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
Gabriela Rojas Castro

Proofreader
José Daniel Gutiérrez-Mendoza

Layout Artist
David Mauricio Valero

For additional information contact
revii_bog@unal.edu.co
Bogotá - Colombia
2024

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

Ingeniería Química / Alimentos / Ambiental

Sistemas / Ingeniería de Computación

Ingeniería Mecanica / Mecatronica / Ciencia de Materiales

Ingeniería Civil / Sanitaria

Ingeniería Industrial

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e103068)

Research Article/Chemical Engineering

http://doi.org/10.15446/ing.investig.103068

Thermal Degradation Analysis of Mamoncillo (*Melicoccus bijugatus*) Waste: Thermal Behaviors, Kinetics, and Thermodynamics

Análisis de degradación térmica de residuos de mamoncillo (*Melicoccus bijugatus*): comportamiento térmico, cinético y termodinámico

Andrés Felipe Rojas-González[1] and Francisco Javier Velasco-Sarria[2]

ABSTRACT

This research studied the thermal conversion characteristics, kinetics, and thermodynamics of *mamoncillo* peels and seeds using non-isothermal thermogravimetric analysis. Kinetic analysis was performed using the Kissinger-Akahira-Sunose, Flynn-Wall-Ozawa, Starink, and Friedman methods. The reaction kinetic models were obtained by means of the master-plots method for 18 different empirical reaction models, calculating the enthalpy, Gibbs free energy, and entropy as thermodynamics parameters. It was found that the average activation energy for *mamoncillo* peels and seeds was 238,71 and 197,60 kJ/mol, respectively. The frequency factor was found to be between 10^9 and 10^{31} s^{-1} for *mamoncillo* peels and between 10^9 and 10^{34} s^{-1} for *mamoncillo* seeds. The average values of ΔH and ΔG were also found to be 233,83 and 192,81 kJ/mol and 164,84 and 162,10 kJ/mol for *mamoncillo* peels and seeds, respectively. The reaction kinetic models regarding the thermal decomposition of *mamoncillo* peels were found to be described by the contracting cylinder (R2) and third-order (F3) models, while those for *mamoncillo* seeds can be described by the second-order (F2) and contracting sphere (R3) models. It was concluded that the pyrolysis process of *mamoncillo* waste can be described by a complex reaction mechanism, and that these wastes have thermal properties with the potential to produce bioenergy.

Keywords: tkinetic models, mamoncillo wastes, pyrolysis, thermodynamic analysis

RESUMEN

En este estudio se investigaron las características de conversión térmica, cinéticas y termodinámicas de las semillas y cáscaras de mamoncillo utilizando análisis termogravimétrico no isotérmico. El análisis cinético se realizó empleando los métodos de Kissinger-Akahira-Sunose, Flynn-Wall-Ozawa, Starink y Friedman. Los modelos cinéticos de reacción se obtuvieron mediante el método de gráficas maestras para 18 modelos de reacción empíricos diferentes, y, como parámetros termodinámicos, se calcularon la entalpía, la energía libre de Gibbs y la entropía. Se encontró que la energía de activación promedio para las cáscaras y las semillas de mamoncillo fue de 238,71 y 197,60 kJ/mol respectivamente. El factor de frecuencia estuvo entre 10^9 y 10^{31} s^{-1} para las cáscaras de mamoncillo y entre 10^9 y 10^{34} s^{-1} para las semillas de mamoncillo. También se encontró que el valor promedio de ΔH y ΔG estaba entre 233,83 y 192,81 kJ/mol y 164,84 y 162,10 kJ/mol para las cáscaras y las semillas respectivamente. Se encontró que los modelos cinéticos de reacción para la descomposición térmica de cáscaras de mamoncillo se pueden describir mediante los modelos cilindro de contracción (R2) y de tercer orden (F3), mientras los de las semillas se pueden describir por medio de los modelos de segundo orden (F2) y esfera de contracción (R3). Se concluyó que el proceso de pirólisis de los residuos de mamoncillo se puede describir utilizando un mecanismo de reacción complejo, y que estos residuos presentan propiedades térmicas con potencial para producir bioenergía.

Palabras clave: túneles, método de elementos finitos, Mohr-Coulomb, elástico-elastoplástico

Received: February 1st, 2023
Accepted: July 21st, 2023

Introduction

The growing global demand for electricity, thermal energy, and fuels, the deterioration of the environment, and the depletion of fossil fuel reserves have led to the search for alternative energy sources, mainly of lignocellulosic origin. One of these alternatives is lignocellulosic biomass waste (*i.e.*, agro-industrial waste), which is characterized by being a low-cost, highly available, renewable, and sustainable material (Pacheco *et al.*, 2022). Agro-industrial waste includes fruit byproducts, which mainly consist of peels and seeds. Fruit byproducts are obtained from direct consumption and/or processing to obtain juices, jellies, wines, pulps, *etc.* (Lam *et al.*, 2016). This biomass is composed of cellulose (32-45%), hemicellulose (19-25%), lignin (14-26%), extractives, and ash (Rony *et al.*, 2019).

---

[1] PhD in Engineering – Chemical Engineering, Universidad del Valle, Colombia. Affiliation: Associate professor, Universidad Nacional de Colombia, Chemical Engineering Department, Colombia. E-mail: anfrojasgo@unal.edu.co

[2] PhD in Engineering, Universidad del Valle, Colombia. Affiliation: Technical director of the Fuels Combustion Laboratory, Universidad del Valle, Colombia. E-mail: francisco.velasco@correounivalle.edu.co

One of the fruits produced and directly consumed in the intertropical zone of America is *mamoncillo*, also known as *mamón* or *quenepa* (Calderón *et al.*, 2021). This fruit generates peels and seeds as waste, which represent 65% of the whole fruit. *Mamoncillo* seeds have been studied to determine their polyphenol content and antioxidant capacity and as a source of starch (Moo-Huchin *et al.*, 2020). These seeds have been used for the extraction of dyes with applications in fabric dyeing (Vejar *et al.*, 2016), and studies have been conducted on their total flavonoids content, total phenolic compounds and, antioxidant activity have been evaluated (Can-Cauich *et al.*, 2017).

Lignocellulosic biomass can be converted to solid, liquid, and gaseous products, which are employed to obtain fuels or value-added products (Kumar *et al.*, 2020). These products can be obtained through several technological routes, which can be categorized as *thermochemical*, *biochemical*, and *physicochemical* (Bridgwater, 2012). Thermochemical conversion processes can be further subdivided into combustion, gasification, liquefaction, and pyrolysis. They are used to transform biomasses into bio-oil, gaseous fuel, and biochar (Emiola-Sadiq *et al.*, 2021). Pyrolysis is the simplest thermochemical conversion process to produce biochar, bio-oils, and syngas, which has industrial and ecological importance and plays an important role in the development of bioenergy systems (Bensidhom *et al.*, 2021). This conversion process is considered to be efficient and low-cost, and it is classified as slow, fast, and flash pyrolysis depending on the heating rate and residence time (Kan *et al.*, 2012). Fast and flash pyrolysis are used to transform biomass into bio-oil, while slow pyrolysis transforms biomass into gaseous fuel and charcoal (Gogoi *et al.*, 2018).

To better understand pyrolysis and its operation, it is essential to know the kinetic and thermodynamic parameters involved, as well as the thermal characteristics of the biomass (Yang *et al.*, 2019; Pawar *et al.*, 2021). Thermogravimetric analysis (TGA) is a powerful technique to understand the biomass pyrolysis process. This technique gives detailed information about reaction mechanisms, stability, reactivity, thermodynamic parameters, and decomposition kinetics (Emiola-Sadiq *et al.*, 2021). TGA can be categorized as isothermal or non-isothermal. Non-isothermal analysis is carried out for linear heating rates, which involves heating the biomass from room temperature to a desired temperature at a desired constant heating rate (Mishra and Bhaske, 2014). The data obtained from non-isothermal TGA can be handled using isoconversional (model-free methods) or model-fitting methods (Gogoi *et al.*, 2018). The most commonly employed model-free techniques are the Kissinger-Akahira-Sunose (KAS), Flynn-Wall-Ozawa (FWO), Starink, and Friedman methods. These are used to calculate kinetic parameters as recommended by the Kinetics Committee of the International Confederation for Thermal Analysis and Calorimetry (ICTAC) (Vyazovkin *et al.*, 2011). To predict mechanisms or kinetic models, the ICTAC recommends the Coats-Redfern integral method and master-plots associated with Criado method (Santos *et al.*, 2020).

This study presents the thermal, kinetic, and thermodynamic analysis of *mamoncillo* (*Melicoccus bijugatus*) waste (peels and seeds) by means of pyrolysis using non-isothermal thermogravimetry at three heating rates (10, 20, and 40 °C/min). The kinetic parameters (activation energy, E_a ; pre-exponential factor, A_a) of *mamoncillo* waste were determined via three integral isoconversional methods (KAS, FWO, and Starink) and a differential isoconversional method (Friedman). The mechanism or kinetic model for the thermal degradation reaction of the *mamoncillo* waste was established using the master-plots method and 18 different reaction mechanisms. Other thermodynamic parameters were determined, such as enthalpy (ΔH), Gibbs free energy (ΔG), and entropy (ΔS).

Methods

Materials

Mamoncillo peels and seeds were collected from fruits and vegetables stores in the city of Manizales, Colombia. These wastes were reduced in size to dimensions of less than 1 cm. Then, they were dried at 45 °C until they reached constant weight in order to avoid the loss of substances with low molecular weight. Afterwards, they were reduced in size using a disc mill to obtain a powder of less than 250 µm (-60 mesh). Finally, the powdered samples were stored in plastic bags and placed in a desiccator to prevent moisture absorption.

Feedstock characterization

The powdered *mamoncillo* wastes were characterized for proximate, ultimate, and higher heating value (HHV) analysis. These tests were carried out in triplicate, and the standard deviation of the data was determined. Proximate, moisture, ash, and volatile matter analyses were carried out according to the ASTM-E871-82 (2019), ASTM E872-82 (2019), and ASTM E1577-11 standards. The fixed carbon content was determined by difference. The elemental composition of carbon, hydrogen, and nitrogen was quantified using a LECO-CHN628 analyzer, while the amount of sulfur was obtained using a LECO-S632 analyzer. The oxygen content was determined by difference, and the HHV was obtained using an SDACM3100 bomb calorimeter.

Thermogravimetric analysis

Thermal analysis experiments were performed in a TA-Instrument-Q600 simultaneous TGA-DSC thermogravimetric analyzer. Pyrolysis was carried out at three heating rates (10, 20, and 40 °C/min), heating between room temperature and 900 °C in an inert atmosphere (nitrogen) and with a gas flow of 100 ml/min. The experiments were carried out at a high nitrogen flow rate to eliminate some secondary reactions (Chen *et al.*, 2017). To reduce the limiting steps of the degradation reaction due to mass and heat transfer, a sample size of 15-16 mg was maintained. Meanwhile, to reduce diffusion limitations, the samples were maintained at a particle size of less than 250 µm. This analysis was carried out in duplicate. The TGA results were used to obtain kinetic and thermodynamic parameters.

Kinetic analysis

In this work, the kinetic parameters and the reaction mechanism for the thermal degradation of *mamoncillo* peels and seeds were determined. The thermal decomposition of lignocellulosic biomass is a heterogeneous process because the biomass components undergo thermochemical reactions in the solid state, with very complex reaction mechanisms (Santos *et al.*, 2020). Therefore, determining the pyrolysis kinetics of biomass requires the use of different reaction kinetic models.

