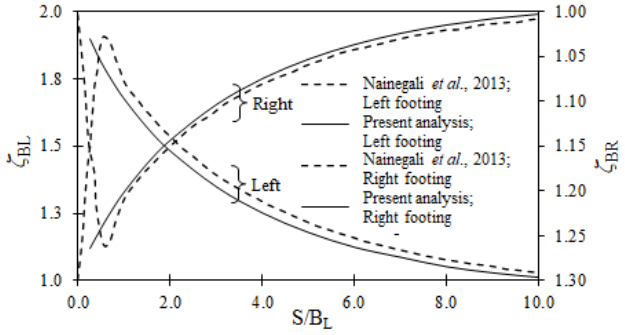
Table 3. Mechanical properties of soil considered for validation

Parameters	γ (kg/m³)	E (kPa)	ν	c (kPa)	ϕ	ψ
Nainegali *et al*. (2013)	-	$30x10^3$	0,3	-	-	-
Nainegali *et al*. (2018)	1 700	$30x10^3$	0,3	0,1	$36,5^0$	$0,1^0$

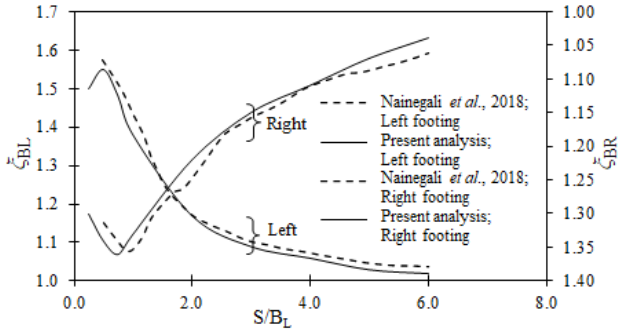
Source: Nainegali *et al*. (2013, 2018)

The adopted model was validated by adopting the soil properties presented in Table 3 (Nainegali *et al*., 2013, 2018). The results of the comparison are presented in Figures 2a and 2b regarding Nainegali *et al*. (2013) and Nainegali *et al*. (2018), respectively. The variation was found to be similar and almost matching. However, a slight

difference was observed in our results, given the difference in the discretization scheme and the size of the elements considered. The differences in the computed interference factor at S/B_L = 0,50 for the left and right footings when compared to that of Nainegali *et al*. (2013) are 5 and 1%, respectively. Similarly, these differences, in comparison with Nainegali *et al*. (2018), are 1,5% and 1,9%, respectively. Overall, it can be seen that this finite element analysis model can be used reliably to address the aforementioned problem.



a) Comparison with Nainegali *et al*. (2013)



b) Comparison with Nainegali *et al*. (2018)
Figure 2. Results comparison with Nainegali *et al*. (2013, 2018)
Source: Authors

Results

A rigorous finite element analysis was performed to investigate the effect of interference on closely placed footings while considering asymmetry with respect to dimension (width), emphasizing the embedment of the footings in a geosynthetic reinforced soil medium. The analysis results regarding non-dimensional interference factors associated with the ultimate bearing capacity and the settlement were interpreted. The interference factor for the ultimate bearing capacity was estimated as defined in Equations (1a) and (1b) for the left and right footings, respectively. Similarly, the interference factor for the settlement was estimated according to the definitions given in Equations (2a) and (2b), for the left and right footings.

The permissible settlement (δ^p) is the maximum settlement allowed, irrespective of the footing type and size. As per the Indian Standard Code (IS 1904-1986), the maximum permissible settlement is 50 mm for isolated foundations in sand and hard clay for steel and reinforced concrete structures. The permissible bearing pressure (q^p) is the pressure corresponding to the δ^p assumed for an isolated footing. q^p (which corresponds to δ^p = 50mm) for different footing widths, depth of embedment, and soil friction angles was estimated, in the case of an unreinforced soil medium, by Ekbote and Nainegali (2021) (Table 4). These values are used to evaluate the interference factor for the settlement, as defined in Equations (2a) and (2b).

$$\xi_L^r = \frac{\text{Ultimate bearing capacity of the left footing in presence of the right footing embedded in reinforced soil medium}}{\text{Ultimate bearing capacity of an isolated footing identical to that of the left footing embedded in an unreinforced soil medium}} \quad (1a)$$

$$\xi_R^r = \frac{\text{Ultimate bearing capacity of the right footing in presence of the left footing embedded in reinforced soil medium}}{\text{Ultimate bearing capacity of an isolated footing identical to that of the right footing embedded in an unreinforced soil medium}} \quad (1b)$$

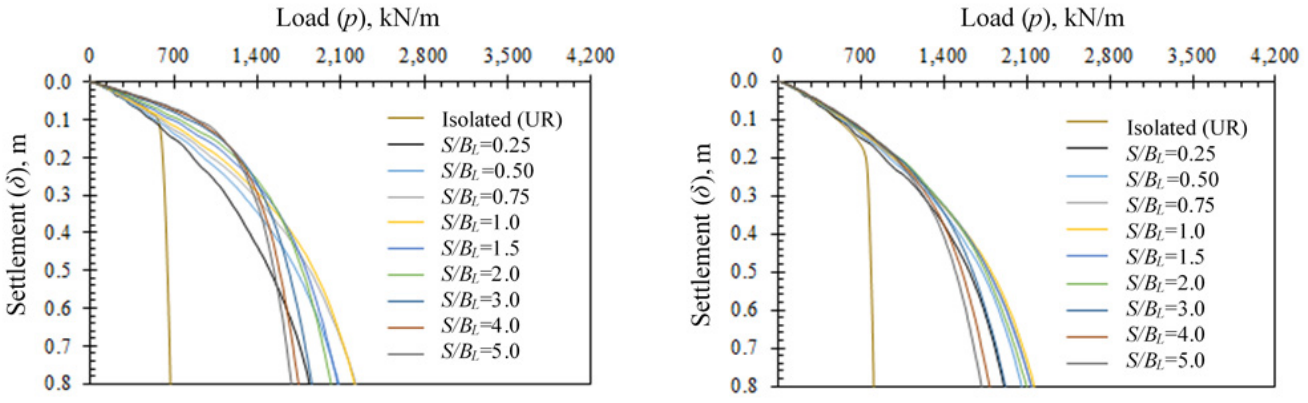
$$\zeta_L^r = \frac{\text{Settlement of the left footing embedded in reinforced soil medium (corresponding to permissible bearing pressure of an isolated footing same in size of the left footing, embedded in unreinforced soil medium)}}{\text{Permissible settlement}} \quad (2a)$$

$$\zeta_R^r = \frac{\text{Settlement of the right footing embedded in reinforced soil medium (corresponding to permissible bearing pressure of an isolated footing same in size of the right footing, embedded in unreinforced soil medium)}}{\text{Permissible settlement}} \quad (2b)$$

Table 4. Permissible bearing pressure q^p corresponding to the permissible settlement δ^p for an isolated footing

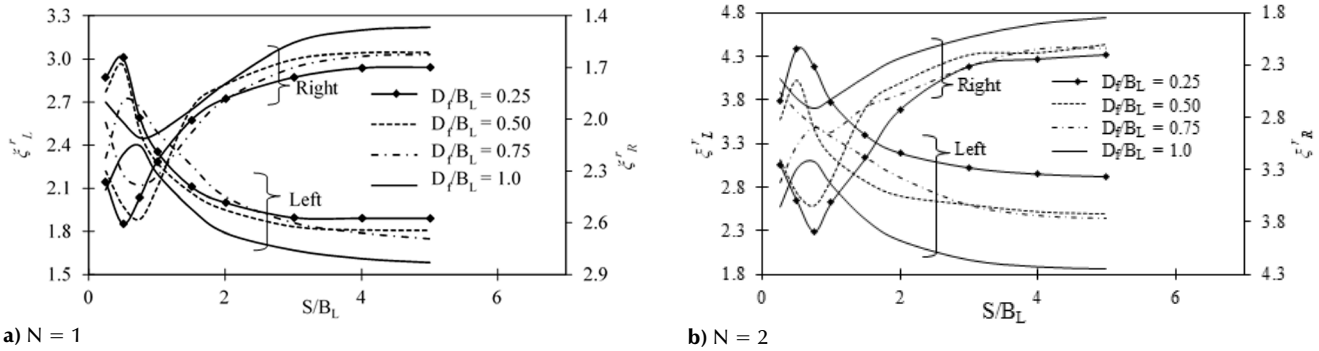
B	D_f/B_L	ϕ, (°)	q^p, kN/m	D_f/B_L	ϕ, (°)	q^p, kN/m
1,0	0,25	25	158,4	0,75	25	246,4
		30	279,3		30	362,6
		35	427,0		35	465,1
		40	528,3		40	521,4
	0,50	25	208,4	1,00	25	254,5
		30	337,1		30	369,6
		35	444,6		35	471,2
		40	533,4		40	492,9
1,5	0,25	25	185,1	0,75	25	229,7
		30	290,7		30	321,1
		35	373,8		35	396,0
		40	439,1		40	434,8
	0,50	25	210,9	1,00	25	233,5
		30	304,3		30	325,3
		35	380,7		35	399,9
		40	439,6		40	426,6
2,0	0,25	25	190,8	0,75	25	215,8
		30	272,8		30	294,3
		35	336,1		35	355,5
		40	386,8		40	386,1
	0,50	25	201,5	1,00	25	192,8
		30	279,5		30	298,4
		35	341,5		35	357,8
		40	388,3		40	375,6

Source: Ekbote and Nainegali (2021)



a) Left footing **b)** Right footing

Figure 3. p-δ plots for interfering asymmetric footings embedded in reinforced soil. D_f/B_L = 1,0, ϕ = 30°, B_R = 2,0B_L, N = 2.
Source: Authors



a) N = 1 **b)** N = 2

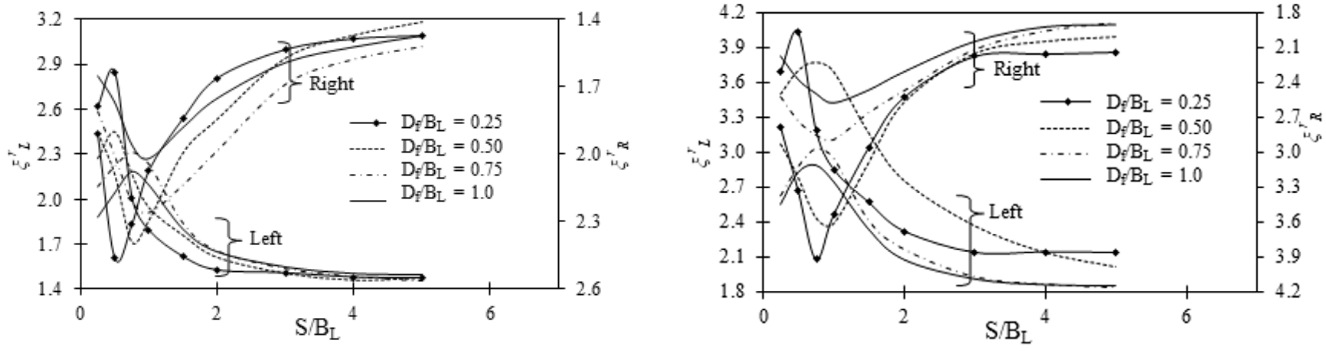
Figure 4. Variation of ξ_L^r and ξ_R^r with the S/B_L ratio for various D_f/B_L values in soil with ϕ = 30°: B_R = 1,5B_L
Source: Authors

The load-settlement (p-δ) response of the left and right asymmetrical interfering footings placed at different spacings (S/B_L ratio) for the case of B_R = 2,0B_L, embedded at depth, and D_f/B_L=1,0 in soil a having friction angle of ϕ=30° and reinforced with a single layer (N = 2) of reinforcement are illustrated in Figures 3a and 3b. The p-δ curve of the corresponding identical isolated footing embedded in an unreinforced soil bed (Isolated (UR)) is also shown in the figures for comparison. It can be noted that the response of asymmetrical interfering footings differs when compared to that of isolated ones. An explanation for the increased bearing capacity with a decreased footing spacing can be found in Stuart (1962). The load-carrying capacity increases with an increase in the number of reinforcement layers. The enhanced shear strength of the reinforced soil bed, which also increases with the number of reinforcing layers, is responsible for the higher load bearing capacity. According to Shukla (2021), the friction between the soil and the reinforcement layers primarily drives the reinforced soil's composite system. Additionally, the presence of or increase in reinforcement layers transforms the mechanism of soil shear failure into one of general shear failure (Guido *et al.*, 1985). From Figure 3, it can be stated that the p-δ curves of the interfering footings for spacings greater than 3,0 are more or less overlapping. This indicates that the interfering footings are starting to act as an isolated one. It cannot be ignored that there is a significant effect due to the asymmetry of the base (width of the adjacent base), *i.e.*, there is a behavioral difference in p-δ for both the left and the right footings. For the smaller left footing, the p-δ curves for different spacings are more distributed than those of the larger right footing.

By obtaining the bearing capacity from the load-settlement curves, the interference factors associated with the ultimate bearing capacity for the left footing (ξ_L^r) and the right footing (ξ_R^r) were evaluated for the range of parameters varied in the analysis, using Equations (1a) and (1b), respectively. The load-settlement curves obtained for the interfering footings do not show a definitive failure point. Therefore, in order to obtain the ultimate bearing capacity, a double-tangent method was used, referring to Naderi and Hataf (2014), Ghosh *et al.* (2015), and Saha Roy and Deb (2018). The variations in ξ_L^r and ξ_R^r with spacing (S/B_L ratio) for footing embedment at different depths (D_f/B_L ratio) in soil having a friction angle of ϕ = 30° for the case of B_R = 1,5B_L are presented in Figures 4a and 4b for single (N = 1) and double (N = 2) layers of reinforcement. The case of B_R = 2,0B_L is illustrated in Figures 5a and 5b. Similarly, Figures 6a and 6b, for N = 1 and 2, present these variations for the case of B_R

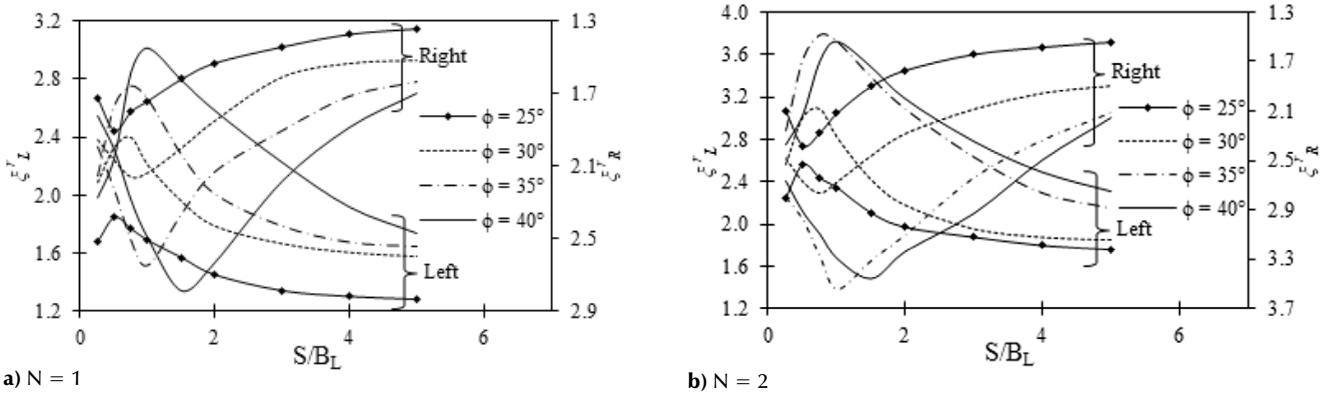
$= 1{,}5B_L$, footings embedded at $D_f/B_L = 1{,}0$, and different soil friction angles. The same is illustrated in Figures 7a and 7b for the case of $B_R = 2{,}0B_L$.

It can be observed, in Figures 4a, 4b, 5a, and 5b, that the footings embedded at smaller depths have higher interference factors. The interference factors of each footing are presented on either side of the figures. The peak interference factors ($\xi_L^{r,max}$ and $\xi_R^{r,max}$) decrease progressively with an increased embedment depth of the footings. Moreover, the peak interference factors increase with the soil friction angle. The spread of the plastic zone beneath the footing grows with an increase in soil friction angle, which is often related to an increasing shear strength. Additionally, when the soil friction angle increases, the failure pattern expands in the lateral and downward directions with increasing scope (Kouzer and Kumar, 2010; Yang *et al*., 2017). Furthermore, Figures 6a, 6b, 7a, and 7b show that the spacings at which the peak occurs increase for the right footing when compared to the left footing.



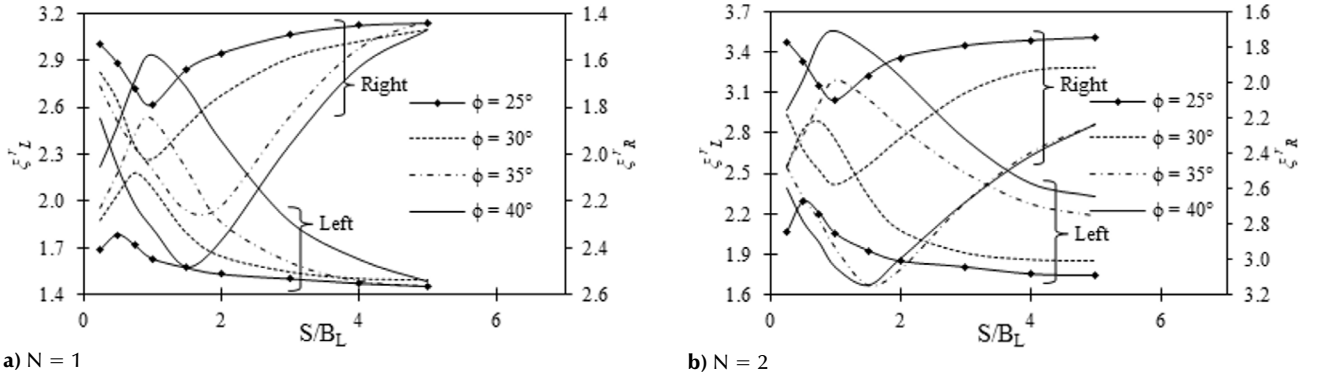
a) N = 1 **b)** N = 2

Figure 5. Variation of ξ_L^r and ξ_R^r with the S/B_L ratio for various D_f/B_L values in soil with $\phi = 30°$: $B_R = 2{,}0B_L$
Source: Authors



a) N = 1 **b)** N = 2

Figure 6. Variation of ξ_L^r and ξ_R^r with the S/B_L ratio for various soil friction angles at $D_f/B_L = 1{,}0$: $B_R = 1{,}5B_L$
Source: Authors



a) N = 1 **b)** N = 2

Figure 7. Variation of ξ_L^r and ξ_R^r with the S/B_L ratio for various soil friction angles at $D_f/B_L = 1{,}0$: $B_R = 2{,}0B_L$
Source: Authors

In Figures 4-7, knowing the interference factor, the percent increase in the bearing capacity of the interfering footings can be obtained in comparison with the corresponding isolated footing, *i.e.*, [(ξ_L^r or ξ_R^r – 1,0) x 100]. As an example, it can be noted that, for D_f/B_L = 1,0 for the left footing when B_R = 2,0B_L, placed with one layer of reinforcement (N = 1) and ϕ = 25°, 30°, 35°, and 40° the percent increase in the bearing capacity of the interfering footings at S/B_L = 0,50 is 78, 104, 118, and 148%, respectively, compared to their corresponding isolated footing. In the case of a specified angle of soil friction and spacing ratio, according to Equation (1), the interference factor increases with a decrease in the depth of the footing embedment. However, the load-carrying capacity of the soil increases with an increase in the embedment depth of the footing, due to an increased shearing zone and enhanced overburdening. The bearing capacity predominantly increases with an increase in the embedment depth for both isolated and interfering footings. As an example, consider a specified ϕ = 30° and S/B_L = 0,5 for the case B_R = 2,0B_L. The bearing capacity of the left footing when N = 1 increases from 104 to 185%, with a decrease in the depth of footing embedment from D_f/B_L = 1,0 to 0,25. The improvement in the ultimate bearing capacity for symmetrical interfering footings in a soil medium reinforced with N-layers of geosynthetic material is evident in earlier research on surface footings (Ghazavi and Lavasan, 2008; Lavasan *et al*., 2017; Biswas and Ghosh, 2018) and embedded footings (Ekbote and Nainegali, 2019a), which holds true in the case of asymmetric footings. However, $\xi_L^{r,max}$ is higher than $\xi_R^{r,max}$. The peak interference factors for N = 2 are shown in Table 5 for various embedment depths, soil friction angles, and footing widths. For the footings embedded at D_f/B_L = 1,0, the percentage difference in $\xi_L^{r,max}$ and $\xi_R^{r,max}$ for ϕ = 25°, 30°, 35° and 40° for B_R = 1,5B_L is 8,6, 12,3, 7,7, and 8,6%, respectively. Similarly, for B_R = 2,0B_L, these values are 8,3, 10,7, 1,6, and 11,5%, respectively. The percentage decrease in $\xi_L^{r,max}$ for N = 2, B_R = 2,0B_L when the embedment depth is increased from 0,25 to 1,0 for ϕ = 25°, 30°, 35° and 40° is 43,23, 39,44, 53,75, and 42,41%, respectively, and the percentage decrease in $\xi_R^{r,max}$ for ϕ = 25°, 30°, 35°, and 40° is 73, 52, 40, and 42%.

The settlement analysis of asymmetrical interfering footings embedded in a reinforced soil medium was carried out by evaluating the settlement at the bearing pressure corresponding to the permissible settlement of an isolated footing (50 mm) embedded in an unreinforced soil medium. The permissible bearing pressures corresponding to the permissible settlement for given values regarding the right footing width, embedment depth, and soil friction angle is shown in Table 4. Accordingly, the interference factors for the settlement of the left footing (ζ_L^r) are calculated using Equation (2a) and those for the right footing (ζ_R^r) using Equation (2b). The variations of ζ_L^r and ζ_R^r with spacings for N = 1 and 2 and the unreinforced case (N = 0) are presented in Figures 8a, 8b, 8c, and 8d for footings embedded at depths D_f/B_L = 0,25, 0,50, 0,75, and 1,0, respectively, with B_R = 1,5B_L and ϕ = 30°. The magnitude of the settlement interference factors of the left footings (ζ_L^r) are presented on the left-hand side of the figures, and those of the right footings (ζ_R^r) are presented on the right-hand side with values in reverse order. The width of the right footing (B_R = 2,0B_L) is shown in Figures 9a to d. Similarly, the variations of ζ_L^r and ζ_R^r with spacings for different soil friction angles (ϕ = 25°, 30°, 35°, and 40°) and D_f/B_L = 1,0 when B_R = 1,5B_L and B_R = 2,0B_L are presented in Figures 10a to d and 11a to d. The settlement interference factor continuously decreases with increased spacing between the footings for all the soil friction angles and embedding depths considered as the arching. The confinement effect of soil diminishes with the increasing proximity of the footings. The maximum settlement ($\zeta_L^{r,max}$ and $\zeta_R^{r,max}$) occurs at the minimum spacing considered (S/B_L = 0,25). However, the magnitude of the settlement decreases with an increase in both the embedment depth and the friction angle. The settlement was found to decrease further with the introduction of reinforcement layers, which is due to an increase in the shear strength of the soil.

The percentage decrease for the maximum settlement, given a footing width B_R = 1,5B_L, with N = 1 and ϕ = 30°, when the footing embedment depth is increased from D_f/B_L = 0,25 to 1,0, is 21% for the left footing and 20,7% for the right one. Similarly, the percentage decrease in the

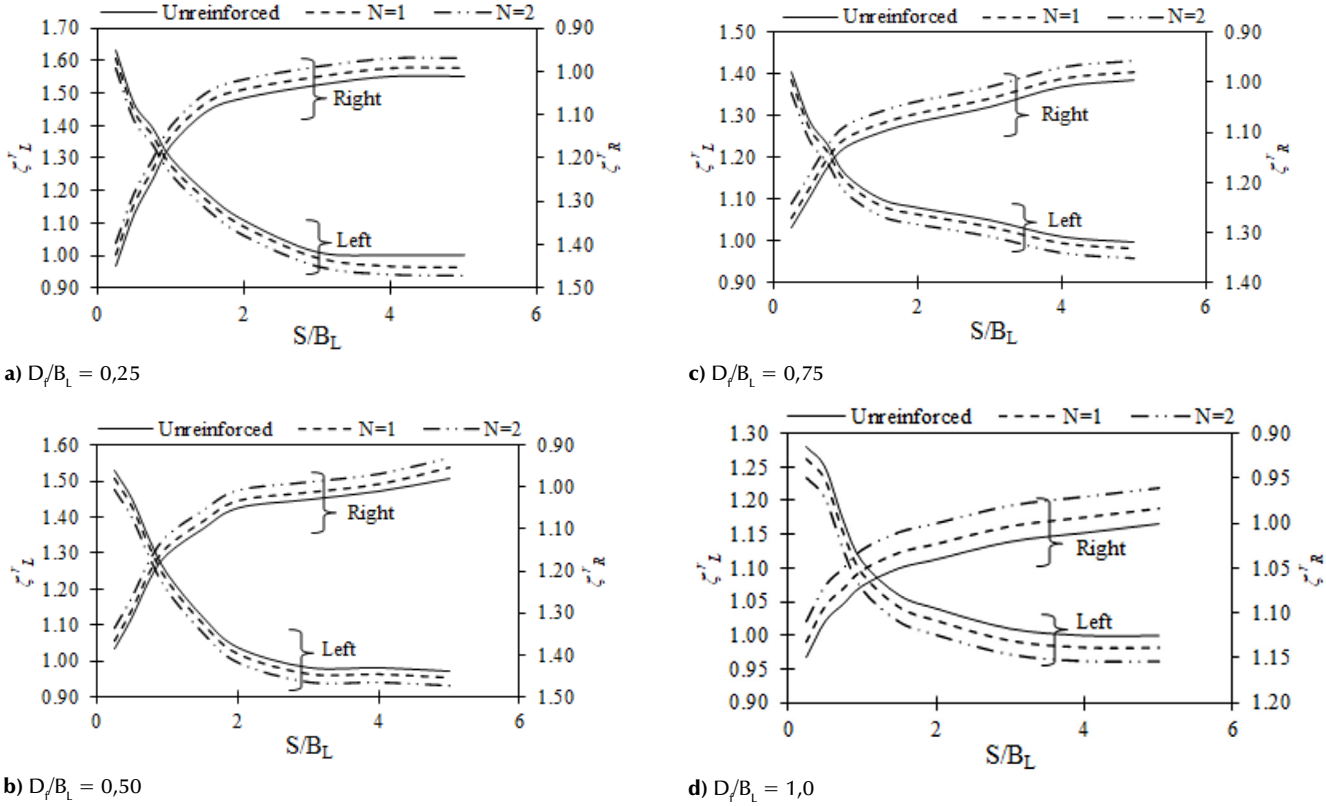
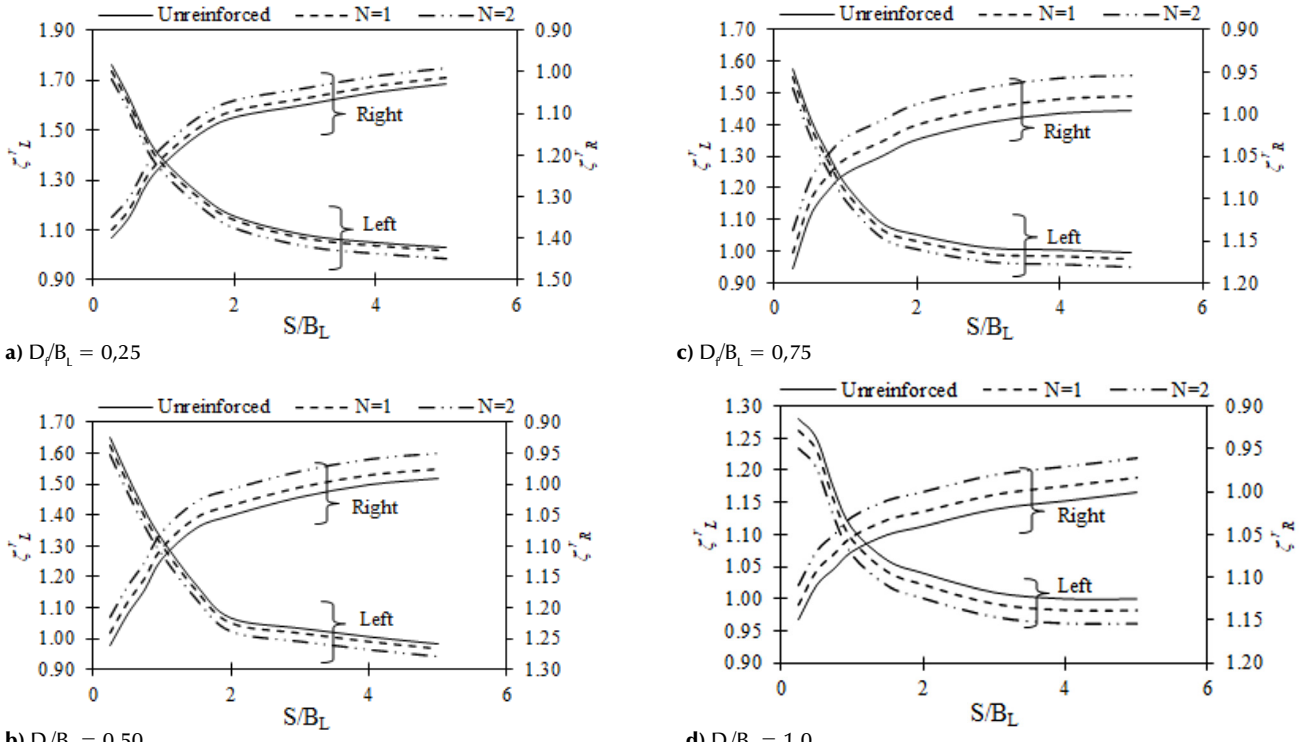


Figure 8. Variations in ζ_L^r and ζ_R^r with S/B_L ratio for N = 0 (unreinforced), 1, and 2 at different embedment depths in soil with ϕ = 30°: B_R = 1,5B_L
Source: Authors

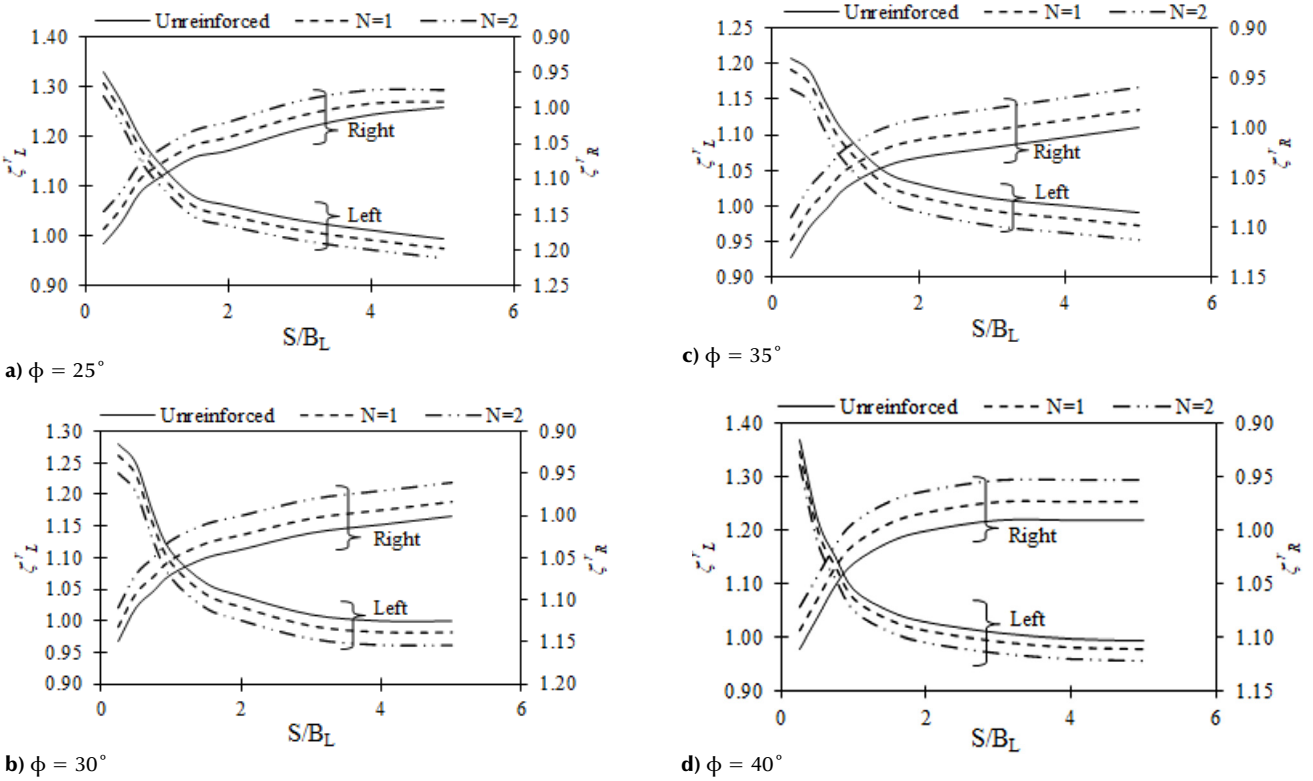
maximum settlement when the embedment depth is increased from D_f/B_L = 0,25 to 1,0 for $B_R = 2{,}0B_L$, N = 1, and ϕ = 30° is 22,3% for the left footing and 20,9% for the right one. Furthermore, the decreases in the settlement when the soil friction angle is increased from ϕ = 25° to 40° for D_f/B_L = 1,0, $B_R = 1{,}5B_L$, and N = 1, for the left and right footings, respectively. These values (when $B_R = 2{,}0B_L$) are 14% for the left footing and 7,2% for the right one. Similarly, the decrease in the settlement when the soil friction angle is increased from ϕ = 25° to 40° for D_f/B_L = 1,0, $B_R = 1{,}5B_L$, and N = 2 is 11,1% (left footing) and 5,8% (right footing). For $B_R = 2{,}0B_L$, these values are 13,9 and 7,2%, respectively.



a) D_f/B_L = 0,25 **c)** D_f/B_L = 0,75

b) D_f/B_L = 0,50 **d)** D_f/B_L = 1,0

Figure 9. Variations in ζ_L^r and ζ_R^r with S/BL ratio for N = 0 (unreinforced), 1, and 2 at different embedment depths in soil with ϕ = 30°: BR = 2,0BL
Source: Authors



a) ϕ = 25° **c)** ϕ = 35°

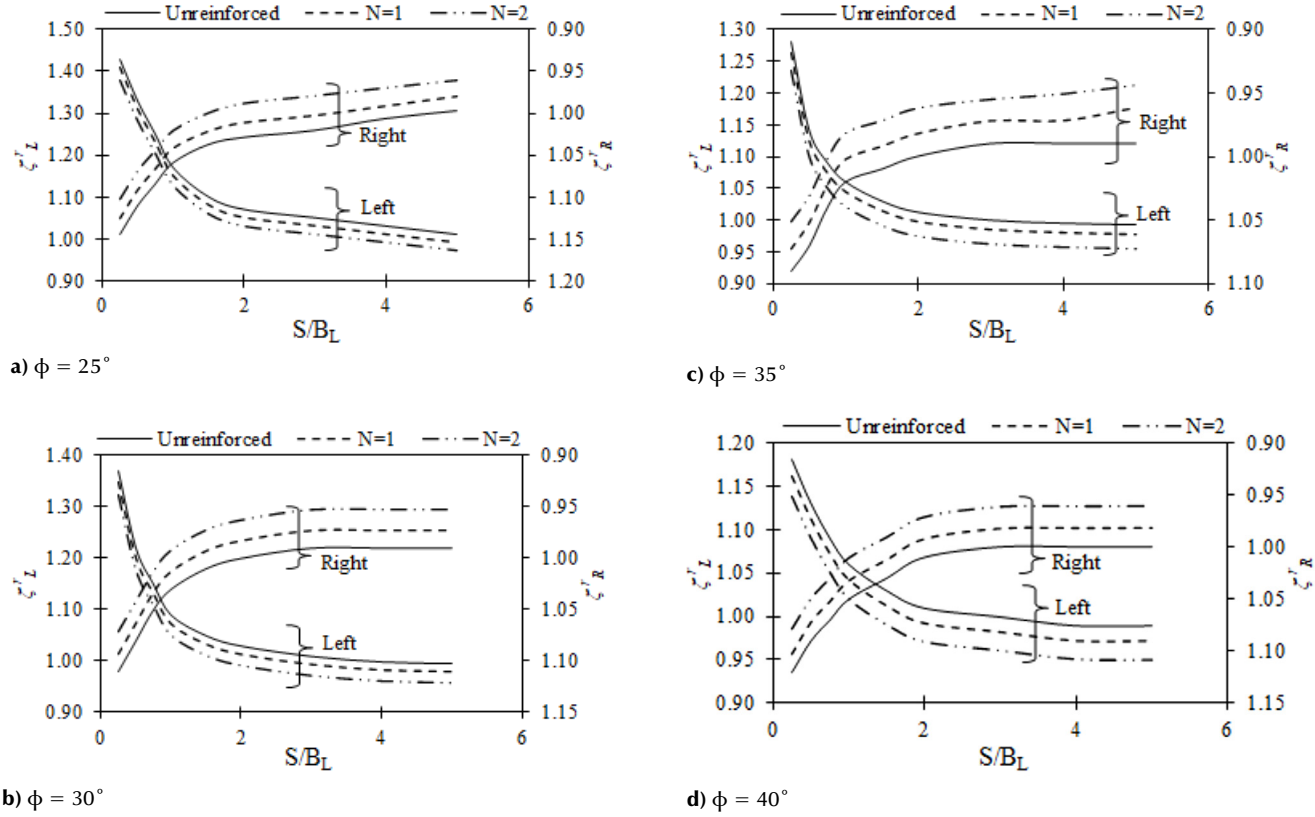
b) ϕ = 30° **d)** ϕ = 40°

Figure 10. Variations in ζ_L^r and ζ_R^r with the S/B_L ratio for N = 0 (unreinforced), 1, and 2 in soil with different ϕ, with D_f/B_L = 1,0: $B_R = 1{,}5B_L$
Source: Authors

Table 5. Peak interference factors ($\xi_L^{r,max}$ and $\xi_R^{r,max}$) for different D_f/B_L, ϕ, and B_R/B_L values for two layers of reinforcement

D_f/B_L	ϕ	$B_R = 1,5B_L$		$B_R = 2,0B_L$		D_f/B_L	ϕ	$B_R = 1,5B_L$		$B_R = 2,0B_L$	
		$\xi_L^{r,max}$	$\xi_R^{r,max}$	$\xi_L^{r,max}$	$\xi_R^{r,max}$			$\xi_L^{r,max}$	$\xi_R^{r,max}$	$\xi_L^{r,max}$	$\xi_R^{r,max}$
0,25	25°	2,61	2,67	3,28	3,64	0,75	25°	2,69	2,65	2,62	2,21
	30°	4,39	3,89	4,03	3,92		30°	3,48	2,95	3,03	2,89
	35°	3,97	3,71	4,92	4,41		35°	3,89	3,59	3,78	3,36
	40°	5,14	4,97	5,07	4,42		40°	3,78	3,68	3,61	3,47
0,50	25°	2,92	2,70	2,67	2,44	1,0	25°	2,56	2,34	2,29	2,10
	30°	4,02	3,64	3,77	3,62		30°	3,09	2,71	2,89	2,58
	35°	4,00	3,68	3,89	3,73		35°	3,78	3,49	3,20	3,15
	40°	3,76	3,73	3,83	3,76		40°	3,72	3,40	3,56	3,15

Source: Authors



a) ϕ = 25°

c) ϕ = 35°

b) ϕ = 30°

d) ϕ = 40°

Figure 11. Variations in ζ_L^r and ζ_R^r with the S/B_L ratio for N = 0 (unreinforced), 1, and 2 in soil with different ϕ and D_f/B_L = 1,0: $B_R = 2,0B_L$
Source: Authors

The failure of the two closely placed asymmetrical footings can be visualized through the total displacement contour plots and is presented for $B_R = 2,0B_L$ and $D_f/B_L = 1,0$ in a soil medium (ϕ = 30°) with a double layer of geosynthetic reinforcement (Figures 12a, 12b, 12c, and 12d) for footings spaced at S/B_L = 0,25, 0,50, 1,0, and 5,0. The plots correspond to the ultimate bearing capacity. The displacement vectors are also shown over the plots to represent the direction and magnitude. Figures 12a to d show a general form of shear failure. It can be observed (Figures 12a and 2b) that the displacement vectors in between the two footings are absent, *i.e.*, the soil heave is not observed. This implies that adjacent footings with very close spacings (S/B_L = 0,25 and 0,50) act as one unit and relate to the single footing having a greater width (B_L+S/B_L+B_R). Furthermore, with an increase in spacing (Figure 12c), the heave between the footings is observed. However, this coincides with the passive zones of the close footings, and, at far away spacings (Figure 12d), the footings start behaving as independent ones. When the soil medium is reinforced, the zone between the two footings gets more confined than in unreinforced soil. Thus, the soil starts to densify until the shear failure occurs. The load carrying capacity was found to increase with a decrease in the settlement. The failure in the reinforced soil medium was assumed to be related to shear given a slip between the soil and the reinforcement (Basudhar *et al.*, 2008).