The so-called *model-free methods* are based on determining E_a without knowing the reaction's kinetic model (Chen *et al.*, 2017; Liu *et al.*, 2020). They are also known as *isoconversional methods* because i) E_a is obtained as a function of conversion, and ii), as an isoconversional principle, at a particular conversion, the reaction rate for the thermal decomposition of biomass is function of temperature (non-isothermal analysis) (Singh *et al.*, 2020a). Among the most used isoconversional methods to calculate kinetic parameters are the KAS, FWO, and Starink integral methods, in addition to the Friedman differential method (Singh *et al.*, 2020a). The Equations of these methods are presented below:

$$ln\left(\frac{\beta}{T_\alpha^2}\right)=ln\left[\frac{A_aR}{E_ag(\alpha)}\right]-\frac{E_a}{RT_\alpha} \qquad (1)$$

FWO method (Santos *et al.*, 2020):

$$ln(\beta)=ln\left[\frac{A_aR}{E_ag(\alpha)}\right]-5,331-1,052\left(\frac{E_a}{RT_\alpha}\right) \qquad (2)$$

Starink method (Singh *et al.*, 2020a):

$$ln\left(\frac{\beta}{T_\alpha^{1,92}}\right)=C-1.008\left(\frac{E_a}{RT_\alpha}\right) \qquad (3)$$

Friedman method (Gogoi *et al.*, 2018):

$$ln\left(\beta\frac{d\alpha}{dT}\right)=ln\left[A_af(\alpha)\right]-\frac{E_a}{RT_\alpha} \qquad (4)$$

where α is the fractional conversion; E_a, A_a, and T_a are the activation energy, pre-exponential factor, and absolute temperature for a value α, respectively; β is the heating rate; *R* is the universal gas constant (8,314 J/kmol.K); $g(\alpha)$ is a function that describes a reaction model in its integral form; $f(\alpha)$ is a function that describes a reaction model in its differential form; and *C* is a constant. Table 1 presents the algebraic expression of 18 different reaction models, which are a function of $f(\alpha)$ and $g(\alpha)$.

Here, *f*(0,5) and *g*(0,5) are functions that describe various reaction models in differential and integral form, respectively. They are evaluated at a fractional conversion of 0,5. $T_{0,5}$ is the absolute temperature at a conversion value of 0,5. The *Z*(α) master-plot consists of theoretical curves from different reaction kinetic models for solid-state degradation reactions, which are independent of the kinetic parameters (Wang *et al.*, 2019; Singh *et al.*, 2021). In these curves, the integral and differential forms of the reaction models are combined. These curves are used to identify the most appropriate theoretical kinetic model that describes an experimental reaction process.

To predict the mechanisms or kinetic model of thermal degradation reactions, the recommendations given by the ICTAC (Vyazovkin *et al.*, 2011) were followed, which involve using the master-plots method in association with the Criado method (Santos *et al.*, 2020). To this effect, the following Equation was used (Gogoi *et al.*, 2018; Santos *et al.*, 2020; Aboulkas *et al.*, 2010):

$$\frac{Z(\alpha)}{Z(0,5)}=\frac{f(\alpha)g(\alpha)}{f(0.5)g(0,5)}=\left(\frac{T_\alpha}{T_{0,5}}\right)^2\left[\frac{(d\alpha/dt)_\alpha}{(d\alpha/dt)_{0,5}}\right] \qquad (5)$$

The theoretical master-plots were constructed by plotting for each kinetic model, while the experimental master-plot was obtained by plotting with experimental data (Santos *et al.*, 2020). Both terms of Equation (5) were plotted as a function of fractional conversion. The coincidence of a theoretical master-plot with the experimental one indicates that the thermal degradation of biomass is carried out by means of the reaction kinetic model given by the theoretical master curve (Aboulkas *et al.*, 2010).

$$\frac{f(\alpha)g(\alpha)}{f(0,5)g(0,5)} \qquad \left(\frac{T_\alpha}{T_{0,5}}\right)^2\left[\frac{(d\alpha/dt)_\alpha}{(d\alpha/dt)_{0,5}}\right]$$

In this research, the four isoconversional methods (KAS, FWO, Starink, and Friedman) were used to determine E_a. A_a was calculated by means of Equation (6), while the reaction mechanism or kinetic model (*f*(α)) of the thermal decomposition for *mamoncillo* waste was obtained using master-plots.

Thermodynamic parameters

The thermodynamic analysis of the mamoncillo waste was carried out by calculating the enthalpy (ΔH), Gibbs free

Table 1. Thermal degradation kinetic models in the solid state

Reaction Models	*f(α)*	*g(α)*
Nucleation models		
Power law (P2)	$2\alpha^{1/2}$	$\alpha^{1/2}$
Power law (P3)	$3\alpha^{2/3}$	$\alpha^{1/3}$
Power law (P2/3)	$2/3\alpha^{-1/2}$	$\alpha^{3/2}$
Power law (P4)	$4\alpha^{3/4}$	$\alpha^{1/4}$
Avarami-Erofeev (A1)	$(3/2)(1-\alpha)[-\ln(1-\alpha)]^{1/3}$	$[-\ln(1-\alpha)]^{2/3}$
Avarami-Erofeev (A2)	$2(1-\alpha)[-\ln(1-\alpha)]^{1/2}$	$[-\ln(1-\alpha)]^{1/2}$
Avarami-Erofeev (A3)	$3(1-\alpha)[-\ln(1-\alpha)]^{2/3}$	$[-\ln(1-\alpha)]^{1/3}$
Avarami-Erofeev (A4)	$4(1-\alpha)[-\ln(1-\alpha)]^{3/4}$	$[-\ln(1-\alpha)]^{1/4}$
Diffusion models		
One-dimensional diffusion (D1)	$1/(2\alpha)$	α^2
Two dimensional (D2) - (Valensi model)	$-[1/\ln(1-\alpha)]$	$[(1-\alpha)\ln(1-\alpha)]+\alpha$
Diffusion control (D3) - (Jander model)	$[3(1-\alpha)^{2/3}]/[2(1-(1-\alpha)^{1/3})]$	$[1-(1-\alpha)^{1/3}]^2$
Diffusion control (D4) - (Ginstling-Brounshtein model)	$(3/2)[((1-\alpha)^{-1/3}-1)]^{-1}$	$1-(2/3)\alpha-(1-\alpha)^{2/3}$
Reaction order models		
Zero order (F0)	1	α
First order (F1) – (Mampel Model)	$(1-\alpha)$	$-\ln(1-\alpha)$
Second order (F2)	$(1-\alpha)^2$	$[1/(1-\alpha)]-1$
Third order (F3)	$(1-\alpha)^3$	$(1/2)[(1-\alpha)^{-2}-1]$
Geometrical contraction models		
Contracting cylinder (R2)	$2(1-\alpha)^{1/2}$	$1-(1-\alpha)^{1/2}$
Contracting sphere (R3)	$3(1-\alpha)^{2/3}$	$1-(1-\alpha)^{1/3}$

Source: Emiola-Sadiq *et al.* (2021), Mishra and Bhaske (2014), Vyazovkin *et al.* (2011)

energy (ΔG), and entropy (ΔS) via the KAS, FWO, Starink, and Friedman methods. These parameters, together with E_a and A_a, allow better understanding the pyrolysis process of biomass. This analysis is fundamental to establishing the feasibility and efficiency of pyrolysis, as well as for the calculation of energy consumption or requirements (Singh *et al.*, 2020a). A, ΔH, ΔG, and ΔS, are given by the following Equations (Singh *et al.*, 2021; Raza *et al.*, 2022):

$$A_a = \frac{\beta E_a}{RT_m^2} e^{\left(\frac{E_a}{RT_m}\right)} \tag{6}$$

$$\Delta H = E_a - RT_m \tag{7}$$

$$\Delta G = E_a + RT_m \, ln\left(\frac{k_B T_m}{h\, A_a}\right) \tag{8}$$

$$\Delta S = \frac{\Delta H - \Delta G}{T_m} \tag{9}$$

where E_a is the activation energy obtained via different isoconversional methods; β denotes the heating rates; *R* is the universal gas constant (8,314 J/kmol.K); T_m is the peak temperature in the DTG curve; k_b is the Boltzmann constant (1,3819 x 10^{-23} J/K); and *h* is the Planck constant (6,6269 x 10^{-34} J.s) (Kumar *et al.*, 2020; Santos *et al.*, 2020).

Results and discussion

Waste characterization

The results obtained for the proximate analysis on a dry basis, the ultimate analysis on a dry ash free basis, the atomic ratios, and the HHV of *mamoncillo* wastes are shown in Table 2. Note that the total moisture, ash, and fixed carbon content are higher in the peels than in the seeds. *Mamoncillo* peels have four times more ash content than the seeds. The peels can cause issues related to slags, corrosion, or clogging in the equipment (Chen *et al.*, 2018). Additionally, a high ash content can cause poor heat generation and reduced energy conversion (Singh *et al.*, 2020a). The high volatile matter contents in the seeds should facilitate ignition at low temperatures and benefit the formation of gaseous and liquid products during the pyrolysis process (Santos *et al.*, 2020). These wastes have the same fixed carbon content, and they are in the range required for efficient burning, which is between 15 and 25% (Santos *et al.*, 2020). The empirical formulas for the peels and seeds, as obtained from the ultimate analysis, are $C_{20}H_{29,5}O_{13,7}N_{0,20}S_{0,04}$ and $C_{20}H_{29,4}O_{13,9}N_{0,14}S_{0,03}$, respectively. Both wastes have the same empirical formula. These similar results lead to similar values for the the atomic ratios H/C and O/C and the HHV. However, the little difference in HHV is due to the fact that the *mamoncillo* peels have a higher ash content (Rojas and Flórez, 2019).

Table 2. Proximate and ultimate analysis of *mamoncillo* waste

	Peels	*Seeds*
Proximate analysis, %		
Total moisture	66,99±2,23	59,54±1,78
Ash*	5,23±0,51	1,33±0,12
Volatile mater*	77,12±1,75	83,54±1,92
Fixed carbon*	17,65±0,83	15,13±0,89
Ultimate analysis, %**		
Carbon – C	48,73±1,07	48,55±1,15
Hydrogen – H	6,04±0,11	6,00±0,13
Oxygen – O	44,37±0,92	44,89±1,01
Nitrogen – N	0,58±0,02	0,29±0,01
Sulfur – S	0,39±0,01	0,18±0,01
Atomic ratio H/C	1,49±0,07	1,40±0,09
Atomic ratio O/C	0,68±0,03	0,69±0,02
High heating value-HHV, MJ/kg*	18,08±0,79	19,09±0,94

*Dry basis. **Dry ash free basis.
Source: Authors

Thermogravimetric analysis

Figure 1 shows the TG/DTG curves for the peels (Figure 1a) and the seeds (Figure 1b) at three heating rates. Note that thermal decomposition of *mamoncillo* waste can be described using three main zones (Santos *et al*., 2020). The first zone indicates the drying of wastes, and it is called the *passive zone*, where the total moisture is removed (Wang *et al*., 2019; Singh *et al*., 2021). This zone is between 30 and 170 °C for peels and 30 and 190 °C for seeds. In the second zone, the *main active pyrolysis zone* (Emiola-Sadiq *et al*., 2021), the elimination of volatile matter takes place, which happens between 170 and 650 °C for peels and 190 and 590 °C for seeds. In the last zone, there is char formation, with temperatures of >650 and >590 °C for peels and seeds, respectively. The peak temperatures for the three different heating rates are 303,22, 314,50, and 326,20 °C for peels, and 291,49, 303,17, and 315,56 °C for seeds. These results indicate that *mamoncillo* seeds are pyrolytically more reactive than *mamoncillo* peels.

Kinetic analysis

The kinetic parameters corresponding to the pyrolysis of *mamoncillo* waste were evaluated by calculating E_a and A_a via the KAS, FWO, Starink, and Friedman isoconversional methods. To obtain the kinetic properties, only the second zone in the DTG curve was considered, as it is characterized by the highest mass loss has the fraction that is considered to be pyrolysable (Mehmood *et al*., 2017).

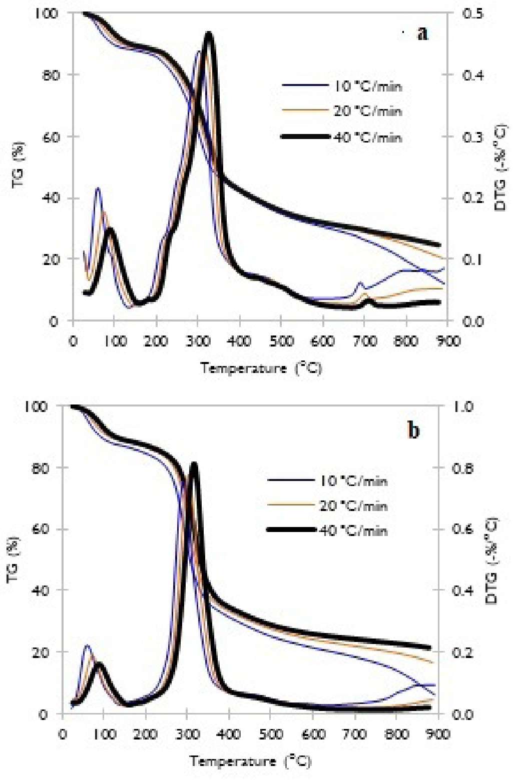


Figure 1. TG/DTG curves at different heating rates for a) *mamoncillo* peels and b) *mamoncillo* seeds
Source: Authors

Figures 2a and 2b show the variations of E_a with conversion (α=0,05-0,80) as estimated via the isoconversional methods for *mamoncillo* peels and seeds. Regarding mamoncillo peels, it was found that E_a varies with conversion ranges of 132,94-365,40, 134,20-357,21, 132,21-362,92 and 152,24-578,27 kJ/mol for the KAS, FWO, Starink, and Friedman methods, respectively. As for the seeds, E_a varies with conversion ranges of 120,07-389,62, 122,03-380,32, 119,44-386,94, and 135,13-491,09 kJ/mol. The average E_a for the peels are 240,15, 237,36, 238,63, and 296,84 kJ/mol for these methods. For the *mamoncillo* seeds, the average E_a are 198,22, 197,49, 197,10, and 228,25 kJ/mol. It was also found that, for the KAS, FWO, and Starink methods, the activation energy deviation from the average E_a was between 29 and 31% for the peels and from 33 to 35% for the seeds. For the Friedman method, these deviations were 47 and 42% for peels and seeds, respectively. These results indicate that the integral methods were more accurate than the differential method (Vyazovkin *et al*., 2011).

During the pyrolysis of *mamoncillo* waste, E_a showed little variations with regard to the KAS, FWO, and Starink methods. For the peels, the average values obtained were 240,15, 237,36, and 238,63 kJ/mol, respectively. For the seeds, these values were 198,22, 197,49, and 197,10 kJ/mol. Meanwhile, the Friedman method reported significant variations, with values of 296,84 and 228,25 kJ/mol for the peels and seeds, respectively. These results indicate that *mamoncillo* seeds are more reactive than *mamoncillo* peels during the pyrolysis process (Singh *et al*., 2020a). Moreover, Figure 2 shows that conversions of up to 0,20 and low E_a values (184 kJ/mol for peels and 161 kJ/mol for seeds) are needed to break the weaker bonds between molecules and remove light components. This indicates the start of the pyrolysis reaction (Pawar *et al*., 2021). For conversions between 0,20 and 0,70, the value of E_a increases to 270 kJ/mol for peels and 201 kJ/mol for seeds. This may be mainly due to cellulose as well as hemicellulose pyrolysis (Santos *et al*., 2020). For conversions greater than 0,70, a higher E_a was found, which is due to lignin pyrolysis (Pawar *et al*., 2021).

As mentioned earlier, model-free methods can generally be split into two categories: differential and integral. These are used for estimating kinetic parameters (pre-exponential factor, activation energy, and reaction order) (Vyazovkin *et al*., 2011). The activation energy may show differences depending on the method used for calculation. These differences are due to the inherent and unavoidable inaccuracy and imprecision of differential methods. Therefore, they can sometimes exhibit numerical instability when compared to integral methods; when the reaction heat varies noticeably with the temperature program and when they are applied to differential data, their accuracy can be limited due to the difficulty in determining the baseline (Vyazovkin *et al*., 2011; Cai *et al*., 2018). The Friedman method is the most general isoconversional differential technique, and it requires derivative conversion data, which leads it to be numerically unstable and sensitive to noise

(Cai *et al.*, 2018). In fact, in this work, the Friedman method showed a higher average E_a than KAS, FWO, and Starink for both *mamoncillo* wastes. This behavior can also be explained by the fact that the Friedman method normally depends on the instantaneous rate of biomass conversion without any pre-assumptions, which it is not possible in integral methods (Pawar *et al.*, 2021). The authors used this method for the sake of comparison, but they do not recommend using it.

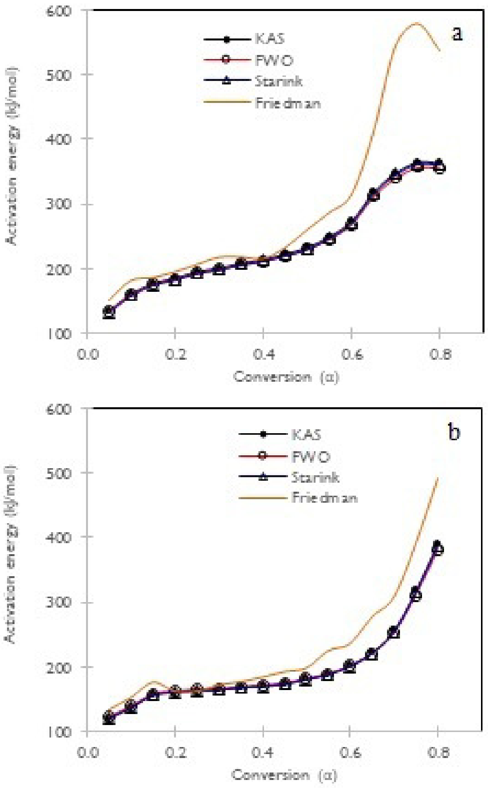


Figure 2. Activation energy at different conversions using the KAS, FWO, Starink and Friedman methods for. a) *Mamoncillo* peels, b) *mamoncillo* seeds.
Source: Authors

Table 3 presents the variations in A_a with conversion (α=0,05-0,80) at 10 °C/min as estimated via isoconversional methods for *mamoncillo* waste. This factor, also called the *frequency factor*, measures the frequency of active collision that occurs between the reactant molecules. It is used to explain the chemistry of the pyrolysis reaction, and it serves to optimize the experimental conditions of pyrolytic processes (Santos *et al.*, 2020). It is known that, for $A_a < 10^9$ s^{-1}, only surface reactions are taking place, while $A_a > 10^9$ s^{-1} indicates that i) there is a complex reaction, which does not depend on the contact surface area; ii) a high collision of molecules is required; and iii) a high E_a is needed for biomass pyrolysis (Kaur *et al.*, 2017). In our study, the A_a values calculated via the KAS, FWO, and Starink methods were found to be between 10^9 and 10^{31} s^{-1} for peels and between 10^9 and 10^{34} s^{-1} for seeds. It was found that A_a varies with conversion as E_a changes, *i.e.*, A_a increases when the conversion is higher.

It was also found that A_a decreases with the increase in heating rate. This indicates that, at low conversions (<0,20), the reactivity of the thermal decomposition of *mamoncillo* wastes is low, with the highest values at high conversion (Singh *et al.*, 2020a).

Reaction kinetic models

We predicted the reaction model during the pyrolysis of mamoncillo waste using 18 reaction kinetic models (solid-state mechanisms) with the corresponding experimental and theoretical master-plot curves. The experimental and theoretical curves at 10 °C/min are presented in Figure 3.

Table 3. Pre-exponential factor (A_a) for *mamoncillo* wastes *vs.* the degree of conversion at 10 °C/min

	A_a, s^{-1} - Peels			A_a, s^{-1} - Seeds		
α	**KAS**	**FWO**	**Starink**	**KAS**	**FWO**	**Starink**
0,05	8,96x10^{9}	1,18x10^{10}	7,65x10^{9}	9,68x10^{9}	1,49x10^{9}	8,43x10^{9}
0,10	2,83x10^{12}	2,98x10^{12}	2,32x10^{12}	4,06x10^{10}	5,80x10^{10}	3,45x10^{10}
0,15	1,00x10^{14}	9,25x10^{13}	7,98x10^{13}	3,11x10^{12}	3,76x10^{12}	2,56x10^{12}
0,20	5,60x10^{14}	4,95x10^{14}	4,42x10^{14}	7,89x10^{12}	9,40x10^{12}	6,46x10^{12}
0,25	5,25x10^{15}	4,30x10^{15}	4,07x10^{15}	1,43x10^{13}	1,69x10^{13}	1,17x10^{13}
0,30	1,65x10^{16}	1,32x10^{16}	1,27x10^{16}	1,85x10^{13}	2,20x10^{13}	1,51x10^{13}
0,35	9,13 x10^{16}	6,86x10^{16}	6,95x10^{16}	3,50x10^{13}	4,09x10^{13}	3,69x10^{13}
0,40	2,91x10^{17}	2,12x10^{17}	2,20x10^{17}	6,25x10^{13}	7,20x10^{13}	5,05x10^{13}
0,45	1,92x10^{18}	1,30x10^{18}	1,43x10^{18}	1,33x10^{14}	1,49x10^{14}	1,07x10^{14}
0,50	1,61x10^{19}	1,00x10^{19}	1,18x10^{19}	6,62x10^{14}	6,98x10^{14}	5,25x10^{14}
0,55	5,70x10^{20}	3,03x10^{20}	4,06x10^{20}	3,62x10^{15}	3,56x10^{15}	2,84x10^{15}
0,60	8,86x10^{22}	3,75x10^{22}	6,09x10^{22}	5,97x10^{16}	5,21x10^{16}	4,59x10^{16}
0,65	1,25x10^{27}	3,39x10^{26}	7,99x10^{26}	4,56x10^{18}	3,28x10^{18}	3,39x10^{18}
0,70	8,03x10^{29}	1,64x10^{29}	4,87x10^{29}	6,37x10^{21}	3,29x10^{21}	4,48x10^{21}
0,75	2,80x10^{31}	5,10x10^{30}	1,66x10^{31}	4,85x10^{27}	1,33x10^{27}	3,07x10^{27}
0,80	2,33x10^{31}	4,63x10^{30}	1,39x10^{31}	2,72x10^{34}	3,66x10^{33}	1,53x10^{34}

Source: Authors

Regarding the *mamoncillo* peels, it can be seen (Figure 3a) that, for conversions from 0,05 to 0,50, the experimental curve almost overlaps the R2 model or the contracting cylinder model, and, for conversion ranges between 0,50 and 0,80 the experimental curve is close to the F3 model or the third-order reaction model. It can be concluded that the decomposition of *mamoncillo* peels in each conversion range (0,05-0,50 and 0,50-0,80) can be described by a single reaction model. Meanwhile, regarding the *mamoncillo* seeds (Figure 3b), note that, for conversions between 0,05 and 0,50, the experimental curve is close to the F2 model or the second-order reaction model, and, for conversion ranges of 0,50-0,80, the experimental curve almost overlaps the R3 model or the contracting sphere model. This indicates that the pyrolysis of *mamoncillo* seeds can be described by a complex reaction mechanism.