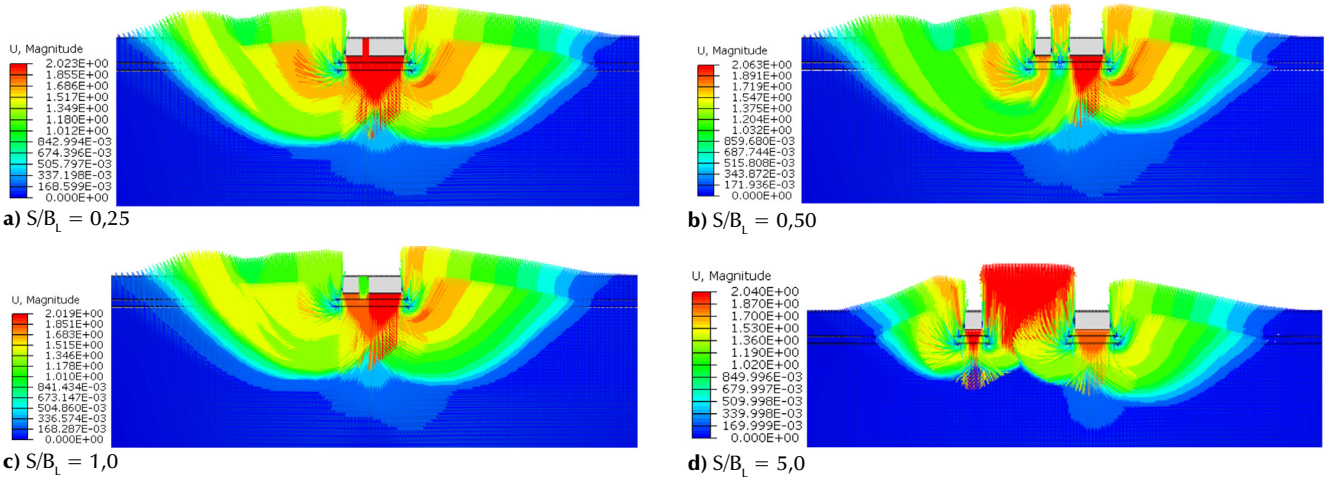


Figure 12. Total displacement contour plot for asymmetrical interfering footings on reinforced soil with $\phi = 30^{\circ}$ for $D_f/B_L = 1,0$, $B_R/B_L = 2,0$, and N = 2 at different S/B_L
Source: Authors

Conclusions

The performance of two closely spaced asymmetric footings embedded in a cohesionless reinforced foundation soil medium was studied using finite element analysis. Asymmetry was considered only with respect to the width of the footings, and thus the entire domain was used in the analysis while adopting appropriate boundary conditions. A sensitivity analysis for the domain and element size was performed, and the finite element model was validated with benchmark problems reported in the literature. The influence of the interference phenomenon on the load-settlement response, the ultimate bearing capacity, the settlement, and the displacement pattern was analyzed by varying several contributing parameters. Observations on the ultimate bearing capacity and the settlement were presented in terms of non-dimensional interference factors defined for this purpose. This study leads to the following conclusions:

- The effect of the interference phenomenon is significant, and its influence is asymmetry, that is, the influence of the large footing on the small adjacent one is predominant.
- The depth of footing embedment and the soil friction angle have a noteworthy influence on the interference effect.
- For a given depth of embedment, the peak interference factor associated with the ultimate bearing capacity increases with an increase in the soil friction angle. It also decreases with an increase in the embedment depth of the footing for a given value of soil friction angle.
- The interference factors associated with the ultimate bearing capacity for the reinforced soil medium (ξ_L^r and ξ_R^r) vary similarly to those of unreinforced soil with spacing between the footings.
- In comparison with the unreinforced soil medium, the settlement analysis for the reinforced soil medium (which considered the corresponding permissible limit) was found to follow a similar variation regarding the interference factors (ξ_L^r and ξ_R^r) and the spacing between the footings. However, the magnitudes were found to decrease, which accounts for the reduced settlement. Reductions in the settlement magnitude of 1,5 and 4,5% were noted for N = 1 and N = 2, respectively.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (original draft, writing, review, and editing), data curation.

References

Al-Ashou, M. O., Sulaiman, R. M., and Mandal, J. N. (1994). Effect of number of reinforcing layers on the interference between footings on reinforced sand. *Indian Geotechnical Journal, 24*(3), 285-301.

Alzabeebee, S. (2020). Numerical analysis of the interference of two active machine foundations. *Geotechnical and Geological Engineering, 38*(5), 5043-5059. https://doi.org/10.1007/s10706-020-01347-w

Alzabeebee, S. (2022). Interference of surface and embedded three strip footings in undrained condition. *Transportation Infrastructure Geotechnology, 9*(2), 250-267. https://doi.org/10.1007/s40515-021-00172-9

Basudhar, P. K., Dixit, P. M., Gharpure, A., and Deb, K. (2008). Finite element analysis of geotextile-reinforced sand-bed subjected to strip loading. *Geotextiles and Geomembranes, 26*(1), 91-99. https://doi.org/10.1016/j.geotexmem.2007.04.002

Basudhar, P. K., Saha, S., and Deb, K. (2007). Circular footings resting on geotextile-reinforced sand bed. *Geotextiles and Geomembranes, 25*(6), 377-384. https://doi.org/10.1016/j.geotexmem.2006.09.003

Biswas, N., and Ghosh, P. (2018). Interaction of adjacent strip footings on reinforced soil using upper-bound limit analysis. *Geosynthetics International, 25*(6), 599-611. https://doi.org/10.1680/jgein.18.00020

Ekbote, A. G., and Nainegali, L. (2019a). Interference of two closely spaced footings embedded in unreinforced and reinforced soil medium: a finite element approach using ABAQUS. *Arabian Journal of Geosciences, 12*(22), 1-21. https://doi.org/10.1007/s12517-019-4868-0

Ekbote, A. G., and Nainegali, L. (2019b). Interference of two closely spaced strip footings embedded in cohesionless fibre-reinforced foundation soil bed. In American Society of Civil Engineers, *Geo-Congress 2019: Foundations* (pp. 454-464). American Society of Civil Engineers. https://doi.org/10.1061/9780784482094.041

Ekbote, A. G., and Nainegali, L. (2021). Finite element analysis of two nearby interfering asymmetric footings embedded in cohesionless foundation medium. *Geomechanics and Geoengineering, 16*(4), 263-276. https://doi.org/10.1080/17486025.2019.1664776

Ekbote, A. G., Nainegali, L., Rajhans, P., and Deepak, M. S. (2022). Behavioural assessment on influence of adjacently placed strip footings at different embedment level. *Architecture, Civil Engineering, Environment, 15*(4), 93-103. https://doi.org/10.2478/ACEE-2022-0041

Erickson, H. L., and Drescher, A. (2002). Bearing capacity of circular footings. *Journal of Geotechnical and Geoenvironmental Engineering, 128*(1), 38-43. http://ojps.aip.org/gto

Fazeli Dehkordi, P., Ghazavi, M., Ganjian, N., and Karim, U. F. A. (2021). Parametric study from laboratory tests on twin circular footings on geocell-reinforced sand. *Scientia Iranica, 28*(1), 96-108. https://doi.org/10.24200/SCI.2019.51471.2208

Fuentes, W., Duque, J., Lascarro, C., and Gil, M. (2019). Study of the bearing capacity of closely spaced square foundations on granular soils. *Geotechnical and Geological Engineering, 37*(3), 1401-1410. https://doi.org/10.1007/s10706-018-0694-5

Ghazavi, M., and Lavasan, A. A. (2008). Interference effect of shallow foundations constructed on sand reinforced with geosynthetics. *Geotextiles and Geomembranes, 26*(5), 404-415. https://doi.org/10.1016/j.geotexmem.2008.02.003

Ghosh, P., and Kumar, P. (2009). Interference effect of two nearby strip footings on reinforced sand. *Contemporary Engineering Sciences, 2*(12), 577-592. http://m-hikari.com/ces/ces2009/ces9-12-2009/ghoshCES9-12-2009.pdf

Ghosh, P., and Sharma, A. (2010). Interference effect of two nearby strip footings on layered soil: theory of elasticity approach. *Acta Geotechnica, 5*(3), 189-198. https://doi.org/10.1007/s11440-010-0123-2

Ghosh, P., Basudhar, P. K., Srinivasan, V., and Kunal, K. (2015). Experimental studies on interference of two angular footings resting on surface of two-layer cohesionless soil deposit. *International Journal of Geotechnical Engineering, 9*(4), 422-433. https://doi.org/10.1179/1939787914Y.0000000080

Ghosh, P., Rajesh, S., and Sai Chand, J. (2017). Linear and nonlinear elastic analysis of closely spaced strip foundations using Pasternak model. *Frontiers of Structural and Civil Engineering, 11*(2), 228-243. https://doi.org/10.1007/s11440-010-0123-2

Graham, J., Raymond, G. P., and Suppiah, A. (1984). Bearing capacity of three closely-spaced footings on sand. *Geotechnique, 34*(2), 173-181. https://doi.org/10.1680/geot.1984.34.2.173

Griffiths, D. V., Fenton, G. A., and Manoharan, N. (2006). Undrained bearing capacity of two-strip footings on spatially random soil. *International Journal of Geomechanics, 6*(6), 421-427. https://doi.org/10.1061/(ASCE)1532-3641(2006)6:6(42

Guido, V. A., Biesiadecki, G. L., and Sullivan, M. J. (1985). *Bearing capacity of a geotextile-reinforced foundation* [Conference presentation]. Proceedings of the Eleventh International Conference on Soil Mechanics and Foundation Engineering, San Francisco, CA, USA. https://www.issmge.org/uploads/publications/1/34/1985_03_0134.pdf

Gupta, A., Lakshman, G. K., and Sitharam, T. (2018). *Interference of square footings on geocell reinforced clay bed: Experimental and numerical studies* [Conference presentation]. 3rd World Congress on Civil, Structural, and Environmental Engineering (CSEE'18), Budapest, Hungary. https://doi.org/10.11159/icgre18.142

Hansen, B. J. (1970). *A revised and extended formula for bearing capacity. No. 28*. Danish Geotechnical Institute. https://trid.trb.org/view/125129

Hill, R. (1950). *The mathematical theory of plasticity*. Oxford University Press. https://global.oup.com/academic/product/the-mathematical-theory-of-plasticity-9780198503675?cc=in&lang=en&

IS 1904-1986 (Reaffirmed 2006). *Indian standard code of practice for design and construction of foundations in soils: General requirements*. Bureau of Indian Standards. https://law.resource.org/pub/in/bis/S03/is.1904.1986.pdf

Javid, A. H., Fahimifar, A., and Imani, M. (2015). Numerical investigation on the bearing capacity of two interfering strip footings resting on a rock mass. *Computers and Geotechnics, 69*, 514-528. https://doi.org/10.1016/j.compgeo.2015.06.005

Keawsawasvong, S., and Ukritchon, B. (2022). Design equation for stability of a circular tunnel in anisotropic and heterogeneous clay. *Underground Space, 7*(1), 76-93. https://doi.org/10.1016/j.undsp.2021.05.003

Keawsawasvong, S., Thongchom, C., and Likitlersuang, S. (2021). Bearing capacity of strip footing on Hoek-Brown rock mass subjected to eccentric and inclined loading. *Transportation Infrastructure Geotechnology, 8*, 189-202. https://doi.org/10.1007/s40515-020-00133-8

Khing, K. H., Das, B. M., Yen, S. C., Puri, V. K., and Cook, E. E. (1992). Interference effect of two closely-spaced shallow strip foundations on geogrid-reinforced sand. *Geotechnical and Geological Engineering, 10*(4), 257-271. https://doi.org/10.1007/BF00880704

Kouzer, K. M., and Kumar, J. (2008). Ultimate bearing capacity of equally spaced multiple strip footings on cohesionless soils without surcharge. *International Journal for Numerical and Analytical Methods in Geomechanics, 32*(11), 1417-1426. https://doi.org/10.1002/nag.677

Kouzer, K. M., and Kumar, J. (2010). Ultimate bearing capacity of a footing considering the interference of an existing footing on sand. *Geotechnical and Geological Engineering, 28*(4), 457-470. https://doi.org/10.1007/s10706-010-9305-9

Kumar, A., and Saran, S. (2003a). Closely spaced footings on geogrid-reinforced sand. *Journal of Geotechnical and Geoenvironmental Engineering, 129*(7), 660-664. https://doi.org/10.1061/(ASCE)1090-0241(2003)129:7(660)

Kumar, A., and Saran, S. (2003b). Closely spaced strip footings on reinforced sand. *Geotechnical Engineering, 34*(3), 177-186.

Kumar, A., and Saran, S. (2004). Closely spaced rectangular footings on reinforced sand. *Geotechnical and Geological Engineering, 22*(4), 497-524. https://doi.org/10.1023/B:GEGE.0000047041.52563.49

Kumar, J., and Bhattacharya, P. (2010). Bearing capacity of interfering multiple strip footings by using lower bound finite elements limit analysis. *Computers and Geotechnics, 37*(5), 731-736. https://doi.org/10.1016/j.compgeo.2010.05.002

Kumar, J., and Bhattacharya, P. (2013). Bearing capacity of two interfering strip footings from lower bound finite elements limit analysis. *International Journal for Numerical and Analytical Methods in Geomechanics, 37*(5), 441-452. https://doi.org/10.1002/nag.1104

Kumar, J., and Bhoi, M. (2010). Effect of interference of strip footings and strip anchors on their elastic settlements. *International Journal of Geotechnical Engineering, 4*(2), 289-297. https://doi.org/10.3328/IJGE.2010.04.02.289-297

Kumar, J., and Ghosh, P. (2007). Ultimate bearing capacity of two interfering rough strip footings. *International Journal of Geomechanics, 7*(1), 53-62. https://doi.org/10.1061/(ASCE)1532-3641(2007)7:1(5

Kumar, J., and Ghosh, P. (2007a). Upper bound limit analysis for finding interference effect of two nearby strip footings on sand. *Geotechnical and Geological Engineering, 25*(5), 499-507. https://doi.org/10.1007/s10706-007-9124-9

Kumar, J., and Kouzer, K. M. (2008). Bearing capacity of two interfering footings. *International Journal for Numerical and Analytical Methods in Geomechanics, 32*(3), 251-264. https://doi.org/10.1002/nag.625

Lavasan, A. A., and Ghazavi, M. (2012). Behavior of closely spaced square and circular footings on reinforced sand. *Soils and Foundations, 52*(1), 160-167. https://doi.org/10.1016/j.sandf.2012.01.006

Lavasan, A. A., and Ghazavi, M. (2016). Failure mechanism and soil deformation pattern of soil beneath interfering square footings. *Journal of Numerical Methods in Civil Engineering, 1*(2), 48-56. https://doi.org/10.29252/NMCE.1.2.48

Lavasan, A. A., Ghazavi, M., and Schanz, T. (2017). Analysis of interfering circular footings on reinforced soil by physical and numerical approaches considering strain-dependent stiffness. *International Journal of Geomechanics, 17*(11), 04017096. https://doi.org/10.1061/(ASCE)GM.1943-5622.000099

Lavasan, A. A., Ghazavi, M., von Blumenthal, A., and Schanz, T. (2018). Bearing capacity of interfering strip footings. *Journal of Geotechnical and Geoenvironmental Engineering, 144*(3), 04018003. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001824

Lee, J., and Eun, J. (2009). Estimation of bearing capacity for multiple footings in sand. *Computers and Geotechnics, 36*(6), 1000-1008. https://doi.org/10.1016/j.compgeo.2009.03.009

Lee, J., Eun, J., Prezzi, M., and Salgado, R. (2008). Strain influence diagrams for settlement estimation of both isolated and multiple footings in sand. *Journal of Geotechnical and Geoenvironmental Engineering, 134*(4), 417-427. https://doi.org/10.1061/(ASCE)1090-0241(2008)134:4(417)

Mabrouki, A., Benmeddour, D., Frank, R., and Mellas, M. (2010). Numerical study of the bearing capacity for two interfering strip footings on sands. *Computers and Geotechnics, 37*(4), 431-439. https://doi.org/10.1016/j.compgeo.2009.12.007

Meyerhof, G. G. (1963). Some recent research on the bearing capacity of foundations. *Canadian Geotechnical Journal, 1*(1), 16-26. https://doi.org/10.1139/t63-003

Naderi, E., and Hataf, N. (2014). Model testing and numerical investigation of interference effect of closely spaced ring and circular footings on reinforced sand. *Geotextiles and Geomembranes, 42*(3), 191-200. https://doi.org/10.1016/j.geotexmem.2013.12.010

Nainegali, L. S., Basudhar, P. K., and Ghosh, P. (2013). Interference of two asymmetric closely spaced strip footings resting on non-homogeneous and linearly elastic soil bed. *International Journal of Geomechanics, 13*(6), 840-851. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000290

Nainegali, L., Basudhar, P. K., and Ghosh, P. (2018). Interference of strip footings resting on non-linearly elastic foundation bed: a finite element analysis. Iranian *Journal of Science and Technology, Transactions of Civil Engineering, 42*(2), 199-206. https://doi.org/10.1007/s40996-018-0094-3

Nainegali, L., Basudhar, P. K., and Ghosh, P. (2021). Interference of proposed footing with an existing footing resting on non-linearly elastic dense and loose cohesionless soil bed. *European Journal of Environmental and Civil Engineering, 25*(14), 2574-2591. https://doi.org/10.1080/19648189.2019.1638311

Nazzal, M. D., Abu-Farsakh, M. Y., and Mohammad, L. N. (2010). Implementation of a critical state two-surface model to evaluate the response of geosynthetic reinforced pavements. *International Journal of Geomechanics, 10*(5), 202-212. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000058

Noorzad, R., and Manavirad, E. (2014). Bearing capacity of two close strip footings on soft clay reinforced with geotextile. *Arabian Journal of Geosciences, 7*(2), 623-639. https://doi.org/10.1007/s12517-012-0771-7

Paikaray, B., Das, S. K., and Mohapatra, B. G. (2021). Effect of reinforcement layout on interference effect of square footings on reinforced crusher dust. *International Journal of Geotechnical Engineering, 15*(10), 1268-1277. https://doi.org/10.1080/19386362.2020.1712531

Potts, D.M. and Zdravkovic, L. (1999). *Finite element analysis in geotechnical engineering: Theory/application*. Thomas Telford. https://www.icevirtuallibrary.com/doi/book/10.1680/feaiget.27534

Prandtl, L. (1920). Ü̈ber die Härte plastischer Körper. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1920*(1920), 74-85. https://eudml.org/doc/59075

Saha Roy, S., and Deb, K. (2018). Closely spaced rectangular footings on sand over soft clay with geogrid at the interface. *Geosynthetics International, 25*(4), 412-426. https://doi.org/10.1680/jgein.18.00025

Sekhar, J. C., Sasmal, S. K., and Behera, R. N. (2020). Effect of interference on ultimate bearing capacity of strip footings resting over multi-layered-reinforced soil. *Australian Journal of Multi-Disciplinary Engineering, 16*(1), 31-42. https://doi.org/10.1080/14488388.2020.1733739

Shiau, J., Chudal, B., Mahalingasivam, K., and Keawsawasvong, S. (2021). Pipeline burst-related ground stability in blowout condition. *Transportation Geotechnics, 29*, 100587. https://doi.org/10.1016/j.trgeo.2021.100587

Shokoohi, M., Veiskarami, M., and Hataf, N. (2019). A numerical and analytical study on the bearing capacity of two neighboring shallow strip foundations on sand. Iranian *Journal of Science and Technology, Transactions of Civil Engineering, 43*(1), 591-602. https://doi.org/10.1007/s40996-018-0189-x

Shukla, S. K. (Ed.). (2021). *ICE handbook of geosynthetic engineering: Geosynthetics and their applications*. ICE Publishing. https://www.icevirtuallibrary.com/doi/abs/10.1680/hge.41752

Skempton, A.W. (1951). *The bearing capacity of clay*. Building Research Congress. https://www.u-cursos.cl/ingenieria/2009/2/CI52Q/1/material_docente/bajar?id_material=244428

Stuart, J. G. (1962). Interference between foundations, with special reference to surface footings in sand. *Geotechnique, 12*(1), 15-22. https://doi.org/10.1680/geot.1962.12.1.15

Swain, A., and Ghosh, P. (2015). Experimental study on dynamic interference effect of two closely spaced machine foundations. *Canadian Geotechnical Journal, 53*(2), 196-209. https://doi.org/10.1139/cgj-2014-0462

Terzaghi, K. (1943). *Theoretical Soil Mechanics*. John Wiley and Sons. https://doi.org/10.1002/9780470172766

Vesić, A. S. (1973). Analysis of ultimate loads of shallow foundations. *Journal of the Soil Mechanics and Foundations Division, 99*(1), 45-73. https://doi.org/10.1061/JSFEAQ.0001846

Vivek, P. (2011). Static and dynamic interference of strip footings in layered soil [Master's dissertation, Indian Institute of Technology]. https://www.iitk.ac.in/ce/geotech/geotech_lab_iitkanpur_017.htm

Yang, F., Zheng, X. C., Sun, X. L., and Zhao, L. H. (2017). Upper-bound analysis of Nγ and failure mechanisms of multiple equally spaced strip footings. *International Journal of Geomechanics, 17*(9), 06017016. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000984

Yodsomjai, W., Keawsawasvong, S., and Likitlersuang, S. (2021). Stability of unsupported conical slopes in Hoek-Brown rock masses. *Transportation Infrastructure Geotechnology, 8*, 279-295. https://doi.org/10.1007/s40515-020-00137-4

Zidan, A. F., and Mohamed, M. (2019). Numerical analysis of bearing capacity of multiple strip footing on unreinforced and reinforced sand beds. *SN Applied Sciences, 1*(11), 1-13. https://doi.org/10.1007/s42452-019-1520-2

Research Article / Industrial Engineering

http://doi.org/10.15446/ing.investig.101868

Lean Six Sigma Tools for Efficient Milking Processes in Small-Scale Dairy Farms

Herramientas lean *six sigma* para un proceso de ordeño eficiente en granjas lecheras de pequeña escala

Eduardo G. Satolo[1], Guilherme A. Ussuna[2], and Priscilla A. B. Mac-Lean[3]

ABSTRACT

This research paper aims to use lean six sigma tools to handle milk in small Brazilian dairy estates. The search for efficiency in this process preponderates in all productive sectors. However, it is still not exploited in the dairy chain, specifically by small-scale producers, although milking is the main income for many of them. In Brazil, small producers are the main suppliers of the chain, and maintaining their competitiveness requires specific actions. Therefore, through the action-research method, four small producers in the center-west of the state of São Paulo, Brazil, were involved in identifying problems in their milking management process, investigating and implementing practical solutions. To this effect, lean tools such as flowcharts, Ishikawa diagrams, value stream mapping, and PDCA cycles were employed, and six feasible improvements were defined. The successful implementation of said improvements led to the acquisition of knowledge by small producers, who received financial returns, i.e., an increase in the amount paid per liter of milk. This research shows that, through interdisciplinarity, positive solutions to problems of different and complex natures, such as those in the dairy chain, can be obtained.

Keywords: rural management, small entrepreneurs, lean six sigma production, dairy farming, productivity

RESUMEN

Este artículo de investigación tiene por objetivo utilizar herramientas lean six sigma en el manejo de la leche en propiedades pequeñas de Brasil. La búsqueda de la eficiencia en este proceso es preponderante en todos los sectores productivos. Sin embargo, esta aún no se ha explotado en la cadena láctea, específicamente en el caso de productores de pequeña escala, si bien el ordeño representa el principal ingreso de muchos de ellos. En Brasil, los productores pequeños son los principales proveedores de la cadena, y mantener su competitividad requiere acciones específicas. Por lo tanto, mediante el método de investigación-acción, se involucró a cuatro pequeños productores del centro-oeste del estado de São Paulo, Brasil, investigando e implementando soluciones prácticas. Para este fin, se emplearon herramientas lean como diagramas de flujo, diagramas Ishikawa, mapas de flujo de valor, y ciclos PDCA, definiendo seis mejoras viables. La implementación exitosa de dichas mejoras llevó a la adquisición de conocimientos por parte de los pequeños productores, que recibieron rendimientos financieros, i.e., un aumento en la cantidad pagada por litro de leche. Esta investigación muestra que, a través de la interdisciplinaridad, se pueden obtener soluciones a problemas de naturalezas distintas y complejas, como es el caso de aquellos en la cadena láctea.

Palabras clave: gestión rural, pequeños empresarios, producción lean six sigma, ganadería lechera, mejora de la productividad

Received: May 24th, 2022
Accepted: June 30th, 2023

Introduction

In the dairy chain, studies aimed at improving managerial or organizational processes are commonly carried out in the dairy, whose theoretical basis is lean six sigma (LSS). As an example, the study by Powell *et al*. (2017) focused on the environmental sustainability gains from the use of LSS. Foshchii and Krasnokutska (2021) conducted a study on the feasibility of adopting lean production in a Ukrainian dairy. Finally, Trubetskaya *et al*. (2023) used the LSS in association with the ISO 5001 standard to optimize the use and consumption of energy in the dairy processing plant. In all these studies, the positive impacts of using LSS in processing plants were observed.

LSS is regarded as the latest generation of improvement methodologies that have a global outreach (Samanta *et al*., 2023). It is the result of merging lean theory and six sigma theory, which sparked revolutions in administrative strategies and were the object of tireless research throughout the 20th century (Patel and Patel, 2021), and their integration was considered to be essential (Francescatto *et al*.., 2023). LSS helps reduce the non-value-added activities, waste, defects, and nonconformities encountered during processes, and it

[1] PhD, Methodist University of Piracicaba (UNIMEP), Piracicaba, Brazil. Affiliation: Professor at São Paulo State University (UNESP), Tupã, Brazil. E-mail: eduardo.satolo@unesp.br.

[2] MSc., São Paulo State University (UNESP), Tupã, Brazil. E-mail: ussuna@gmail.com.

[3] PhD, Universidade de São Paulo (USP), Brazil. Affiliation: Professor at São Paulo State University (UNESP), Tupã, Brazil. E-mail: priscilla.mac-lean@unesp.br.

addresses issues related to waste and process flow with a focus on reducing product and process variation (Patel and Patel, 2021).

Considering the flexibility of LSS, Francescatto *et al.* (2023) highlighted the versatility of its application in different areas, which is noted in the high number of recent articles published on the topic, with a diversity of approaches and foci. However, no field studies have been identified which are applied to dairy producers in order to improve their organizational environment – in this case rural land – using administrative tools that contribute to efficiency gains.

According to the FAO (2021), it is estimated that 895 million people worldwide (14% of the population) depend directly on dairy farming. Thus, studies that promote the improvement of management and processes are necessary because of their socioeconomic impact.

Therefore, the research question of this study is

How can the use of lean six sigma tools provide milking management efficiency in small dairy farms?

To answer this question, our study applied the concepts inherent to LSS theory in Brazilian small-scale estates via the action-research methodology. We chose small-scale dairy producers because milk production in Brazil plays an important social role, as it is typically carried out in family farms and contributes significantly to the retention of rural workers (Casali *et al.*, 2020; Tramontini *et al.*, 2021; Tonet *et al.*, 2023). Dairy production in the country involves more than 1,5 million farmer families (IBGE, 2017), with low production scales and diverse strategies, posing challenges to the development of production systems (Bodenmüller Filho *et al.*, 2010; Lima *et al.*, 2023). In all Brazilian municipalities and states (Kuwahara *et al.*, 2018), each rural property has its own characteristics, with a clear availability of resources and decision-making difficulties (Koerich *et al.*, 2019).

Here lies the novelty and contribution of this research: it inserts a theory consolidated in other fields of the industry into this agent of the dairy chain, demonstrating that its applicability brings positive and promising results, especially with small-scale producers, who have a significant role in the economy.

Materials and methods

Based on the concepts proposed by Coughlan and Coghlan (2002), this framed within an action-research approach, as it solves a collective problem with the cooperation and participation of those involved, producing knowledge and solving a practical problem. The steps for conducting this research are depicted in Figure 1.

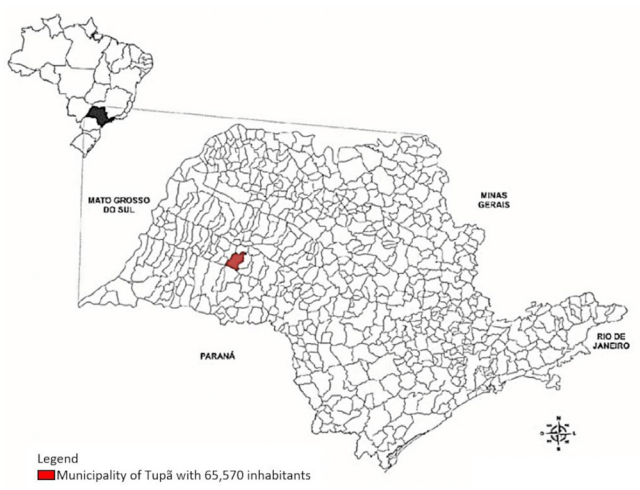


Figure 1. Steps for conducting action-research in this work
Source: Authors

This study was carried out in four small-scale dairy farms, in the municipality of Tupã, São Paulo state, Brazil, at the coordinates 21°56′ S, 50°30 W and at a 530 m altitude (Figure 2).

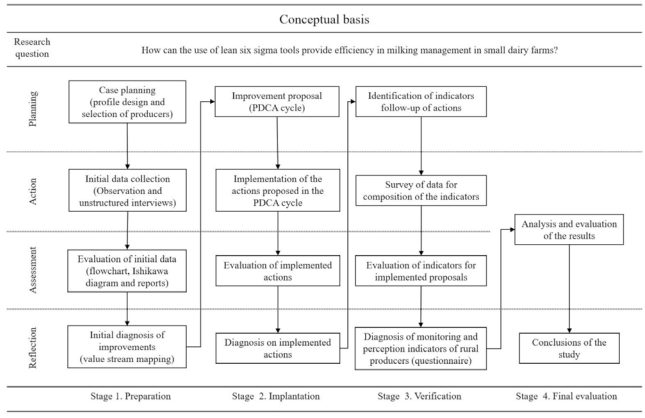


Figure 2. Location of the municipality of Tupã, Brazil
Source: IBGE (2021)

The main production source of the four small-scale farms (PA, PB, PC, and PD) is milk. They have been actively involved in dairy farming for at least 20 years. PA and PC are managed by women and PB and PD by men. The milk produce is collected by dairies located up to 60 km away from the estates.

The farms have a minimum of 15 dairy heads (be it in the lactation process or not) and a history of participation in training together with rural dairy programs. The farmers signed a consent form before the beginning of the research. The phases between data collection and the implementation of the actions took place between September 2018 and December 2019, with a total of 44 visits to the estates.

The characteristics of these small-scale producers are shown in Table 1.

Table 1. Zootechnical indices of the four dairy farmer families participating in the study

Zootechnical indices	Property (P)			
	A	B	C	D
Hectares (ha) for milk production	10,50	3,50	8,50	10
Total weight of animals* (Kg)	37,080	18,360	44,330	10,215
A.U. ** of the property	82,40	40,80	98,51	22,70
A.U./ha	7.84	11.65	11.58	2.27
Total number of cows*** (unit)	56	27	50	15
Total number of lactating cows (unit)	30	21	27	10
% of lactating cows	53.57	77.77	54	66.66
Total milk production per year (kg/year)	154,758	54,750	147,200	36,500
Total milk production per year/ ha (kg/ha.year)	14,738	15,642	17,317	3,650

Source: Authors

Data collection considered the status of the property and was performed through visits to the estates, using direct observation techniques and unstructured interviews to collect relevant information not available in the environment, which had value for our work and the application of lean tools, *i.e.*, flowcharts, Ishikawa diagrams, value stream mapping, and reporting.

Based on the collected data, an analysis of the current issues with regard to the milking process was carried out, with the subsequent elaboration of a value stream mapping. From this scenario, in partnership with small-scale producers, a corrective action plan was built through the PDCA cycle.