Thermodynamic parameters

The thermodynamic parameters ΔH, ΔG, and ΔS involved in the thermal decomposition of mamoncillo wastes were calculated using the activation energy obtained via the KAS, FWO, Starink, and Friedman methods. The values of ΔH, ΔG, and ΔS were calculated using Equations (7) to (9). The changes in ΔH, ΔG and ΔS with the conversion value at 10 °C/min are reported in Tables 4-9, respectively. ΔH is the minimum energy required by the biomass to form products during pyrolysis (Singh *et al.*, 2021; Kaur *et al.*, 2017). Thereupon, ΔH evaluates the total energy consumed in the conversion of *mamoncillo* waste into products (Santos *et al.*, 2020). ΔG represents the increase in the system's total energy for the formation of activated complexes (Maia and Morais, 2016), and it can be used to determine whether the reactions are spontaneous. A negative ΔG value indicates the occurrence of spontaneous reactions, while a positive value denotes a nonspontaneous reaction (Chen *et al.*, 2023). Meanwhile, ΔS is associated with the measure of randomness or disorder of energy and matter in a system, as well as to the formation of new chemical compounds (Kumar *et al.*, 2020).

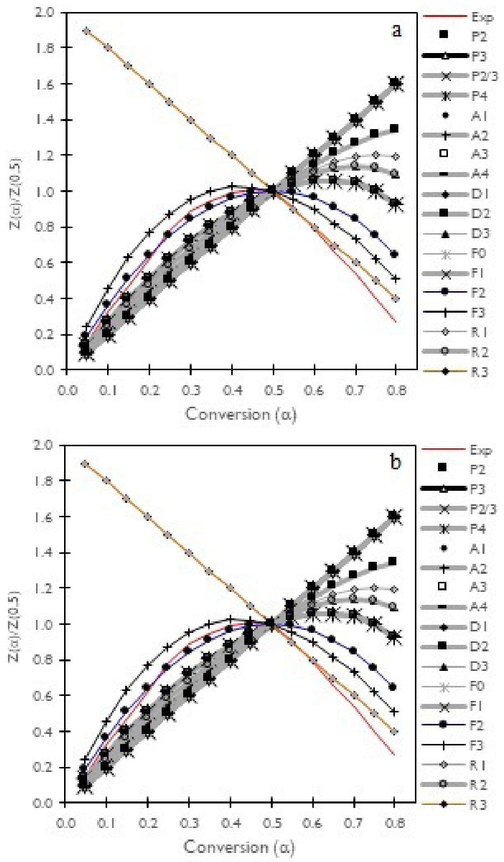


Figure 3. Master-plots for different kinetic models and experimental data at 10 °C/min for a) mamoncillo peels and b) *mamoncillo* seeds
Source: Authors

In Tables 4 and 5, it can be seen that, for both wastes, i) ΔH increases with conversion for all methods except for Friedman; ii) the average values are similar for the KAS, FWO, and Starink methods; iii) the H values for the Friedman method are higher than those of KAS, FWO, and Starink; iv) the ΔH values for *mamoncillo* seeds are lower than those reported for *mamoncillo* peels; and v) all ΔH values are positive.

Positive ΔH values indicate that the thermal degradation of *mamoncillo* wastes took place via endothermic reactions. This means that an external source of heat is required to transform biomass into gas, oil, and charcoal during the pyrolysis process (Chen *et al.*, 2017; Kumar *et al.*, 2020). It was found that the difference between the values of E_a and ΔH at each conversion level is less than 5 kJ/mol. This small difference indicates the formation of low-energy activated complexes in the product-forming potential barrier (Santos *et al.*, 2020). Therefore, the pyrolysis of *mamoncillo* wastes for energy generation is viable (Kumar *et al.*, 2020; Santos *et al.*, 2020). Moreover, for the *mamoncillo* peels, the average ΔH values at three heating rates (10, 20, and 40 °C/min) for the KAS, FWO, Starink, and Friedman methods were 235,27 ± 0,10, 232,47 ± 0,10, 233,74 ± 0,10, and 291,95 ± 0,10 kJ/mol, respectively. For *mamoncillo* seeds, these values were 193,42 ± 0,10, 192,70 ± 0,10, 192,30 ± 0,10, and 223,46 ± 0,10 kJ/mol, respectively. These results show that the effect of the heating rate on ΔH is negligible. Similar results were reported for banana leaves (Singh *et al.*, 2020a) and acai seeds (Santos *et al.*, 2020). It can be seen that *mamoncillo* seeds consume less total energy than *mamoncillo* peels during their conversion into products.

Table 4. ΔH variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔH, kJ/mol				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	128,14	129,41	127,42	147,45
0,10	154,84	155,08	153,92	178,07
0,15	171,45	171,09	170,40	182,49
0,20	179,49	178,92	178,39	191,49
0,25	189,95	189,02	188,77	202,62
0,30	195,32	194,26	194,10	214,00
0.,35	203,32	201,99	202,04	214,42
0,40	208,76	207,26	207,44	211,76
0,45	217,60	215,77	216,22	228,82
0,50	227,59	225,36	226,13	256,15
0,55	244,34	241,37	242,75	282,66
0,60	268,09	264,04	266,32	310,44
0,65	313,15	306,98	311,02	406,55
0,70	343,68	336,18	341.,2	539,85
0,75	360,48	352,42	357,99	573,48
0,80	359,61	351,96	357,15	532,53
Average	**235,36**	**232,57**	**233,84**	**292,05**

Source: Authors

The variations in ΔG with the conversion level for *mamoncillo* peels and seeds as obtained via the KAS, FWO, Starink, and Friedman methods are shown in Tables 6 and 7. The average ΔG values for *mamoncillo* waste are in the range between 162 and 164 kJ/mol. The ΔG values are similar for both wastes. However, the average ΔG values the peels were slightly higher than those obtained for the seeds, which means that, for thermal decomposition, peels require more heat. Similar results have been reported for the pyrolysis of agricultural residues (Chen *et al.*, 2017). In addition, ΔG was found to decrease with conversions of 0,05-0,80, implying that the energy of the reaction system decreased when the pyrolysis started (Yuan *et al.*, 2017). It was also found that, for both wastes, ΔG is positive, which indicates that the thermal decomposition reaction is nonspontaneous. Therefore, this reaction requires energy consumption for the chemical bonds to break. This result suggests that the reactivity of the thermal degradation was low and that the pyrolysis process requires an external energy source to obtain the activated complex (Chen *et al.*, 2017).

Table 5. ΔH variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔH, kJ/mol				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	115,37	117,33	114,75	130,43
0,10	132,30	133,91	131,56	149,20
0,15	152,03	152,90	151,14	172,37
0,20	156,27	157,08	155,36	158,44
0,25	158,99	159,76	158,06	158,40
0,30	160,17	160,96	159,24	168,43
0,35	163,08	163,79	163,32	173,13
0,40	165,72	166,37	164,75	180,41
0,45	169,16	169,71	168,16	188,74
0,50	176,51	176,76	175,46	194,48
0,55	184,29	184,22	183,18	220,71
0,60	197,14	196,52	195,93	232,08
0,65	217,05	215,54	215,69	273,03
0,70	250,39	247,34	248,77	303,86
0,75	312,94	306,94	310,82	386,87
0,80	384,93	375,62	382,25	486,36
Average	**193,52**	**192,80**	**192,40**	**223,56**

Source: Authors

Table 6. ΔG variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔG, kJ/mol				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	167,45	167,40	167,47	349,76
0,10	166,57	166,56	166,60	379,58
0,15	166,09	166,10	166,12	382,14
0,20	165,88	165,90	165,91	390,54
0,25	165,62	165,64	165,64	401,05
0,30	165,49	165,51	165,51	411,65
0,35	165,30	165,33	165,33	411,21
0,40	165,17	165,21	165,20	407,75
0,45	164,98	165,02	165,01	424,66
0,50	164,77	164,82	164,80	452,14
0,55	164,44	164,49	164,47	478,84
0,60	164,00	164,07	164,03	506,77
0,65	163,27	163,36	163,30	604,22
0,70	162,83	162,93	162,86	740,10
0,75	162,60	162,71	162,63	776,34
0,80	162,61	162,71	162,62	736,82
Average	**164,82**	**164,86**	**164,85**	**490,85**

Source: Authors

Tables 8 and 9 present the variations in ΔS with the conversion level for *mamoncillo* peels and seeds as obtained via the KAS, FWO, Starink, and Friedman methods. ΔS is negative at conversion of up to 0,10 for the peels, and the seeds exhibit negative values up to a fractional conversion of 0,30. Low ΔS values indicate that these wastes undergo small physical and chemical charges during thermal degradation, transitioning to a new condition close to thermodynamic equilibrium (Dhyani *et al.*, 2017), increasing the pyrolysis reaction time (Singh *et al.*, 2020b). For high ΔS values, the reactivity is high, the reaction time decreases, and the process is far from thermodynamic equilibrium (Singh *et al.*, 2021; Mallick *et al.*, 2018). Here, ΔS increased when the fractional conversion increased for the KAS, FWO, and Starink methods and both wastes, implying that the pyrolysis was far from reaching equilibrium. The ΔS values calculated via the Friedman method are all negative for *mamoncillo* waste. ΔS followed similar trend as ΔH and E_a, which increased with the increasing fractional conversion. This can be attributed to the fact that the reaction rate increased for the 0,05-0,80 conversions (Singh *et al.*, 2021).