With the planning of corrective actions and the active support of the researchers, the actions to be implemented at the time were defined, considering the situation of each producer, the financial situation of the property, and the time of execution of the action. This stage of the research took place from August to December 2019, with 18 visits to the estates.

As a final stage called *post-intervention scenario*, a questionnaire with ten questions (Appendix A) was applied, which identified the impact of the interventions on the efficiency of the small-scale producers' management of the milking process, which is addressed in the discussion section.

Results

1st stage – Problem identification from the process flowchart

To identify the problems, the ideal flowchart based on Rosa *et al*. (2014) was employed. From the observation of the milking process, the steps and failures of the milking management process were determined (Table 2).

Table 2. Synthesized flowchart with the results regarding efficient management-milking practices by small-scale producers

#	Good practices in handling-milking (stages)	Are they performed?			
		PA	PB	PC	PD
1	Driving the cows to the milking place calmly, without running, screaming, or hitting	Yes	Yes	Yes	Yes
2	Accommodation in the corrals or waiting room for the cows to rest for a few minutes	No	Yes	Yes	Yes
3	Conducting the milking line, following the appropriate order, *i.e.*, 1st) first-calf cows, 2nd) cows that never had mastitis; 3rd) who have had mastitis but are cured, 4th) cows with mastitis	No	No	Yes	No
4	Driving the cows to the milking parlor and placing them on the kennel or chains	Yes	Yes	Yes	Yes
5	Attachment of straps (ropes on the hind legs) *Optional, use only when necessary	Yes	Yes	Yes	No
6	Handwashing	No	No	Yes	Yes
7	Placing the hands on the leg or udder before the teats to avoid scaring the animal	Yes	No	Yes	Yes
8	Cleaning the ceilings with running water (being careful of not soiling the udder)	Yes	No	Yes	Yes
9	Withdrawal of the first three jets of milk from the breasts in the mug with a black background	Yes	Yes	Yes	Yes
10	Analysis of the presence of lumps in the milk in the mug	Yes	Yes	Yes	Yes
11	Pre-dipping application (used to disinfect teats) via a non-return applicator cup	Yes	No	Yes	Yes
12	Waiting 30 seconds	Yes	Yes	Yes	Yes
13	Cleaning of the teats with disposable paper	Yes	No	Yes	Yes
14	Disposing of the pre-dipping solution in a garbage bag	No	No	Yes	Yes
15	Placing the teat cups on the cow	Yes	Yes	Yes	Yes
16	Removing the teat cups only after exhaustion of milk from the udder/ cutting the vacuum for mechanical milking	Yes	Yes	Yes	Yes
17	Application of post-dipping on the teats	Yes	No	Yes	Yes
18	Keeping the animal upright after milking to close the sphincter	Yes	Yes	Yes	Yes
19	Sanitized mechanical milking	Yes	Yes	Yes	Yes
20	*Transportation of milk to the tank can occur in step 21 or after 16.	Yes	Yes	Yes	Yes
21	Cleaning the milking place	Yes	Yes	Yes	Yes

Source: Authors

2nd stage – Elaboration of Ishikawa diagram for problem consolidation

The results of the flow diagram and direct observation allowed using the Ishikawa diagram tool. The Ishikawa fishbone diagram is a tool to discover the roots of a problem and to generate ideas for improvement (Souliotis *et al*., 2018). The causes are associated with six possible categories, called *6M's*, which are machine, manpower, methods, material, maintenance, and mother nature. These are summarized in Figure 3.

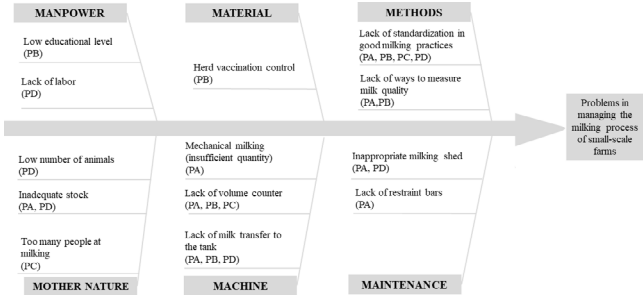


Figure 3. Representation of the Ishikawa diagram for management problems in the milking process of small-scale farms
Source: Authors

3rd stage - Value stream mapping (VSM), highlighting points for improvement

The problems identified in the Ishikawa diagram were the subject of studies and proposals for improvement by the researchers together with the small-scale producers. To this effect, the elaboration of the VSM) allowed to visually represent the problems in the milking process (Figure 4) for the four estates. The VSM is a comprehensive tool for ensuring continuous improvement activities and yields effective results for waste elimination activities, in addition to showing value- and non-value-added tasks (Jamil *et al*., 2020).

Gray signs represent such problems, and, in the literature, they are described as opportunities for the organization to perform concrete or continuous improvement processes, which are called *improvement explosions*.

4th stage – Preparation of the action plan through the PDCA cycle

The proposals for improvements were presented and discussed with the producers, where, out of the 13 problems previously identified, six showed a feasible solution. Constraints, including high investments, execution times, and the decision of producers meant that other issues were not the target of the study at the time. Through the PDCA cycle, actions on viable improvements were determined (Table 3). The PDCA cycle is characterized by a focus on continuous improvement regarding products and processes, through gradual changes that are carried out in four phases and drive the development of the company (Silva *et al*., 2017).

5th stage – Implementing the action plan

Once the planning of measures had been performed through the PDCA cycle, the project was implemented. Each improvement is detailed in sequence.

Improvement 1 – Insufficient knowledge of efficient milking management practices

The milking process is a fundamental step for obtaining high-quality dairy products, and factors related to animal hygiene, milking equipment, and milkers lead to milk contamination (Simões *et al*., 2017). Furthermore, the implementation of good practices allows producers to comply with Brazilian regulations, meeting international hygiene, economic, social, and environmental requirements in favor of the productive chain (MAPA, 2018). Therefore, it was decided to initiate the improvements through training in connection with this process. To this effect, the manual of good milking management practices developed by Rosa *et al*. (2014) was used.

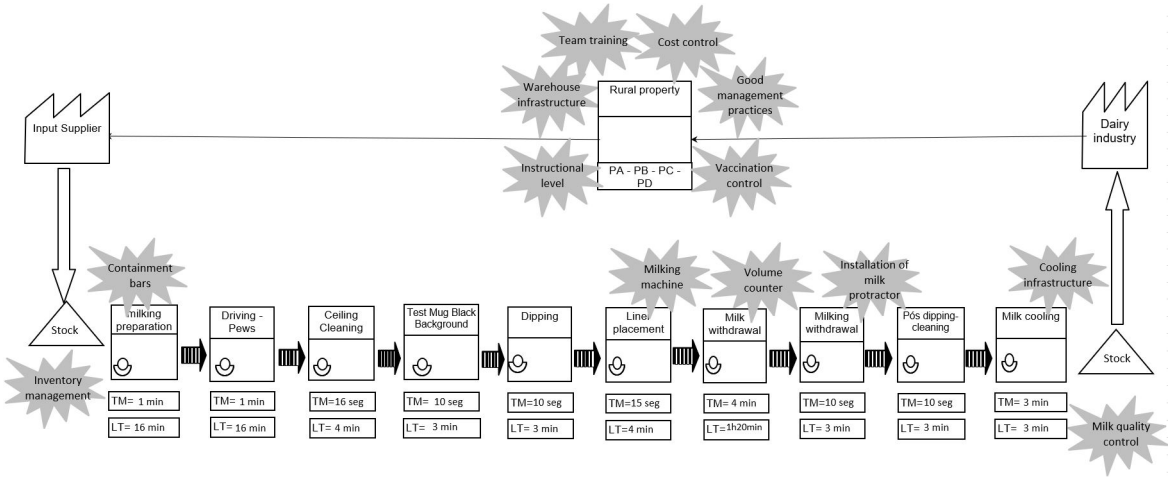


Figure 4. Value stream mapping and identification of issues in the process
Source: Authors

Improvement 2 – Absence of or low control of the volume of milk per animal

In this study, the absence of or low control of the volume produced per animal was mapped for three of the studied estates, which was an unclear problem and was solved by creating a control table (Appendix B).

Through the exposure of producers to difficulties in managing daily control, weekly control was employed to compare productive relationships (constant or with positive and/or negative oscillations) within milking. The need to note the number of dry cows was highlighted because it is important to periodically monitor the entry and exit of cows in this cycle.

Improvement 3 – Insufficient production cost control

As a common factor in all the studied small-scale dairy farms, insufficient knowledge of the production costs and profit margins has become a deficiency and a recurring problem for dairy farming in Brazil, as highlighted by the 2018 IFCN study. This creates problems in terms of the competitiveness of the Brazilian milk market.

Supported by a technical worksheet, *i.e.*, a table for calculating the cost of milk production in Embrapa's family farming (Appendix C) elaborated by Tupy *et al*. (2002), as well as by the collection of information from the small-scale producers, the production costs and the profitability of milk culture were measured.

Table 3. PDCA for the estates

#	Problem	Plan	Do	Check	Act	Estates
1	Lack of knowledge regarding good milking management practices	Training producers on good milking management practices Elaborating a visual map with the associated steps	Training	Fortnightly checking performance regarding good milking management practices	Conducting additional training if necessary	PA PB PC PD
2	Absence of or poor control of the volume of milk per animal	Elaborating a daily production volume control sheet Training producers	Implementing the control sheet Training	Checking the use of spreadsheets fortnightly	Conducting additional training if necessary	PA PB PC
3	Absence of production cost control	Elaborating milk cost control spreadsheets Train producers	Implement dairy cost control spreadsheets Training	Checking the use of spreadsheets fortnightly Checking the impact on financial reality	Conducting additional training if necessary	PA PB PC PD
4	Lack of mechanical milking equipment	Design an acquisition plan for mechanical milking equipment Determining needs for structural adjustments Training producers	Purchasing mechanical milking equipment Making structural adjustments Training	Monitoring the acquisition process Follow-up of the the installation process Checking the use of new equipment weekly	Conducting additional training if necessary	PA
5	Absence of vaccination control	Elaborating vaccination control sheets Training producers	Implementing vaccination control sheets Training	Checking the use of spreadsheets fortnightly	Conducting additional training if necessary	PB
66	Structural problems	Designing the installation of containment bars Change the cooling tank location	New warehouse construction project for the tank	Monitoring the construction process	Building and validating the effectiveness of the new location	PA
		Determining needs for structural adjustments Designing the renovations of the raw material warehouse Designing a new space for the cooling tank Training producers	Renovating the shed Training producers for the new environment	Monitoring the shed renovation process Tracking training results	Conducting additional training if necessary	PD

Source: Authors

3. The values presented in this article were converted based on the average Brazilian real quotation in October 2020.

Improvement 4 – Insufficient equipment for mechanical milking

One of the dairy farms (PA) had a problem with the agility of milk extraction due to the high number of animals and the idleness of professionals during the period of milk extraction. Thus, the need to acquire mechanized milking and improve the milking environment was identified. Mechanized milking was budgeted, and a purchase for U$ 550,00[3] was carried out.

Improvement 5 – Absence of vaccination control

The inappropriate use of drugs to control diseases is a problem that affects dairy farmers in various locations, and it has been the subject of studies from different perspectives (Sabapara *et al*., 2012; Krömker and Leimbach, 2017; Madkar *et al*. 2020). The high cost of veterinary drugs, the inadequate knowledge of the disease and its control, and the unavailability of veterinarians (Madkar *et al*., 2020) are reasons that sometimes lead small-scale dairy producers to apply uncontrolled medication to their animals.

A vaccination and medication control spreadsheet (Appendix D) and the training of small-scale producers on its correct use were the methods used to solve this problem, which, although simple, have a significant impact on the dairy culture.

Improvement 6 – Infrastructure issues

The inadequate structural aspects of the PA's milking shed led to small improvements (known as *kaizens* in lean thinking), given the absence of containment bars in the milking environment and long-distance restrictions to be covered in order to dump the milk next to the cooling tank, which could result in sanitary and logistical losses for the producer. For this producer, the adaptation of the cooling tank had several advantages, as a milking system was also installed (improvement 4). The study of the producer's milking area entailed structural adjustments (Figure 5), including the construction of containment bars and an attached area for the installation of the cooling tank.

PD needed major changes (called *kaikaku improvements*) to meet the requirements of the legislation in terms of the storage of raw material and the transfer of milk from the collection point to the cooler, as well as the general conservation of its infrastructure. To this effect, some improvements were proposed, including the establishment of a large and increased space for the storage of feed and other related equipment, the creation of a separate environment close to the milking environment to house the cooling tank, and the organization of tools and medicines associated with milk culture (5S technique). Such improvements were budgeted at U$ 5 000,00 and carried out by a contractor.

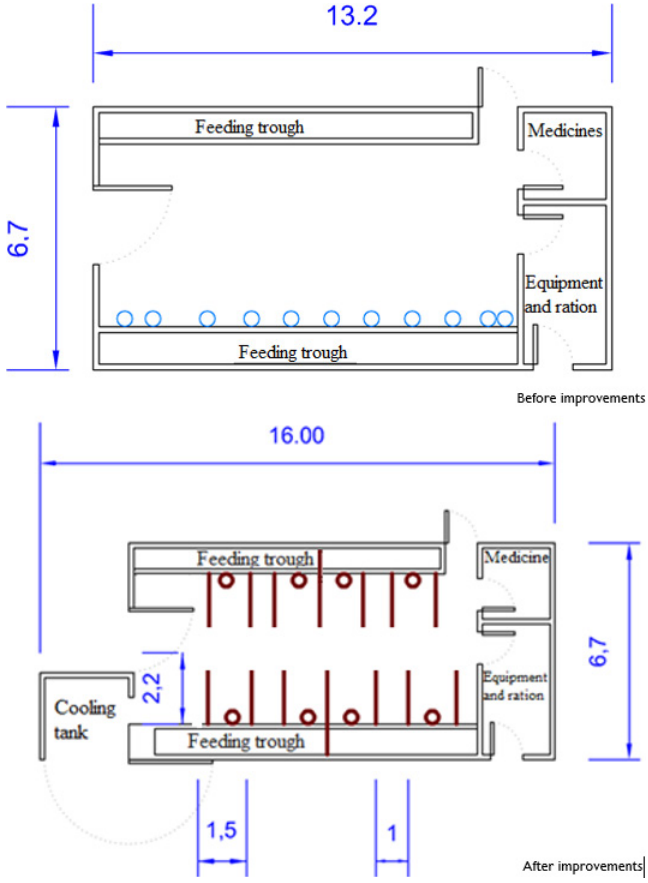


Figure 5. Floor plan of the PA milking shed (measures in meters)
Source: Authors

Discussion

According to FAO and IFD (2013), milk production systems worldwide must combine profitability with the responsibility to protect human and animal health and well-being, the environment, and productive and economic relationships. Milk producers, as the primary producers in the supply chain, must have the opportunity to add value to milk by adopting methods that meet the sanitary demands of the government, processing industries, and consumers (FAO and IDF, 2013). As shown in this study, these requirements can be fulfilled through lean production and its associated tools, which provide added value to the milk product. In this regard, Satolo *et al*. (2017) stressed that, in an agro-industrial production environment, as in other productive sectors, lean production tools should be selected according to organizational needs.

This study focused on the milking process. The use of the steps in the *Guide to Good Milking Management Practices* (Rosa *et al*., 2014) for the detection of problems and the subsequent training of small producers yielded positive results from the perspective of the latter. For these producers, the flowchart and the Ishikawa diagram allowed for a correct analysis of the milking process and the identification of improvements since, according to producer

A, "among many identified problems, mainly structural ones are consistent with reality," which confirms producer C's point of view: "yes, these problems are faithfully consistent with our daily lives".

Simões *et al*. (2015) argued that the use of best management practices reduces the presence of microorganisms, improves the health of animals' mammary glands, and, in conjunction with good processing practices, leads to a better nutritional quality for consumers, reducing the risks associated with raw milk and its products (Fagnani *et al*., 2021), which, for Züge (2015), ensurs that ownership will continue or become viable under the economic, social, and environmental triad.

Good results were obtained regarding the control of the volume of milk per animal (improvement 2), which is necessary for a productive dairy culture, especially for small-scale producers. Over time, the use of production data becomes an important means for decision-making. The generated data can be used in predictions, automated decisions, models trained from data, or any type of data visualization that provides insights (Bichler *et al*., 2017).

For dairy producers, controlling the volume of milk per animal allows estimating the future production, and, by setting the correct concentrate volume, over- or under-feeding is avoided, thus meeting the animal's needs, which results in a better production efficiency and reduced feed costs.

Knowing the volume of milk produced can also favor the reproduction of individuals with greater productive potential, helping genetically improve the herd and discarding the less productive animals.

In this regard, which underscores the importance of data control and knowledge, Vicario and Coleman (2020) stated that, if companies cannot manage and process the data to their advantage, they will be outperformed by competitors.

In the same context, improvement 3 was applied, which deals with cost control.

The cost control table was implemented in 50% of the producers, as the other participants declared they had it under control. The knowledge of the expenses and income of a company is the basis for its economic sustainability (Niţă and Stefea, 2014), and it is no different for a small rural enterprise. In this way, the knowledge acquired from improvement 3 allows producers to measure the technical-economic performance of their production in the future, as well as to diagnose the cost of the dietary supplementation method and compare the costs of other dairy producers.

For the small-scale producers participating in this study, the calculated cost of milk production identified an average profit margin of USD 0,044 per liter, a value of USD 0,025 compared to national studies (IEA, 2018), and USD 0,074 compared to international studies (IFCN, 2018).

In addition, detailing the production expenses allows planning for future improvement actions. The costs related to cattle feed are worth noting, which were 67,5%, a rate higher than that identified by Gabler *et al*. (2000) (60%).

This point confirms the studies by Lopes and Lopes (1999), which, two decades ago, underlined the importance of electronic tables or field notebooks as an essential means of controlling financial costs, given that they facilitate the grouping of expenditure and help milk quality extension teams and producers to make future decisions (Paixão *et al*., 2017).

The use of control sheets also brought benefits along with improvement 5, by allowing an effective management in controlling the vaccination of cattle. Vaccination is the most efficient and cost-effective strategy for preventing the appearance of clinical cases in herds (Groenendaal *et al*., 2015), and it may or may not be mandatory. For producers, these measures should be regarded as a form of investment, as the incidence of diseases sometimes causes economic losses greater than prevention (Grunitzky *et al*., 2020) and can compromise their production. Van Schaik *et al*. (1996) indicated that vaccination, in addition to the improvement of the clinical signs of dairy cattle, brought financial benefits 9,5 times higher than the cost of vaccination. Management procedures for the vaccination of dairy cattle, in turn, must be done safely and without causing stress or suffering to the animals; the lack of knowledge and information, according to Madkar *et al*. (2020), can affect the growth, production, and reproduction performance of dairy animals, as well as the milk quality.

Improvements in physical facilities and equipment took place in two phases of the study. In improvement 4, mechanical milking was installed. According to Hansen and Strte (2020), milking systems have gained popularity in medium-sized or family-owned dairy farms, as they allow for flexible working hours, in addition to improving production efficiency and increasing business sustainability. Paliy (2017) highlighted that mechanical milking contributes to the efficiency of the milking process through conditional stimulation and the unconditional reflexes of dairy cows, contributing to improving the yield per animal per hour of production according to Hogeveen and Ouweltjes (2003).

As for improvement 6, changes were made to the infrastructure of the milking shed, which has a direct impact on animal welfare. It is noteworthy that improvements in infrastructure do not lead to high investment. Sometimes, with low investments, important issues are addressed, providing efficiency to the production process. According to Grunitzky *et al*. (2020), concerns about the use of rational management in production systems have been neglected for a long time, and this has happened in in Brazil. However, this aspect should be put in a prominent position due to the growing consumer demand for quality products that come from systems that value animals' quality of life. The arrangement of facilities in a dairy farm, especially within

the milking parlor, has a direct impact on the production, on animal welfare, and on the longevity of animals within the production process.

Finally, it is noteworthy that the improvements have resulted in positive returns for small-scale producers, namely improved milk quality, which has led to a USD 0,06 increase in the amount paid to PC. Furthermore, the producers involved have reported that improvements are necessary to maintain the dairy activity in the estates, and that "the learning relationship with the university is important to us, although there is an exchange relationship, the product always wins with your support on this scale" (PC). At this point, it is possible to observe the producers' awareness of the importance of science, and that proximity to the university promotes the advancement of culture.

Conclusion

The use of LSS tools has proved to be extremely valuable in identifying problems and proposing and implementing improvements in milking management by small-scale dairy farms. In this study, the implementation of improvements generated gains for producers in different aspects: (i) management: the implementation of control spreadsheets allowed the producers to be aware of their dairy farm and make decisions, highlighting actions related to feed costs; (ii) the milking environment: structural improvements and the installation of equipment provided quality to animal welfare while also allowing rural producers to carry out their daily activities; and (iii) financial, *i.e.*. the improvement in the price paid per liter of milk delivered. The implementation of improvements associated with the return of producers showed that, through interdisciplinarity, *i.e.*, the interconnection of different areas of knowledge, positive solutions to different types of problems can be achieved. Thus, this article creates a legacy for the milk production chain, with a study that can be replicated in different Brazilian and world realities, revealing a means of implementing a solution to problems that affect the main process of obtaining raw milk.

Acknowledgements

This Project was funded by the Brazilian agency called *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), through grant number 424722/2018-6, as well as by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES), grant number CAPES/DS nº 88881.593696/2020-01.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

Bichler, M., Heinzl, A., and van der Aalst, W. M. P. (2017). Business analytics and data science: Once again? *Business & Information Systems Engineering, 59*, 77-79. http://doi.org/10.1007/s12599-016-0461-1

Bodenmueller Filho, A., Damasceno, J. C., Previdelli, I. T. S., Santana, R. G., Ramos, C. E. C. O., and dos Santos, G.T. (2010). Typology of production systems based on the milk characteristics, *Revista Brasileira de Zootecnia, 39*, 1832-1839. http://doi.org/10.1590/S1516-35982010000800028

Casali, M., de Mendonça, B. S., de Brito, M. M., dos Santos, M. G. R., Lima, P. G. L., Siqueira, T. T. S., Damasceno, J. C., and Bánkuti, F. I. (2020). Information asymmetry among dairy producers in Paraná, Brazil. *Semina: Ciências Agrárias, 41*, 293-304. http://doi.org/10.5433/1679-0359.2020v41n1p293

Coughlan, P., and Coghlan, D. (2002). Action research for operations management. *International Journal of Operations & Production Management, 22*, 220-240. http://doi.org/10.1108/01443570210417515

Fagnani, R., Nero, L. A., and Rosolem, C. P. (2021). Why knowledge is the best way to reduce the risks associated with raw milk and raw milk products. *Journal of Dairy Research, 88*, 238-243. http://doi.org/10.1017/S002202992100039X

FAO (2021). *Crops and livestock products – FAOSTAT*. https://www.fao.org/faostat/en/#data

FAO and IDF (2013). *Guide to good dairy farming practice. Animal production and health guidelines.* http://www.fao.org/3/ba0027e/ba0027e00.htm

Foshchii, M., and Krasnokutska, N. (2021). Features of the lean production implementation at the enterprises of Ukraine's dairy industry. *Three Seas Economic Journal*, *2*, 48-53. http://doi.org/10.30525/2661-5150/2021-2-8

Francescatto, M., Neuenfeldt Júnior, A., Kubota, F.I., Guimarães, G. and de Oliveira, B. (2023). Lean Six Sigma case studies literature overview: Critical success factors and difficulties. *International Journal of Productivity and Performance Management*, *72*, 1-23. http://doi.org/10.1108/IJPPM-12-2021-0681

Gabler, M. T., Tozer, P. R. and Heinrichs, A. J. (2000). Development of a cost analysis spreadsheet for calculating the costs to raise a replacement dairy heifer. *Journal of Dairy Science, 83*, 1104-1109. https://doi.org/10.3168/jds.S0022-0302(00)74975-7

Groenendaal, H., Zagmutt, F. J., Patton, E. A., and Wells, S. J. (2015). Cost-benefit analysis of vaccination against Mycobacterium avium ssp. paratuberculosis in dairy cattle, given its cross-reactivity with tuberculosis tests. *Journal of Dairy Science, 98*, 6070-6084. http://doi.org/10.3168/jds.2014-8914

Grunitzky, L., Centenaro, J. R., Oliveira, A. G., Cheffer, I. M., and Braz, P. H. (2020). Vaccination in dairy cattle: An animal welfare practice known to producers? *PUBVET, 14*, 1-4. http://doi.org/10.31533/pubvet.v14n6a582.1-4

Hansen, B.G. and Stræte, E.P. (2020). Dairy farmers' job satisfaction and the influence of automatic milking systems. *NJAS - Wageningen Journal of Life Sciences, 92*, 100328. http://doi.org/10.1016/j.njas.2020.100328

Hogeveen, H., and Ouweltjes, W. (2003). Sensors and management support in high-technology milking. *Journal of Animal Science, 81*, 1-10. http://doi.org/10.2527/2003.81suppl_31x

IBGE (2017). *Censo Agropecuário 2017*. https://sidra.ibge.gov.br/pesquisa/censo-agropecuario/censo-agropecuario-2017

IBGE (2021). *Cidades e Estados*. https://www.ibge.gov.br/cidades-e-estados/sp/tupa.html

IEA (2018). *Valor da produção: estatísticas de produção dos principais produtos agrícolas paulistas*. http://www.iea.sp.gov.br/out/verTexto.php?codTexto=14543

IFCN (2018). *The IFCN Dairy Report 2018: For a better understanding of dairy world*. https://ifcndairy.org/wp-content/uploads/2018/10/Dairy-Report-Article-2018.pdf

Jamil, N., Gholami, H., Saman, M. Z. M., Streimikiene, D., Sharif, S. and Zakuan, N. (2020). DMAIC-based approach to sustainable value stream mapping: Towards a sustainable manufacturing system. *Economic Research*, *33*, 331-360. http://doi.org/10.1080/1331677X.2020.1715236

Koerich, G., Damasceno, J. C., Bánkuti, F. I., Parré, J. L. and Santos, G. T. (2019). Influence of forage production area, concentrate supply, and workforce on productive results in milk production systems. *Revista Brasileira de Zootecnia*, *48*, e20170177. http://doi.org/10.1590/rbz4820170177

Krömker, V., and Leimbach, S. (2017). Mastitis treatment –Reduction in antibiotic usage in dairy cows. *Reproduction in Domestic Anim*als, *52*, 21-29. http://doi.org/10.1111/rda.13032

Kuwahara, K. C., Damasceno, J. C., Bánkuti, F. I., Prizon, R. C., Rossoni, D. F., and Eckstein, I. I. (2018). Sustainability and typology of dairy production systems. *Semina: Ciências Agrárias*, *39*, 2081-2092. http://doi.org/10.5433/1679-0359.2018v39n5p2081

Lima, P. G. L., Bánkuti, F. I., Damasceno, J. C., dos Santos, G. T., Borges, J. A. R., and Ferreira, F. C. (2023). Factors influencing concentrate feeding: Dairy farmers' perceptions of dairy production system characteristics and market relations. *Animal - Open Space*, *2*, 100041-8. http://doi.org/10.1016/j.anopes.2023.100041

Lopes, M. A., and Lopes, D. C. F. (1999). Desenvolvimento de um sistema computacional para cálculo do custo de produção do leite de cabras. *Revista Brasileira de Agroinformática, 2*, 1-12. https://infoagro.deinfo.uepg.br/artigos/pdf/info_030.pdf

Madkar, A. R., Dutt, T., Boro, P., and Bharti, P. K. (2020). Health care managemental practices followed by dairy owners in western Maharashtra. *Journal of Entomology and Zoology*, *8*, 417-419. https://www.entomoljournal.com/archives/2020/vol8issue6/PartF/8-5-151-830.pdf

MAPA (2018). *Instrução normativa n° 76, de 26 de novembro de 2018*. https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=30/11/2018&jornal=515&pagina=9

Niţă, C. G., and Ştefea, P. (2014). Cost control for business sustainability. *Procedia-Social and Behavioral Sciences*, *124*, 307-311. http://doi.org/10.1016/j.sbspro.2014.02.490

Paixão, M. G., Lopes, M. A., Costa, G. M., Souza, G. N., Abreu, L. R., and Pinto, S. M. (2017). Milk quality and financial management at different scales of production on dairy farms located in the south of Minas Gerais state, Brazil. *Rev. Ceres, 64*, 213-221. http://doi.org/10.1590/0034-737X201764030001

Paliy, A. P. (2017). Innovations in the establishment physiology technologies milking high-productive cows. *Scientific Messenger of LNU of Veterinary Medicine and Biotechnologies. Series: Agricultural Sciences, 19*(74), 12-14. http://doi.org/10.15421/nvlvet74

Patel, A. S., and Patel, K. M. (2021). Critical review of literature on Lean Six Sigma methodology. *International Journal of Lean Six Sigma*, *12*, 627-674. http://doi.org/10.1108/IJLSS-04-2020-0043

Powell, D., Lundeby, S., Chabada, L., and Dreyer, H. (2017). Lean Six Sigma and environmental sustainability: The case of a Norwegian dairy producer. *International Journal of Lean Six Sigma*, *8*, 53-64. http://doi.org/10.1108/IJLSS-06-2015-0024.

Rosa, M. S., Costa, M. J. R. P., Sant'Anna, A. C., and Madureira, A. P. (2014). *Boas práticas de manejo - ordenha*. Funep.

Sabapara, G. P., Desai, P. M., Singh, R., and Kharadi, V. B. (2012). Constraints of tribal dairy animal owners of South Gujarat. *Indian Journal of Animal Sciences, 82*, 538. http://epubs.icar.org.in/ejournal/index.php/IJAnS/article/view/17713/8540

Samanta, M., Virmani, N., Singh, R. K., Haque, S. N., and Jamshed, M. (2023). Analysis of critical success factors for successful integration of lean six sigma and Industry 4.0 for organizational excellence. *The TQM Journal*, *2022*, 0215. http://doi.org/10.1108/TQM-07-2022-0215

Satolo, E. G., Hiraga, L. E. S., Goes, G. A., and Lourenzani, W. L. (2017). Lean production in agribusiness organizations: multiple case studies in a developing country. *International Journal of Lean Six Sigma, 8*, 335-358. http://doi.org/10.1108/IJLSS-03-2016-0012

Silva, A. S., Medeiros, C. F., and Vieira, R. K. (2017). Cleaner production and PDCA cycle: Practical application for reducing the Cans Loss Index in a beverage company. *Journal of Cleaner Production*, *150*, 324-338. http://doi.org/10.1016/j.jclepro.2017.03.033

Simões, G. H., Pozza, M. S. P., Zambom, M. A., Lange, M. J., and Neumann, M. E. (2015). Dairy production system type and critical points of contamination. *Semina: Ciências Agrárias*, *36*, 3923-3934. http://doi.org/10.5433/1679-0359.2015v36n6p3923

Souliotis, A., Giazitzi, K., and Boskou, G. (2018). A tool to benchmark the food safety management systems in Greece. *Benchmarking: An International Journal, 25*, 3206-3224. http://doi.org/10.1108/BIJ-02-2017-0028

Tonet, R. M., Bánkuti, F. I., Damasceno, J. C., da Silva Siqueira, T. T., Bouroullec, M. D. M., and Loddi, M. M. (2023). Typology of Brazilian dairy farms based on vulnerability characteristics. *Animal-Open Space*, *2*, 100040. http://doi.org/10.1016/j.anopes.2023.100040

Tramontini, R. C. M., Bánkuti, F. I., Pozza, M. S. S., Massuda, E. M., Damasceno, J. C., Dias, A. M., Ítavo, L. C. V., and Santos, G. T. (2021). Typology of dairy production systems based on management strategies in Paraná State, Brazil. *Tropical Animal Science Journal*, *44*, 123-130. http://doi.org/10.5398/tasj.2021.44.1.123

Trubetskaya, A., McDermott, O., and McGovern, S. (2023). Implementation of an ISO 50001 energy management system

using Lean Six Sigma in an Irish dairy: a case study. *The TQM Journal*, 35, 1-24. http://doi.org/10.1108/TQM-08-2022-0252

Tupy, O., Manzano, A., Esteves, S. N., Novaes, N. J., Camargo, A. C., Freitas, A. R., and Machado, R. (2002). Planilha para cálculo do custo de produção de leite na agricultura familiar. *Circular Técnica Embrapa, 32*, 1-11.

van Schaik, G., Kalis, C. H., Benedictus, G., Dijkhuizen, A. A., and Huirne, R. B. (1996). Cost-benefit analysis of vaccination against paratuberculosis in dairy cattle. *Veterinary Record*, *139*, 624-627

Vicario, G., and Coleman, S. (2020). A review of data science in business and industry and a future view. *Applied Stochastic Models in Business and Industry*, *36*, 6-18. http://doi.org/10.1002/asmb.2488

Züge R. M. (2015). Boas práticas na produção de leite. In. P. do C. Martins, G. A. Piccinini, E. E. B. Krug, C. E. Martins, and F. C. F. Lopes (Eds.), *Sustentabilidade ambiental, social e econômica da cadeia produtiva do leite* (pp. 415-432). Embrapa. https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1021902/1/Livro2015Sustentabilidadecompleto.pdf

Appendices

Appendix A - Final project return questionnaire

1. Do the problems identified on your property match reality?
2. Was the support given by the researcher to identify the problems adequate?
3. Can the suggested improvements be achieved in time?
4. Do you see improvement in the post-execution milking process?
5. Did you manage to assess the improvement in any milk quality indicator with the changes made?
6. Were there any difficulties in implementing the improvements? If so, which ones?
7. Do you intend to continue with the improvements that have not yet been implemented?
8. Did the research bring satisfactory results?
9. Was the support given by the researcher to implement the actions adequate?
10. How could researchers make this research better? Please make suggestions.

Appendix B – Control table for the property's milk volume

Month:_____					
DAIRY CONTROL – DATE _____/_____/_____					
Cow		**Producion (liters)**			**Comments**
No.	**Name**	**1st Milking**	**2nd Milking**	**Total**	

Appendix C – Producer worksheet for production cost control

Milk production cost worksheet	
Producer name	
Property Name	
County	
Month:	Year:

#	**Items**	**U$**
1	Labor expenses	
2	Food expenses	
3	Other expenses	
4	Capital costs [(1) + (2) + (3)] x 0.20	
5	Financial expenses	
6	TOTAL COST (1) + (2) + (3) + (4) + (5)	
7	TOTAL REVENUE	
8	Milk sale	
9	Other sales	
10	NET PROFIT	
11	Total milk product (liters)	
12	Milk sold (liters)	
13	Cost/liter of milk produced	
14	Value received per liter of milk	
15	Profit/liter of milk produced	
16	Number of cows in the herd (unit)	
17	Milk production per cow in the herd (liters)	

Appendix D – Producer worksheet for vaccine control

COW TREATMENT - MEDICINES / VACCINES					
Animal name and number	Treatment date	Treatment reason	Drug name	Grace period (hours/days/ milkings)	Date of return to milking

Research Article / Education in Engineering https://doi.org/10.15446/ing.investig.98094

A Hybrid-Flipped Classroom Approach: Students' Perception and Performance Assessment

Un enfoque de aula invertida híbrida: evaluación de la opinión y el rendimiento de los estudiantes

Bengisu Yalcinkaya Gokdogan[1], Remziye Busra Coruk[2], Mohamed Benzaghta[3], and Ali Kara[4]

ABSTRACT

This study presents an improved hybrid-flipped classroom (hybrid-FC) education method based on technology-enhanced learning (TEL) along with diluted classes for a course on probability and random processes in engineering. The proposed system was implemented with the participation of two student groups who alternated weekly between attending face-to-face activities and fully online classes as a sanitary measure during the pandemic. The education model was combined with the flipped classroom (FC) approach in order to improve the quality of learning and address the negative effects of remote education. Before the lessons, the students studied the course material, filled a question form, and then took a low-stake online quiz. Then, the students attended a session where the questions reported in the forms were discussed, and they took an online problem-solving session followed by an individual quiz. Class sessions were available to both online and face-to-face students, as well as in the form of video recordings for anyone who missed lessons. Qualitatively and quantitatively, the proposed education method proved to be more effective and comprehensive than conventional online methodologies. The students' performances were evaluated via quizzes and exams measuring the achievement of the course learning outcomes (CLOs). Weekly pre/post-tests were applied to examine the students' progress in each topic. Midterm and final exams were planned to measure the level of success for all course topics. Additionally, the students' perception was assessed with questionnaires and face-to-face interviews. A performance assessment showed an apparent increase in the success rate, and the students' perception was found to be positive.