Table 7. ΔG variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔG, kJ/mol				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	164,23	164,15	164,25	332,46
0,10	163,61	163,55	163,63	347,15
0,15	162,98	162,95	163,00	367,27
0,20	162,85	162,83	162,88	350,56
0,25	162,77	162,75	162,80	349,12
0,30	162,74	162,71	162,76	358,58
0,35	162,66	162,64	162,65	362,74
0,40	162,58	162,56	162,61	369,80
0,45	162,49	162,47	162,52	378,17
0,50	162,29	162,29	162,32	383,90
0,55	162,10	162,10	162,12	411,22
0,60	161,79	161,80	161,82	423,15
0,65	161,35	161,38	161,37	465,50
0,70	160,69	160,74	160,72	497,33
0,75	159,66	159,75	159,69	581,91
0,80	158,70	158,81	158,73	684,78
Average	**162,09**	**162,09**	**162,12**	**416,48**

Source: Authors

The knowledge of thermogravimetric, kinetic, and thermodynamic parameters is fundamental in determining the feasibility and efficiency of the pyrolysis process, designing pyrolyzers, determining the energy balance, and calculating the energy consumption or requirements (Kumar *et al.*, 2020, Singh *et al.*, 2020a). These parameters can be used to determine the optimal operating conditions to obtain a specific pyrolysis product (bio-oil, biochar, and syngas) (Li *et al.*, 2023) from *mamoncillo* wastes. Therefore, this study can contribute to understanding *mamoncillo* waste pyrolysis and its future applications. For example, it was found that these wastes have an ignition temperature (Figure 1) between 210 °C (peels) and 250 °C (seeds), as well as high volatile matter contents (between 77 and 84% on a dry basis). This indicates that the pyrolysis of these wastes can produce more bio-oil than biochar and syngas (Nawaz *et al.*, 2021). On the other hand, by controlling the operating conditions, it is possible to produce a higher proportion of biochar than bio-oil and syngas, or a higher proportion of syngas than bio-oil and biochar. These wastes have different applications. Bio-oil can be used as solid biofuel for cooking purposes and domestic heating, or as a solvent for extracting important chemical compounds. Biochar can be used as a soil and water conditioner (it is a good biosorbent), as a solid fuel for cooking purposes and domestic heating, or as a supercapacitor (Nawaz *et al.*, 2021; Li *et al.*, 2023). On the other hand, condensable gases or synthesis gas (syngas) can be used as a biofuel for combustion engines or boilers (Nawaz *et al.*, 2021).

Table 8. ΔS variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔS, J/mol K				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	-68,19	-65,91	-69,50	-351,02
0,10	-20,34	-19,92	-22,00	-349,63
0,15	9,30	8,65	7,43	-346,40
0,20	23,62	22,59	21,65	-345,35
0,25	42,22	40,56	40,12	-344,27
0,30	51,76	49,87	49,59	-342,92
0,35	65,97	63,60	63,70	-341,43
0,40	75,62	72,96	73,28	-340,04
0,45	91,29	88,05	88,84	-339,79
0,50	108,99	105,04	106,41	-340,05
0,55	138,63	133,38	135,82	-340,36
0,60	180,59	173,45	177,46	-340,64
0,65	260,04	249,18	256,30	-342,95
0,70	313,78	300,58	309,62	-347,43
0,75	343,31	329,15	338,95	-351,95
0,80	341,78	328,34	337,46	-354,44

Source: Authors

Conclusions

The kinetics involved in the thermal degradation of mamoncillo peels and seeds was determined by means of TGA experiments at three heating rates (10, 20, and 40 °C/min). The E_a and the frequency factor were calculated using four isoconversional methods (KAS, FWO, Starink, and Friedman). The E_a was found to be slightly constant in the 0,15-0,50 conversion range, but E_a increased with increases between 0,50 to 0,80, suggesting that the studied pyrolysis was a multistep process. The frequency factor values were found to be between 10^9 and 10^{34} s^{-1}, which indicates that thermal degradation occurred under complex reactions. The optimal kinetic models for the pyrolysis of *mamoncillo* peels were R2, or the contracting cylinder model, in the conversion range of 0,05-0,50 and the F3, or the third-order reaction model, in the conversion range of 0,50-0,80. Meanwhile, for *mamoncillo* seeds, the optimal kinetic models were F2, or the second-order reaction model, in the conversion range of 0,05-0,50 and R3, or the contracting sphere model, in the conversion range of 0,50-0,80. This indicates that the pyrolysis of *mamoncillo* wastes can be characterized by a complex reaction mechanism. The results obtained for the thermodynamic parameters (ΔH, ΔG, and ΔS) suggest that *mamoncillo* waste has properties associated with a potential use as a bioenergy source. The positive ΔH, positive ΔG, and negative ΔS values confirm that the studied pyrolysis process is endothermic and nonspontaneous.

Table 9. ΔS variations with conversion at 10 °C/min for *mamoncillo* peels for the KAS, FWO, Starink, and Friedman methods

ΔS, J/mol K				
α	**KAS**	**FWO**	**Starink**	**Friedman**
0,05	-86,52	-82,91	-87,67	-357,80
0,10	-55,45	-52,50	-56,80	-350,58
0,15	-19,39	-17,80	-21,00	-345,16
0,20	-11,64	-10,18	-13,31	-340,25
0,25	-6,69	-5,30	-8,39	-337,79
0,30	-4,54	-3,11	-6,25	-336,77
0,35	0,75	2,05	1,19	-335,80
0,40	5,56	6,74	3,79	-335,41
0,45	11,82	12,81	10,01	-335,49
0,50	25,18	25,63	23,27	-335,46
0,55	39,30	39,18	37,29	-337,39
0,60	62,62	61,49	60,42	-338,39
0,65	98,66	95,92	96,19	-340,87
0,70	158,87	153,37	155,94	-342,65
0,75	271,46	260,68	267,66	-345,44
0,80	400,65	383,98	395,86	-351,41

Source: Authors

Acknowledgements

The authors are grateful for the financial support provided by Universidad Nacional de Colombia, as well as for the technical support provided by Universidad del Valle during this study.

CRedit author statement

Andrés Felipe Rojas-González conceived the idea and did the background research, collected the data, developed the workflow, performed the assessment, supervised the research, and provided critical feedback. Francisco Javier Velasco-Sarria helped to do the background research, collected literature data, conducted the formal analysis, collected experimental data, and provided critical feedback. All authors contributed to writing the manuscript and approved its definitive version for publication.

Conflicts of interest

The authors declare that they have no conflict of interest that could influence the contents of this manuscript.

References

Aboulkas, A., El harfi, K., and El Bouadili, A. (2010). Thermal degradation behaviors of polyethylene and polypropylene. Part I: Pyrolysis kinetics and mechanisms. *Energy Conversion Management*, *51*, 1363-1369. https://doi.org/10.1016/j.enconman.2009.12.017

Bensidhom, G., Trabelsi, B. H., Mahmood, M. A., and Ceylan, S. (2021). Insights into pyrolite feedstock potential of date palm industry wastes: Kinetic study and product characterization. *Fuel*, *285*, 119096. https://doi.org/10.1016/j.fuel.2021.119096

Bridgwater, A. V. (2012). Review of fast pyrolysis of biomass and product upgrading. *Biomass and Bioenergy*, *38*, 68-94. https://doi.org/10.1016/j.biombioe.2011.01.048

Calderón, J. C., Sachdev, H., Thepanondh, S., and Quintanilla, Y. A. (2021). Morphological and physico-chemical characterization of fruit of *melicoccus bijugatus jacq*. *Bangladesh Journal Botanic*, *50*(2), 387-394. https://doi.org/10.3329/bjb.v50i2.54096

Cai, J., Xu, D., Dong, Z., Yu, X., Yang, Y., and Banks, S. W. (2018). Processing thermogravimetric analysis data for isoconversional kinetic analysis of lignocellulosic biomass pyrolysis: Case study of corn stalk. *Renewable and Sustainable Energy Reviews*, *82*, 2705-2715. https://dx.doi.org/10.1016/j.rser.2017.09.113

Can-Cauich, C. A., Sauri-Duch, E., Betancur-Ancona, D., Chel-Guerrero, L., González-Aguilar, G. A., Cuevas-Glory, L. F., Pérez-Pacheco, E., and Moo-Huchin, V. M. (2017). Tropical fruit peel powders as functional ingredients: Evaluation of their bioactive compounds and antioxidant activity. *Journal of Functional Foods*, *37*, 501-506. https://doi.org/doi:10.1016/j.jff.2017.08.028

Chen, D., Shuang, E., and Liu, L. (2018). Analysis of pyrolysis characteristics and kinetics of sweet sorghum bagasse and cotton stalk. *Journal of Thermal Analysis and Calorimetry*, *131*, 1899-1909. https://doi.org/10.1007/s10973-017-6585-9

Chen, J., Wang, Y., Lang, X., Ren, X., and Fan, S. (2017). Evaluation of agricultural residues pyrolysis under non-isothermal conditions: Thermal behaviors, kinetics, and thermodynamics. *Bioresource Technology*, *241*, 340-348. https://doi.org/10.1016/j.biortech.2017.05.036

Chen, X., Cai, D., Yang, Y., Sun, Y., Wang, B., Yao, Z., Jin, M., Liu, J., Reinmoller, M., Badshah, S. L., and Magdziarz, A. (2023). Pyrolysis kinetics of bio-based polyurethane: Evaluating the kinetic parameters, thermodynamic parameters, and complementary product gas analysis using TG/FTIR and TG/GC-MS. *Renewable Energy*, *205*, 490-498. https://doi.org/10.1016/j.renene.2023.01.078

Dhyani, V., Kumar, L., and Bhaskar, T. (2017). Thermal decomposition kinetics of sorghum straw via thermogravimetric analysis. *Bioresource Technology*, *245*, 1122-1129. https://doi.org/10.1016/j.biortech.2017.08.189

Emiola-Sadiq, T., Zhang, L., and Dalai, A.K. (2021). Thermal and kinetic studies on biomass degradation via thermogravimetric analysis: A combination of model-fitting and model-free approach. *ACS Omega*, *6*, 22233-22247. https://doi.org/10.1021acsomega.1c02937

Gogoi, M., Konwar, K., Bhuyan, N., Borah, R. C., Kalita, A. C., Nath, H. P., and Saikia, N. (2018). Assessments of pyrolysis kinetics and mechanisms of biomass residues using thermogravimetry. *Bioresource Technology Reports*, *44*, 0-49. https://doi.org/10.1016/j.biteb.2018.08.016

Kan, T., Strezov, V., and Evans, T.J. (2012). Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters. *Renewable and Sustainable Energy Reviews*, *57*, 1126-1140. https://doi.org/10.1016/j.rser.2015.12.185

Kaur, R., Gera, P., Jha, M. K., and Bhaskar, T. (2017). Pyrolysis kinetics and thermodynamic parameters of castor (*Ricinus communis*) residue using thermogravimetric analysis. *Bioresource Technology*, *250*, 422-428. https://doi.org/10.1016/j.biortech.2017.11.077

Kumar, M., Shukla, S. K., Upadhyay, S. N., and Mishra, P. K. (2020). Analysis of thermal degradation of banana (*Musa balbisiana*) trunk biomass waste using iso-conversional models. *Bioresource Technology*, *310*, 123393, https://doi.org/10.1016/j.biortech.2020.123393

Lam, S. S., Liew, R. K., Lim, X. Y., Ani, F. N., and Jusoh, A. (2016). Fruit waste as feedstock for pyrolysis technique. *International Biodeterioration & Biodegradation*, *113*, 325-333. https://doi.org/10.1016/j.ibiod.2016.02.021

Li, Y., Wan, Y., Chai, M., Li, Ch, Nishu, L., Yellezuome, D., and Liu, R. (2023). Pyrolysis kinetics and thermodynamic parameters of bamboo residues and its three main components using thermogravimetric analysis. *Biomass and Bioenergy*, *170*, 106705. https://doi.org/10.1016/j.biombioe.2023.106705

Liu, H., Chen, B., and Wang, C. (2020). Pyrolysis kinetics study of biomass waste using Shuffled Complex Evaluation algorithm. *Fuel Processing Technology.*, *208*, 106509. https://doi.org/10.1016/j.fuproc.2020.106509

Maia, A. D., and Morais, L. C. (2016). Kinetic parameters of red pepper waste as biomass to solid biofuel. *Bioresource Technology*, *204*, 157-163. https://doi.org/10.1016/j.biortech.2015.12.055

Mallick, D., Poddar, M. K., Mahanta, P., and Moholkar, V. S. (2018). Discernment of synergism in pyrolysis of biomass blends using thermogravimetric analysis. *Bioresource Technology*, *261*, 294-305. https://doi.org/10.1016/j.biortech.2018.04.011

Mehmood, M. A., Ye, G., Luo, H., Liu, C., Malik, S., Afzal, I., Xu, J., and Ahmad, M. S. (2017). Pyrolysis, and kinetic analyses of Camel grass (*Cymbopogon schoenanthus*) for bioenergy. *Bioresource Technology*, *228*, 18-24. https://doi.org/10.1016/j.biortech.2016.12.096

Mishra, G., and Bhasker, T. (2014). Non isothermal model-free kinetics for pyrolysis of rice straw. *Bioresource Technology*, *169*, 614-621. https://doi.org/10.1016/j.biortech.2014.07.045

Moo-Huchin, V. M., Ac-Chim, D.M., Chim-Chi, Y. A., Ríos-Soberanis, C. R., Ramos, G., Yee-Madeira, H. T., Ortiz-Fernández, A., Estrada-León, R. J., and Pérez-Pacheco, E. (2020). Huaya (*Melicoccus bijugatus*) seed four as a new source of starch: Physicochemical, morphological, thermal, and functional characterization. *Journal of Food Measurement and Characterization*, *14*, 3299-3309. https://doi.org/10.1007/s11694-020-00573-3

Nawaz, A., Misha, R. K., Sabbarwal, S., and Kumar P. (2021). Studies of physicochemical characterization and pyrolysis behavior of low-value waste biomass using thermogravimetric analyzer: Evaluation of kinetic and thermodynamic parameters. *Bioresource Technology Reports*, *16*, 100858. https://doi.org/10.1016/j.biteb.2021.100858

Pacheco, J. I., Lucca, F. A., Goncalves, W. D., Chacón, G., and Sousa, V. (2022). Influence of biomass waste from agro-industries on obtaining energetic gases assisted by chronoamperometric process. *International Journal Hydrogen Energy*, *47*(2), 735-746. https://doi.org/10.1016/j.ijhydene.2021.10.045

Pawar, A., Panwar, N. L., Jain, S., Jain, N. K., and Gupta, T. (2021). Thermal degradation of coconut husk waste biomass under non-isothermal condition. *Biomass Conversion and Biorefinery*, *13*, 7613-7622. https://doi.org/10.1007/s13399-021-01657-w

Raza, M., Abu-Jdayil, B., Al-Marzouqi, A.H., and Inayat, A. (2022). Kinetic and thermodynamic analyses of date palm surface fibers pyrolysis using Coats-Redfern method. *Renewable Energy*, 163, 67-77. https://doi.org/10.1016/j.renene.2021.10.065

Rojas, A. F., and Flórez, C. (2019). Valorización de residuos de frutas para combustión y pirólisis. *Revista Politécnica, 15*(28), 42-53. https://doi.org/10.33571/rpolitec.v15n28a4

Rony, A. H., Koug, L., Lu, W., Dejam, M., Adidharma, H., Gasem, K. A. M., Zheng, Y., Norton, U., and Fan, M. (2019). Kinetics, thermodynamics, and physical characterization of corn stover (Zea mays) for solar biomass pyrolysis potential analysis. *Bioresource Technology, 284*, 466-473. https://doi.org/10.1016/j.biortech.2019.03.049

Santos, V. O., Queiroz, L. S., Araujo, R. O., Ribeiro, F. C. P., Guimaraes, M. N., Costa, C. E. F., Chaar, J. S., and Souza, L. K. C. (2020). Pyrolysis of acai seed biomass Kinetics and thermodynamic parameters using thermogravimetric analysis. *Bioresource Technology Reports, 12*, 100553. https://doi.org/10.1016/j.biteb.2020.100553

Singh, R. K., Patil, T., Pandey, D., and Sawarkar, A. N. (2020a). Pyrolysis of banana leaves biomass: Physico-chemical characterization, thermal decomposition behavior, kinetic and thermodynamic analyses. *Bioresource Technology*, *310*, 123464. https://doi.org/10.1016/j.biortech.2020.123464

Singh, R. K., Patil, T., and Sawarkar, A. N. (2020b). Pyrolysis of garlic husk biomass: Physico-chemical characterization, thermodynamic and kinetic analyses. *Bioresource Technology Reports, 12*, 100558. https://doi.org/10.1016/j.biteb.2020.100558

Singh, R. K., Patil, T., Pandey, D., and Sawarkar, A. N. (2021). Pyrolysis of mustard oil residue: A kinetic and thermodynamic study. *Bioresource Technology*, *339*, 125631. https://doi.org/10.1016/j.biortech.2021.125631

Vejar, A. G., Tolosa, B., Parra, J. W., and Rodríguez-Ordoñez, D. R. (2016). Uso de la cáscara de mamón (*Melicoccus bijugatus*) para el teñido de telas. *Avances en Química*, *11*(3), 123-128. https://www.redalyc.org/articulo.oa?id=93349879004

Vyazovkin, S., Burnham, A. K., Criado, J. M., Pérez-Marqueda, L. A., Popescu, C., and Sbirrazzuoli, N. (2011). ICTAC Kinetics Committee recommendations for performing kinetic computations on thermal analysis data. *Thermochemical Acta*, *520*, 1-19. https://doi.org/10.1016/j.tca.2011.03.034

Wang, C., Wang, X., Jiang, X., Li, F., Lei, Y., and Lin, Q. (2019). The thermal behavior and kinetics of co-combustion between sewage sludge and wheat straw. *Fuel Processing Technology*, *189*, 1-14. https://doi.org/10.1016/j.fuproc.2019.02.024

Yang, H., Ji, G., Clough, P.T., Xu, X., and Zhao, M. (2019). Kinetics of catalytic biomass pyrolysis using Ni-based functional materials. *Fuel Processing Technology*, 195, 106145. https://doi.org/10.1016/j.fuproc.2019.106145

Yuan, X., He, T., Cao, H., and Yuan, Q. (2017). Cattle manure pyrolysis process: kinetic and thermodynamic analysis with isoconversional methods. *Renewable Energy*, *107*, 489-496. https://doi.org/10.1016/j.renene.2017.02.026

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e106683)

Research Article/Chemical, Food, and Environmental Engineering http://doi.org/10.15446/ing.investig.106683

Hydrodeoxygenation of Anisole via Cu supported on Zeolite: HZSM-5, MOR, and Indonesian Activated Natural Zeolite

Hidrodesoxigenación de anisol mediante Cu soportado en zeolita: HZSM-5, MOR y zeolita natural activada de Indonesia

Khoirina Dwi Nugrahaningtyas[1], Aji Indo Sabiilagusti[2], Fitria Rahmawati[3], Eddy Heraldy[4], and Yuniawan Hidayat[5]

ABSTRACT

The conversion of biomass waste into an alternative energy source requires effective and efficient hydrodeoxygenation (HDO) catalysts. This research aimed to synthesize a bifunctional zeolite-based catalyst for anisole conversion into BTX. The noble metal Cu was impregnated on HZSM-5, mordenite, and Indonesian activated natural zeolite (ANZ) to form HDO catalysts. X-ray fluorescence (XRF), X-ray diffraction (XRD), surface area and pore profile analysis, Fourier transform infrared analysis, ammonia-temperature programmed desorption (NH_3-TPD), pyridine gravimetry, morphology, and scanning electron microscopy-energy dispersion elemental mapping (SEM-EDX) were used to determine the catalyst's properties. The HDO reaction test used anisole as a model compound in a semi-flow reactor with hydrogen gas at 350 and 500 ˚C for 1 h. Copper nanocrystals were found on the surface of the zeolites in several metal phase types, including Cu, Cu_2O, CuO, and $Cu(OH)_2$. Due to the copper bonds inside the zeolite pores, the internal pore surface area decreased. The acidity also decreased since it is strongly related to the surface area. At 350 ˚C, Cu was found to be less active. However, at 500 ˚C, copper activity increased, leading to an increase in anisole conversion and BTX selectivity. The catalyst with the highest anisole conversion and BTX selectivity was Cu/HZSM-5 (*i.e.*, 53,28 and 13,06% v, respectively).

Keywords: acidity, BTX, copper, metal phase, surface area

RESUMEN

La conversión de residuos de biomasa en una fuente de energía alternativa requiere catalizadores de hidrodesoxigenación (HDO) efectivos y eficientes. Esta investigación tuvo como objetivo sintetizar un catalizador bifuncional a base de zeolita para la conversión de anisol en BTX. El metal noble Cu se impregnó en HZSM-5, mordenita y zeolita natural activada de Indonesia (ANZ) para formar catalizadores HDO. Se empleó fluorescencia de rayos X (XRF), difracción de rayos X (XRD), análisis de área superficial y del perfil de los poros, análisis infrarrojo por transformada de Fourier, desorción programada a temperatura de amoníaco (NH_3-TPD), gravimetría de piridina, morfología y mapeo elemental de dispersión de energía por microscopía electrónica de barrido (SEM-EDX) para determinar las propiedades del catalizador. La prueba de reacción HDO utilizó anisol como compuesto modelo en un reactor de semiflujo con gas hidrógeno a 350 y 500 ˚C durante 1 h. Se encontraron nanocristales de cobre en la superficie de las zeolitas en varios tipos de fases metálicas, incluyendo Cu, Cu_2O, CuO y $Cu(OH)_2$. Debido a los enlaces de cobre dentro de los poros de la zeolita, el área de la superficie del poro interno disminuyó. La acidez también se redujo, pues está estrechamente relacionada con la superficie. A 350 ˚C, se encontró que el Cu era menos activo. Sin embargo, a 500 ˚C, la actividad del cobre aumentó, lo que provocó un aumento en la conversión de anisol y la selectividad de BTX. El catalizador con la mayor conversión de anisol y selectividad de BTX fue Cu/HZSM-5 (*i.e.*, 53,28 y 13,06 % v respectivamente).