Keywords: engineering education, flipped classroom, hybrid education, probability and random processes, statistical analysis

RESUMEN

Este estudio presenta un método educativo mejorado de aula invertida híbrida (hybrid-FC) basado en el aprendizaje mejorado por tecnología (TEL) junto con clases diluidas para un curso sobre probabilidad y procesos aleatorios. El sistema propuesto se implementó con la participación de dos grupos de estudiantes que alternaban semanalmente entre asistir a actividades presenciales y totalmente en línea como una medida sanitaria durante la pandemia. El modelo educativo se combinó con el enfoque de aula invertida (FC) para mejorar la calidad del aprendizaje y hacer frente a los efectos negativos de la educación a distancia. Antes de asistir a clase, los estudiantes estudiaban el material del curso, completaban un formulario de preguntas y luego tomaban un quiz en línea de bajo impacto. Luego, los estudiantes asistían a una sesión en la que se discutían las respuestas en los formularios, y tomaban una sesión en línea de resolución de problemas, seguida de un cuestionario individual. Las clases estaban disponibles tanto para los estudiantes en línea como para aquellos que asistían de forma presencial; también había grabaciones de video para quien faltara a clases. Cualitativa y cuantitativamente, el método educativo propuesto demostró ser más efectivo y completo que los métodos online convencionales. El desempeño de los estudiantes se evaluó mediante cuestionarios y exámenes que medían el logro de los resultados de aprendizaje del curso (CLO). Se realizaron exámenes previos y posteriores semanales para examinar el progreso de los estudiantes en cada tema. Se planificaron exámenes parciales y finales para medir el nivel de éxito de todos los temas del curso. Además, la percepción de los estudiantes se evaluó con cuestionarios y entrevistas presenciales. La evaluación del desempeño mostró un aumento aparente en la tasa de éxito, y se encontró que la opinión de los estudiantes era positiva.

Palabras clave: educación en ingeniería, aula invertida, educación híbrida, probabilidad y procesos aleatorios, análisis estadístico

Received: September 1st, 2021
Accepted: February 15th, 2023

[1] Electrical and Electronics engineer, Atilim University, Turkey. PhD candidate in Electrical and Electronics Engineering, Atilim University, Turkey. Affiliation: Research/Teaching Assistant, Atilim University, Turkey. Email: bengisu.yalcinkaya@atilim.edu.tr

[2] Electrical and Electronics engineer, Atilim University, Turkey. PhD candidate in Electrical and Electronics Engineering, Atilim University, Turkey. Affiliation: Research/Teaching Assistant, Atilim University, Turkey. Email: busra.tezel@atilim.edu.tr

[3] Department of Information and Communication Technologies, University Pompeu Fabra, Barcelona, Spain. PhD candidate in the Department of Information and Communication Technologies, University Pompeu Fabra, Barcelona, Spain. Affiliation: Research/Teaching Assistant, Department of Information and Communication Technologies, University Pompeu Fabra, Barcelona, Spain. Email: mohamed.benzaghta@upf.edu

[4] Electrical and Electronics engineer, Erciyes University, Turkey. PhD in Electrical and Electronics Engineering, Hacettepe University, Turkey. Affiliation: Professor, Gazi University, Turkey. Email: akara@gazi.edu.tr

Introduction

Improving current education systems by adopting technological developments is becoming a necessity for most universities. An effective education is of great importance, especially in the field of engineering and applied education (Rodríguez *et al*., 2011). There are several methods that can assist and improve traditional teaching techniques (Merideno *et al*., 2015). For example, web-based learning, mobile learning techniques, computer-based simulation tools, and remote access laboratories are becoming very efficient educational materials (Gomez *et al*., 2014; Coruk *et al*., 2020; Kesim *et al*., 2012).

Measuring the effectiveness of new education methods is also an important aspect that must be taken into account. It was reported by several studies that teaching with the help of some student-centered methods such as remote teaching practices, group-based problem-solving, and project-based learning is much more effective than employing classical teaching methods (Cigdemoglu *et al*., 2014).

Due to the recent pandemic, education processes in universities have been greatly affected. As courses were planned for face-to-face tutoring, the closing of universities made it a necessity for instructors and students to rapidly adapt to remote teaching methods (Carrillo *et al*., 2020; Moorhouse, 2020); instructors had to change to online teaching, which required them to implement a number of digital tools and several new resources to teach the course concepts, solve problems, and grade students (Almaiah *et al*., 2020; Assunção Flores *et al*., 2020; König, 2020; Watermeyer *et al*., 2021). Additionally, the effects of pandemic conditions, the transition to remote education, and the use of digital assistive technologies on students have been analyzed in the literature (Iglesias-Pradas *et al*., 2021).

However, the extent to which instructors have successfully mastered online education approaches remains unknown (König, 2020). Moreover, there are several other shortcomings of moving towards fully online teaching, such as inexperience, difficulties associated with poor infrastructure (limited Internet access in rural areas), an uncomfortable environment at home, and the lack of support and mentoring for students during exams (Zhang *et al*., 2020). Therefore, it is essential to go beyond emergency online teaching practices and come up with new teaching methods that enable sufficient learning outcomes while following careful planning and cautious instructional designs (Hodges *et al*., 2020).

The flipped classroom (FC) approach is a course delivery education method that has recently gained popularity (Chiquito *et al*., 2020). In this method, the classical in-class and out-of-class activities exchange places. In general, the course lecture delivery is performed in terms of self-study materials, such as presentations and video lectures. Moreover, there can be some supplementary evaluation exams, *i.e.*, short online quizzes to track student discipline and to encourage lecture collaboration between instructors and students (Öncel *et al*., 2019).

Studies on the FC method have presented promising improvements in student performance when compared to traditional methods (Bishop *et al*., 2013; Kim *et al*., 2014; Mason *et al*., 2013; O'Flaherty *et al*., 2015; Öncel *et al*., 2019). In García-Ramírez (2019), the implementation and improvement of the FC model in a pavements course are evaluated using the students' final grades. Recently, a bibliometric review method was used to map the conceptual, intellectual, and social structure of research development regarding flipped learning (Al Mamun, 2021). In addition, a pedagogical FC model with an online learning management platform was proposed as an alternative for virtual teaching (Vilchez-Sandoval *et al*., 2021).

The pandemic era has shown that even well-practiced education methods – including FC – have been mostly unsuccessful under current conditions. Considering the difficulties that may arise in the education process, improving new alternative approaches is an important issue. During the pandemic, most schools switched to remote education. However, it has become a great necessity to develop novel approaches that both minimize any problems encountered with regard to remote education and comply with the health measures in effect. Developing a system where students partially attend face-to-face classes while some technology-enhanced educational methods are implemented may provide solutions to the problems that arose during these times. A hybrid education system, where pandemic cautions are implemented even though students are partially allowed to continue face-to-face education, has not been studied in the literature. Enhancing this hybrid method with improved learning approaches such as FC can pave the way for effective learning under such challenging conditions. It is also critical to measure the success of said method, as well as the perception of students.

This study presents an efficient and comprehensive education method of teaching Probability and Random Processes (PRP) based on the flipped or inverted classroom technique and technology-enhanced learning (TEL), combined with a hybrid (online and in-class) method to address the challenges of the pandemic. The proposed hybrid-FC approach aims to ensure effective learning in the pandemic era, during which most schools have switched to remote education in order to maintain the rigid health measures implemented by the authorities.

This study aimed to increase student performance and minimize the negative effects of remote learning with the improved hybrid-FC education model in a basic engineering course. Extensive measurements were conducted to demonstrate that the proposed method is quite practical and efficient, and it can be implemented as an improved version of the FC approach, not only in the pandemic era, but also in the face of other conditions such as overpopulated classes or schools.

Probability and random processes (PRP) was selected in this work because it is one of the most important courses; it is a must in the commonly overpopulated sophomore classes of the undergraduate curricula of Electrical (EE), Electrical and Electronics (EEE), Electrical and Computer (ECE), and Computer Engineering (CE) departments. This course serves as a foundation for a variety of subjects, specifically signal analysis, signal processing in communication applications, and biomedical and control systems in EEE. The topics of the course are considered to be an integral part of engineering curricula by most engineering accreditation organizations, as is the case of the ABET (Accreditation Board of Engineering and Technology). On the other hand, most students tend to find this course difficult to understand; they sometimes fail to develop an interest in pursuing careers in these areas of EE, as well as in other similar disciplines.

In light of the above, a practical system was first developed in which students could alternately attend face-to-face classes. In this study, two groups attended classes biweekly. This, in order to allow them to continue their formal education and keep the social aspects of campus life active to a certain extent. Then, the FC approach, which has been verified in previous research, was combined with the proposed diluted or de-populated classes methodology to improve the students' learning process. To evaluate the effectiveness of the model, qualitative (interviews), and quantitative (questionnaires, quizzes, and exams) assessment methods were employed. Unlike many studies in the literature, along with the students' perspectives, their achievements in the course were taken into account.

Course activities framed within the FC approach were created in two sections, *i.e.*, before and during the lesson hour. Before the class, the students studied the course materials, which included lecture notes, recorded course videos, and supplementary videos as out-of-class activities. In addition, the students' knowledge was evaluated via online quizzes (pre-tests). The lesson hour was divided into three sections: a discussion with the instructor, problem-solving accompanied by the instructor, and individual quizzes (post-tests). The proposed hybrid-FC approach was implemented with the alternated participation of two student groups. Thus, half of the class attended in-class activities, while the others took online lessons. This, as a measure against crowded classes in the context of a pandemic. Both online and in-class students attended the class sessions at the same time. Additionally, the lessons were video-recorded and were made available to those who had missed the classes or for later study. The students' perception of the hybrid-FC education method was assessed in detail. To this effect, questionnaires and face-to-face interviews were applied several times during the semester. Furthermore, the students' performance was also evaluated using their quizzes and exam results, which were calculated based on the Course Learning Outcomes (CLOs) and their regular participation. The effectiveness of the hybrid-FC approach was also evaluated according to these results.

Related literature

During the first two years of their undergraduate engineering studies, students are usually taught to deal with deterministic problems in the form of mathematics and physics courses. This causes many students to have difficulties in the PRP course, as they tend not to have a good probabilistic intuition (Sheikh, 2019).

From class observations, it was reported by Erjongmanee (2019) that students tend to perform poorly in this course due to a lack of reflection on mathematical concepts and their applications. Thus, the traditional technique used in teaching probability has been replaced by alternative approaches. Another technique used by Erjongmanee (2019) is to enhance application skills when dealing with probability concepts. In this method, students are further engaged in class activities by doing exercises.

The study by Nascimento *et al*. (2016) aimed at evaluating students' performance in written exams in order to identify their difficulties when dealing with probability courses. The authors suggested that adopting a double-blind peer review process can enhance the students' understanding of the topics.

The textbooks and examples used in teaching probability to EEE/ECE undergraduates were characterized and assessed in Nagy *et al*. (2008). The results show that there is a conflict between the preferences of students and those of the instructor. However, there were some examples that were appreciated by both.

Modern visualization and simulation tools such as Mathematica and MATLAB offer an interactive way to develop probabilistic intuition. Sheikh (2019) presented an approach to teach PRP to EEE/ECE juniors, which was based on using MATLAB simulations to emphasize real-world problems from areas related to signal processing and communications. Similarly, Mezhennaya *et al*. (2018) studied the influence of using computer algebra systems (CAS), along with supplementary educational materials such as computer modeling patterns and complementary tasks created in Mathematica.

A different method for teaching probability to graduate students in engineering departments was proposed by Richards *et al*. (2012). The delivery method for this graduate course switched from real-time education methods to a fully asynchronous approach. In their study, the authors explored the impact of this new method on students' perceptions and learning outcomes.

Method

The authors of this paper have been studying the concepts of *flipped classroom* (FC), *collaborative learning*, and *remote laboratories* in engineering education for more than

ten years. When the pandemic started, this study started with the following research questions based on previous experiences:

- Can we develop a hybrid educational method, encompassing flipped classrooms (FC) and diluted or depopulated classes in order to improve learning processes in pandemic era?
- Can we measure the effectiveness of this method and its applicability to meet the requirements of different contexts such as overpopulated classes and schools in engineering education both during and after the pandemic era?

This section describes the improved hybrid-FC structure of the PRP course. With the concept of *FC*, the differences in the perspective of the students towards this course and the change in their success rates are measured.

Course structure

In this section, previous developments regarding the traditional FC structure and the proposed hybrid-FC approach are presented. PRP is an obligatory undergraduate course in the EEE curriculum. It is a one-term course for second-grade students. The course consists of three hours of theoretical teaching per week. A total of 90 students take the course every semester.

A traditional FC educational method similar to that of Öncel *et al*. (2019) has been used in PRP courses in Atilim University for several years. However, innovations are added in this study, and course activities are carried out in a hybrid fashion with diluted classes and employing TEL. A hybrid education procedure means that in-class course activities are divided into smaller sessions (in terms of a number of students), and sessions are carried out face-to-face on campus, as well as online using real-time educational software. This proposed hybrid-FC educational method can be regarded as a solution to the spread of the pandemic and its effect on continuing face-to-face teaching activities.

Hybrid-flipped classroom

The FC activities were divided into two categories: before and during the lesson hour each week. Before class, the students had to complete pre-class activities by accessing the online course materials. After studying the materials, the students were expected to fill in the *Ask Questions to the Instructor* online form. This form guides the instructor on which points to focus on during the discussion hour. Furthermore, in this way, students attend the lectures well prepared. The online pre-test was carried out with the aim of following up on student's participation and encouraging them to take interest in the lesson. During the lesson hour, a discussion was first held using a hybrid system. The questions previously posted on the *Ask Questions to the Instructor* form were discussed, as well as some additional ones. Thus, discussion provided students with the opportunity to easily ask questions seeking to better understand the topic. Additionally, students gained practice with the problem-solving hours, consolidating the knowledge they acquired from theory. Finally, the post-test activity evaluated how well the students had learned the topic. Post-tests were also carried out in a hybrid fashion. The hybrid-FC method was implemented with an odd number of students online and an even number in-class for one week, which was alternated on a weekly basis. In addition to all of these activities, a midterm and a final exam were held face-to-face on campus.

The hybrid-FC activities conducted in the studied year were different to those of previous years (*i.e.*, the traditional FC structure) in order to add innovation and overcome the shortcomings of fully online learning methods. These differences can be listed as follows:

- Instead of using well-known online lecture videos from other institutions, course videos created by the instructors and based on the students' needs and feedback from previous years were used.
- Supplementary videos elaborated by other instructors were presented to provide students with a better understanding and to teach the topics more comprehensively.
- Problem-solving hours were implemented to fill the gap between theory and practice, which was one of the main difficulties that students had in previous years.
- Discussion hours and post-tests were carried out in a hybrid fashion with diluted classes (online and in-class).
- Individual-specific quizzes (post-tests) were included in the hybrid-FC structure in order to evaluate the performance of the students both face-to-face and online on a weekly basis.
- The proposed method used TEL more effectively when compared to classical FC implementations.
- Online course platforms and digital course materials were provided to be used in hybrid-FC education. The online parts of the discussion and problem-solving hours were conducted via Zoom. Live online and recorded video lessons, supplementary videos, lecture notes, additional course materials, the *Ask Questions to the Instructor* form, and the online part of the pre-tests/post-tests were made available on the Moodle platform. Through Moodle, students could access all the materials they needed. Interaction between the students and the online course platforms used is another key parameter of the learning process. The contribution of these digital platforms to learning was also thoroughly considered in the effectiveness measurements.

Although FC activities were conducted weekly, questionnaires and interviews were periodically applied with the aim get feedback from the students.

Assessment methods

Assessment methods based on both student perception and performance were essential for improving the course and

measuring the effectiveness of the hybrid-FC model. This study employed qualitative (interviews) and quantitative (anonymous class and university-wide questionnaires, as well as the evaluation of both quizzes/exams and CLOs) assessment methods. While quantitative assessment provides the opportunity to analyze numerical data, interviews can be interpreted using qualitative assessment. The implemented university-wide questionnaires and CLOs were also compared against the previous year in order to analyze the students' performance.

Face-to-face interviews containing open-ended questions were held twice, in the middle and at the end of the semester. A total of 40 students were selected, considering different Cumulative Grade Point Averages (CGPA), skill levels, and lesson participation. Attention was paid to the fact that 30 of the students attended both interviews, whereas 10 were randomly selected. Thus, a variety of student perspectives was considered. The questions covered a wide range of aspects regarding the course, such as students' opinions about the workload, the content of the course, the adequacy of the course materials, the comprehensibility of the topics, and the contribution of the course method to the learning process.

Questionnaires: To obtain the students' opinions about the hybrid-FC structure, two different questionnaire types were designed: anonymous class and university-wide questionnaires. The questions employed a 5-point Likert scale and were open-ended. Anonymous class questionnaires were applied twice in the middle and at the end of the semester. The common questions of the two questionnaires were elaborated based on the content, quality, and contribution to the learning process of the proposed hybrid-FC course activities. These course activities include course videos, supplementary videos, discussion hours, problem-solving hours, pre-tests/post-tests. In addition to the common ones, specific questions were included in both questionnaires with the aim to evaluate the students' perspectives on a variety of aspects. In the first questionnaire, information about the students and their workload, background, opinions about the course, and future projects were requested. The second questionnaire included questions aimed at evaluating the course and the students' knowledge levels. These questions are shown in Table 1. Additionally, the university-wide questionnaire was implemented online and anonymously in order for the students to evaluate the instructor and the course at the end of the semester.

Quizzes and exams: While weekly pre/post-tests were conducted to track the students' progress in each topic, midterm and final exams were carried out to measure the overall student success. Pre-tests were performed before the lesson hour with the aim to encourage the participants to study the topic with the provided lecture notes, video lessons, and supplementary videos. These tests included 10 multiple-choice questions about basic concepts. Post-tests were performed to evaluate the knowledge acquired by the students during the lesson hours. These tests included one or two open-ended questions.

The midterm and final exams (*i.e.*, the main exams) were conducted face-to-face while complying with the pandemic sanitary measures. The success of the students and the effectiveness of the hybrid-FC education method were evaluated by calculating CLOs with the results obtained from the pre-test/post-tests and the exams at the end of the semester.

Table 1. Anonymous in-class questionnaires

	1st questionnaire	**2nd questionnaire**
Different questions	Position goal and sectoral preference after graduation	How useful did you find course activities? (1: Useless, 5: Useful)
	Any previous work or experience on PRP before taking the course?	Have you spent enough time to learn the topics? (1: Could not spend enough time, 5: Absolutely spent enough time)
	Hours to be spent weekly on this course	Evaluate the difficulty of the subjects in the course (1: Easy to learn, 5: Difficult to learn)
	Expectations from the course	Do you think you have acquired enough knowledge in the course? (1: Certainly did not learn, 5: Absolutely learned)
	CGPA and gender	N/A
Common questions	Content of activity* Quality of activity * Contribution of activity* to your learning process. How many activities** of PRP course did you attend? Any other suggestions? Comment briefly.	

*Activity: Course videos, supplementary videos, discussion hours, problem-solving hours, pre-tests/post-tests.
**Activities: Discussion hours, problem-solving hours, pre-tests/post-tests.
Source: Authors

Course Learning Outcomes (*CLOs*)*:* CLOs determine what abilities the students will have when they have successfully completed a course. CLOs are defined under the MUDEK accreditation in effect in the EEE Department (national accreditation organization endorsed by ABET). The goal is to provide the same standards every year, matching the exam questions with the CLOs. The success level of the students in the course is defined in terms of the CLOs. This provides the opportunity to compare the students' success levels interannually.

Results and discussion

This section presents the results of the quantitative/qualitative assessments conducted to measure the effectiveness of the hybrid-FC education method. The students' perspectives

and performance were analyzed and compared to those of previous years. The results of the questionnaires were analyzed via a t-test. The students' success levels were evaluated in terms of the CLOs.

Interviews

Some students stated that they spent more than six hours a week, so the workload was heavy. On the other hand, some students stated that, even though they spent a lot of time per week on this course, this workload was normal, considering that there was no need to study heavily before the main exams. Moreover, they found the online course platforms used in the lectures to be user-friendly, easy to access, and supportive in their learning. They added that, when these course materials are adequately studied, there is no need to attend the discussion hour. Some other answers to the questions are discussed below:

- Since the hybrid-FC education method is new, the students underwent an adaptation process. However, they stated that this adaptation required a shorter time than other courses that are completely online.
- It was stated that the example questions given in the course videos had a great effect on learning.
- Most of the students stated that, even though they attended the course face-to-face, the recorded course videos made a great contribution to their understanding of the topics they missed.
- While some students reinforced their knowledge with these videos, others found them more complex and difficult.
- The students declared that, if they attended the discussion hour without having studied the course materials, this activity was inefficient.
- The students stated that the problem-solving hours were the most effective activity with regard to understanding the topics, and they suggested solving more problems. This was found to be more efficient since it took place before the post-test.
- The students believe that the pre-tests/post-tests are useful for weekly work, but the number of quizzes is high. They noted similarities in the difficulty of the online and in-class quizzes.

Improvements were made to the videos, as there were comments regarding poor sound quality. The post-test durations were adjusted. The students emphasized that the hybrid-FC activities were carried out in the best way during the pandemic. Most of them stated that they had no difficulty in following the lectures; they did not experience any negative effects associated with reduced direct contact and procrastination, which are the common deficiencies of a fully online education.

Questionnaires

A total of two anonymous in-class questionnaires were carried out during the semester. In the first questionnaire, general information about the students was asked in the form of open-ended questions, and questions about course activities using 5-point Likert scales were included. Accordingly, 38% of students had a CGPA above 3, 36% of them had between 2,5 and 3, 24% scored between 2 and 2,5, and the remainder had less than 2 out of 4. 24% of students spent 1-2 hours per week, 55% of them spent 2-4 hours, and 21% spent more than 4 hours per week in this course. 91% of the students answered "No" and 9% of them answered "Yes" to the question "Have you had any previous work or experience on PRP before taking the course?"

Some questions were common to both questionnaires, with the aim to detect any changes in the students' opinions. The mean and standard deviation (SD) results are given in Table 2.

Table 2. Analysis of the common questions in the questionnaires

#	ITEM ASKED		1st		2nd	
			Mean	SD	Mean	SD
1	Course videos	Content	3,37	1,04	3,71	0,91
		Quality	3,47	1,04	3,82	1,00
		Contribution	3,28	1,12	3,49	1,09
2	Supplementary videos	Content	3,25	1,15	3,42	1,05
		Quality	3,22	1,20	3,53	1,00
		Contribution	3,14	1,25	3,16	1,23
3	Discussion hours	Content	3,57	1,08	3,65	0,89
		Quality	3,65	1,13	3,95	0,85
		Contribution	3,44	1,01	3,52	1,14
4	Problem-solving hours	Content	4,04	1,08	4,27	0,95
		Quality	3,60	1,19	4,16	0,90
		Contribution	4,02	1,14	4,25	0,90
5	Pre-tests	Content	3,44	0,91	3,76	0,94
		Contribution	3,74	1,16	3,91	1,01
6	Post-tests	Content	3,51	1,10	3,62	1,10
		Contribution	3,60	1,23	4,00	0,91

Source: Authors

The results show a moderately positive perception of the pro-posed method. The overall mean score of the activities increased from 3,56 to 3,73. At the end of the semester, some additional questions were asked in the second questionnaire, aiming to further understand the students' opinions. The results of four questions (81 students) answered with a 5-point Likert scale are shown in Figure 1. For example, the question "How useful did you find course activities?" was rated with 1 by four students, with 3 by 28 students, and with 5 by 12 students.

Students evaluated whether they had acquired enough knowledge at the end of the semester, with a 3,17 mean score. Since 37% of the answers have 4-5 points, the results can be considered positive. The students graded the difficulty

of the course with a mean score of 3,78. This course seems to be quite difficult for students, which is why alternative approaches such as FC have emerged.

FC activities minimize the difficulty of this course, in light of the fact that the students rated their knowledge level with a high score. The students spent enough time on this course (3,4 mean score), and they found its activities useful (3,3 mean score). Since 45% of the students rated this question with 4-5 points, this research finds hybrid-FC activities to contribute positively to the learning process.

The results were analyzed with a t-test to determine the difference between several questionnaires. The degree of coincidence in the results is determined by the p-value. If the p-values are less than 0,1, the results are reliable (Lakens, 2017). As an example, p-values were calculated as 0,08 for "Content", 0,001 for "Quality", and 0,09 for "Contribution to the learning process" regarding the "Problem-solving hours". Therefore, the answers given to these questions were reliable. Similarly, the p-value of the "Contribution to the learning process" for "Post-tests" was calculated as 0,02. The fact that the two most important activities have high mean values was statistically proven by the t-test, as well as the reliability of the results, which constitutes a significant result.

The university-wide questionnaire is applied every year to students of all courses and faculties. In this type of questionnaire, the instructor and the course are evaluated by the students. Table 3 shows the results (mean score between 1 and 5) obtained for the previous (59 students) and the studied years (69 students).

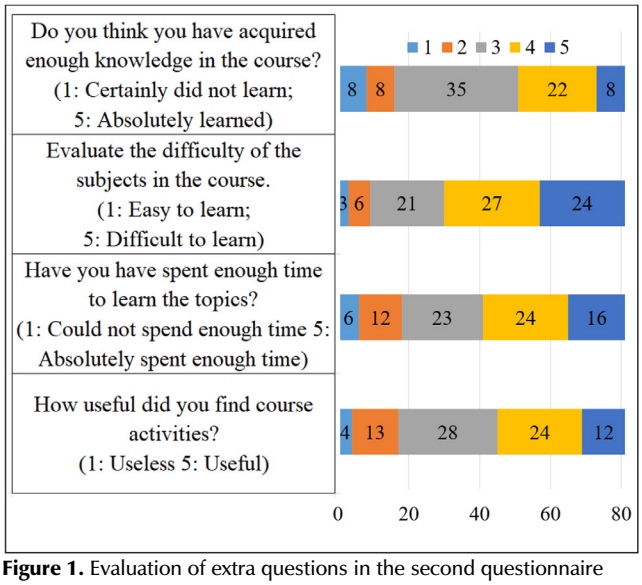


Figure 1. Evaluation of extra questions in the second questionnaire
Source: Authors

Quizzes

Seven topics were covered, with course activities for one. Pre-tests/post-tests were used to evaluate students before and after the lesson hour. In the pre-tests, more than 70% of the students scored over 80/100. In three subjects, 80% of the students achieved this level of success. The average grade in the post-tests of five topics was about 70, and it was about 50 points in two topics. These differences led to differences in the CLO results.

Table 3. University-wide questionnaire results

Assessments	Previous Year	Current Year
The course activities contributed to the learning process.	3,38	4,07
The course gave me knowledge, skills, and abilities	3,23	3,90
The course activities met the demands of the students.	3,11	3,77
Satisfied with the course activities	3,25	4,03

Source: Authors

Course Learning Outcomes (CLOs)

Quiz and exam questions were matched with the CLOs, and their success scores were calculated by the instructor in order to determine the students' performance. With these results, the improved hybrid-FC method used in the studied year was compared against the FC education approach employed in the previous year. The results are shown in Table 4.

It was observed that the success of the students increased in comparison with that of the previous year. The highest increase was reported for CLO-5, which increased from 2,4 to 3,4. Only the CLO-2 result slightly decreased. It can be interpreted that this decline occurred because the topics covered by CLO-2 were taught at the beginning of the semester, while the students were adapting to the inverted classroom activities. Generally speaking, the positive effect of the hybrid-FC education method is confirmed by these results.

Table 4. Course learning outcomes

#	Outcome	Previous Year	Current Year
1	Describe the fundamental concepts of the theory of probability, including conditional probability, random variables, probability distributions, and expectation.	2,4	2,6
2	Identify the probability mass function, the probability distribution function, the cumulative distribution function, expected value, variance, and standard deviation.	2,6	2,4
3	Perform a wide variety of probability calculations and derivations.	2,2	2,4
4	Solve practical problems consisting of genuine data through probabilistic reasoning.	2,3	3,0
5	Identify random processes and their characterization, the autocorrelation function, and the cross-correlation function.	2,4	3,4
6	Discuss the results of statistical analyses through graphical and verbal means.	2,5	3,3

Source: Authors

For the sake of comparison, in Rodríguez-Paz *et al*. (2021), the flipped classroom approach was applied to both online and face-to-face education during an engineering course. Instead of diluted classes, according to the spread of the pandemic, lessons were taken fully online or fully face-to-face. In our study, online and in-class sessions were held in parallel, with two student groups alternately participating in the sessions on a weekly basis. This helps students to adapt to changing conditions while keeping social life active, giving them the opportunity to interact with classmates, which reduces the negative effects of online learning, as well as pandemic stress. In the study by Rodríguez-Paz *et al*. (2021), a survey was applied to the students in order to measure the effectiveness of the proposed method. It was reported that 42,90% of the students agreed that they would not prefer to take online courses. When the students were asked to examine their own performances, 46,90% did not find any difference between online or in-class learning. 39,30% of the students said that their marks got better, which is consistent with our questionnaire results (Table 3). In our study, the effectiveness of the proposed educational method was measured using not only surveys, but also additional comprehensive quantitative and qualitative techniques in order to obtain more reliable results.

For further comparison, another study (Öncel *et al*., 2019) presented an FC approach for Communication Systems, which is also a core course in the EE curriculum. The efficiency of the proposed method was also measured through both qualitative and quantitative methods. Along with the interviews and surveys, CLO scores showed that flipped learning (FC) is more efficient than traditional learning methods. In this study, as complementary research, further improvements are presented in order for the FC approach to meet the requirements of effective learning under challenging conditions such as a pandemic. Our method can be applied in overpopulated engineering classes or schools, especially in developing countries.

Conclusions

This study presents an improved hybrid-FC education method and measures its effectiveness in a PRP course held during the pandemic at the EEE department of Atilim University. Therefore, based on extensive measurements, it aims to demonstrate that the proposed method is quite effective and may work under various conditions, including populated classes and schools and the pandemic era. Students' perspectives and performance were evaluated with quantitative/qualitative measurement methods according to the feedback received. The answers given to the questions in the interviews were interpreted, the results of the two questionnaire types were statistically analyzed (t-test), and students' achievement levels were evaluated in terms of the course learning outcomes, which were calculated based on the quiz and exam results. By comparing the results with those of the previous year, a significant increase was observed. Based on the obtained results, the proposed education model can be adapted to most classes in engineering, with the aim to improve learning processes under challenging conditions. The proposed method also improves the students' performance while minimizing the negative effects of remote education, even if rigid health measures are taken. To the authors' knowledge, this is the first applicable method of its kind in engineering education.

As a conclusion, the improved hybrid-FC education method based on TEL with diluted classes benefits students' perspectives and performance in PRP, which students generally define as a difficult course. The proposed method was introduced during pandemic days, when alternative educational methods became necessary. Thus, several negative effects of remote learning were eliminated while complying with all procedures aimed at mitigating the spread of the pandemic. Even though the proposed method and its effectiveness are confirmed for the PRP course in this paper, its structure can be adapted to other courses and engineering disciplines. Our approach could provide solutions to the negative effects of fully online learning while preserving its educational and systematic benefits. This study may guide researchers and instructors to follow such an approach for a variety of courses during pandemic times, as well as after lockdowns. However, it should be noted that the proposed hybrid-FC method cannot replace the classical, fully face-to-face teaching models. Nevertheless, the results show that it is applicable to major engineering courses in case of a pandemic or in overpopulated classes or schools, which is very common in developing countries.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (originaldraft, writing, re-view, and editing), data curation

References

Al Mamun, M. A., Azad, M. A. K., and Boyle, M. (2022). Review of flipped learning in engineering education: Scientific mapping and research horizon. *Education and Information Technologies, 27*, 1261-1286. https://doi.org/10.1007/s10639-021-10630-z

Almaiah, M. A., Al-Khasawneh, A., and Althunibat, A. (2020). Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. *Education and Information Technologies, 25*, 5261-5280. https://doi.org/10.1007/s10639-020-10219-y

Assunção Flores, M., and Gago, M. (2020). Teacher education in times of COVID-19 pandemic in Portugal: national, institutional and pedagogical responses. *Journal of Education for Teaching, 46*(4), 507-516. https://doi.org/10.1080/02607476.2020.1799709

Bishop, J., and Verleger, M. A. (2013). The flipped classroom: A survey of the research. In ASEE (Eds.), *2013 ASEE Annual Conference & Exposition* (pp. 23.1200.1-23.1200.18). ASEE. https://peer.asee.org/22585

Carrillo, C., and Flores, M. A. (2020). COVID-19 and teacher education: A literature review of online teaching and learning practices. *European Journal of Teacher Education, 43*(4), 466-487. https://doi.org/10.1080/02619768.2020.1821184

Chiquito, M., Castedo, R., Santos, A. P., López, L. M., and Alarcón, C. (2020). Flipped classroom in engineering: The influence of gender. *Computer Applications in Engineering Education, 28*(1), 80-89. https://doi.org/10.1002/cae.22176

Cigdemoglu, C., Kapusuz, K. Y., and Kara, A. (2014). Heterogeneity in classes: cooperative problem-solving activities through cooperative learning. *Croatian Journal of Education, 16*(4), 999-1029. https://doi.org/10.15516/cje.v16i4.1019

Coruk, R. B., Yalcinkaya, B., and Kara, A. (2020). On the design and effectiveness of Simulink-based educational material for a communication systems course. *Computer Applications in Engineering Education, 28*(6), 1641-1651. https://doi.org/10.1002/cae.22344

Erjongmanee, S. (2019). Alternative approach to teach probability and statistics for college engineering students. In IEEE (Eds.), *2019 IEEE Global Engineering Education Conference* (*EDUCON*) (pp. 931-936). IEEE. https://doi.org/10.1109/EDUCON.2019.8725172

Hodges, C., Moore, S., Lockee, B., Trust, T., and Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review, 27*(1), 1-12. http://hdl.handle.net/10919/104648

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., and Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior, 119*, 106713. https://doi.org/10.1016/j.chb.2021.106713

García-Ramírez, Y. (2019). It is not enough to flip your classroom. A case study in the course of Pavements in Civil Engineering. *Ingeniería e Investigación, 39*(3), 62-69. https://doi.org/10.15446/ing.investig.v39n3.81426

Gómez, J., León, E., Cubides, C., Rodríguez, A., Mahecha, J., and Rubiano, J. C. (2014). Evolution of teaching and evaluation methodologies: The experience in the computer programming course at the Universidad Nacional de Colombia. *Ingeniería e Investigación, 34*(2), 85-89. https://doi.org/10.15446/ing.investig.v34n2.41276

Kesim, M., and Ozarslan, Y. (2012). Augmented reality in education: current technologies and the potential for education. *Procedia – Social and Behavioral Sciences, 47*, 297-302. https://doi.org/10.1016/j.sbspro.2012.06.654

Kim, M. K., Kim, S. M., Khera, O., and Getman, J. (2014). The experience of three flipped classrooms in an urban university: an exploration of design principles. *The Internet and Higher Education, 22*, 37-50. https://doi.org/10.1016/j.iheduc.2014.04.003

König, J., Jäger-Biela, D. J., and Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *European Journal of Teacher Education, 43*(4), 608-622. https://doi.org/10.1080/02619768.2020.1809650

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. Social psychological and personality science, 8(4), 355-362. https://doi.org/10.1177/1948550617697177

Mason, G. S., Shuman, T. R., and Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education, 56*(4), 430-435. https://doi.org/10.1109/TE.2013.2249066

Merideno, I., Antón, R., and Prada, J. G. (2015). The influence of a non-linear lecturing approach on student attention: Implementation and assessment. *Ingeniería e Investigación, 35*(3), 115-120. https://doi.org/10.15446/ing.investig.v35n3.49644

Mezhennaya, N. M., and Pugachev, O. V. (2018). On the results of using interactive education methods in teaching Probability Theory. *Problems of Education in the 21st Century, 76*(5), 678. https://ceeol.com/search/article-detail?id=941840

Moorhouse, B. L. (2020). Adaptations to a face-to-face initial teacher education course 'forced' online due to the COVID-19 pandemic. *Journal of Education for Teaching, 46*(4), 609-611. https://doi.org/10.1080/02607476.2020.1755205

Nagy, G., and Sikdar, B. (2008). Classification and evaluation of examples for teaching probability to electrical engineering students. *IEEE Transactions on Education, 51*(4), 476-483. https://doi.org/10.1109/TE.2007.914942

Nascimento, M. M., Morais, E., and Martins, J. A. (2016). Skyfall: Representations in probability problems in Engineering. In IEEE (Eds.), *2016 2nd International Conference of the Portuguese Society for Engineering Education* (*CISPEE*) (pp. 1-4). IEEE. https://doi.org/10.1109/CISPEE.2016.7777741

O'Flaherty, J., and Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education, 25*, 85-95. https://doi.org/10.1016/j.iheduc.2015.02.002

Öncel, A. F., and Kara, A. (2019). A flipped classroom in communication systems: Student perception and performance assessments. *The International Journal of Electrical Engineering & Education, 56*(3), 208-221. https://doi.org/10.1177/0020720918788718

Richards, L. G., and Donohue, S. K. (2012). Managing student outcomes in a totally asynchronous learning environment: Lessons learned. In IEEE (Eds.), *2012 Frontiers in Education Conference Proceedings* (pp. 1-6). IEEE. https://doi.org/10.1109/FIE.2012.6462471

Rodríguez, J. C. C., and Contreras, J. S. (2011). The quality of engineering education is a key factor in development. *Ingeniería e Investigación, 31*(1), 40-50. https://doi.org/10.15446/ing.investig.v31n1SUP.27927

Rodríguez-Paz, M. X., González-Mendivil, J. A., Zamora-Hernández, I., and Sánchez, B. (2021). A hybrid and flexible teaching model for engineering courses suitable for pandemic conditions towards the new normality. In IEEE (Eds.), *2021 IEEE Global Engineering Education Conference* (*EDUCON*) (pp. 381-387). IEEE. https://doi.org/10.1109/EDUCON46332.2021.9454014

Sheikh, W. (2019). An intuitive, application-based, simulation-driven approach to teaching probability and random

processes. *The International Journal of Electrical Engineering & Education, 2019*, 0020720919866405. https://doi.org/10.1177/0020720919866405

Vilchez-Sandoval, J., Llulluy-Nunez, D., and Lara-Herrera, J. (2021). Work in progress: Flipped classroom as a pedagogical model in virtual education in networking courses with the Moodle Learning Management System against COVID 19. In IEEE (Eds.), *2021 IEEE World Conference on Engineering Education* (*EDUNINE*) (pp. 1-3). IEEE. https://doi.org/10.1109/EDUNINE51952.2021.9429101

Watermeyer, R., Crick, T., Knight, C., and Goodall, J. (2021). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education, 81*, 623-641. https://doi.org/10.1007/s10734-020-00561-y

Zhang, W., Wang, Y., Yang, L., and Wang, C. (2020). Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak. *Journal of Risk and Financial Management*, *13*(3), 55. https://doi.org/10.3390/jrfm13030055

Research Article / Education in Engineering https://doi.org/10.15446/ing.investig.101349

Adapting to Remote Learning during COVID-19: An Engineering Education Approach

Adaptación para el aprendizaje a distancia durante el COVID-19: un enfoque de educación en ingeniería

Rafael Granillo-Macías[1]

ABSTRACT

With a focus on the problems faced by the education sector, this article shows the main technological tools used by teachers for remote teaching in an engineering education program during the first year of the COVID-19 pandemic. Through a case study in a Latin American university positioned in the QS World University Ranking, the tools most frequently used for communication, storing information, messaging, and sharing learning videos are described. This study examines engineering's areas of knowledge, considering that learning them poses a challenge for teaching in virtual environments due to the need to carry out laboratory practices. Using statistical methods, the relationships between the areas of the education program, the teachers' sociodemographic variables, and the technological tools used to offer online classes during the lockdown are described. The results show that, from the perspective of the teacher and for the field of engineering, the tools with the most significant effect, in sometimes essential for teaching during the pandemic, were those related to the Google platforms, which were used in 75% of the cases for video communication and storage. Meanwhile, with 44 and 56% WhatsApp and Moodle, were the most used for messaging and sharing contents and learning materials.