Palabras clave: acidez, BTX, cobre, fase metálica, área superficial

Received: January 10th, 2023
Accepted: July 28th, 2023

---

[1] Chemistry Department, Universitas Gadjah Mada, Indonesia. Dr., Universitas Gadjah Mada, Indonesia. Affiliation: Full professor, Sebelas Maret University, Indonesia. E-mail: khoirinadwi@staff.uns.ac.id

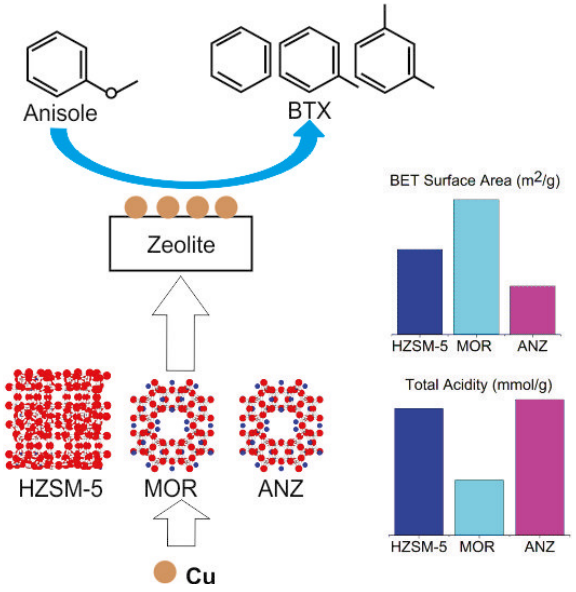
[2] Chemistry Department, Sebelas Maret University, Indonesia. M.Si-c Sebelas Maret University, Indonesia. Affiliation: Master student, Sebelas Maret University, Indonesia. Sebelas Maret University. E-mail: aji.sabiila@student.uns.ac.id

[3] Chemistry Department, Universitas Brawijaya, Indonesia. M.Si, Universitas Gadjah Mada, Indonesia. Dr, Institut Teknologi Bandung, Indonesia. Affiliation: Full Professor, Sebelas Maret University, Indonesia. E-mail: fitria@mipa.uns.ac.id

[4] Chemistry Department, Universitas Gadjah Mada, Indonesia. Dr., Universitas Gadjah Mada, Indonesia. Affiliation: Full professor, Sebelas Maret University, Indonesia. E-mail: eheraldy@mipa.uns.ac.id

[5] Chemistry Department, Universitas Gadjah Mada, Indonesia. Dr., Universitas Gadjah Mada, Indonesia. Affiliation: Associate professor, Sebelas Maret University, Indonesia. E-mail: yuniawan.hidayat@mipa.uns.ac.id

Graphical abstract



Introduction

Biomass-based fuels have been extensively studied to solve supply shortages and environmental issues [1], [2]. Non-edible sources such as lignin may be a promising option for overcoming this issue, as it is easily found in many types of wood, including cheap wood and paper industry waste [3], [4]. The main liquid output of rapid lignin pyrolysis is bio-oil. This product has a low calorific value and is corrosive, so it cannot directly replace petroleum [5]. Therefore, catalytic hydrodeoxygenation (HDO) is required to remove the oxygen content from pyrolysis products in order to meet the requirements for transportation fuels or those for conventional fossil fuel additives. Fast pyrolysis followed by HDO is the preferred method for converting lignin to aromatic hydrocarbons as fuel fractions, such as benzene, toluene, and xylene (BTX) [1], [6], [7].

Transition metal-supported catalysts have excellent performance in the HDO of biomass-derived oils to produce aromatic hydrocarbons [2], [5]. However, Cu-supported catalysts for HDO have not received much attention [4], [8]. This is intriguing because Cu has good adsorption capabilities as per the Sabatier principle, which are almost equal to those of noble metals such as Pt, Pd, and Ru[9]. By comparing it to Ni, Co, and Mo, [4] discovered that Cu-supported HZSM-5 has the highest liquid product in the HDO of anisole, a model compound of lignin-derived oil. [8] found that depositing Cu into the pores of ZSM-5 gradually reduced the surface area of the catalyst. Surprisingly, a 5% (w/w) Cu loading increased the acidity of ZSM-5, but an excessive increase in the Cu concentration (above 10% w/w) dramatically reduced the acidity and the catalyst activity.

Supports for HDO have been extensively researched, including inert materials (such as SiO_2 and Al_2O_3), activated carbon, metal oxides, and acid materials such as zeolite. Zeolite is the most promising of these materials due to its high thermal stability, shape selectivity, and adjustable acidity [4]. Because of its high acidity and small pore size, HZSM-5 has a high selectivity for converting oxygenated compounds to aromatic hydrocarbons [10]-[12]. However, HZSM-5 has a hard time in converting large oxygenated compounds [13], and excessive metal concentrations easily block its pores [8]. Studying ideal supports or identifying alternate supports with width pores is a challenging task. Therefore, this study was dedicated to synthesizing and comparing HDO catalyst-based copper with various carriers (*i.e.*, HZSM-5, MOR, and ANZ).

Experimentation

Materials

The support materials used for catalyst synthesis were HZSM-5 (SiO_2/Al_2O_3= 26) supplied by ACS Material, H-Mordenite HS-690 (SiO_2/Al_2O_3= 200, MOR) supplied by Wako Chemical USA, and Indonesian Natural zeolite (NZ).

Synthesizing Cu-supported zeolite

We produced Indonesian-activated natural zeolite (ANZ) while following our prior studies [14], [15]. The bi-functional catalysts were produced using wet impregnation. Several zeolites, ammonia, and $Cu(NO_3)_2.3H_2O$ (p.a, E Merck) were blended and then refluxed for 16 hours at 30 °C and then at 80 °C for 4 hours [16], [17]. Then, the impregnated zeolite was dried in a rotary evaporator (IKA RV 05 Basic) at a pressure of 72 mBar and at 40 °C. The dry powder was then calcined for 3 hours at 550 °C in N_2 gas and then reduced for 2 hours at 400 °C in H_2 gas. The catalysts were labeled as Cux/Z, where x represents the concentration of Cu (0, 4, or 8) and Z denotes the zeolite (HZSM-5, MOR, or ANZ).

Characterization

Except for MOR, which involved a BRUKER S2 Ranger, the elemental composition of the catalyst was determined using PANalytical's Minipal 4 X-ray fluorescence (XRF) instrument. Simultaneously, X-ray diffraction spectroscopy (XRD, Rigaku Miniflex) was used to determine the metal phase composition. Crystallinity and size were determined using the Scherrer equation. A gas sorption analyzer (GSA, NOVA 1200e) was used to evaluate catalyst texture.

Catalyst acidity was measured via the gravimetric method, with pyridine as a probe molecule [18], and the results were confirmed using Fourier-transform infrared spectroscopy (FTIR, Shimadzu 8201 PC) and NH_3 temperature-programmed desorption (NH_3-TPD, ChemiSorb 2750 Micromeritics). The NH_3-TPD instrument ran at 10 °C/min between 100 and 700 °C and was then held for 10 min at 700 °C. Desorbed NH_3 was detected using a thermal conductivity detector (TCD). The peak area was then calculated using standard linear regression in order to determine the catalyst's acidity.

A scanning electron microscope (SEM, JEOL JSM 6510 LA) was used to evaluate the morphology of the catalyst. Each catalyst specimen was prepared using a gold coating. SEM was carried out using the FEI Ltd Inspect-S50. An additional analysis via energy-dispersive X-rays (EDX, EDAX Ltd.) was performed to identify the elemental composition and its location on the catalyst surface.

Catalyst performance evaluation

The catalytic performance of anisole HDO was studied using a flow reactor (EDULAB). For each test, several grams of catalyst were placed in the reactor and reduced *in situ* in H2 gas at a rate of 15 mL/min for 2 h. Furthermore, the reactor was heated to 350 °C before the anisole flowed at a rate of 10 mL/min with H_2 as a carrier gas. The liquid fraction was obtained after the outflow stream passed through the ice condenser.

The liquid fraction was determined via GC-MS analysis (Shimadzu GCMS-QP2010 SE), while the mass spectra were obtained in the electron ionization (EI) mode and in the m/z range of 50-600. Peak mass was then identified and quantified according to the Wiley library database.

Results and discussion

Copper content

Theoretically, MOR has a larger pore dimension than HZSM-5, resulting in an excellent adsorption capacity. The pore structure of MOR is composed of a 12-membered ring with a dimension of 6,5 x 7 Å [19], while HZSM-5 is composed of a 10-membered ring with a diameter of 5-6 Å [20]. ANZ, on the other hand, is primarily made up of a MOR framework. Thus, it appears that its capabilities are somewhat MOR-like. Given that MOR had the greatest adsorption capacity, it is not surprising that the Cu4/MOR catalyst demonstrated the best impregnation efficiency, which was 94% (Table 1).

Table 1. Catalyst composition as obtained via XRF

Composition (% w/w)			
Catalysts	Si	Al	Cu
HZSM-5	51,6	2,67	-
Cu4/HZSM-5	42,4	0,43	2,29
Cu8/HZSM-5	35,5	0,99	4,71
MOR[21]	44,17	0,74	-
Cu4/MOR[21]	39,6	0,64	3,74
Cu8/MOR[21]	36,1	0,6	5,52
ANZ	32,3	3,83	-
Cu4/ANZ	32,9	3,97	3,61
Cu8/ANZ	31,6	3,78	4,58

Metal phase composition

All the catalysts show diffraction patterns similar to their parent zeolite (Fig. 1). New Cu crystal peaks appear on the catalysts after impregnation. The parent zeolite characteristic peaks remain, indicating that copper impregnation did not harm or change the zeolite structure.

The crystal size is greatly influenced by the catalyst's composition. All catalysts exhibit high crystallinity (up to 96,12%), with varying crystal sizes. The crystal size does not change significantly after impregnation, except for the Cu4/MOR and Cu4/ANZ catalysts. $Cu(OH)_2$ phases are found in Cu4/MOR catalysts, resulting in large crystal sizes. Meanwhile, the Cu4/ANZ catalyst show the largest crystal size, given the appearance of an unknown crystal at 35, 52, and 77°.

Copper was discovered in a variety of metal phases, including Cu, CuO, Cu_2O, and $Cu(OH)_2$. Despite the absence of copper oxide and hydroxide (Fig. 1), the calculation test reveals that they are matched to catalysts. These findings are debatable, but copper oxide and hydroxide may have low intensities in the diffraction pattern of the catalysts. Furthermore, these findings can help to explain the various catalyst colors. Cu/HZSM-5 and Cu/ANZ are dark reddish gray due to their crystal composition. The gray coloration of the catalyst is associated with HZSM-5 and ANZ as the parent zeolite, whereas Cu, Cu_2O, and CuO add brown, red, and black to the catalysts, respectively. MOR is a bright white substance, so Cu4/MOR is pale green, which is associated with its $Cu(OH)_2$ content. Meanwhile, Cu8/MOR is reddish brown due to Cu and Cu_2O.

The copper metal phase has a distinct diffraction pattern corresponding to ICSD #53246, indicating that the reduction process successfully converted copper ions into copper metal. The copper crystal peak intensities increase as the copper content increases. The Rietica software with Le bail methods was used to further study the phase composition and the molar percentage of the catalysts (Table 2).

The Cu metal phase was found in all catalysts after impregnation, indicating a successful reduction process. However, Cu/HZSM-5 and Cu/ANZ contain CuO and Cu_2O, which indicates that copper was not completely reduced. Cu/MOR appears to be better reduced; here, no CuO was obtained. This work cannot explain the formation of $Cu(OH)_2$ in Cu4/MOR because copper hydroxide should be easily converted to copper oxide at low temperatures (150 °C) [22].