Keywords: SARS-CoV-2, engineering education, Latin America, learning during the pandemic

RESUMEN

Con énfasis en las problemáticas que enfrentó el sector educativo, en este artículo se muestran las principales herramientas tecnológicas utilizadas por los docentes para la enseñanza remota en un programa educativo de ingeniería durante el primer año de la pandemia del COVID-19. A través de un caso de estudio en una universidad latinoamericana posicionada en el QS World University Ranking, se describen las herramientas que se utilizaron con mayor frecuencia para la comunicación, el almacenamiento de información, la mensajería y el envío de materiales de aprendizaje y videos. Este estudio examina las áreas de conocimiento de la ingeniería, considerando que el aprendizaje de estas representa un reto para la enseñanza en entornos virtuales debido a la necesidad de realizar prácticas de laboratorio. Utilizando métodos estadísticos, se describen las relaciones entre las áreas del programa educativo, las variables sociodemográficas del docente y las herramientas tecnológicas utilizadas para dar clases en línea durante el confinamiento. Los resultados muestran que, desde la perspectiva del docente y para el campo de la ingeniería, las herramientas con el efecto más significativo y que incluso resultan esenciales para la enseñanza en pandemia fueron aquellas relacionadas con las plataformas de Google, siendo utilizadas en un 75 % de los casos para la comunicación y el almacenamiento. Entretanto, con 44 y 56 %, WhatsApp y Moodle fueron las de mayor uso para enviar mensajes y compartir contenido y materiales de aprendizaje.

Palabras clave: SARS-CoV-2, educación en ingeniería, América Latina, aprendizaje en pandemia

Received: February 28th, 2022
Accepted: March 4th, 2023

Introduction

The SARS-CoV-2 virus, first reported at the end of December 2019 in Wuhan, China, has had a severe impact worldwide. The imminent spread of the highly infectious disease (COVID-19) caused by this coronavirus led to taking drastic measures in different countries, which included lockdowns. Efforts by multiple governments to stop the viral outbreak included strict protocols and regulations to facilitate social distancing (Bozkurt and Sharma, 2020).

According to data from UNESCO (2020), more than 1,5 billion students of all ages were affected due to the closing of schools and universities. This amounts to 90% of the students enrolled around the world. In light of the above, most schools, colleges, and universities sought to minimize the impact on education, transforming teaching by employing different approaches and alternatives and migrating to online environments with the help of digital tools and educational websites (Hassan *et al*., 2020; Mohmmed *et al*., 2020).

[1] Industrial Engineering, Universidad Autónoma del Estado de Hidalgo - Escuela Superior de Ciudad Sahagún, México. PhD in Logistics and Supply Chain Management, Universidad Popular Autónoma de Puebla, México. Affiliation: Full professor, Universidad Autónoma del Estado de Hidalgo – Escuela Superior de Ciudad Sahagún, México. Email: rafaelgm@uaeh.edu.mx

The measures taken by the countries of Latin America were the suspension of face-to-face classes at all education levels, which gave rise to the deployment of remote learning modalities through a variety of formats and platforms (CEPAL-UNESCO, 2020). Teaching was redesigned, in most cases, through the digitization of the content offered to students through various digital tools, including synchronous sessions in virtual environments (videoconferences) at the same time as face-to-face sessions would take place.

As the primary support for learning during the pandemic, the integration of information and communication technologies (ICTs) has influenced various fields of study, bringing significant changes to teaching processes (Ramya and Poongodi, 2021). In the process of educating during this kind of emergencies, ICT tools have shown their broad usefulness for teaching, playing a prominent role in advancing the learning of teachers and students, specifically in areas such as engineering (Kavitha *et al*., 2021).

The COVID-19 pandemic brought many challenges to higher education regarding teaching, learning, and research collaborations between government and education institutions (Mok *et al*., 2021). Engineering education is a specific sector of higher education where practical exercises strongly determine the effectiveness of the teaching and learning process, which includes experiments and laboratory classes. In addition, technical skills are crucial for engineers, according to the paradigm of *learning by doing*. Therefore, the students of engineering courses need access to laboratories, specialized equipment, and instruments in order to carry out practical experiments to acquire the technical and social skills associated with their area of specialization (Ożadowicz, 2020).

Before the pandemic, the expectations regarding the quantity and quality of education in engineering had experienced a significant increase (Ramya and Poongodi, 2021) due to the need for a technological transformation aimed at providing solutions for problems within the so-called *Industry 4.0*.

Focusing on the problems faced by the education sector due to COVID-19, and highlighting the challenges faced by engineering education in this context, this research raises two initial questions:

- What strategies have been implemented for online engineering education COVID-19?
- What technological tools have been used for remote teaching regarding the different subjects that make up an education program in engineering?

The objective is to identify, from the sociodemographic perspective of the teacher, the current technological tool with the most significant effect on the teaching of engineering in the face of catastrophic events.

Through a case study in a Latin American university positioned in the QS World University Ranking, this article shows some of the main technological tools used during the first year of the COVID-19 pandemic for remote teaching in an engineering education program.

This document is organized as follows. Literature review section outlines some of the proposals for learning engineering made during the pandemic by different universities and education centers at an international level. Method section describes the methodology used for this research, which involves a case study where the teacher identifies the technological tools used from March 2020 to March 2021. Results section presents the analyses carried out. Finally, the last section presents the discussion and conclusions of this research work.

Literature review

The exponential development of technology has allowed online education to become a viable alternative with high potential to change the paradigm of traditional learning (face-to-face conferences in a classroom) towards an education based on technological resources that promote interaction, collaboration, and learning in virtual environments (Ferreira *et al*., 2018). Currently, the introduction of new tools and learning methods based on approaches such as the so-called *TechTeach* seeks to improve the quality of teaching and increase motivation towards learning. Flipped classrooms, gamification, soft skills training, and quizzes are some TechTeach tools that have been successfully implemented in the teaching-learning processes of higher education during the pandemic (Portela, 2020).

Specifically, the teaching of engineering via online tools has been reported in the fields of electricity, computer science, and information technology (Martínez *et al*., 2019). However, although online education is not unknown to teachers, there are still opportunities for improvement. Some of the main weaknesses identified in this modality of teaching are, as pointed out by Dhawan (2020), Syauqi *et al*. (2020), and Cojocariu *et al*. (2014), the lack of two-way interaction, a possible loss of direct communication, isolation, the need to include a large number of didactic resources, and an excess of theoretical content that does not provide students with real practice.

Given the exceptional and unprecedented circumstances caused by the COVID-19 pandemic, in a context in which schools closed their doors in 185 countries, the opportunity to contribute to solving the weaknesses of online teaching arose (Espino-Díaz *et al*., 2020), revealing a need to train teachers in the use of technological tools and to adapt the different contents of the study plans to the new context of the health crisis (Asgari *et al*., 2021).

Aiming to provide a general overview of the subject, some proposals made by different universities and education centers worldwide were identified by reviewing the literature published in databases and academic journals.

Through a survey designed to collect perceptions about problems, pedagogical strategies, and student-teacher interactions during the teaching process, Fardoun *et al*. (2020) present an exploratory study conducted in Ibero-America. These authors conclude that the most used technological and material resources for monitoring and evaluating of the students in terms of the strategies and activities that support the teaching process are blogs, portfolios, forums, collaborative work, and, to a lesser extent, videos made by the students.

Among the strategies used for teaching, according to Culzoni *et al*. (2021), different faculties of engineering and technological universities in Argentina used platforms or virtual classrooms to respond to their students' queries, as well as information systems, videoconferences and free software for subjects in the area of exact sciences.

In Spain, some universities that offered educational programs in engineering replaced the traditional exam with a sequence of remote collaborative learning activities in order to foster students' engineering and thinking skills as well as transversal competencies (Ripoll *et al*., 2021).

García-Alberti *et al*. (2021) analyze an education program in civil engineering offered in universities in Spain and Peru. This study describes teaching and evaluation strategies including technological tools, such as virtual classrooms and videoconferences.

Through a methodological proposal to provide tools for remote education in electrical engineering, Evstatiev and Hristova (2020) present a case study of two universities in Bulgaria with regard to the adaptation of engineering classes to the context of COVID-19. The methodology of this study includes several phases: 1) an analysis of needs (including technological aspects), 2) the preparation of educational materials, 3) the selection of teaching methods, and 4) and their evaluation.

Mohd Zulkefli (2020) presents a study carried out on students of an education program in computer science in Oman. This study seeks to evaluate the effectiveness of Google Classroom as a learning tool for e-learning. The author suggests that most students perceived Google Classroom as positive due to its functionality and mobility. It is also concluded that the challenges faced by students when using Google Classroom are similar to the problems that are addressed in other e-learning platforms such as Moodle.

To identify the challenges experienced by students enrolled in different engineering departments during online classes in the spring of 2020, Asgari *et al*. (2021) conducted an observational study based on questionnaires applied in a university in the United States. The results show the multiple issues that negatively influence online engineering education, such as 1) technical problems, 2) learning challenges, 3) security and privacy issues, and 4) the lack of practical training in different technological tools for effective communication. These authors recommend strategies for students, teachers, and administrators to close the gap between the improvement of online engineering education and the tools and technology used for this purpose.

In India, Hassan *et al*. (2020) studied the teachers' perspective on the learning model based on online platforms. They analyzed the challenges and problems that arose in the migration of education to virtual environments, and they identified the tools used for teaching. The results obtained from a sample of teachers from different institutes in India indicate that a large percentage of teachers (87,5%) used WhatsApp to communicate and give instructions, while 56,9% of teachers used WhatsApp in combination with the Zoom videoconference platform. 37,5% of the teachers used Google Classroom and WhatsApp. Studies such as that conducted by Sahoo *et al*. (2021) analyze the consequences of the pandemic on the teaching-learning system, confirming the effects of COVID-19 on higher education institutions.

In Indonesia, Suryaman *et al*. (2020) conducted a study on the learning profile of engineering students during the COVID-19 pandemic, using quantitative and qualitative methods for data collection. The authors point out that, from the students' perspective, online learning applications must be accessible and combined with various applications. Among the tools for online learning, the most useful were Google Classroom, followed by WhatsApp and Zoom.

With the aim to suggest activities for teachers to implement new tools regardless of the learning modality, Kanetaki *et al*. (2022a) analyze the effectiveness and student participation patterns regarding videos disseminated through YouTube as asynchronous support for engineering laboratories.

Ożadowicz (2020) describes a case study carried out in a technical university, discussing the tools and methods for remote learning introduced during the COVID-19 lockdown. Tools such as videoconferences, webinars (lectures and laboratory classes), and mind maps are reported in this study as the main activities carried out during remote learning.

García Martín and García Martín's study on the use of technological tools in the Spanish education sector (2021) indicates that, during the lockdown, teachers mainly used digital devices or applications for instructional design based on electronic documents, spreadsheets, and Google or Prezi presentations.

Using a hybrid machine learning model that includes statistical analyses, Kanetaki *et al*. (2022b) explore the prediction of grades for a population of students in an engineering course during the first semester of the 2020 academic year.

Methodology

This research employed a hypothesis-generating case study approach, with a quantitative and qualitative methodology based on the proposals of Eisenhardt (1989), Stake (1999), and

Yin (2003). The stages of this work are shown in Figure 1 and which include 1) the formulation of objectives and hypotheses, 2) a description of the context and the participants, 3) data collection and recording, 4) the analysis and validation of information, and 5) general conclusions and implications.

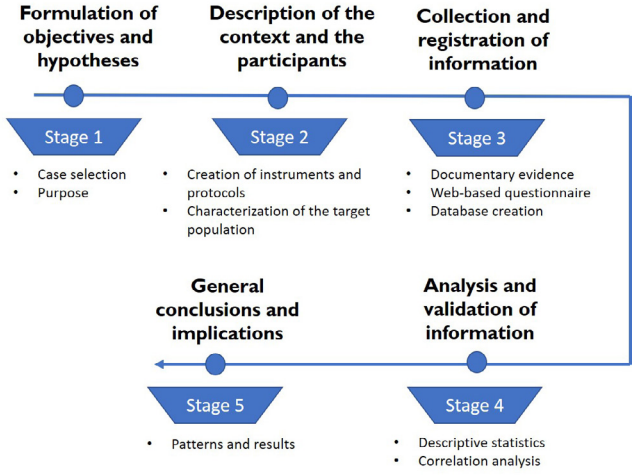


Figure 1. Methodological approach
Source: Authors

According to Eisenhardt (1989), Yin (2003), and Iglesias-Pradas *et al*. (2021) the case study methodology is used as a research strategy with the purpose of understanding the dynamics and complexity of a particular population, which also allows understanding the meaning of an experience and building theories from the early stages of a project.

For the data record corresponding to stage 3, this study considered that the selection of a specific technological tool corresponds to the instructional design used by the teacher, according to the principle of academic freedom and autonomy promoted by the analyzed institution.

Objective and hypothesis

The objective was to show the technological tools used for teaching and learning during the first year of the pandemic, considering the different subjects taught in an engineering program at a public university in Latin America. The following hypotheses were formulated:

H1. The use of technological tools for remote teaching differs depending on the area of knowledge.

H2. A teacher's use of technological tools for remote teaching is influenced by sociodemographic variables such as education level, gender, and type of contract.

Description of the context and the participants

This study was carried out at a Latin American university during the 2020 academic year, selecting a faculty located in an industrial zone with a high demand for engineering professionals. This university offers different study programs in Mexico, distributed among its other faculties, including the educational program in Industrial Engineering, on which this study was carried out. The university under study is ranked 801-1000 in the QS Global World University Ranking of 2022 and 171-180 in the LatAm University Ranking (Quacquarelli Symonds, 2021).

At the beginning of the pandemic, and for 12 months (corresponding to two semesters of the 2020 academic year), the 46 teachers of the analyzed education program were asked to record the technological tools used with their students during social distancing. The number of respondents corresponds to the total population of teachers assigned to this educational program, out of which 48% have postgraduate studies. They are 19 women and 27 men. In addition, the teachers are categorized according to the number of hours/classes they teach in person, that is, a full-time teacher (full dedication) has a workload greater than 10 hours per week, while a part-time teacher has a workload between 5 and 10 hours per week. There is a third category of teacher with minimum workload that includes those who teach between 1 and 4 hours per week.

Among this university's tools for teaching in digital environments, the most outstanding are 1) a technological platform developed in Moodle, called *Plataforma Garza*, 2) an open electronic repository, 3) databases of scientific journals, 4) a YouTube channel, 5) the G Suite for Education, and 5) an iTunes U channel.

The curriculum of this education program is of the schooling type (face-to-face modality) divided in nine semesters made up of theoretical hours and teaching practices, where the student takes different subjects on basic sciences, arts and languages, manufacturing, environmental engineering, and quality, among others.

The subjects that make up this program are divided by areas of knowledge, as shown in Table 1.

According to the characteristics of the instructional design and the learning objectives, each subject has a primary focus, either towards practice, theory, or both. These subjects are classified within the areas of basic sciences (BS), industrial engineering (IE), engineering sciences (ES), economic and administrative sciences (EAS), social sciences and humanities (SSH), and complementary courses (CC).

Data collection and recording

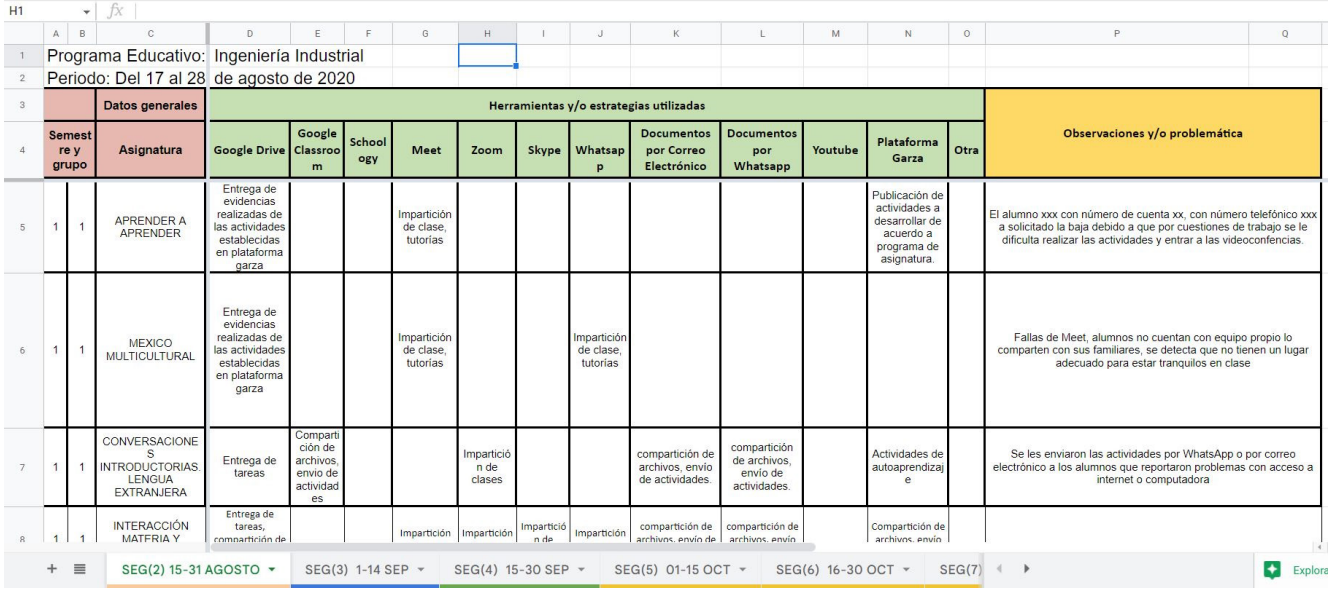
To collect information, a quantitative study with a descriptive and correlational approach was carried out in the form of a survey. The selected questionnaire technique had the objective of providing information on the different perceptions of the participating teachers. For this case, a list of dichotomous, open-ended, and simple-selection questions was designed.

Figure 2 shows an example of the instrument used to record information from the teachers' perspective.

Table 1. Areas of knowledge and subjects

Knowledge area	Subjects	Approach	
		Theoretical	Practical
Basic sciences	Matter and energy interaction, Precalculus, General chemistry, Differential and integral calculus, Statistics and probability, Differential equations, Inferential statistics, Mechanics, Thermofluids, Linear algebra	•	•
Industrial engineering	Work studies, Sustainable development and environment, Maintenance management, Environmental assessment, Processes, Production ergonomics, Quality fundamentals, Programming and simulation topics, Production management, Plant siting and layout, Quality management, Ergonomics, Logistics and supply chain, Reengineering, World-class tools	•	•
Engineering sciences	Introduction to systems theory, Computer technology, Materials studies, Industrial drawing, Electrical machines and electronic devices, Manufacturing, Safety and hygiene, Numerical control, CAD-CAM, Computer manufacturing	•	•
Economic and administrative sciences	Accounting and costs, Administration and strategic planning, Economic and financial analysis, Human resources administration, Project formulation and evaluation, Organizational development	•	
Social sciences and humanities	Learning to learn, Multicultural Mexico, Fundamentals of research methodology, Leadership and collaborative work	•	
Complementary courses	Languages, Arts, Physical and emotional health	•	•

Source: Authors



1	Programa Educativo:	Ingeniería Industrial													
2	Periodo: Del 17 al 28	de agosto de 2020													
3		Datos generales	Herramientas y/o estrategias utilizadas												Observaciones y/o problemática
4	Semestre y grupo	Asignatura	Google Drive	Google Classroom	Schoology	Meet	Zoom	Skype	Whatsapp	Documentos por Correo Electrónico	Documentos por Whatsapp	Youtube	Plataforma Garza	Otra	
5	1 1	APRENDER A APRENDER	Entrega de evidencias realizadas de las actividades establecidas en plataforma garza			Impartición de clase, tutorías							Publicación de actividades a desarrollar de acuerdo a programa de asignatura.		El alumno xxx con número de cuenta xx, con número telefónico xxx a solicitado la baja debido a que por cuestiones de trabajo se le dificulta realizar las actividades y entrar a las videoconfencias.
6	1 1	MEXICO MULTICULTURAL	Entrega de evidencias realizadas de las actividades establecidas en plataforma garza			Impartición de clase, tutorías			Impartición de clase, tutorías						Fallas de Meet, alumnos no cuentan con equipo propio lo comparten con sus familiares, se detecta que no tienen un lugar adecuado para estar tranquilos en clase
7	1 1	CONVERSACIONES INTRODUCTORIAS. LENGUA EXTRANJERA	Entrega de tareas	Comparticón de archivos, envio de actividades			Impartición de clases			compartición de archivos, envio de actividades.	compartición de archivos, envio de actividades.		Actividades de autoaprendizaje		Se les enviaron las actividades por WhatsApp o por correo electrónico a los alumnos que reportaron problemas con acceso a internet o computadora
8	1 1	INTERACCIÓN MATERIA Y	Entrega de tareas, compartición de			Impartición	Impartición	Impartición	Impartición	compartición de archivos, envio de	compartición de archivos, envio		Comparticón de archivos, envio		

SEG(2) 15-31 AGOSTO | SEG(3) 1-14 SEP | SEG(4) 15-30 SEP | SEG(5) 01-15 OCT | SEG(6) 16-30 OCT | SEG(7)

Figure 2. Information record
Source: Authors

The data collection instrument (Figure 2) was shared via the Google Documents Web platform, which could be accessed with the institutional email account of each participant. Technological tools used were classified as 1) communication applications based on video calls, 2) storage applications, 3) messaging and document sharing applications, and 4) content and learning materials. In this Google document, teachers were asked to indicate the technological tool used for teaching, delivering assignments, sharing files, sending activities, and teamwork, among others. In total, teachers could select between 12 technological tools or indicate others. Thus, several options were provided, which included WhatsApp, Zoom, Google Classroom, and YouTube.

Additionally, teachers were asked to record, some of the main problems they encountered during their remote sessions.

In total, during the first year of the COVID-19 pandemic, records of 1 037 sessions were obtained for the 341 students who make up this education program.

Analysis and validation of information

In this stage, descriptive statistical analyses were carried out to learn about the main technological tools used by teachers in the different subjects and areas of knowledge. The validity and reliability of the information obtained from the questionnaires were evaluated using the SPSS software.

The observed frequencies were compared to the expected ones via Chi-square statistics, considering Equation (1):

$$X^2 = \sum_i \sum_j \frac{\left(n_{ij} - m_{ij}\right)^2}{m_{ij}} \quad (1)$$

where n indicates the observed frequencies and m the expected ones, with i being any row and j any column within an array of two qualitative variables, to test the hypothesis that the classification criteria are independent (*i.e.*, technological tools, area of knowledge, and sociodemographic variables). In this analysis, $p<0,05$ was considered to be statistically significant value.

With the previous stages, and using the information obtained through descriptive statistics and correlational analysis, the general conclusions and implications for this case study were drawn.

Results

Descriptive statistics

The data obtained from teachers with regard to the first year of the pandemic (March 2020 – March 2021) were collected, organized, and analyzed using SPSS version 21. In the first analysis, descriptive statistics related to the use of the tools were performed. As shown in Table 2, Google Meet was the primary tool for communication via video calls, with WhatsApp and Zoom occupying the second and third place, respectively. The highest percentages of use for Google Meet were reported for Material and energy interaction, General chemistry, Differential equations, and Inferential statistics, subjects taught by four professors (three part-time professors, and one full-time professor).

Concerning storage tools, Google Drive was the most commonly used, with 71%, in addition to WhatsApp for messaging and sharing documents. The highest percentages reported for Google Drive corresponded the subjects of Matter and energy interaction, Chemistry, and Fundamentals of research methodology. Meanwhile, WhatsApp was used more frequently in Languages, Materials studies, and Calculus, subjects taught by four part-time professors.

Table 2. Technological tools used during the pandemic

Classification	Tool	Frequency	Percentage	Accumulated percentage
Communication via video calls	Google Meet	644	62,1	62,1
	WhatsApp	204	19,7	81,8
	Zoom	146	14,1	95,9
	Hangouts	26	2,5	98,4
	Skype	11	1,1	99,4
	None	6	0,6	100
Storage	Google Drive	736	71	71
	None	301	29	100
Messaging and document sharing	WhatsApp	429	41,4	41,4
	Email	384	37,0	78,4
	None	224	21,6	100
Content and learning materials	Moodle	550	53	53
	None	365	35,2	88,2
	Google Classroom	122	11,8	100
Video content	None	757	73	73
	YouTube	251	24,2	97,2
	Other	29	2,8	100

Source: Authors

As shown in Table 2, the content tool of choice for learning materials was Moodle, as Plataforma Garza was the most widely used tool in all subjects. In the case of video content, most teachers did not use a specific device. YouTube was mainly used in specific Language subjects, which were taught by four part-time teachers.

Google Forms, Quizizz, Khan Academy, Idroo, Pixton, Genially, and Jamboard were other options reported in particular cases by some teachers. The subjects that reported very little use of technological tools were Arts, Sustainable development, Accounting and Costs, and Numerical control, which were taught by part-time teachers.

Areas of knowledge

Subsequently, the descriptive statistics of the technological tools used in the subjects grouped by areas of knowledge were carried, aiming out to understand the relationship between these categorical variables (Table 3). These results are shown below.

Communication via video calls: The most used technological tool for communication via video calls was Google Meet, with 81% in the IE area, 67% in SSH, 63,2% in EAS, 61,4% in BS, 52% in ES, and 54,5% in CC. Zoom and WhatsApp ranked second as video call communication tools, mainly

in the areas of CC, ES, and BS, with 29,9, 17,9, and 10,5%, respectively.

When the Chi-square test regarding the relationship between the video call communication tools and the areas of knowledge was performed, an asymptotic significance level of 0,00 was obtained (considering a value of α=0,05). Therefore, statistically sepaking, there is a relationship between the two variables.

Storage: Of the tools used for storage, the most used in SSH, CC, BS, IE, and EAS was Google Drive, with 69,1, 71,9, 74,4, 76,8, and 80,5%, respectively. This tool was predominant for teachers, except for the ES knowledge area, where 44,7% of the participants did not use any storage tool. When evaluating the relationship between technological storage tools and the areas of knowledge, an asymptotic significance level of 0,00 was obtained, with a value of α=0,05. Therefore, statistically speaking, it can be considered that there is a relationship between both variables.

Messaging and document sharing: The primary tool used for messaging and sharing documents was WhatsApp in the CC, BS, and ES areas, with 48,2, 47,4, and 36,8%, respectively. Meanwhile, email was the most widely used tool in SSH with 39,4%, as well as in EAS with 46%. The very little of the IE area used a tool for messaging and sharing documents (37,5%). Regarding the relationship between these tools for messaging and sharing documents and the areas of knowledge, an asymptotic significance level of 0,00 was obtained, with a value of α=0,05, so, statistically speaking, there is a relationship between these variables.

Content tools and learning materials: Moodle was the main tool in SSH, CC, BS, IE, EAS, with 64,9, 43,3, 54, 59,5, and 60,9%, respectively. Google Classroom was also used in all areas. However, in most cases, ES did not report the use of a specific tool for learning.

According to the Chi-square test performed, an asymptotic significance level of 0,00 was obtained (considering a value of α=0,05), so, statistically speaking, it can be assumed that there is a relationship between the tools used and the areas of knowledge.

Video content: In all areas of knowledge, the use of video tools was not predominant. The most relevant participation of YouTube tool was reported in BS, EAS, and CC with 30,9, 28,7, and 26,8%, respectively. Regarding the relationship between this tool and the areas of knowledge, an asymptotic significance level of 0,00 was obtained (considering a value of α=0,05). Thus, there is a relationship between these variables.

Note that, in all cases, there is a statistical relationship between the technological tool used by the teacher and the area of knowledge.

As shown in Table 3, the results obtained indicate that Google Drive, Google Meet, WhatsApp, Moodle, and YouTube exhibited a higher frequency of use in Basic Sciences and Industrial Engineering. Zoom was also used by teachers in Complementary Courses such as Languages, Arts, and Occupational health.

In contrast, the teachers who taught courses in the areas of Social Sciences and Economic and Administrative Sciences were the ones who reported the lowest use of technological tools to offer remote learning courses.

Sociodemographic variables

To verify the H2 hypothesis raised at the beginning of this research, a second analysis was carried out which considered the qualitative variables of the teachers. For this second analysis, statistical association measurements were also used, accompanied by a significance test regarding the sociodemographic variables of education level, gender, and contract type (Table 4).

Communication via video calls: An independence hypothesis test was carried out, finding that the communication tools are indeed related to sociodemographic variables. On the other hand, based on the statistical value of Pearson's Chi-square test, the null hypothesis of independence between the categorical variables can be rejected.

Storage: According to the statistical value of Pearson's Chi-square test, the contract type and gender are independent

Table 3. Descriptive statistics: technological tools and areas of knowledge

Areas of knowledge	Communication via video call						Storage		Messaging and document sharing			Contents and learning materials			Video content		
	Google Meet	Zoom	Hangout	Skype	WhatsApp	None	Google Drive	None	WhatsApp	Email	None	Moodle	Google Classroom	None	YouTube	Other	None
SSH	63	3	1	3	24	0	65	29	34	37	23	61	16	17	22	5	67
CC	122	67	0	3	32	0	161	63	108	86	30	97	42	85	60	5	159
BS	175	30	9	2	69	0	212	73	135	116	34	154	29	102	88	16	181
IE	136	5	0	2	25	0	129	39	55	50	63	100	19	49	22	0	146
EAS	55	9	0	0	23	0	70	17	32	40	15	53	10	24	25	3	59
ES	93	32	16	1	31	6	99	80	65	55	59	85	6	88	34	0	145

Source: Authors

Table 4. Analysis of sociodemographic variables

Variables	Communication via video call		Storage		Messaging and document sharing		Contents and learning materials		Video content	
	x^2	g	x^2	g	x^2	g	x^2	g	x^2	g
Contract type	28,24	0,00	0,06	0,96	11,17	0,02	38,28	0,00	21,56	0,00
Gender	47,29	0,00	0,00	0,99	2,01	0,36	39,80	0,00	2,21	0,33
Education level	20,09	0,00	6,60	0,01	1,65	0,43	2,356	0,308	6,36	0,04

X^2 = Pearson's Chi-square; g = asymptotic significance (bilateral)

Source: Authors

of the storage tool used by the teachers. However, the academic degree is indeed related, so the independence hypothesis is rejected.

Messaging and document sharing: The education level and gender are independent of the tools used in this regard. According to the independence hypothesis test, there is a relationship between the type of contract and use of messaging and document sharing tools.

Content tools and learning materials: The independence hypothesis test results indicate that the type of contract and gender are related to teachers' content tools and learning materials. In the opposite case, the academic degree variable is independent. An asymptotic significance level higher than α=0,05 was obtained.

Video content tools: It was found that the contract type and the education level are related to the use of these tools. Gender was found to be independent due to the value obtained in the Chi-square test.

Discussion

Due to COVID-19, severe challenges emerged in higher education, which can be analyzed from the teacher and the students' perspective, as well as from the learning contents. The strategies applied for the continuity of the teaching and learning processes during the pandemic were mainly based on different open technological platforms.

In this case study, it can be observed that, for the field of engineering, the most valuable tools were related to Google platforms, as is the case of Meet and Drive. This, in addition to technological tools such as Moodle for content and WhatsApp for communication and messaging between teachers and students.

The use of digital technologies associated with Industry 4.0 requirements and trends are also factors that must be considered regarding the learning competencies taught in virtual environments and, in general, for the graduation profile of engineering students (González-Hernández and Granillo-Macías, 2020).

Although technological tools for learning are diverse and constantly growing, there is still a significant gap between them and the students. In the records provided by the teachers regarding the main problems that occurred during the delivery of remote sessions, the students' lack of response to messages and video call sessions was recurrent. This was mainly caused issues involving Internet access, limitations regarding economic and technological resources, or illness derived from COVID-19. Another complication reported by the teachers involved getting the students to focus and pay attention during remote sessions.

The different areas of knowledge in engineering also pose a challenge for teaching in virtual environments. In most cases, the technological tool used by the teacher was related to the area of knowledge. For example, the teachers used tools such as Google Meet and WhatsApp in areas like BS or IE, and, in other areas related to complementary courses, they used tools such as Zoom and YouTube.

In particular, platforms such as YouTube were little explored by the participants of this study, although this tool, according to Kanetaki *et al.* (2023), has proven to be effective for learning in engineering.

The results reveal that, from the perspective of the teacher, for the field of engineering, the tools with the most significant effect, which in some cases are essential for teaching during a pandemic, were those related to the G Suite, WhatsApp, and Moodle.

The relationships between the studied sociodemographic variables and the tools used by teachers stand out. The contract type is dependent on communication tools, messaging and document sharing, content and learning materials, and video content. In contrast, the education level is independent of the messaging and document sharing and content and materials categories.

Conclusions

Based on the results obtained, the strategies applied for the continuity of teaching and learning activities during the quarantine were based on different technological platforms freely accessible for students and teachers. In this study, the areas of knowledge in an engineering program and their relationship with teachers' technological tools were analyzed. It was determined that, although there are areas that can be brought to virtual environments through different technological tools, there are other areas that, due to their orientation towards laboratory practices, require the use of various tools and communication methods.

As future work, the need for future training courses in technological tools is proposed. This would allow teachers to be more flexible and adapted to the particular needs of each area of knowledge. Additionally, it is necessary to promote plans for the expansion and consolidation of knowledge in specific tools for teaching in virtual environments, with a focus on engineering education, from the basic to the complementary areas.

Given the current scenario regarding education (and especially engineering programs), universities must make additional efforts to ensure that online learning continues to meet the rigorous requirements of national higher education accrediting bodies such as the Inter-Institutional Committees for the Evaluation of Higher Education (CIEES) and the Engineering Teaching Accreditation Council (CACEI) of Mexico, as well as international entities such as the Accreditation Board for Engineering and Technology (ABET).

Acknowledgments

The author would like to express their gratitude to the National Council of Science and Technology (CONACYT) and the Autonomous University of the State of Hidalgo, Mexico.

CRediT author statement

Granillo-Macías, R.: conceptualization, data curation, methodology, validation, writing (original draft, review, editing.

References

Asgari, S., Trajkovic, J., Rahmani, M., Zhang, W., Lo, R. C., and Sciortino, A. (2021). An observational study of engineering online education during the COVID-19 pandemic. *PLOS ONE, 16*(4), e0250041. https://doi.org/10.1371/journal.pone.0250041

Bozkurt, A., and Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education, 15*(1), i-vi. https://doi.org/10.5281/zenodo.3778083

CEPAL-UNESCO (2020). *Informe COVID-19 CEPAL-UNESCO*. https://repositorio.cepal.org/bitstream/handle/11362/45904/1/S2000510_es.pdf

Cojocariu, V.-M., Lazar, I., Nedeff, V., and Lazar, G. (2014). SWOT anlysis of e-learning educational services from the perspective of their beneficiaries. *Procedia – Social and Behavioral Sciences, 116*, 1999-2003. https://doi.org/10.1016/j.sbspro.2014.01.510

Culzoni, C., Panigatti, C., and Bergesse, C. B. (2021). El proceso de enseñanza y aprendizaje en asignaturas de matemática para ingeniería durante el aislamiento por Covid 19. *Revista Iberoamericana de Tecnología En Educación y Educación En Tecnología, 28*, e9. https://doi.org/10.24215/18509959.28.e9

Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems, 49*(1), 5-22. https://doi.org/10.1177/0047239520934018

Eisenhardt, K. M. (1989). *Building theories from case study research*. *The Academy of Management Review, 14*(4), 532. https://doi.org/10.2307/258557

Evstatiev, B. I., and Hristova, T. V. (2020). *Adaptation of electrical engineering education to the COVID-19 situation: Method and results* [Conference presentation]. 2020 IEEE 26th International Symposium for Design and Technology in Electronic Packaging (SIITME), Pitesti, Romania. https://doi.org/10.1109/SIITME50350.2020.9292142

Fardoun, H., González, C., Collazos, C. A., and Yousef, M. (2020). Estudio exploratorio en Iberoamérica sobre procesos de enseñanza-aprendizaje y propuesta de evaluación en tiempos de pandemia. *Education in the Knowledge Society (EKS), 21*, 9. https://doi.org/10.14201/eks.23537

Ferreira, J., Behrens, M., Torres, P., and Marriott, R. (2018). The necessary knowledge for online education: Teaching and learning to produce knowledge. *EURASIA Journal of Mathematics, Science and Technology Education, 14*(6), 86463. https://doi.org/10.29333/ejmste/86463

García Martín, J., and García Martín, S. (2021). Uso de herramientas digitales para la docencia en España durante la pandemia COVID-19. *Revista Española de Educación Comparada, 38*, 151. https://doi.org/10.5944/reec.38.2021.27816

García-Alberti, M., Suárez, F., Chiyón, I., and Mosquera Feijoo, J. C. (2021). Challenges and experiences of online evaluation in courses of civil engineering during the lockdown learning due to the COVID-19 Pandemic. *Education Sciences, 11*(2), 59. https://doi.org/10.3390/educsci11020059

González-Hernández, I. J., and Granillo-Macías, R. (2020). Competencias del ingeniero industrial en la Industria 4.0. *Revista Electrónica de Investigación Educativa, 22*, 1-14. https://doi.org/10.24320/redie.2020.22.e30.2750

Hassan, M. M., Mirza, T., and Hussain, M. W. (2020). *A* critical review by teachers on the online teaching-learning during the COVID-19. *International Journal of Education and Management Engineering, 10*(6), 17-27. https://doi.org/10.5815/ijeme.2020.05.03

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., and Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior, 119*, 106713. https://doi.org/10.1016/j.chb.2021.106713

Kanetaki, Z., Stergiou, C., Bekas, G., Jacques, S., Troussas, C., Sgouropoulou, C., and Ouahabi, A. (2022a). Acquiring, analyzing and interpreting knowledge data for sustainable engineering education: An experimental study using YouTube. *Electronics, 11*(14), 2210. https://doi.org/10.3390/electronics11142210

Kanetaki, Z., Stergiou, C., Bekas, G., Troussas, C., and Sgouropoulou, C. (2022b). A hybrid machine learning model for grade prediction in online engineering education. International Journal of Engineering Pedagogy (IJEP), 12(3), 4-24. https://doi.org/10.3991/ijep.v12i3.23873

Kanetaki, Z., Stergiou, C., Troussas, C., and Sgouropoulou, C. (2023). Developing novel learning spaces through social media channels for sustainable CAD engineering education. In A. Krouska, C. Troussas, and J. Caro (Eds.), *Novel & Intelligent Digital Systems: Proceedings of the 2nd International Conference* (*NiDS 2022*) (pp. 359-371). Springer. https://doi.org/10.1007/978-3-031-17601-2_35

Kavitha, M., Jayasankar, T., Maheswara Venkatesh, P., Mani, G., Bharatiraja, C., and Twala, B. (2021). COVID-19 disease diagnosis using smart deep learning techniques. *Journal of Applied Science and Engineering, 24*(3), 271-277. https://doi.org/10.6180/jase.202106_24(3).0001

Martínez, P. J., Aguilar F. J., Ortiz, M. (2019). Transitioning from face-to-face to blended and full online learning engineering master's program. *IEEE Transactions on Education 63*, 2-9. https://doi.org/10.1109/TE.2019.2925320

Mohd Zulkefli, N. A. (2020). Evaluating e-Learning Google Classroom tools for computer science subjects during COVID-19 Pandemic. *International Journal of Advanced Trends in Computer Science and Engineering, 9*(4), 6251-6258. https://doi.org/10.30534/ijatcse/2020/304942020

Mohmmed, A. O., Khidhir, B. A., Nazeer, A., and Vijayan, V. J. (2020). Emergency remote teaching during coronavirus pandemic: The current trend and future directive at Middle East College Oman. *Innovative Infrastructure Solutions, 5*(3), 72. https://doi.org/10.1007/s41062-020-00326-7

Mok, K. H., Xiong, W., Ke, G., and Cheung, J. O. W. (2021). Impact of COVID-19 pandemic on international higher education and student mobility: Student perspectives from mainland China and Hong Kong. *International Journal of Educational Research, 105*, 101718. https://doi.org/10.1016/j.ijer.2020.101718

Ożadowicz, A. (2020). Modified blended learning in engineering higher education during the COVID-19 lockdown – Building automation courses case study. *Education Sciences, 10*(10), 292. https://doi.org/10.3390/educsci10100292

Portela, F. (2020). TechTeach – An innovative method to increase the students engagement at classrooms. *Information, 11*(10), 483. http://dx.doi.org/10.3390/info11100483

Quacquarelli Symonds (2021). *QS Top Universities*. http://www.topuniversities.com/

Ramya, D., and Poongodi, O. T. (2021). A study on the usage of Information Communication Technology tools in the teaching-learning process of engineering education. *Journal of Applied Science and Engineering, 25*(2), 321-326. https://doi.org/10.6180/jase.202204_25(2).0017

Ripoll, V., Godino-Ojer, M., and Calzada, J. (2021). Teaching chemical engineering to biotechnology students in the time of covid-19: assessment of the adaptation to digitalization. *Education for Chemical Engineers, 34*, 21-32. https://doi.org/10.1016/j.ece.2020.11.001

Sahoo, K. K., Muduli, K. K., Luhach, A. K., and Poonia, R. C. (2021). Pandemic COVID-19: An empirical analysis of impact on Indian higher education system. *Journal of Statistics and Management Systems, 24*(2), 341-355. https://doi.org/10.1080/09720510.2021.1875571

Stake, R. E. (1999). *Investigación con estudio de casos*. Ediciones Morata.

Suryaman, H., Kusnan, and Mubarok, H. (2020). Profile of online learning in building engineering education study program during the COVID-19 pandemic. *International Journal of Recent Educational Research, 1*(2), 63-77. https://doi.org/10.46245/ijorer.v1i2.42

Syauqi, K., Munadi, S., and Triyono, M. B. (2020). Students' perceptions toward vocational education on online learning during the COVID-19 pandemic. *International Journal of Evaluation and Research in Education*, *9*(4), 881. https://doi.org/10.11591/ijere.v9i4.20766

UNESCO (2020). *COVID-19 education response*. https://en.unesco.org/covid19/educationresponse/globalcoalition

Yin, R. (2003). *Case study research: Design and methods* (3rd ed.). Sage.

Research article / Educational Engineering

http://doi.org/10.15446/ing.investig.101817

Mask Detection and Categorization during the COVID-19 Pandemic Using Deep Convolutional Neural Network

Detección y categorización de máscaras durante la pandemia del COVID-19 utilizando una red neuronal convolucional profunda

Kamil Dimililer[1,3], and Devrim Kayali[2]

ABSTRACT

With COVID-19 spreading all over the world and restricting our daily lives, the use of face masks has become very important, as it is an efficient way of slowing down the spread of the virus and an important piece to continue our daily tasks until vaccination is completed. People have been fighting this disease for a long time, and they are bored with the precautions, so they act carelessly. In this case, automatic detection systems are very important to keep the situation under control. In this research, deep learning models are trained with as little input data as possible in order to obtain an accurate face mask-wearing condition classification. These classes are mask-correct, mask wrong, and no mask, which refers to proper face mask use, improper face mask use, and no mask use, respectively. DenseNets, EfficientNets, InceptionResNetV2, InceptionV3, MobileNets, NasNets, ResNets, VGG16, VGG19, and Xception are the networks used in this study. The highest accuracy was obtained by the InceptionResNetV2 and Xception networks, with 99,6%. When other performance parameters are taken into consideration, the Xception network is a step forward. VGG16 and VGG19 also show an accuracy rate over 99%, with 99,1 and 99,4%, respectively. These two networks also had higher FPS and the two lowest initialization times during implementation. A comparison with recent studies was also carried out to evaluate the obtained accuracy. It was found that a higher accuracy can be obtained with the possible minimum input size.

Keywords: COVID-19, mask detection, deep learning, classification

RESUMEN

Con el COVID-19 extendiéndose por todo el mundo y restringiendo nuestra vida diaria, el uso de mascarillas se ha vuelto muy importante, pues es una forma eficiente de frenar la propagación del virus, y una pieza importante para continuar con nuestras tareas diarias hasta que se complete la vacunación. La gente ha estado luchando contra la enfermedad durante mucho tiempo y se aburre con las precauciones, por lo que actúa con descuido. En este caso, los sistemas de detección automática son muy importantes para mantener la situación bajo control. En esta investigación se entrenan modelos de aprendizaje profundo con el mínimo de datos de entrada posibles para obtener una clasificación precisa de las condiciones de uso de las mascarillas. Estas clases son maskcorrect, maskwrong y nomask, que se refieren al uso adecuado, a un uso inadecuado y al no uso de la mascarilla facial, respectivamente. DenseNets, EfficientNets, InceptionResNetV2, InceptionV3, MobileNets, NasNets, ResNets, VGG16, VGG19 y Xception son las redes utilizadas en este estudio. La mayor precisión la obtuvieron las redes InceptionResNetV2 y Xception, con un 99,6 %. Cuando se tienen en cuenta otros parámetros de rendimiento, la red Xception un paso adelante. VGG16 y VGG19 presentan una tasa de precisión superior al 99 %, con 99,1 y 99,4 % respectivamente. Estas dos redes también presentaron FPS más altos y los dos tiempos de inicialización más bajos en la implementación. También se realizó una comparación con estudios recientes para evaluar la precisión obtenida. Se encontró que se puede obtener una mayor precisión con el mínimo tamaño de entrada posible.

Palabras clave: COVID-19, detección de mascarillas, aprendizaje profundo, clasificación

Received: March 25th 2022
Accepted: October 31st 2022

Introduction

Our lifestyles started changing with the advent of COVID-19, a very dangerous virus that quickly spread all over the world. Due to its high infection rate and severity, precautions were taken at a global level, and restrictions were put into effect. The use of face masks was one of the most important and effective ways for slowing down the infection from the very beginning of the pandemic (Cheng *et al*., 2020).

Despite the fact that vaccination continues all over the world, the use of masks is still important, since sufficient vaccination has not yet been reached and the virus can mutate and change. As we are trying to return to our daily lives while coexisting with the disease, we still have to take some precautions to avoid any rapid spread. This means that serious control measures are needed, especially in public places. Automated control is key to obtain fast, reliable, and continuously working systems. Artificial intelligence (AI) makes it possible to obtain such systems in many different areas. AI-based image processing has been applied in various cases (Dimililer, 2022; Dimililer and Kayalı, 2021;

[1]Department of Electrical and Electronic Engineering. Near East University. Nicosia, North Cyprus, Mersin 10 Turkey. E-mail: kamil.dimililer@neu.edu.tr
[2]Department of Electrical and Electronic Engineering. Near East University. Nicosia, North Cyprus, Mersin 10 Turkey. E-mail: devrim.kayali@ktemo.org
[3]Applied Artificial Intelligence Research Centre. Near East University. Nicosia, North Cyprus, Mersin 10 Turkey.

Kayali *et al*., 2022) including health sciences (Dimililer, 2017; Dimililer *et al*., 2020; Dimililer *et al*., 2016; Dimililer and Zarrouk, 2017). In the case of convolutional neural networks (CNNs), their architecture internally covers the image preparation phase.

As an example, a smart door project was studied by Baluprithviraj *et al.* (2021), which is integrated into a mobile app. The AI-based device detects whether a person is wearing a face mask and sends an alert message to the mobile app.

Pandey (2020) also worked on a system that detects whether an individual is wearing a face mask. The system raises alarm when a person without a mask is detected.

Besides mask detection, studies on automatic COVID-19 detection and survival analysis have also been carried out. Narin *et al*.. (2021) used X-ray images and deep CNNs for the automatic detection of COVID-19. Three different binary classifications were implemented to discriminate COVID-19 from normal (healthy) subjects and viral and bacterial pneumonia. For survival analysis, Atlam *et al.* (2021) used the Cox regression model and deep learning. They presented two different systems called Cox_Covid_19 and Deep_Cox_Covid_19 to predict the survival probability and the most important symptoms affecting it. Cox_Covid_19 is based on Cox regression, and the enhanced version, Deep_Cox_Covid_19, is obtained by combining an auto-encoder deep neural network and Cox regression.

These kinds of systems reduce the need for manpower and help authorities to maintain control. In light of the above, the contributions of this research can be summarized as follows:

- The detection and classification of three different mask-wearing conditions.
- A comparison regarding the training time and accuracy of the networks with their minimum input sizes and gray-scale images.
- A comparison of the resulting output model sizes and FPS of the networks, which play an important role when selecting a suitable model for an application.
- A performance comparison with recent studies in the literature that involve two- and three-class datasets.
- The obtention of high accuracy on three-class datasets, even with the lowest possible model input size.

Literature Review

Suresh *et al.* (2021) studied a MobileNet network for face mask detection in public areas. This system detects if an individual has a face mask. If said mask is not detected, the authorities are notified and their face is captured.

Chavda *et al.* (2021) proposed a CNN-based real-time face mask detector. The first stage involved face detection, and NasNetMobile, DenseNet121, and MobileNetV2 networks were trained for the second stage. According to the experimental results, the NasNetMobile network was selected as the most suitable network for real-time application in terms of both accuracy and average interference time.

Mohan *et al.* (2021) proposed a tiny CNN architecture that can be used on devices with constrained resources. The proposed architecture was compared to the SqueezeNet and the Modified SqueezeNet (a smaller version of the former). According to the results, the proposed method outperformed the SqueezeNet and the Modified SqueezeNet in both size and accuracy.

Amin *et al.* proposed a real-time deep learning-based method for crowd counting and face mask detection.

Snyder and Husari (2021) developed an automated detection model and implemented it on a mobile robot named *Thor*. The developed method consists of three components: the detection of human existence, the detection and extraction of the human face, and the classification of masked/unmasked faces. In the first component, ResNet50 is used with Feature Pyramid Network (FPN). The second component detects and extracts human faces with multi-task convolutional neural networks (MT-CNNs). The last component consists of a trained CNN classifier for the detection phase.

Khamlae *et al.* developed and trained a CNN with a dataset consisting of 848 images with a 416x416 resolution. An accuracy rate of 81% was obtained with the trained model, which was then implemented at the front gate of a campus building.

Kodali and Dhanekula (2021) proposed a deep learning-based network for face mask detection. This method consists of two parts. The first part, the pre-processing step, converts the RGB image to a grayscale one and resizes and normalizes it. Then, the proposed CNN classifies the face images with and without face masks. The validation accuracy of the proposed model is 96%.

Pinki and Garg (2020) developed a deep-learning model for a face mask detection system by fine-tuning MobileNetV2. The developed model can be integrated into systems in public areas to identify whether the people are wearing masks. If a person is not wearing a mask, a notification is sent to the relevant authorities.

Sen and Patidar (2020) developed a deep learning-based method using MobileNetV2 and building four fully connected layers on top of it. The developed model detects people with or without face masks from images and video streams with an accuracy of 79,24%.

Yadav (2020) proposed a system that uses both deep learning and geometric techniques for detection, tracking, and validation purposes. The system performs real-time monitoring to detect face masks and social distancing between people.

Shete and Milosavljvec (2020) also studied a system that detects both face masks and social distancing, using DSF, YOLOv3, and ResNet classifier pre-trained models. This system provides the number of non-violating and violating people as an output with a confidence score of 100%.

Srinivasan *et al.* (2021) proposed a system that uses Dual Shot Face Detector (DSFD), Density-based spatial clustering of applications with noise (DBSCAN), YOLOv3, and MobileNetV2 for an effective solution regarding face

mask and social distancing detection. According to an evaluation of its performance, the system obtained accuracy and F1 scores of 91,2 and 90,79%, respectively.

Meivel *et al.* (2021) proposed a real-time social distancing measurement and face mask detection system using Matlab. RCNN, Fast RCNN, and Faster RCNN algorithms were chosen to this effect.

Wang *et al.* (2021) proposed a face mask detection solution dubbed *Web-based efficient AI recognition of masks* (WearMask). It is an in-browser, serverless, edge-computing-based application that requires no software installation and can be used by any device with an internet connection and access to web browsers. This framework integrates deep learning (YOLO), a high-performance neural network inference computing framework (NCNN), and a stack-based virtual machine (WebAssembly).

Dey *et al.* (*2021*) proposed a multi-phase deep learning-based model for face mask detection in images and video streams, which was named *MobileNet Mask*. According to their experimental results, MobileNet Mask achieved about 93% accuracy with 770 validation samples, and about 100% with 276 validation samples.

Naufal *et al.* (2021) conducted a comparative study on face mask detection using support vector machines (SVM), k-nearest neighbors (KNN), and deep CNNs (DCNN). Although CNN required a longer execution time compared to KNN and SVM, it reported the best average performance, with an accuracy of 96,83%.

Vijitkunsawat and Chantngarm (2020) also conducted a comparative study on the performance of SVM, KNN, and MobileNet for real-time face mask detection. According to the results, MobileNet has the best accuracy with regard to both images and real-time video inputs.

Singh *et al.* (2021) studied and implemented two different real-time face mask detection models. One of them involved YOLOv4, and the second one used MobileNetV2 and a CNN. A comparison of the experimental results showed that MobileNetV2 and CNN yielded better results, with an accuracy of 98%, in comaprison with the accuracy of YOLOv4 (88,92%).

Sivakumar *et al.* made a deep learning based study on face mask detection by using Viola-Jones algorithm and the Convolution Neural Networks (CNN).

Nagrath *et al.* (2021) proposed SSDMNV2, an approach based on OpenCV and deep learning. In this method, a Single Shot Multibox Detector is used as a face detector, and the MobilenetV2 network is used as a classifier. The accuracy and F1 score of the proposed system are 92,64 and 93%, respectively.

Sanjaya and Rakhmawan (2020) also used MobileNetV2 to develop a face mask detection algorithm in order to fight COVID-19. Their model can discriminate people with and without a face mask with an accuracy of 96,85%.

Sakshi *et al.* (2021) also used MobileNetV2 to implement a face mask detection model to be used both on static pictures and real-time videos.

Venkateswarlu *et al.* (2020) used MobileNet and a global pooling block to develop a face mask detection model. The global pooling layer flattens the feature vector, and a fully connected dense layer and the softmax layer utilize classification. The system was evaluated with different datasets, and the obtained accuracies were 99 and 100%.

Rudraraju *et al.* (2020) proposed a face mask detection system to control various entrances of a facility using fog computing. Fog nodes are used for processing the video streams of the entrances. Haar-cascade-classifiers are used to detect faces in the frames, and each fog node consists of two MobileNet models. The first model classifies whether the person is wearing a mask, and, if this model gives the output that the person is wearing a face mask, then the second model classifies whether the mask-wearing condition is proper. As a result of this two-level classifier, the person is allowed to enter the facility if they are properly wearing a mask. The accuracy of this system is about 90%.

Hussain *et al.* (2021) proposed a Smart Screening and Disinfection Walkthrough Gate (SSDWG), an IoT-based control and monitoring system. This system performs rapid screening, includes temperature measuring, and stores the record of suspected individuals. The screening system uses a deep learning model for face mask detection and the classification of people who are wearing face mask their properly, improperly, or wear no mask. A transfer learning approach was used on CNN, ResNet-50, Inception v3, MobileNetV2, and VGG-16. Results showed that VGG-16 achieved the highest accuracy, with 99,81%, followed by the MobileNetV2 with 99,6%.

Adusumalli *et al.* (2021) studied a system that detects whether people are wearing face masks. The system is based on OpenCV and MobileNetV2. It also sends message to the person if they are not wearing a mask, and their face is stored in the database.

Das *et al.* (2020) used OpenCV and CNNs in their research to develop a face mask detection system. They used two different datasets, and the obtained accuracies were 95,77 and 94,58%.

Aydemir *et al.* (2022) used pre-trained DenseNet201 and ResNet101 for feature extraction. Then, they used an improved RelieF selector to choose features in order to train an SVM classifier. They had three cases in their research: mask *vs.* no mask *vs.* improper mask; mask *vs.* no mask and improper mask; and mask *vs.* no mask. They obtained 95,95, 97,49, and 100% accuracy for these cases, respectively.

Wu *et al.* (2022) proposed an automatic facemask recognition and detection framework called FMD-Yolo. Im-Res2Net-101 was used to extract features, which is a modified Res2Net structure. Then, an aggregation network En-PAN was used to merge low-level details and high-level semantic information. According to their experimental results, at the intersection over union (IoU) level 0,5, FMD-Yolo had average precisions of 92 and 88,4% for two different datasets.

Mar-Cupido *et al.* (2022) conducted a study on classifying face masks used by people according to their type. They used the pre-trained models ResNet101v2, ResNet152v2, MobileNetv2, and NasNetMobile to classify the classes KN95, N95, cloth, surgical, and no mask. According to their results, ResNet101v2 and ResNet152v2 both had higher accuracies (98 and 97,37%, respectively) in comparison

with NasNetMobile and MobileNetv2, both with 93,24% accuracy.

Agarwal *et al.* (2022) proposed a hybrid model with CNNs and SVMs. They used the medical mask dataset (MDD) and the real-world masked face recognition dataset (RMFD) for training and evaluating the proposed model. The obtained accuracy was 99,11%.

Crespo *et al.* (2022) carried out a study on both two- and three-class face mask-wearing conditions using deep learning while focusing on ResNets. The results showed that ResNet-18 had the best accuracies for both two and three classes, with 99,6 and 99,2% respectively.

Jayaswal and Dixit (2022) proposed a framework that uses a Single Shot Multibox Detector and Inception V3 (SSDIV3). They also proposed two versions of synthesized face mask datasets and compared SSDIV3 against VGG16, VGG19, Xception, and MobileNet. Their experimental results showed the higher accuracy of their proposed system, with 98%.

Singh *et al.* (2022) proposed a framework that uses cloud computing, fog computing, and deep neural networks. It is a smart and scalable system that can detect mask wearing and distancing violations. When the proposed framework was tested on video, it achieved a 98% accuracy.

Materials and methods

The goal of this study is to obtain an accurate face mask detection model that classifies three mask-wearing conditions with as little data as possible. This work selected models included in Keras, namely DenseNets, EfficientNets, InceptionResNetV2, InceptionV3, NasNets, MobileNets, ResNets, VGG16, VGG19, and Xception. These networks have different architectures, and some of them come in different versions that vary in parameters and depth. Moreover, these networks have differences in their default input shape, minimum input shape, and input data scale requirements. To train every network with minimal data, input shapes, and input data scale, requirements are examined and pre-processing is performed in order to obtain the necessary shapes.

Dataset

The dataset used in this research was created by the authors based on the Labeled Faces in the Wild (LFW) dataset (Huang et al., 2007). OpenCV was used for face and eyes detection in the LFW dataset. Once the faces were extracted from the images, face masks were added to the faces using eyes as reference points. To obtain the three classes needed, the first condition was to place the face mask properly covering the mouth and the nose of the person. In the second condition, the face mask was placed at a lower position, aiming to simulate improper face mask-wearing. For the last condition, the images were left untouched (no mask class). Figure 1 shows examples of the classes in the dataset.

Figure 1. Dataset Classes. Rows represent the maskcorrect, maskwrong, and nomask classes respectively.
Source: Authors

Pre-processing

The pre-processing stage was applied differently for each network depending on their properties. Each of the selected networks has a default input shape, and some of them also need the input data to be scaled between specific values for training. Since the goal was to train the network with as little data as possible, the minimum input shapes of each network were examined as shown in Table 1. The second most important thing was the rescaling of the input data in order to meet the network requirements. To this effect, each network's input scale requirements were found (Table 2). Finally, the image was converted to grayscale in order to train with only one channel. Although these network inputs are generally three-channel images, grayscale images were used to reduce the input data. To summarize: first, the input image was converted to grayscale; then, the input image was resized to the minimum input size of the network; and the input data were rescaled according to the network's input scale requirements. For networks with no input scale requirements, the input data were scaled between 0 and 1.

Network models

As mentioned at the beginning of this section, the networks selected for this study are DenseNets, EfficientNets, InceptionResNetV2, InceptionV3, MobileNets, NasNets, ResNets, VGG16, VGG19, and Xception.

Regarding the DenseNets, the models used in this research were DenseNet121, DenseNet169, and DenseNet201. The default input shape of these networks is (224, 224, 3), which means the expected input, is a three-channel image of size 224x224. These networks can be trained with smaller images, but the size should not be smaller than 32x32. The EfficientNet models used in this research were EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7. The inputs expected by EfficientNet networks are the float pixel values of the images between 0 and 255. The minimum image size supported by the EfficientNets is also 32x32.

The InceptionResNetV2 model has some differences when compared to EfficientNets and DenseNets. This network expects a default input shape of (299, 299, 3), and the input data need to be scaled between -1 and 1. 75x75 is the minimum image size accepted by this network.

InceptionV3 network has a default input shape of (299, 299, 3), and the input data scale requirement is between -1 and 1, like InceptionResNetV2. InceptionV3 has a minimum input image size of 75x75.

Table 1. Minimum input shapes of the networks

Network	Minimum input shape
DenseNet 121-169-201	32x32
EfficientNet B0-B7	32x32
InceptionResNetV2	75x75
InceptionV3	75x75
MobileNet - MobileNetV2	32x32
NasNetMobile - NasNetLarge	32x32
ResNet 50-101-152	32x32
ResNet 50V2-101V2-152V2	32x32
VGG16	32x32
VGG19	32x32
Xception	71x71

Source: Authors

The MobileNet and MobileNetV2 networks have a default input shape of (224, 224, 3), and the minimum supported image size is 32x32, like DenseNets. However, these networks require the input data to be scaled between -1 and 1.

NasNets includes NasNetLarge and NasNetMobile, whose input requirements are just like DenseNets, with a default input shape of (224, 224, 3) and a minimum image size of 32x32. As their names explain, NasNetLarge is a larger network, with more parameters to train, and NasNetMobile is the lightweight version of the network, with fewer parameters to train that can however be implemented on more devices.

ResNet networks can be grouped into two: ResNets and ResNetV2s. ResNets have a default input shape of (224, 224, 3), and the minimum input size is 32x32. ResNetV2s also have the same properties, but their input scale requirement is between -1 and 1. This research used ResNet50, ResNet101, ResNet152, ResNet50V2, ResNet101V2, and ResNet152V2.

VGG16 and VGG19 are also networks that have a minimum input size of 32x32 and a default input shape of (224, 224, 3). The last network used was Xception, which has a minimum input size of 71x71 and a default input shape of (299, 299, 3). This network also requires the input data to be scaled between -1 and 1.

All of the networks described above have the option of loading pre-trained weights from a path or from ImageNet. However, in this research, pre-trained weights were not used, which means that the weights were randomly initialized.

Training

For the training process, training parameters were selected after several experiments and used for all the networks. A value of patience parameter was determined in order to track the learning process of the networks, aiming to make sure that the learning process continued as long as the model continued to improve. To this effect, the patience value was set at 100 epochs by experimenting, and validation loss was tracked during the whole training process of each network. This means that each network continued with the training process as long as the validation loss of the network kept

Table 2. Input scale requirements of the networks. (-) means that there is no requirement

Network	Input Scale Requirement
DenseNet 121-169-201	-
EfficientNet B0-B7	between 0 and 255
InceptionResNetV2	between -1 and 1
InceptionV3	between -1 and 1
MobileNet - MobileNetV2	between -1 and 1
NasNetMobile - NasNetLarge	-
ResNet 50-101-152	-
ResNet 50V2-101V2-152V2	between -1 and 1
VGG16	-
VGG19	-
Xception	between -1 and 1

Source: Authors

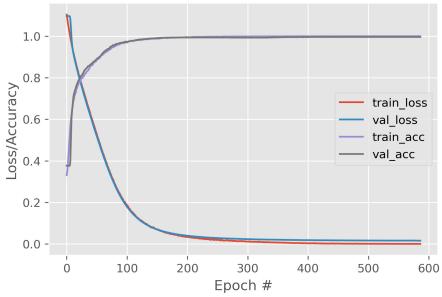


Figure 2. Xception training graph
Source: Authors

decreasing. In other words, the network continued to learn. However, if the network stopped learning and the validation loss did not decrease for the selected patience value of 100 epochs, the training process would automatically stop, and the best weights of the training process would be restored before saving the model. As an example, Xception's training graph is shown in Figure 3. For comparison, the training of all networks was carried out with the same computer, which had 32GB DDR4 system memory, a 10th Gen. Intel Core i7-10750H processor, and an NVIDIA GeForce RTX 2070 SUPER graphics card with 8GB GDDR6 video memory.

Results and discussion

After the experimental results were obtained, networks with an accuracy above 70% were selected for comparison. ResNet50, ResNet101, and ResNet152, as well as EfficientNets, were not able to obtain a good accuracy with 32x32 grayscale images, so they were not included in the comparison.

After obtaining results with the original dataset for all of the networks, different image preprocessing techniques were evaluated for further improvements to the networks. Contrast stretching was applied to obtain low- and high-contrast images. Furthermore, Discrete Cosine Transform (DCT) was applied, which has improved accuracy by compressing the image in some of the studies. However, in our case, none of these approaches was able to improve the accuracy of the networks. Since grayscale images with a resolution as low as 32x32 were being used, the task became harder for the networks, even with the original images.

Although image preprocessing affects ANN applications in a good way, our experiments showed that it can negatively affect the results for small input data, which is 1024 (32x32x1) in our research.

Table 3 shows the overall performances of the networks included in the comparison.

By comparing the training times of the networks, it can be seen that VGG16 required the shortest time, with 105 minutes. On the contrary, ResNet152V2 reported the highest value, with 607 minutes. The training times of Xception and MobileNet were also very close to that of VGG16, with 108 and 112 minutes, respectively.

While evaluating the networks on the test dataset, their evaluation were also recorded. An 80-20% train-test ratio was used for all of the networks, and the test dataset included 1 579 images. VGG16 reported the best evaluation time, with 2,52 seconds, and VGG19 was the second best (2,68 seconds). Considering that the image size for Xception was 71x71, its evaluation time of 4,05 seconds was also a good result. The highest evaluation time was 14,2 seconds, obtained by the InceptionResNetV2 with a 75x75 test image size.

The size of the network is also an important parameter for implementation. It may not be always possible to implement a model with the highest accuracy on every device because of the size of the network and its memory requirements. By comparing the output model sizes of all networks, it was observed that the MobileNetV2 network has the smallest size, with 29,6 MB. The MobileNet network has the second smallest size with 38,9 MB, and ResNet152V2 and InceptionResNetV2 have the two largest model sizes (679 and 638 MB, respectively).

The *system initialization time* refers to the time in seconds required by the face mask detection system for processing during implementation. When the trained models were implemented, the lowest initialization time was obtained by VGG16, with 5,5 seconds, followed by VGG19 followed with 6,1 seconds. The InceptionResNetV2 network reported the highest initialization time (104,2 seconds). In addition, the FPS of the system was also monitored during implementation. The FPS of the system includes both face detection in the captured video and the classification of the detected faces by the network. MobileNet, VGG16, and VGG19 had the highest FPS, with 14. MobileNetv2 and ResNet50V2 followed them with 13 FPS, and InceptionResNetV2 had the lowest value with 8.

Regarding the accuracies of the networks, InceptionResNetv2 and Xception had the highest value, with 99,6%. VGG19 and VGG16 also showed an accuracy over 99 (with 99,4% and 99,1%, respectively). As for the general performance of the networks, if the highest possible accuracy is desired, the Xception network could be selected, as it has moderate values regarding model size, initialization time, and FPS. If a little lower accuracy can be tolerated, albeit still over 99%, VGG16 or VGG19 are also good options, with higher FPS values and the two lowest initialization times. However, they have slightly higher model sizes, which is also a drawback. Figures 3, 4, and 5 are the confusion matrices of VGG16, VGG19, and Xception.

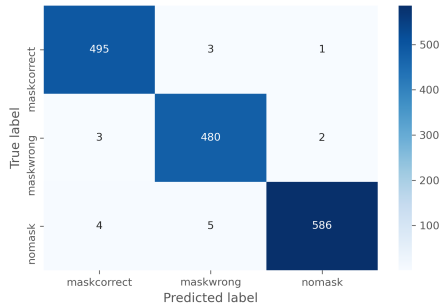


Figure 3. VGG16 confusion matrix
Source: Authors

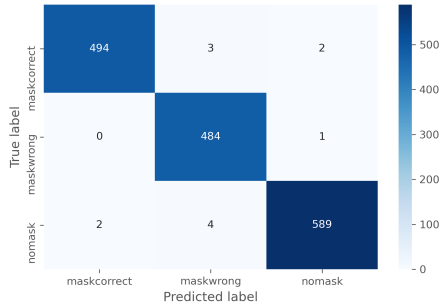


Figure 4. VGG19 confusion matrix
Source: Authors

In recent studies on face mask detection, the maximum accuracy obtained from the experimental results has been also used for comparison.

Agarwal *et al.* (2022) used two different datasets with two classes. They obtained a 99,11% accuracy with a hybrid model with CNNs and SVMs.

Goyal *et al.* (2022) used a two-class custom dataset and a customized CNN architecture with ResNet-10 based on a Single Shot Detector. They obtained a 98% accuracy in their research.

In a study carried out by by Waleed *et al.* (2022) a two-class dataset was also used. With a customized deep CNN model, they obtained a 99,57% accuracy.

Aydemir *et al.* (2022) obtained results in three different cases. The first two cases involved two classes, and the third one used three. The resulting accuracies were 97,49, 100, and 95,95%, respectively.

Guo *et al.* (2022) used a three-class dataset and an improved YOLOv5 in their study, obtaining a mean Average Precision (mAP) of 96,7%.

Han *et al.* (2022) used a customized three-class dataset and a network called *SMD-YOLO*, which is YOLOv4-tiny based. They obtained a mAP of 67,01%.

Eyiokur *et al.* (2021) studied three-class face mask detection and obtained a higher accuracy with an Inception-v3 network (*i.e.*,98,2%).

Table 4 shows the comparison between these studies and our research. It should also be considered that the input sizes of the networks were lower in this research, *i.e.*, the minimum sizes accepted by the networks.

Table 3. Network performance results

Network (Input Size)	Epochs Run	Training time(mins)	Evaluation Time for 1579 test images(s)	Model Size (MB)	System Initialization(s)	FPS	Accuracy
DenseNet121(32x32)	1 748	226	6,13	89,6	60,9	11	0,9690
DenseNet169(32x32)	1 554	267	6,47	157	77,2	10	0,9650
DenseNet201(32x32)	1 681	334	8,05	224	95,4	9	0,9630
InceptionResNetV2(75x75)	682	205	14,2	638	104,2	8	0,9960
InceptionV3(75x75)	1 810	241	6,95	256	46	11	0,9760
MobileNet(32x32)	3 417	112	3,01	38,9	14,9	14	0,8230
MobileNetV2(32x32)	4 118	198	3,69	29,6	26,7	13	0,7300
ResNet50V2(32x32)	1 445	122	4,55	273	27,6	13	0,8980
ResNet101V2(32x32)	1 808	315	6,73	495	49,3	11	0,8820
ResNet152V2(32x32)	2 173	607	6,83	679	74,2	10	0,8800
VGG16(32x32)	2 078	105	2,52	385	5,5	14	0,9910
VGG19(32x32)	2 412	150	2,68	445	6,1	14	0,9940
Xception(71x71)	587	108	4,05	241	24,4	12	0,9960

Source: Authors

Table 4. Comparison with other studies

Reference	Classes	Performance
(Agarwal *et al.*, 2022)	2	99,11% Accuracy
(Goyal et al., 2022)	2	98% Accuracy
(Waleed *et al.*, 2022)	2	99,57% Accuracy
(Aydemir *et al.*, 2022)	2 (Case 1)	97,49% Accuracy
	2 (Case 2)	100% Accuracy
	3 (Case 3)	95,95% Accuracy
(Guo *et al.*, 2022)	3	96,7% mAP
(Han *et al.*, 2022)	3	67,01% mAP
(Eyiokur *et al.*, 2021)	3	98,20% Accuracy
This study	3	99,6% Accuracy

Source: Authors

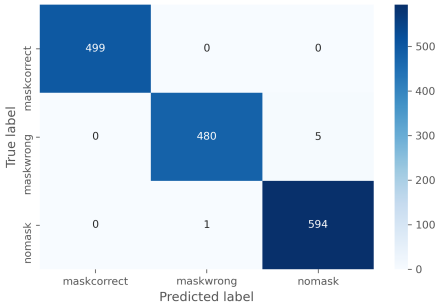


Figure 5. Xception confusion matrix
Source: Authors

Figure 6. Application output
Source: Authors

Conclusion

In this research, deep convolutional neural networks were trained with as little input data as possible to obtain an accurate model for face mask detection and wearing condition classification. A dataset was created by using the Labeled Faces in the Wild (LFW) dataset and adding face masks to it in order to simulate correct, wrong, and no mask wearing conditions. According to our experimental results, EfficientNets and ResNet50, ResNet101, and Resnet152 networks were not able to learn with their minimum input size of 32x32 and grayscale images. Four of the trained networks were able to obtain an accuracy of over 99%, *i.e.*, InceptionResNetv2, Xception, VGG16, and VGG19, with accuracies of 99,6, 99,6, 99,1, and 99,4% respectively. VGG16 and VGG19 also had better FPS when compared to other networks.

The results obtained in this study showed a higher performance for three classes while using fewer input dimensions. This, in comparison with other studies in the literature.

Figure 6 shows the output of our application for three different mask wearing situations. Boundary boxes are placed on detected faces, and the classification results are written with their confidence scores.

For future works, an image pre-processing phase will be added to the process, and a custom architecture will be developed to be tested on various different datasets.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation, writing (original draft, writing, review, and editing), data curation.

References

Adusumalli, H., Kalyani, D., Sri, R. K., Pratapteja, M., and Rao, P. P. (2021). Face mask detection using Opencv. In IEEE (Eds/), *2021 Third International Conference on Intelligent Communication Technologies and Virtual*

Mobile Networks (ICICV) (pp. 1304-1309). IEEE.https://doi.org/10.1109/ICICV50876.2021.9388375

Agarwal, C., Kaur, I., and Yadav, S. (2022). Hybrid CNN-SVM Model for Face Mask Detector to Protect from COVID-19. In M. Gupta, S. Ghatak, A. Gupta, and A. L. Mukherjee (Eds.), *Artificial Intelligence on Medical Data: Proceedings of International Symposium, ISCMM* 2021 (pp. 419-426). Springer. https://doi.org/10.1007/978-981-19-0151-5_35

Amin, P. N., Moghe, S. S., Prabhakar, S. N., and Nehete, C. M. (*2021*). *Deep learning-based face mask detection and crowd counting. In IEEE* (*Eds.*), *2021 6th International Conference for Convergence in Technology* (*I2CT*) (pp. 1-5). IEEE. https://doi.org/10.1109/I2CT51068.2021.9417826

Atlam, M., Torkey, H., El-Fishawy, N., and Salem, H. (2021). Coronavirus disease 2019 (COVID-19): Survival analysis using deep learning and Cox regression model. *Pattern Analysis and Applications, 24*, 993-1005. https://doi.org/10.1007/s10044-021-00958-0

Aydemir, E., Yalcinkaya, M. A., Barua, P. D., Baygin, M., Faust, O., Dogan, S., Chakraborty, S., Tuncer, T., Acharya, U. R. (2022). Hybrid deep feature generation for appropriate face mask use detection. *International Journal of Environmental Research and Public Health*, 19(4), 1939. https://doi.org/10.3390/ijerph19041939

Baluprithviraj, K. N., Bharathi, K. R., Chendhuran, S., and Lokeshwaran, P. (2021). Artificial intelligence based smart door with face mask detection. In IEEE (Eds.), *2021 International Conference on Artificial Intelligence and Smart Systems* (*ICAIS*) (pp. 543-548). IEEE. https://doi.org/10.1109/ICAIS50930.2021.9395807

Chavda, A., Dsouza, J., Badgujar, S., and Damani, A. (2021). Multi-stage CNN architecture for face mask detection. In IEEE (Eds.), *2021 6th International Conference for Convergence in Technology* (*i2ct*) (pp. 1-8). IEEE. https://doi.org/10.1109/I2CT51068.2021.9418207

Cheng, V. C. C., Wong, S. C., Chuang, V. W. M., So, S. Y. C., Chen, J. H. K., Sridhar, S., Kai-Wang To, K., Fuk-Woo C., Fan-Ngai, I. H., Pak-Leung, H., and Yuen, K. Y. (2020). The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *Journal of Infection, 81*(1), 107-114. https://doi.org/10.1016/j.jinf.2020.04.024

Crespo, F., Crespo, A., Sierra-Martínez, L. M., Peluffo-Ordóñez, D. H., and Morocho-Cayamcela, M. E. (2022). A computer vision model to identify the incorrect use of face masks for COVID-19 awareness. *Applied Sciences, 12*(14), 6924. https://doi.org/10.3390/app12146924

Das, A., Ansari, M. W., and Basak, R. (2020). Covid-19 face mask detection using TensorFlow, Keras and OpenCV. In IEEE (Eds.), *2020 IEEE 17th India Council International Conference* (*INDICON*) (pp. 1-5). IEEE. https://doi.org/10.1109/INDICON49873.2020.9342585

Dey, S. K., Howlader, A., and Deb, C. (2020). MobileNet mask: A multi-phase face mask detection model to prevent person-to-person transmission of SARS-CoV-2. In M. S. Kaiser, A. Bandyopadhyay, M. Mahmud and K. Ray (Eds.), *Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020* (pp. 603-613). Springer. https://doi.org/10.1007/978-981-33-4673-4_49

Dimililer, K. (2017). IBFDS: Intelligent bone fracture detection system. *Procedia Computer Science, 120*, 260-267. https://doi.org/10.1016/j.procs.2017.11.237

Dimililer, K. (2022). DCT-based medical image compression using machine learning. *Signal, Image and Video Processing, 16*(1), 55-62. https://doi.org/10.1007/s11760-021-01951-0

Dimililer, K., Dindar, H., and Al-Turjman, F. (2021). Deep learning, machine learning and Internet of Things in geophysical engineering applications: An overview. *Microprocessors and Microsystems, 80*, 103613. https://doi.org/10.1016/j.micpro.2020.103613

Dimililer, K., Ever, Y. K., and Ugur, B. (2016). ILTDS: Intelligent lung tumor detection system on CT images. In J. Corchado-Rodríguez, S. Mitra, S. Thampi, and E. S. El-Alfy (Eds.), *Intelligent Systems Technologies and Applications 2016* (pp. 225-235). Springer. https://doi.org/10.1007/978-3-319-47952-1_17

Dimililer, K., and Kayalı, D. (2021). Image enhancement in healthcare applications: A review. In F. Al-Turjman (Ed.), *Artificial Intelligence and Machine Learning for COVID-19*, (pp. 111-140). Springer. https://doi.org/10.1007/978-3-030-60188-1_6

Dimililer, K., and Zarrouk, S. (2017). ICSPI: Intelligent classification system of pest insects based on image processing and neural arbitration. *Applied Engineering in Agriculture, 33*(4), 453. https://doi.org/10.13031/aea.12161

Eyiokur, F. I., Ekenel, H. K., and Waibel, A. (2022). Unconstrained face mask and face-hand interaction datasets: building a computer vision system to help prevent the transmission of COVID-19. *Signal, Image and Video Processing, 2022*, 1-8. https://doi.org/10.1007/s11760-022-02308-x

Goyal, H., Sidana, K., Singh, C., Jain, A., and Jindal, S. (2022). A real time face mask detection system using convolutional neural network. *Multimedia Tools and Applications, 81*(11), 14999-15015. https://doi.org/10.1007/s11042-022-12166-x

Guo, S., Li, L., Guo, T., Cao, Y., and Li, Y. (2022). Research on mask-wearing detection algorithm based on improved YOLOv5. *Sensors, 22*(13), 4933. https://doi.org/10.3390/s22134933

Han, Z., Huang, H., Fan, Q., Li, Y., Li, Y., and Chen, X. (2022). SMD-YOLO: An efficient and lightweight detection method for mask wearing status during the COVID-19 pandemic. *Computer Methods and Programs in Biomedicine, 221*, 106888. https://doi.org/10.1016/j.cmpb.2022.106888

Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E. (2007). *Labeled faces in the wild: A database for studying face recognition in unconstrained environments* (*tech. rep. No. 07-49*). University of Massachusetts.

Hussain, S., Yu, Y., Ayoub, M., Khan, A., Rehman, R., Wahid, J. A., and Hou, W. (2021). IoT and deep learning-based approach for rapid screening and face mask detection for infection spread control of COVID-19. *Applied Sciences, 11*(8), 3495. https://doi.org/10.3390/app11083495

Jayaswal, R., and Dixit, M. (2023). AI-based face mask detection system: a straightforward proposition to fight with Covid-19 situation. *Multimedia Tools and Applications, 82*, 13241-13273. https://doi.org/10.1007/s11042-022-13697-z

Kayali, D., Olawale, P., Kirsal-Ever, Y., and Dimililer, K. (2022). The effect of compressor-decompressor networks with different image sizes on mask detection using Convolutional Neural Networks-VGG-16. In IEEE (Eds.), 2022 *Innovations in Intelligent Systems and Applications Conference* (*ASYU*) (pp. 1-5). IEEE. https://doi.org/10.1109/ASYU56188.2022.9925317

Khamlae, P., Sookhanaphibarn, K., and Choensawat, W. (2021). An application of deep-learning techniques to face mask detection during the COVID-19 pandemic. In IEEE (Eds.), *2021 IEEE 3rd Global Conference on Life Sciences and Technologies* (*LifeTech*) (pp. 298-299). IEEE. https://doi.org/10.1109/LifeTech52111.2021.9391922

Kodali, R. K., and Dhanekula, R. (2021). Face mask detection using deep learning. In IEEE (Eds.), *2021 International Conference on Computer Communication and Informatics* (*ICCCI*) (pp. 1-5). IEEE. https://doi.org/10.1109/ICCCI50826.2021.9402670

Mar-Cupido, R., García, V., Rivera, G., and Sánchez, J. S. (2022). Deep transfer learning for the recognition of types of face masks as a core measure to prevent the transmission of COVID-19. *Applied Soft Computing, 125*, 109207. https://doi.org/10.1016/j.asoc.2022.109207

Mohan, P., Paul, A. J., and Chirania, A. (2021). A tiny CNN architecture for medical face mask detection for resource-constrained endpoints. In S. Mekhilef, M. Favorskaya, R. K. Pandey, and R. N. Shaw (Eds.), *Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2021* (pp. 657-670). Springer. https://doi.org/10.1007/978-981-16-0749-3_52

Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., and Hemanth, J. (2021). SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable Cities and Society, 66*, 102692. https://doi.org/10.1016/j.scs.2020.102692

Narin, A., Kaya, C., and Pamuk, Z. (2021). Automatic detection of coronavirus disease (covid-19) using X-ray images and deep convolutional neural networks. *Pattern Analysis and Applications, 24*, 1207-1220. https://doi.org/10.1007/s10044-021-00984-y

Naufal, M. F., Kusuma, S. F., Prayuska, Z. A., Yoshua, A. A., Lauwoto, Y. A., Dinata, N. S., and Sugiarto, D. (2021). Comparative analysis of image classification algorithms for face mask detection. *Journal of Information Systems Engineering and Business Intelligence, 7*(1), 56-66. https://doi.org/10.20473/jisebi.7.1.56-66

Pandey, V. (2020). Artificial intelligence based face mask detection system. *International Journal of Innovative Science and Research Technology, 5*(8), 467-470. https://doi.org/10.38124/IJISRT20AUG410

Pinki and Garg, S. (2020) Face mask detection system using deep learning. *International Journal for Modern Trends in Science and Technology, 6*(12), 161-164. https://doi.org/10.46501/IJMTST061231

Rudraraju, S. R., Suryadevara, N. K., and Negi, A. (2020). Face mask detection at the fog computing gateway. In IEEE (Eds.), 2020 *15th Conference on Computer Science and Information Systems* (*FedCSIS*) (pp. 521-524). IEEE. https://doi.org/10.15439/2020F143

Sakshi, S., Gupta, A. K., Yadav, S. S., and Kumar, U. (2021). Face mask detection system using CNN. In IEEE (Eds.), *2021 International Conference on Advance Computing and Innovative Technologies in Engineering* (*ICACITE*) (pp. 212-216). IEEE. https://doi.org/10.1109/ICACITE51222.2021.9404731

Sanjaya, S. A., and Rakhmawan, S. A. (2020). Face mask detection using MobileNetV2 in the era of COVID-19 pandemic. In IEEE (Eds.), *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy* (*ICDABI*) (pp. 1-5). IEEE. https://doi.org/10.1109/ICDABI51230.2020.9325631

Sen, S., and Patidar, H. (2020). Face mask detection system for COVID 19 pandemic precautions using deep learning method. *International Journal of Emerging Technologies and Innovative Research, 7*(10), 16-21. https://www.jetir.org/view?paper=JETIR2010003

Shete, I. (2020). *Social distancing and face mask detection using deep learning and computer vision* [Doctoral dissertation, National College of Ireland]. https://norma.ncirl.ie/id/eprint/4419

Singh, A., Jindal, V., Sandhu, R., and Chang, V. (2022). A scalable framework for smart COVID surveillance in the workplace using deep neural networks and cloud computing. *Expert Systems, 39*(3), e12704. https://doi.org/10.1111/exsy.12704

Nandhis, S., Amarthya, R., Gokul, D., and Jacob, M. S. (2021). Realtime face mask detection using machine learning. In IEEE (Eds.), 2021 *International Conference on System, Computation, Automation and Networking* (*ICSCAN*) (pp. 1-4). IEEE. https://doi.org/10.1109/ICSCAN53069.2021.9526418

SivaKumar, M., Saranprasath, N., Sridharan, N. S., and Praveen, V. S. (2021, May). Comparative analysis of CNN and Viola-Jones for face mask detection. *Journal of Physics: Conference Series, 1916*(1), 012043. https://doi.org/10.1088/1742-6596/1916/1/012043

Snyder, S. E., and Husari, G. (2021). Thor: A deep learning approach for face mask detection to prevent the COVID-19 pandemic. In IEEE (Eds.), *SoutheastCon 2021* (pp. 1-8). IEEE. https://doi.org/10.1109/SoutheastCon45413.2021.9401874

Srinivasan, S., Singh, R. R., Biradar, R. R., and Revathi, S. A. (2021). COVID-19 monitoring system using social distancing and face mask detection on surveillance video datasets. In IEEE (Eds.), *2021 International Conference on Emerging Smart Computing and Informatics* (*ESCI*) (pp. 449-455). IEEE. https://doi.org/10.1109/ESCI50559.2021.9396783

Suresh, K., Palangappa, M. B., and Bhuvan, S. (2021). Face mask detection by using optimistic convolutional neural network. In IEEE (Eds.), *2021 6th International Conference on Inventive Computation Technologies* (*ICICT*) (pp. 1084-1089). IEEE. https://doi.org/10.1109/ICICT50816.2021.9358653

Venkateswarlu, I. B., Kakarla, J., and Prakash, S. (2020). Face mask detection using mobilenet and global pooling block. In IEEE (Eds.), *2020 IEEE 4th Conference on Information and Communication Technology* (*CICT*) (pp. 1-5). IEEE. https://doi.org/10.1109/CICT51604.2020.9312083

Vijitkunsawat, W., and Chantngarm, P. (2020). Study of the performance of machine learning algorithms for face mask detection. In IEEE (Eds.) *2020-5th International Conference on Information technology* (*InCIT*) (pp. 39-43). IEEE. https://doi.org/10.1109/InCIT50588.2020.9310963

Waleed, J., Abbas, T., and Hasan, T. M. (2022). Facemask wearing detection based on deep CNN to control COVID-19 transmission. In IEEE (Eds.), *2022 Muthanna International Conference on Engineering Science and Technology* (*MICEST*) (pp. 158-161). IEEE. https://doi.org/10.1109/MICEST54286.2022.9790197

Wang, Z., Wang, P., Louis, P. C., Wheless, L. E., and Huo, Y. (2021). Wearmask: Fast in-browser face mask detection with serverless edge computing for covid-19. *arXiv preprint*. https://doi.org/10.48550/arXiv.2101.00784

Wu, P., Li, H., Zeng, N., and Li, F. (2022). FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public. *Image and Vision Computing, 117*, 104341. https://doi.org/10.1016/j.imavis.2021.104341

Yadav, S. (2020). Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence. *International Journal for Research in Applied Science and Engineering Technology, 8*(7), 1368-1375. https://doi.org/10.22214/ijraset.2020.30560

Research article/Educational Engineering

http://doi.org/10.15446/ing.investig.103276

A Gender Gap Analysis on Academic Performance in Engineering Students on Admission and Exit Standardized Tests

Análisis de brechas de género en el desempeño académico en pruebas de admisión y egreso en programas de Ingeniería

Luis E. Gallego[1], María A. Casadiego[2]

ABSTRACT

This paper describes a gender-based research work on academic performance in engineering students. This study is based on the results of 9 469 students from Universidad Nacional de Colombia in the college's Admission and exit standard tests (the latter known as Saber Pro Tests). Tools such as descriptive statistics, regression analysis, and data mining are used to estimate both gaps and leaps in scores per gender. These tools are not only used to estimate said gaps, but also to determine whether these gaps are broadened or closed throughout the university education process. The results show that there are still gender gaps in favor of men in the area of mathematics, even in STEM programs. On the contrary, a gap in favor of women is noted in the area of writing, although a decline is also observed in reading comprehension skills for both genders. On the other hand, in terms of the global scores, women improve more than men as a result of their undergraduate experience. Finally, purely disciplinary competencies exhibit a notorious gender gap in favor of men, which should lead to future reforms in this type of programs.

Keywords: gender gap, academic performance, standardized tests, regression analysis, data mining

RESUMEN

Este artículo describe una investigación sobre brechas de género en el desempeño académico de los estudiantes en programas de ingeniería. El estudio está basado en los resultados de 9 469 estudiantes de la Universidad Nacional de Colombia en las pruebas de admisión y egreso (estas últimas conocidas como pruebas Saber Pro). Se utilizan herramientas como la estadística descriptiva, el análisis de regresión y la minería de datos para estimar tanto brechas como saltos de puntaje por género, las cuales se utilizan no solo para estimar tales brechas, sino para determinar si las mismas han sido ampliadas o cerradas a lo largo del proceso de formación universitaria. Los resultados muestran que aún persisten brechas de género en favor de los hombres en el área de matemáticas, incluso en programas STEM. Por el contrario, es notoria una brecha en favor de las mujeres en el área de escritura, aunque también se observa un deterioro en las habilidades de comprensión de lectura para ambos géneros. Por otra parte, en términos de puntajes globales, las mujeres mejoran más que los hombres como resultado de la experiencia del pregrado. Finalmente, las competencias puramente disciplinares exhiben una notoria brecha de género en favor de los hombres, lo que debe llevar a futuras reformas en este tipo de cursos.

Palabras clave: brecha de género, desempeño académico, pruebas estandarizadas, análisis de regresión, minería de datos

Received: June 21st 2022
Accepted: May 18th 2023

Introduction

According to the *2020 Global Gender Gap Report*, gaps in educational achievement still require a decisive effort from worldwide governments to close them. Only 35 countries have achieved gender parity in education and 10% of girls from 15 to 24 years old around world are still illiterate, most of them concentrated in developing countries (WEF, 2020). Consequently, more investment in human capital is still required to improve poor academic performance. In addition, even in countries with high educational achievement, women's abilities are not necessarily aligned with the required skills to make them successful in the so-called *future professions* (WEF, 2020).

Gender gaps are pronounced in college STEM programs, as only 35% of the enrolled students are women, and there is a decreasing trend in their participation in areas such as engineering, natural sciences, mathematics, statistics, and ICTs* (UNESCO, 2019).

As mentioned by Alam (2022), there are three main approaches that explain gender gaps in STEM education: biological, psychological, and sociocultural. On the one hand, the biological approach is based on cognitive patterns, brain functioning (Kucian *et al.*, 2005), and visual-spatial skills (Harris *et al.*, 2021; Reilly *et al.*, 2017; Lawton and Fletcher, 2005). However, much of the recent research argues that gender inequalities in STEM career choices are best explained by relative cognitive strengths rather than cognitive ability levels alone (Tandrayen and

[1] Electrical Engineer, Universidad Nacional de Colombia, Colombia. M.Sc Electrical Engineering, Universidad Nacional de Colombia, Colombia. PhD in Electrical Engineering, Universidad Nacional de Colombia, Colombia. Affiliation: Full professor at the Electrical and Electronics Engineering Department, School of Engineering, Bogotá Campus Universidad Nacional de Colombia. Email: lgallegov@unal.edu.co.

[2] BSc in Literary Studies, Universidad Nacional de Colombia, Colombia. MSc in Literary Studies, Universidad Nacional de Colombia, Colombia. Affiliation: Advisor at the Vice-Ministry of Elementary and Secondary Education, Ministry of Education of Colombia. Email: macasadiegos@unal.edu.co. ORCID:

*Information and Communication Technologies

Gokulsing, 2021). On the other hand, some authors argue that psychological and sociocultural contexts determine the expectancy of women's future participation in STEM education due to (i) gender stereotypes (García-Holgado *et al.*, 2020; Makarova *et al.*, 2019; UNESCO, 2019; García-Holgado *et al.*, 2019; Carlana, 2019; Verdugo-Castro *et al.*, 2018; Riegle-Crumb, 2005), (ii)beliefs regarding abilities (Kuschel *et al.*, 2020; Jungert *et al.*, 2019), and (iii) social interactions (Brenøe and Zolitz, 2020; Raabe *et al.*, 2019).

Colombia has implemented several initiatives to close both gaps, *i.e.*, in learning and egalitarian access to education. Since 2012, public education is free from Elementary to High School. However, (i) socioeconomic background, (ii) regional and ethnic provenance, and (iii) gender are factors that still restrict access to education (MEN and OECD, 2016). In addition, Colombia is not far from the world gender-gap trend in STEM programs. Between 2001 and 2018, more often than men, women chose curricular programs historically associated with female roles, such as health and education sciences. Moreover, even though the gender gap is smaller, this trend was also observed in social and human sciences. In contrast, men are more prone to choosing programs such as engineering, architecture, or urban sciences (CPEM *et al.*, 2020). Consequently, many government and education entities have undertaken some institutional efforts to close this gender gap and improve diversity in engineering education, which may demand a change in some education approaches in the context of the future challenges stated in the Sustainable Development Goals (UNESCO, 2019).

Regarding academic performance, some standardized tests have shown a notorious gender gap in some areas such as science and mathematics in Colombia. For instance, the PISA 2018 math test results showed that the gender gap is 20 points in favor of men, while, in the OECD countries, this value is only 5 points. In contrast, the gender gap in science in the OECD countries is 2 points in favor of women, but, in Colombia, this gap is 12 points in favor of men (Radinger *et al.*, 2018). In addition, Abadía and Bernal (2017) found that women underperform men with similar characteristics in the Colombian High School Exit Examination (known as the *Saber 11 Test*), where personal, family, and school characteristics explain a small part of the gap. However, differences in returns and regional provenance are very determinant. A similar regional approach in the Saber 11 test is found in Cárcamo and Mola (2019). On the other hand, Rincón and Arias (2019) found that, in the College Exit Examination (known as the *Saber Pro Test*), men from Administration areas in Higher Education programs not only score higher in quantitative reasoning but also in English (second language), reading and citizenship competencies, which are language-related areas where women usually score higher than men.

Based on the literature review, gender gap approaches are mainly focused on the Elementary and High School levels, where mathematics and sciences are the most critical areas, as they show increasing gender gaps. This paper hopes to contribute to analyzing gender gaps, this time not only in Higher Education, but also in STEM programs, where closing such gaps is expected, given that the students admitted to these programs are supposedly skillful in math and science. In addition, this study analyzes the change or leap in the academic performance of men and women between the entrance (admission) and exit exams, so that (i) the effect of university education may be estimated and (ii) the gender gap's behavior may be assessed.

This paper is organized as follows. First, a description of the data is presented in terms of the target population and competencies. Next, some methodological approaches are described in terms of the proposed variables and the methods of analysis. Then, the main results are presented in terms of descriptive statistics, regression analysis, and the data mining approach adopted. Finally, the main conclusions of the study are presented.

Data

This study was conducted over a five-year period (2016-2020), based on the scores obtained by 9469 engineering students on the Saber Pro test. The data were collected from two main sources: (i) official records from the ICFES testing agency, and (ii) official admission results from Universidad Nacional de Colombia's Admissions Office. Academic information regarding program enrollment was obtained from ICFES, and socioeconomic information such as strata and public or private schooling was obtained from the official records of the Admissions Office. In addition, students' identities were collected from ICFES, such as as biological sex. It is worth noting that the term *gender* refers to biological sex, as gender identity data were not collected.

Population

The students analyzed came from four different campuses and 31 different engineering programs of Universidad Nacional de Colombia, *i.e.*, Bogotá (9), Medellín (13), Manizales (6) and Palmira (3), covering a wide range of disciplinary areas such as Civil, Chemical, Electrical and Electronics, Mechanical and Mechatronics, Industrial and Computer, Geological, Mining, Agricultural, Environmental, and Petroleum Engineering, among others. A previous study was carried out by Gallego (2021) for the entire population of the same university, using only descriptive statistics. This research is now gender-oriented and exclusively focused on engineering programs. In addition, not only does it use descriptive statistics, but also inferential statistics and data mining approaches.

The gender composition is shown in Figure 1, where an approximate 75/25% is the dominant ratio between men and women on all campuses but Palmira, which shows a more balanced ratio (51/49%). In addition, Figures 2 and 3 show that social strata are concentrated in the low-middle class (strata 2 and 3, with about 75%), as well as a well-balanced ratio between public and private schooling (52/48%). Both the social stratum and the schooling type show no significant gender gaps.

Competencies

In general, the Saber Pro Test evaluates two types of competencies: general and disciplinary. General competencies refer to the skills that must be developed by students regardless of their background field (ICFES, 2021), which are evaluated according to the following modules:

[†]OECD:Organisation for Economic Co-operation and Development

- Quantitative reasoning (Quant)
- English as a Second language test (Eng)
- Written communication (Wrt)
- Reading comprehension (Read)
- Citizenship competencies (Cty)

On the other hand, disciplinary competencies are chosen by each curricular program from a set proposed for each knowledge area. In particular, the knowledge area for this study was engineering, which includes 15 specific disciplinary skills (Figure 8).

Methodology

Variables

Gender Gap Scores (GGS)

In general, the main purpose of this study is to analyze the differences between women and men's scores. These differences are dubbed *Gender Gap Scores* (GGS) and are treated via two different statistical approaches: parametric and non-parametric.

Regarding the parametric approach, GGS are defined as the difference in means between women and men's scores, as shown in Equation (1). A positive/negative GGS value means that women/men scored higher than men/women in the Saber Pro Test. However, even though both means are different, some tests must be run in order to determine whether these differences are statistically significant. This study used a T-test to compare both population means while fulfilling some conditions such as (i) independence, (ii) normality and (iii) homoscedasticity. In this case, the observations were assumed to be independent, given that the exam is taken individually by each student. In addition, normality and homoscedasticity assumptions were respectively verified via Shapiro-Wilks and Fligner-Killeen tests, using the Welch correction in the T-test for the cases exhibiting heteroscedasticity. Finally, a D-Cohen index was calculated to estimate the effect size of the GGS. These estimates are presented in Table 1.

$$GGS = \overline{Score_{Women}} - \overline{Score_{Men}} \tag{1}$$

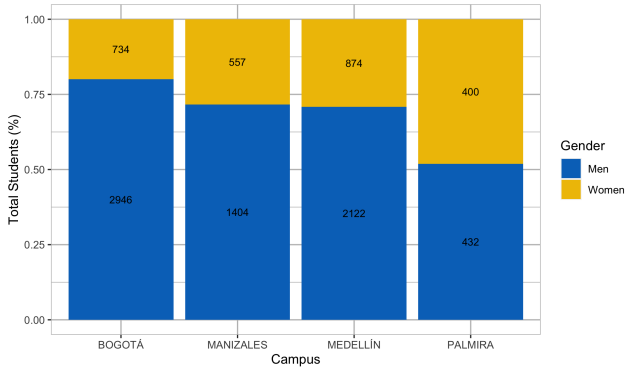


Figure 1. Percent gender distribution and number of students per campus
Source: Authors

Now, to validate the parametric approach, if the T-test applied in samples larger than 30 still showed them to be normally distributed, a bootstrapping method was used to estimate the GGS. Therefore, a resampling with 500 stratified bootstrapped samples (women-men) and a 95% confidence interval was used to estimate the mean of the GGS. These estimates are shown in Table 2. In general, both approaches, parametric and non-parametric, led to almost identical GGS values.

Gender Leap Scores (GLS)

The main purpose of Gender Leap Scores (GLS) is to analyze the changes in academic performance between an entrance and an exit exam, so that the effect of university education can be estimated. In this case, academic performance may increase or decrease as a consequence of obtaining positive or negative GLS, respectively. In this study, the entrance exam corresponds to the University's admission test, while the exit exam is the Saber Pro Test, often taken during the final year of the curriculum.

The scores obtained in both exams must be normalized in order for them to be comparable. Therefore, a max-min normalization was performed while following Equation (2).

$$Score_{esc} = 100 * \frac{(Score - min(score))}{(max(score) - min(score))} \tag{2}$$

where:

- $Score_{esc}$ corresponds to an escalated score that *only* refers to the students that took the Saber Pro Test.
- $Score$ corresponds to the score obtained in the Saber Pro or the admission test.
- $min(score)/max(score)$ corresponds to the minimum/maximum Saber Pro or admission test score only obtained by the students that took the Saber Pro test.

It is important to note that $Score_{esc} \in [0, 100]$ is only relative to the students that took both tests, so a comparison between both (admission and Saber Pro escalated scores) may describe a change in the academic performance by gender (*i.e.*, a GLS).

‡Statistical indicator comparing the difference in means against the joint standard deviation of both samples

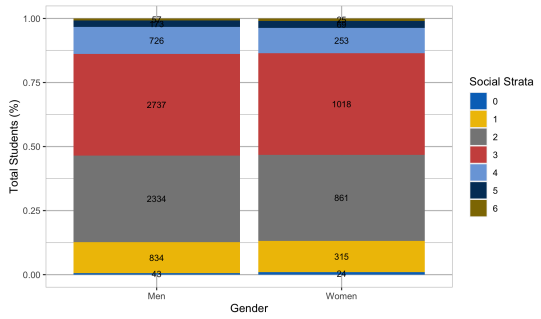


Figure 2. Percent social stratum distribution and number of students per gender
Source: Authors

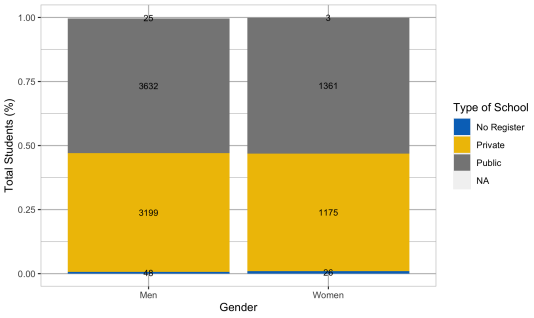


Figure 3. Percent schooling type distribution and number of students per gender
Source: Authors

On the other hand, a plausible comparison between the entrance and exit tests must consider only commonly assessed competencies or skills. In particular, the admission test assesses only the general skills acquired in high-school, such as quantitative reasoning and reading comprehension, but also natural sciences, social sciences, and image analysis. In contrast, the exit exam assesses both general (quantitative reasoning, reading comprehension, written communication, citizenship, and English) and disciplinary competencies (scientific thinking, design, and project formulation, as shown in Figure 8). Therefore, it is not possible to assess a leap in areas that were not specifically assessed in both exams, such as citizenship, English, or written communication. Thus, GLS were calculated only for quantitative reasoning and reading comprehension. In addition, a GLS was calculated for the global score, as a gender gap in global scores was observed in the admission test.

Types of analysis

The types of analysis performed on GGS and GLS were the following:

1. Descriptive statistics analysis: behavior of GGS and GLS score means, confidence intervals, interquartile ranges, and statistical significance tests per skill.

2. Regression analysis

   (a) Correlograms: correlation coefficients between different skill scores in both tests per gender.

   (b) Linear multivariate analysis: linear regressions on global scores explained by the different skill scores per gender. The regressions performed follow Equation (4).

   (c) Logistic regression analysis: logistic regressions on high/low scores or positive/negative leaps regarding gender via Equation (5). The main purpose is to determine the Odds ratios between men and women on the different skill scores.

3. Data mining analysis (association rules): the main purpose of association rules is to discover hidden relations between scores and other dimensions (gender, social strata, and schooling type) based on their frequency of appearance. In this study, a basic Apriori algorithm was applied to obtain simple if-then rules in the form $X \Rightarrow Y$. Only strong rules were analyzed based on support and confidence indicators. Generally speaking, support denotes the frequency of appearance of an itemset X in relation to the total of transactions. In addition, confidence refers to the number of times that if-then rules appear, following Equation (3).

$$Confidence(X \Rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \tag{3}$$

where $X \cup Y$ is an itemset including all items in X and Y. Confidence may be interpreted as a conditional probability $P(Y|X)$, *i.e*, the probability of an item including X while also including Y. In this study, the support and confidence values were set at 0,2 and 0,7, respectively. In other words, rules are only interesting when the antecedent's frequency is higher than 20% and the consequent is observed in 70% of the transactions.

$$GS = \kappa + \alpha_R Read + \alpha_C Cty + \alpha_E Eng + \alpha_Q Quant + \alpha_W Wrt \tag{4}$$

where:

- GS: global Score
- κ: intersect term
- α_R: reading skills score
- α_C: citizenship skills score
- α_E: English skills score
- α_Q: quantitative reasoning score
- α_W: writing skills score

$$log\frac{p}{1-p} = \kappa + \alpha_W Women \tag{5}$$

where:

- p: depending on the analysis, p could be the probability of obtaining a high score (above the average) in any skill or the probability of obtaining a positive leap between the admission and Saber Pro tests (positive GLS).
- κ denotes an intersect term.
- α_W: $log(OddsRatio)$ of being women. In other words, the ratio between the odds of women obtaining a high score or a positive GLS in any skill with respect to men.

Results

Descriptive statistics

Figure 4 shows the boxplot by gender on global scores in the Saber Pro test, where almost identical results are obtained by women and men (similar means and interquartile ranges). Consequently, no meaningful gap in global scores is observed. On the other hand, the evolution in GGS per skill since 2016 seems to be invariant (Figure 5), where the highest GGS are found in quantitative reasoning in favor of men (negative GGS) and writing communication in favor of women (positive GGS). In addition, this behavior is consistent on all campuses, as shown in Figure 6.

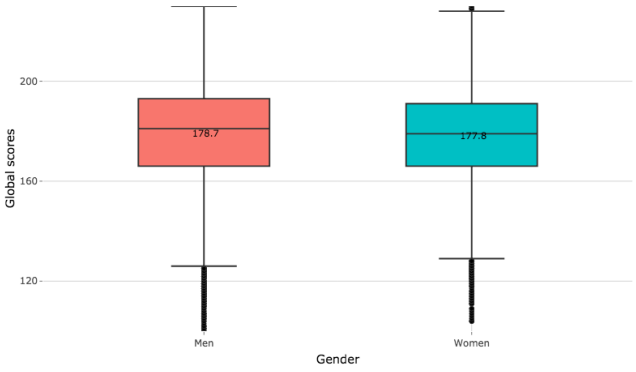


Figure 4. Gender gap on global scores
Source: Authors

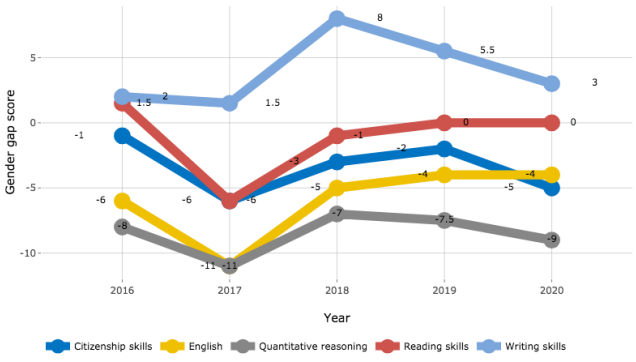


Figure 5. Evolution of gender gap scores per module
Source: Authors

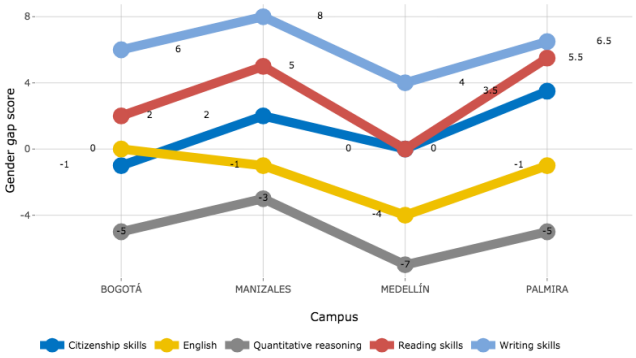


Figure 6. Evolution of gender gap scores per module per campus
Source: Authors

Table 1. GGS statistics per module (parametric approach)

Module	Shapiro-Wilks test, women (p-value)	Shapiro-Wilks test, men (p-value)	Fligner-Killeen homo-cedastic-ity test (p-value)	T-test mean dif-ferences (p-value)	Effect size (D-Cohen)
Citizenship skills	2,07E-19	1,14E-19	1,15E-09	3,04E-03	Negligible
English	2,15E-18	1,70E-18	1,04E-01	1,52E-14	Negligible
Quantitative reasoning	1,04E-16	5,85E-24	1,08E-05	4,26E-44	Small
Reading skills	2,31E-10	1,01E-14	5,09E-05	9,62E-01	Negligible
Writing skills	4,70E-16	5,72E-19	3,70E-04	2,28E-08	Negligible

Source: Authors

Likewise, based on the non-parametric approach described above, Figure 7 and Table 2 show that the GGS per skill are well differentiated, given that the confidence intervals and interquartile ranges are very narrow. Therefore, the GGS behave very consistently. Citizenship skills, English, and quantitative reasoning are in favor of men, especially the latter. In contrast, written skills are in favor of women, whereas there is no gap in the reading comprehension skills.

These GGS results are coherent with some other findings in the international literature about gender-based gaps in academic performance at elementary and secondary schools, where men score higher than women in math but lower in reading tests (Abadía and Bernal, 2017). Consequently, our results suggest that the gender gap in math tests is still present in higher education, and even in the case of STEM programs, which is the main population addressed in this study.

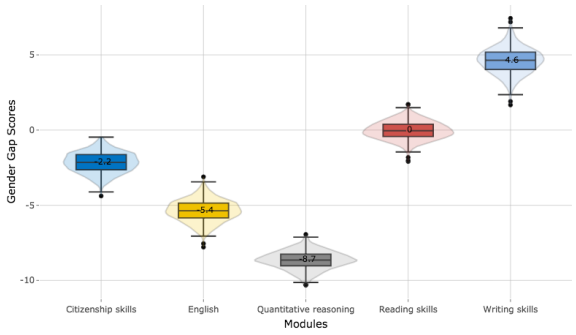


Figure 7. Gender gap per module (bootstrapping)
Source: Authors

In addition, the results shown in Table 1 suggest that the difference in means (between men and women) is statistically significant according to the p-values obtained for the statistical T-test. However, the D-cohen index also suggests that the effect size is almost negligible in most of the skills except for quantitative reasoning, where the gap is considered small.

Table 2. GGS statistics per module using a non-parametric bootstrapping approach - means and confidence intervals (95%)

Module	GGS (mean/conf. interval)
Citizenship skills	-2,2 / -3,4 to -0,8
English	-5,4 / -6,8 to -4
Quantitative reasoning	-8,7 / -9,8 to -7,5
Reading skills	0 / -1,2 to 1,1
Writing skills	4,6 / 3 to 6,3

Source: Authors

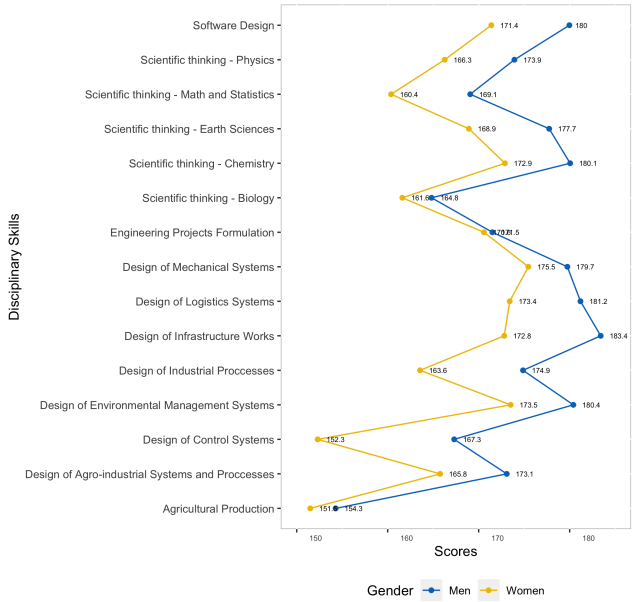


Figure 8. Gender gap for disciplinary skills
Source: Authors

Figure 8 shows the scores in each disciplinary skill by gender. As mentioned earlier, disciplinary skills are chosen by each curricular program. In general, the GGS are in favor of men in almost all disciplinary skills except for agricultural production, engineering project formulation, and scientific thinking in biology. The highest GGS is found in design of control systems, design of industrial processes, and design of infrastructure works.

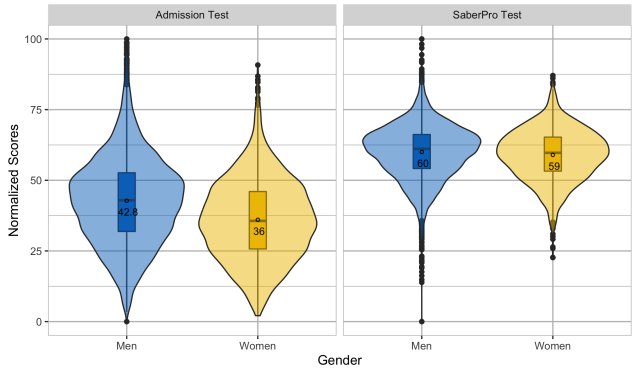


Figure 9. Escalated scores in the admission and Saber Pro tests by gender
Source: Authors

As for the GLS, Figure 9 shows the global escalated scores for both the admission and the Saber Pro tests, where there is a differentiated performance between men and women. Results suggest that the initial GGS gap in the admission test (42,8/36 points) is almost closed in the Saber Pro test (60/59). This behavior is confirmed in Figure 10, where women's leap scores (mean: 22,9) are consistently higher than those of men (mean: 17,3).

Figures 11 and 12 show the GLS scores by social strata and schooling type. As far as social classes are concerned, the results suggest that not only is the leap higher for women in each class, but also that it is higher when the students come from the lower classes. In addition, the gender gap

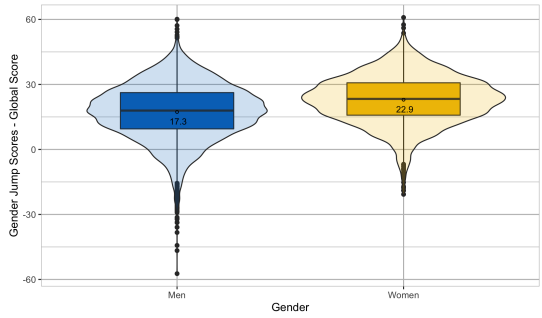


Figure 10. GLS for the global score
Source: Authors

is slightly higher in the lower strata. In terms of schooling type, the results show that the jump is consistently higher for students from public schools, with the gender gap being almost identical for both schooling types. These results are very promising regarding the mission of a public university, showing that not only do women experience the greatest jumps, but also that these jumps are more pronounced for women coming from public schools and lower classes.

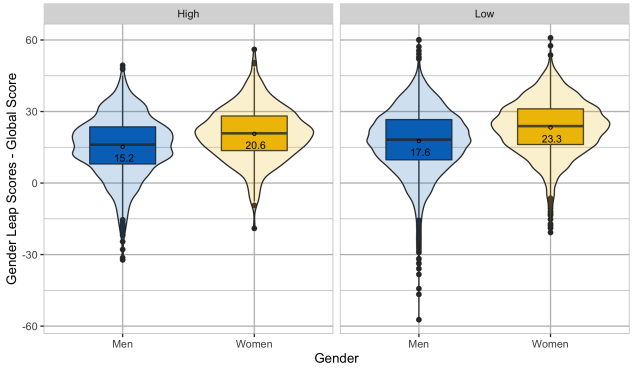


Figure 11. GLS for the global scores by social stratum
Source: Authors

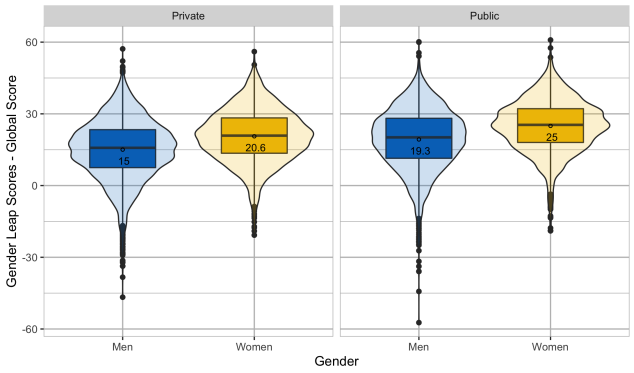


Figure 12. GLS for the global scores by schooling type
Source: Authors

Figure 13 shows the GLS by campus. In general, the greatest jumps are seen in the Palmira campus, where gender parity is significantly higher than on the other campuses. This finding is consistent with previous research showing that academic performance tends to improve in groups with higher proportions of women (Crombie *et al.*, 2003).

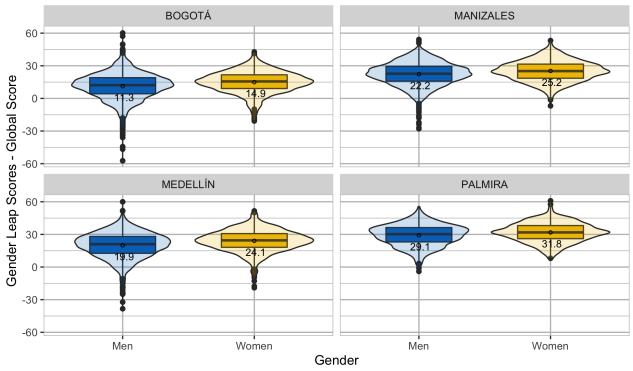


Figure 13. GLS for the global scores by campus
Source: Authors

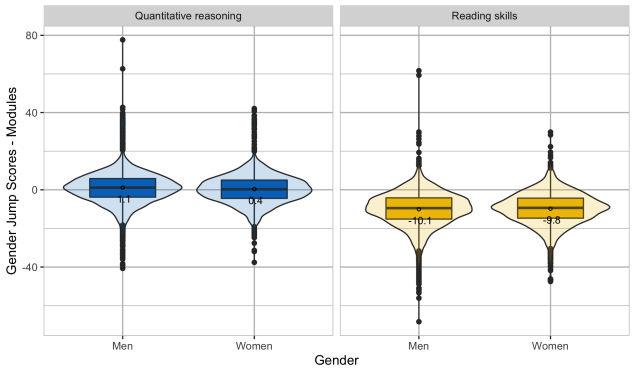


Figure 14. GLS for quantitative reasoning and reading skills
Source: Authors

In contrast, the GLS in quantitative reasoning and reading comprehension are very similar for both genders (Figure 14). Regarding quantitative reasoning, there is no change between men and women, so the gender gap is not only almost identical, but also academic performance remains invariant (GLS close to zero) in the Saber Pro test when compared to the admission exam. On the contrary, the performance in reading comprehension tends to deteriorate along the university experience, as the GLS are negative (GLS close to -10 for both genders). However, performance deteriorates in the same proportion in men and women. These results are confirmed by the statistical tests shown in Table 3, where, for most of the skills, the means for the GLS are statistically different between men and women [§]. However, these differences are negligible for quantitative reasoning and reading skills, as well as small for the global score (D-Cohen result).

Regression analysis

Correlograms

Figures 15 and 16 show correlograms for the escalated scores in each module of the admission test (suffix: Adm) and the Saber Pro Test (suffix: Spro) for men and women. In general, there are moderate associations between the admission and Saber Pro global scores. Therefore, a higher/lower admission score does not necessarily imply higher/lower scores in the exit exam (Saber Pro). Moreover, this association is a bit stronger for women. In other words, there is an effect of the university experience on the academic performance in the exit exam.

Table 3. Statistics for the GLS

Module	Shapiro-Wilks test, women (p-value)	Shapiro-Wilks test, men (p-value)	Fligner-Killeen homo-cedastic-ity test (p-value)	T-test mean dif-ferences (p-value)	Effect Size (D-Cohen)
Quantitative reasoning	2E-14	1E-18	1E-01	1E-04	Negligible
Reading skills	4E-07	3E-11	8E-03	0,16	Negligible
Global score	2E-04	1E-08	9E-12	3E-82	Small

Source: Authors

On the other hand, reading, citizenship and English (second language) skills are the scores that better correlate with the global scores in the Saber Pro test. Surprisingly, quantitative reasoning is not as strongly correlated with the final global score as reading or English skills, even in the case of engineering students. That is to say that, instead of quantitative reasoning, now reading, citizenship, and English are the competencies that make the difference in the global score. In addition, the correlation coefficients (R) remain almost invariant between men and women, as shown in Table 4.

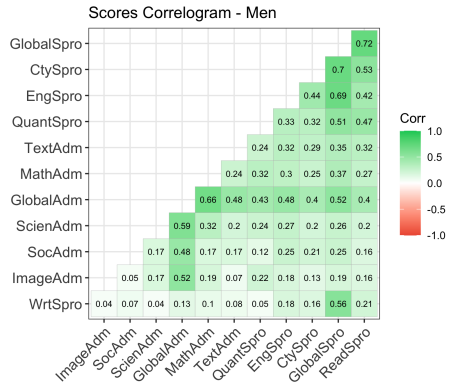


Figure 15. Correlogram for the admission and exit exam scores (men)
Source: Authors

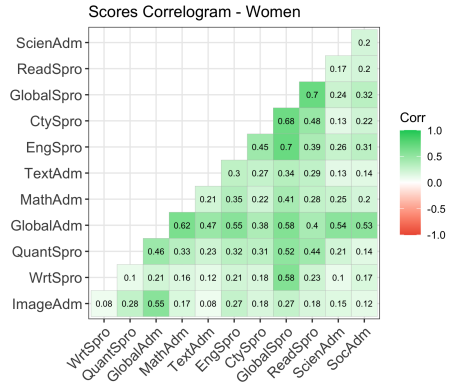


Figure 16. Correlogram for the admission and exit exam scores (women)
Source: Authors

[§]In the case of reading comprehension skills, the p-value is non-conclusive, so there is no statistical evidence that the means for the GLS are different between men and women. This may be explained by the fact that both means are very close,*i.e.*, -10,1 and -9,8 for men and women, respectively.

Table 4. Correlation coefficients (R) for the global score in each skill per gender

Skill	R Men	R Women
Reading	0,72	0,7
Citizenship	0,7	0,68
English	0,69	0,7
Quantitative reasoning	0,51	0,52
Writing	0,56	0,58

Source: Authors

Linear multivariate regressions

As mentioned above, some linear regressions on the Saber Pro global scores per gender were modeled following Equation (4). A linear model was chosen based on the AIC (Akaike Information Criterion) using a step-wise algorithm. The model started with both the admission and Saber Pro module scores. However, the admission scores were neglected, since no significant changes were observed regarding Pearson's R^2 correlation when they were included, going from 0,921 to 0,923. This correlation value not only indicates that 92% of the variance for the global score is explained by the skill scores, but also that this result is coherent with the poor correlation coefficients shown in Figures 15 and 16. Again, the results suggest that there is an effect of the university experience on academic performance in the exit exam.

Table 5 shows the resulting regression coefficients for men and women following Equation (4) . In both cases, R^2 is high enough to explain most of the global score variance (0,922 and 0,919 for men and women, respectively), with all the coefficients being statistically significant ($p_{values} < 2e-16$ for a T-test in both cases) and standard errors of 0,03314 and 0,0325 for men and women, respectively. The regression coefficients remain almost invariant for both genders, and English and citizenship skills are the modules that mostly explain the global scores variance, whereas quantitative reasoning is the skill that most poorly explains it. Moreover, English explains this variance slightly better when the students are women, just as Citizenship skills when the students are men.

Table 5. Regression coefficients by gender, Equation 4

Gender	κ	α_R	α_C	α_E	α_Q	α_W
Men	-0,258	0,294	0,346	0,372	0,166	0,328
Women	-0,269	0,306	0,327	0,395	0,163	0,327

Source: Authors

Logistic regressions

A set of logistic regressions was modeled to estimate the odds ratios between men and women for different skill scores. In general, the modeled logistic regressions followed the formulation shown in Equation 5. However, several sets of regressions were tested for different purposes, as follows:

1. A first set determined the odds ratios of obtaining "high" scores (defined as scores above the average) for any general skill.
2. A second set determined the odds ratios of obtaining "positive" GLS.
3. A third set determined the odds ratios of obtaining higher scores in disciplinary skills when compared to general ones.

Table 6 shows the results for the odd ratios of obtaining a "high" score in any skill. Most of the odds ratios are less than 1, implying that, for most of the skills, a woman's odds of obtaining a high score are only a portion of men's odds. The most critical difference is observed in quantitative reasoning, where a woman's odds of getting a high score are only 55% of men's odds. In other words, men's odds of getting a high score in quantitative reasoning almost duplicate those of women. On the contrary, in the case of writing skills, women's odds of getting a high score are 1,27 times higher than those of men. In addition, most of the obtained odds ratios are statistically significant, except for the case of reading comprehension, where no conclusions can be drawn, as there is no distinction between the odds of women and men, *i.e.*, very close to 1 (0,94).

Table 6. Statistics for the odds ratios of getting a high score in any general skill

Skill	α_W [a]	$Odds_{Ratio}$	p-value
Global score	-0,249	0,78	7,40E-08
Reading	-0,056	0,94	0,2222
Citizenship	-0,140	0,87	0,00269
English	-0,360	0,70	1,40E-14
Quantitative reasoning	-0,59	0,55	2,00E-16
Writing skills	0,237	1,27	4,20E-07

Source: Authors

[a] $(\mathrm{Ln}(Odds_{Ratio}))$

Table 7 shows the results for the odd ratios of obtaining a "positive" GLS score in any skill. Again, a woman's odds of getting a positive GLS score in quantitative reasoning are less than those of men (82%). On the contrary, in the case of the global score, women's odds of getting a positive GLS are 3,53 times higher than men's. Moreover, most of the obtained odds ratios are statistically significant, except for the case of reading comprehension, where no conclusions can be drawn, as there is no distinction between the odds of women and men, *i.e.*, very close to 1 (0,987).

$$ln(\frac{P(Score_{disc} > Score_{gen})}{1 - P(Score_{disc} > Score_{gen})}) = -0,3679 - 0,4624 * Women \quad (6)$$

Finally, Equation 6 shows the final regression coefficients for the odds of getting a higher score in disciplinary skills when compared to general ones. If the students are men, the odds are only 44%, which means that general skill scores are very often higher than disciplinary ones. However, if students are women, the odds decrease down to 14%, which is not only very low but also means that the odds of getting a higher score in disciplinary skills for a woman are much less than those of men (34%). This odds ratio is statistically significant, as shown in Table 8.

Association Rules

This section presents the results for the data mining analysis based on association rules ($X \Rightarrow Y$) using the *Apriori* algorithm, as described in the *Methodology* section. The analysis followed two main approaches:

Table 7. Statistics for the odds ratios of getting a positive GLS per skill

Skill	α_W [a]	$Odds_{Ratio}$	p-value
Quantitative reasoning	-0,192	0,82	3,8E-05
Reading	-0,012	0,98	0,87
Global score	1,261	3,53	2E-16

Source: Authors

[a] (Ln($Odds_{Ratio}$))

Table 8. Statistics for the odds ratios of getting a higher score in disciplinary skills when compared to general ones

Event	α_W [a]	$Odds_{Ratio}$	p-value
$P(Score_{disc} > Score_{gen})$	-0,4624	0,34	2E-16

Source: Authors

[a] (Ln($Odds_{Ratio}$))

- Based on the skills with the highest GGS (quantitative reasoning, English, and writing), association rules were mined, with gender as the rule's consequent. The resulting mined rules are shown in Table 9.
- The exogenous variables of social Stratum (high/low) and schooling type (public/private) were included, and association rules were separately mined for both men and women. A comparison of the support and confidence values obtained is presented for the common mined rules, in order to determine the cases where these rules are stronger. The resulting mined rules are shown in Table 10.

Table 9 confirms some of the results shown in the previous sections. In general, men scoring "high" in quantitative and English but "low" in writing skills is a strong rule, with a confidence greater than 75%. This result is consistent with the high GGS in favor of men for these skills, as was previously analyzed. The fact that men tend to score simultaneously high in the couples [*quantitative*, *English*] and [*English*, *writing*] but also simultaneously low in the couple [*quantitative*, *writing*] were also detected as strong rules. This result is very interesting, as it implies a certain type of association in male engineering students, suggesting that a good performance in quantitative reasoning comes along with a poor performance in writing skills.

Table 10 shows a comparison of the support and confidence values for men and women with regard to the separately mined rules. First of all, the confidence values are almost identical. Therefore, the mined rules are equally confident between men and women. However, the rules are not necessarily equally stronger, as the support values exhibit different values for both genders. As previously mentioned, the social stratum and schooling type were explicitly included in the mined rules with the aim to assess their effect on skill scores. The main results may be stated as follows:

1. Coming from lower classes,
   - Men engineering students score "high" more often than women in quantitative reasoning.
   - Women engineering students score "low" more often than men in second language skills (English). This result is very interesting, given that some other reported works (Corpas, 2013) have found that women score higher in English as a second language. This is not the case for engineering students.
   - Women engineering students score simultaneously "low" more often than men in quantitative reasoning and second language skills (English).
2. Coming from public schools,
   - Women engineering students score simultaneously "low" more often than men in quantitative reasoning and second language skills (English).
3. Coming from lower classes and public schools,
   - Women engineering students score "low" more often than men in quantitative reasoning.

Table 9. Obtained association rules ($X \Rightarrow Y$) using the Apriori algorithm: support (Sup) >0,2 and confidence (Conf) >0,7

Antecedent (X)	Consequent (Y)	Sup	Conf
Quant=High	Sex=Men	0,39	0,78
Eng=High	Sex=Men	0,39	0,76
Wrt=Low	Sex=Men	0,37	0,75
Quant=High & Eng=High	Sex=Men	0,25	0,79
Quant=High & Wrt=Low	Sex=Men	0,20	0,80
Eng=High & Wrt=High	Sex=Men	0,21	0,79

Source: Authors

Table 10. Comparison of support and confidence values for men and women in the obtained association rules ($X \Rightarrow Y$) applying the Apriori algorithm (support >0,2 and confidence >0,7)

Ant[a] (X) ⇒	Cons[b] (Y)	Sup (M/W) [c]	Con(M/W) [d]
(Quant=High)	Strat[e]=Low	0,46/0,33	0,84/0,82
(Eng=Low)	Strat=Low	0,43/0,52	0,93/0,94
(ScTyp[f]=Public & Quant=Low)	Strat=Low	0,27/0,35	0,97/0,98
(Quant=Low & Eng=Low)	Strat=Low	0,26/0,37	0,93/0,95
(ScTyp=Public & Quant=Low)	Eng=Low	0,20/0,29	0,72/0,8
(Quant=Low & Eng=Low)	ScTyp=Public	0,2/0,29	0,74/0,73

Source: Authors

[a] Rule's antecedent.
[b] Rule's consequent.
[c] Rule's support (men/women).
[d] Rule's Confidence (Men/Women).
[e] Social stratum
[f] Schooling type.

Conclusions

Gender inequalities in academic performance at elementary and secondary schools have been widely reported, evidencing that the lack of skills in mathematics and reading may lead to fewer chances of being admitted to STEM programs. In particular, Abadía and Bernal (2017) and Abadía (2017) specifically found that men score better globally and in math and science, as do women in reading, in the high school exit examination known as *Saber 11*. Surprisingly, the results obtained in this study suggest that

the gender gap in math tests is still present in higher education, and more specifically in the addressed STEM programs, which are expected to be math-oriented. One of the possible explanations, regardless of the mathematics background, arises from psychological factors such as women's anxiety, which has been previously reported in standardized tests (De la Rica and Rebollo, 2018). In addition, not only quantitative reasoning exhibited a gender gap, but also other skills including disciplinary competencies.

The highest gender gap score (GGS) in favor of men was found in quantitative reasoning. According to the mathematical analysis performed, this gap is statistically significant, but its effect is small based on the gender score variances (D-cohen). Moreover, there is no improvement from the admission exam scores, since the gender gap remains almost identical in the college exit exam (Saber Pro). In addition, men's odds of getting a high score in quantitative reasoning almost duplicate those of women. On the other hand, the data mining approach is also consistent, as it confirms that men scoring high in quantitative and English are strong rules, with a confidence greater than 75%. Moreover, another strong rule shows that, when students come from lower classes or public schools, women score low more often than men.

In contrast to the results reported in Abadía and Bernal (2017), where men score better than women in high school reading tests, college exit exams show no statistically significant gender gap for engineering students. This may lead to the idea that the gender gap has been closed. However, our results show that reading not only experienced the highest leap from admission tests, but also the most negative, which means that there is a significant weakening in reading skills. This result may lead to deep thoughts about the reading comprehension capabilities in STEM curricular programs that, apparently, are no longer developed since high school.

Another result reported in (Abadía and Bernal, 2017) shows that men score better than women in high school global scores. This is consistent with our findings with regard to the admission tests, where a gender gap in favor of men was also observed. However, this gap is completely closed in the college exit exam, implying that women improve more than men as a result of the college experience. The results suggest that women's odds of improving their admission scores are 3,53 times higher than those of men. In addition, the models for global scores show that the regression coefficients remain almost invariant for both genders where English and citizenship skills are the most significant. In other words, quantitative reasoning is not a determining factor in the global scores of engineering students, since it poorly explains the variances. Consequently, the closing ot the gender gap in global scores is not necessarily related to quantitative reasoning or reading skills.

In terms of the GLS, the results show that females close the gap in the admissions test, which entails higher score jumps. Furthermore, the highest score jumps are observed on the campus where gender parity is significantly higher (Palmira). Additionally, not only do women experience the larger jumps, but these jumps are more pronounced for women from public schools and lower-class backgrounds.

The highest GGS in favor of women was found in writing skills. Mathematically speaking, this gap is also statistically significant, but its effect is considered smaller when compared to the gap in quantitative reasoning. In addition, women's odds of getting a high score in writing skills are 1,27 times higher than those of men. The data mining approach was also consistent, confirming that men scoring low in writing skills is a strong rule, with a confidence greater than 75%. However, another mined association rule showed a very interesting result, implying that a good performance in quantitative reasoning for male students comes with a poor performance in writing skills.

Finally, regarding disciplinary competencies, the results show a significant gender gap in favor of men. This result is very consistent in most of the disciplinary areas of engineering. Moreover, men's odds of getting a higher score in disciplinary skills are only 44% when compared to general skills (quantitative reasoning, reading, writing, *etc.*), which means that general skill scores are very often higher than those of disciplinary skills. In addition, for women students, the odds decrease down to 14%, which is not only very low but also means that the odds of getting a higher score in disciplinary skills for a woman are much less than those of men (34%). A possible explanation for this result may be the reported idea that women's performance is higher at tests that assess knowledge acquisition rather than knowledge application (UNESCO, 2019). In any case, this result may lead to deep thoughts about disciplinary courses in STEM programs, which probably not only need to be improved but also might have to be gender-oriented.

CRediT author statement

All authors: conceptualization, methodology, software, validation, formal analysis, investigation,writing (original draft, writing, re-view, and editing), data curation.

References

Alam, A. (2022, April). Psychological, sociocultural, and biological elucidations for gender gap in STEM education: a call for translation of research into evidence-based interventions. In Alam, A. *Psychological, Sociocultural, and Biological Elucidations for Gender Gap in STEM Education: A Call for Translation of Research into Evidence-Based Interventions. Proceedings of the 2nd International Conference on Sustainability and Equity* (*ICSE-2021*). *Atlantis Highlights in Social S.* https://ssrn.com/abstract=4073986

Abadía, L K., and Bernal, G. (2017). A widening gap? A gender-aased analysis of performance on the Colombian high school exit examination. *Revista de Economía del Rosario, 20*(1), 5-31. https://doi.org/10.12804/revistas.urosario.edu.co/economia/a.6144

Abadía, L. (2017). Gender score gaps of Colombia in PISA test. *Vniversitas Económicas, 17*(8), 1-24.

Brenøe, A. A., and Zolitz, U. (2020). Exposure to more female peers widens the gender gap in stem participation. *Journal of Labor Economics, 38*(4), 1009-1054. https://doi.org/10.1086/706646.

Cárcamo, C., and Mola, J. (2019). Diferencias por sexo en el desempeño académico en Colombia: Un análisis regional.

Economía & Región, 6(1), 133-169. https://revistas.utb.edu.co/index.php/economiayregion/article/view/137

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias, *The Quarterly Journal of Economics, 134*(3), 1163-1224. https://doi.org/10.1093/qje/qjz008

Corpas, M. (2013). Gender differences in reading comprehension achievement in English as a foreign language in Compulsory Secondary Education. *Didactics of Language and Literature. Education, 17*, 67-84. https://tejuelo.unex.es/article/view/2546

CPEM, DANE, and ONU (2020). *Mujeres y hombres: Brechas de género en Colombia*. Consejería Presidencial para la Equidad de la Mujer. https://www.dane.gov.co/files/investigaciones/genero/publicaciones/mujeres-y-hombre-brechas-de-genero-colombia-info\rme.pdf

Crombie, G., Pyke, S.W., Silverthorn, N., Jones, A., and Piccinin, S (2020). Students' perceptions of their classroom participation and instructor as a function of gender and context. *The Journal of Higher Education 74*(1), 51-76. https://doi.org/10.1353/jhe.2003.0001

De la Rica, S., and Rebollo-Sanz, Y. F. (2018). Brechas de género en competencias cognitivas: evidencia internacional. *Cuadernos Económicos De ICE, 95*, 125-149. https://doi.org/10.32796/CICE.2018.95.6645

Gallego, L. (2021). *Saltos y brechas*. Universidad Nacional de Colombia. https://repositorio.unal.edu.co/handle/unal/80700

García-Holgado, A., Mena, J., García-Peñalvo, F. J., Pascual, J., Heikkinen, M., Harmoinen, S., García-Ramos, L., Peñabaena-Niebles, R., and Amores, L. (2020). Gender equality in STEM programs: a proposal to analyse the situation of a university about the gender gap. In IEEE (Eds.), *2020 IEEE Global Engineering Education Conference* (*EDUCON*) (pp. 1824-1830). IEEE.https://doi.org/10.1109/EDUCON45650.2020.9125326

García-Holgado, A., Camacho, A., and García-Peñalvo, F. J. (2019). La brecha de género en el sector STEM en América Latina: Una propuesta europea. In M. L. Sein-Echaluce Lacleta, Á. Fidalgo-Blanco, and F. J. García-Peñalvo (Eds.), *Actas del V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad - CINAIC 2019* (pp. 704-709). Servicio de Publicaciones Universidad de Zaragoza. http://dx.doi.org/10.26754/CINAIC.2019.0143

Harris, D., Lowrie, T., Logan, T., and Hegarty, M. (2021). reasoning, mathematics, and gender: Do spatial constructs differ in their contribution to performance?. *The British Journal of Educational Psychology, 91*(1), 409-441. https://doi.org/10.1111/bjep.12371

Jungert, T., Hubbard, K., Dedic, H., and Rosenfield, S. (2019). Systemizing and the gender gap: examining academic achievement and perseverance in STEM. *European Journal of Psychology of Education, 34*(2), 479-500. https://doi.org/10.1007/s10212-018-0390-0.

Kuschel, K., Ettl, K., Díaz-García, C., and Alsos, G. A. (2020). Stemming the gender gap in STEM entrepreneurship - insights into women's entrepreneurship in science, technology, engineering and mathematics. *International Entrepreneurship and Management Journal, 16*, 1-15. https://doi.org/10.1007/s11365-020-00642-5

ICFES (2021). *Guías de Orientación Saber Pro 2021 - Módulos de Competencias Genéricas - Bogotá Colombia*. https://www.icfes.gov.co/documents/20143/2297330/Guia+de+orientacion+modulos+genericos+Saber-Pro+2021.pdf

Kucian, K., Loenneker, T., Dietrich, T., Martin, E., and Von Aster, M. (2005). Gender differences in brain activation patterns during mental rotation and number related cognitive tasks. *Psychology Science, 47*(1), 112-131.

Lawton, C. A., and Hatcher, D. W. (2005). Gender differences in integration of images in visuospatial memory. *Sex Roles, 53*(9), 717-725. https://doi.org/10.1007/s11199-005-7736-1

Makarova, E., Aeschlimann, B., and Herzog, W. (2019, July). The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. In Frontiers in Education (vol. 4, p. 60). Frontiers Media SA. https://doi.org/10.3389/feduc.2019.00060

MEN and OECD (2016). *La educación en Colombia.* https://www.oecd-ilibrary.org/education/education-in-colombia_9789264250604-en

Raabe, I. J., Boda, Z., and Stadtfeld, C. (2019). The social pipeline: How friend influence and peer exposure widen the STEM gender gap. *Sociology of Education, 92*(2), 105-123. https://doi.org/10.1177/0038040718824095

Radinger, T., Echazarra, A., Guerrero, G., and Valenzuela, J. P. (2018). *OECD reviews of school resources: Colombia*. https://www.oecd.org/education/school/OECD-Reviews-School-Resources-Summary-Colom\bia-Spanish.pdf

Reinking, A., and Martin B. (2018). The gender gap in STEM fields: Theories, movements, and ideas to engage girls in STEM. *Journal of New Approaches in Educational Research, 7*(2), 148-153. https://doi.org/10.7821/naer.2018.7.271

Rincón-Baez, W. U., and Arias-Velandia, N. (2019). Brecha de rendimiento académico por género en Saber Pro en programas de Administración de los departamentos de Colombia. *Panorama, 13*(25), 142-161. https://doi.org/10.15765/pnrm.v13i25.1410

Reilly, D., Neumann, D. L., and Andrews, G. (2017). Gender differences in spatial ability: Implications for STEM education and approaches to reducing the gender gap for parents and educators. In *Visual-Spatial Ability in STEM Education* (pp. 195-224). Springer. https://doi.org/10.1007/978-3-319-44385-0_10

Riegle-Crumb, C. (2005). *The cross-national context of the gender gap in math and science.* In L. V. Hedges and B. Schneider (Eds.) *The Social Organization of Schooling* (pp. 227-243). Russell Sage Foundation.

Tandrayen-Ragoobur, V., and Gokulsing, D. (2021). Gender gap in STEM education and career choices: what matters?. Journal of Applied Research in Higher Education. https: doi.org//10.1108/JARHE-09-2019-0235

UNESCO (2019). *Descifrar el código: la educación de las niñas y las mujeres en ciencias, tecnología, ingeniería y matemáticas* (*STEM*). https://unesdoc.unesco.org/ark:/48223/pf0000366649.

Verdugo-Castro, M., Sánchez-Gómez, C., and García-Holgado, A. (2018, October 24-26). *Gender gap in the STEM sector in pre and university studies of Europe associated with ethnic factors*. In Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18). Association for Computing Machinery, New York, NY, USA, 984-990. https://doi.org/10.1145/3284179.3284348

World Economic Forum (WEF) (2020). *Global gender gap report 2020*. https://www3.weforum.org/docs/WEF_GGGR_2020.pdf