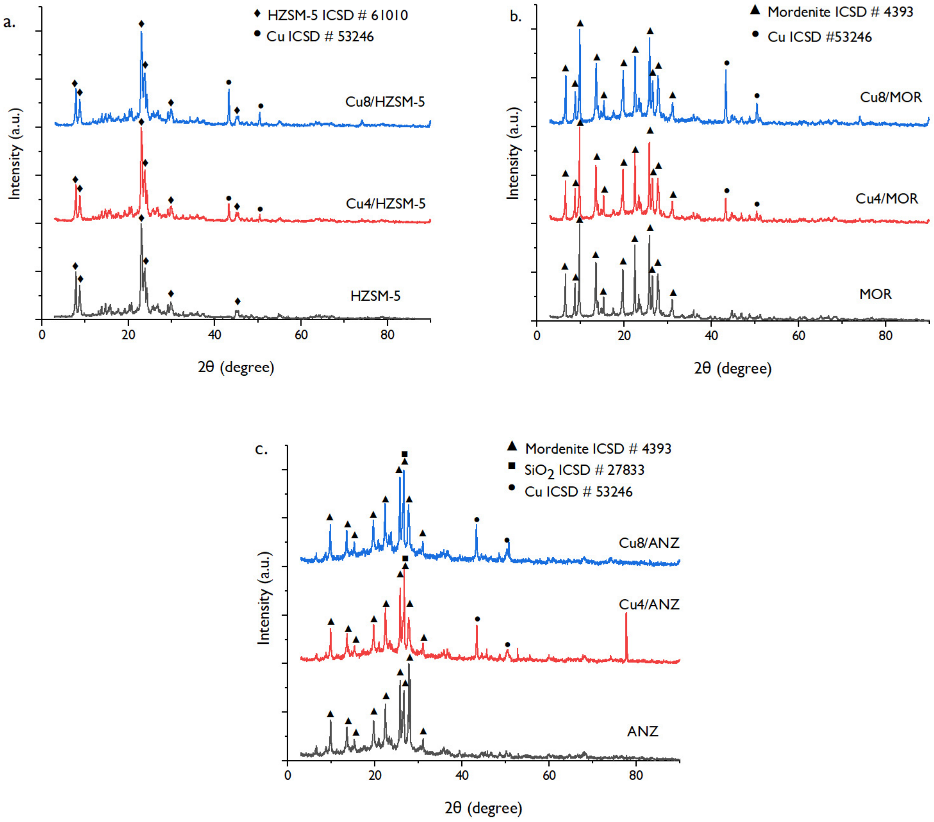


Figure 1. X-ray diffraction pattern of the Cu/HZSM-5, Cu/MOR [21], and Cu/ANZ catalysts

Table 2. Crystal data for Cu/HZSM-5, Cu/MOR, and Cu/ANZ catalysts based on XRD followed by refinement via the Le Bail method

Catalyst	**Crystallinity (%)**	**Crystal size (nm)**	**Molar percentage (%)**								
			Mordenite #4393	**Mordenite #9632**	**HZSM-5 #61010**	**Klipnotilolite #66459**	**SiO_2 #27833**	**Cu #53246**	**$Cu(OH)_2$ #15455**	**Cu_2O #53322**	**CuO #61323**
MOR[21]	96,12	35,73	97,91	2,09							
Cu4/MOR[21]	91,99	40,33	97,20	2,07				0,14	0,48	0,11	
Cu8/MOR[21]	92,30	30,32	97,66	2,08				0,14		0,11	
HZSM-5	93,63	24,40			100						
Cu4/HZSM-5	95,42	28,13			89,18			3,18		2,57	5,07
Cu8/HZSM-5	95,52	26,52			89,21			3,16		2,56	5,07
ANZ	90,12	31,15	98,45			1,55					
Cu4/ANZ	95,49	64,18	97,74			1,53	0,25	0,14		0,11	0,23
Cu8/ANZ	94,45	35,70	97,84			1,54	0,25	0,14			0,23

Textural profile

MOR has the largest surface area and pore volume, as well as the smallest pore size, whereas ANZ has the smallest surface area and pore volume but the largest pore size despite its high structural similarity to MOR (Table 3). The fact that ANZ has a significantly lower Si/Al ratio than MOR, which impacts the surface area and pore profile, may help to explain this discrepancy. HZSM-5 had a middling surface area, pore volume, and pore size when compared to MOR and ANZ.

Table 3. Surface area and pore profile of catalysts based on N_2 isotherms

Catalyst	$S_{Internal}^{a}$ (m^2/g)	S_{BET}^{b} (m^2/g)	V_p^{c} (cc/g)	Pore size^{c} (nm)
HZSM-5	207,1	319,1	0,29	163,3
Cu4/HZSM-5	151,4	304,3	0,26	180,2
Cu8/HZSM-5	141,2	299,2	0,26	181,7
MOR	362,5	506,1	0,33	136,8
Cu4/MOR	244,6	406,0	0,34	173,0
Cu8/MOR	205,1	461,9	0,42	186,9
ANZ	90,5	182,0	0,17	201,5
Cu4/ANZ	77,9	190,7	0,20	224,6
Cu8/ANZ	68,8	207,2	0,19	193,6

[a] Internal surface area calculated using T-plots
[b] BET surface area estimated with relative pressure in the 0,05-0,3 range
[c] Total pore volume and mean pore radius calculated using the BJH model

Typically, copper bonds to the surface of the zeolite pores and blocks them, reducing the surface area and pore volume. This is consistent with the catalysts' internal pores, which progressively close up after impregnation. These findings also suggest that copper was successfully attached to the cavity of zeolites by impregnation. Copper does not wholly block the zeolite's pores, which is beneficial for catalytic reactions.

After impregnation, the surface area of Cu/HZSM-5 catalysts slightly decreases, while that of Cu/ANZ increases. Meanwhile, copper impregnation reduces the surface area of MOR, but Cu8/MOR catalysts have a larger surface area than Cu4/MOR catalysts. These results are confusing but may be clarified by considering the catalysts' increase in external surface area and total pore volume. The zeolite's pores are limited, and some copper prefers to stick to the external surface of the zeolite. Subsequently, copper accumulates and entails roughness, which causes a slight increase in the surface area. Moreover, copper is attached to the pore mouth, thus extending the porosity and slightly increasing the pore volume. Therefore, catalysts exhibit a large surface area after being impregnated with 8% (w/w) rather than 4% (w/w) due to the excess copper on the external surface of the zeolites.

The pore size distribution of the catalysts was evaluated using the BJH method at a 1,5-100 nm radius (Fig. 2). The distribution peaks were found to be broad, which means that the catalysts have various pore sizes. It also suggests that the pores are more slit-shaped than cylindrical.

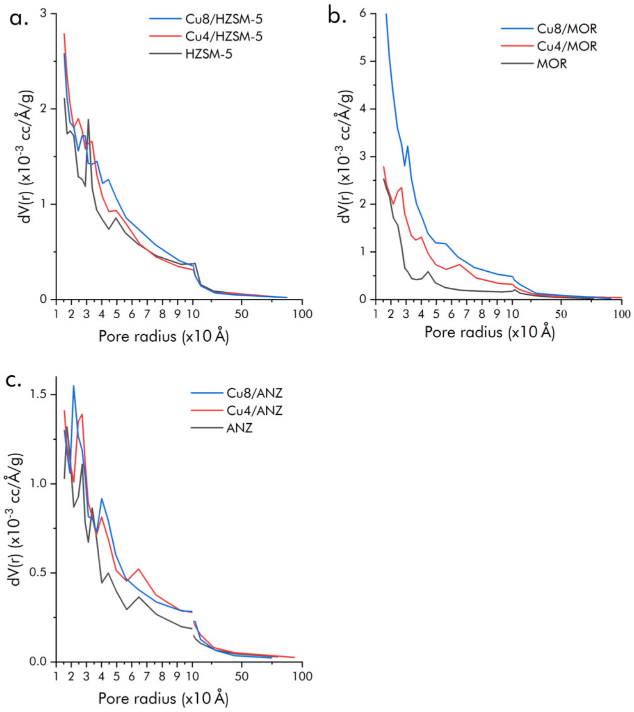


Figure 2. Pore distribution of Cu/HZSM-5, Cu/MOR, and Cu/ANZ catalysts

Functional group and acidity

The FTIR spectra show peaks at 1 200-450 cm^{-1}, which is typical of zeolites. The external T-O (T = tetrahedral Si or Al) exhibits symmetric stretching vibrations at around 1 200 and 1 050 cm^{-1}, as well as asymmetric stretching at around 820 and 790 cm^{-1} (Fig. 3). Meanwhile, the peaks at around 650 and 450 cm^{-1} are attributed to internal T-O symmetric and asymmetric stretching, respectively. The T-O groups appear to be dominant, which indicates that zeolite structure is the main component of the catalysts. However, the T-O vibration pattern seems different for HZSM-5, MOR, and ANZ, due to their different structure and Si/Al ratio. Copper impregnation did not significantly change the T-O functional group, which indicates that the zeolite structure is not ruined by copper, which agrees with XRD results.

The O-H groups are represented by the broad peak of OH stretching at around 3 300-3 600 cm^{-1} and OH bending at around 1 600 cm^{-1}. The O-H group is associated with water molecules or the Bronsted acid site of catalysts. HZSM-5 showed the most significant peak area in the O-H group, indicating that it had the most acidic nature when compared to MOR and ANZ. Furthermore, copper impregnation slightly decreased the peak area of the OH group on the catalysts due to cation exchange. However, acidity evaluation needs further confirmation using more reliable techniques.

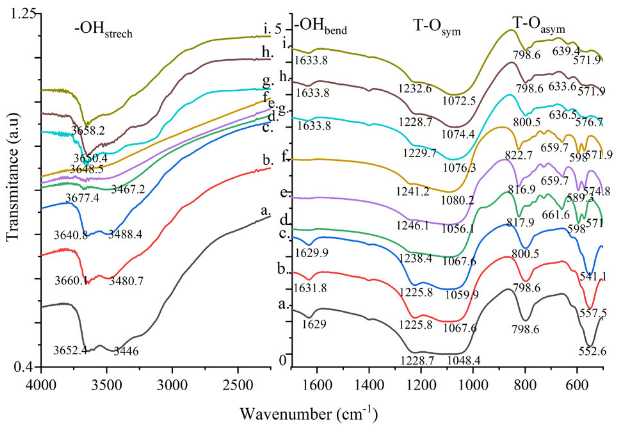


Figure 3. FTIR spectra of a) HZSM-5, b) Cu4/HZSM-5, c) Cu8/HZSM-5, d) MOR, e) Cu4/MOR, f) Cu8/MOR, g) ANZ, h) Cu4/ANZ, and i) Cu8/ANZ

Catalyst acidity was evaluated using pyridine and ammonia as probe molecules (Table 4). Pyridine gravimetry only determines the total acidity, whereas NH_3-TPD can distinguish the acid strength of several sites on the catalysts. All catalysts showed more than one desorption peak in the NH_3-TPD test, indicating several acid sites with different strengths.

Table 4. Acidity of Cu/HZSM-5, Cu/MOR, and Cu/ANZ catalysts based on NH3-TPD and pyridine gravimetry

Catalyst	TPD (mmol/g)[1]				Gravimetry (mmol/g)[2]
	Weak	Medium	Strong	Total	
HZSM-5	0,75	0,68	0,14	1,56	0,85
Cu4/HSZM-5	0,64	0,28	0,43	1,36	0,69
Cu8/HZSM-5	0,52	0,20	0,57	1,28	0,66
MOR	0,19	0,10	0,39	0,68	0,89
Cu4/MOR	0,05	0,29	0,14	0,49	0,59
Cu8/MOR	0,00	0,11	0,23	0,34	0,50
ANZ	0,75	0,00	0,92	1,67	1,08
Cu4/ANZ	0,85	0,00	1,09	1,94	1,23
Cu8/ANZ	0,12	0,00	0,23	0,35	0,94

[1] TPD acidity was calculated based on the area of ammonia desorption peak and a standard linear regression

[2] Gravimetric acidity was calculated based on the difference in weight of the catalyst after pyridine adsorption

The acidity of the catalysts decreased after copper impregnation, except for Cu4/ANZ. [23] reported that impregnation with 0-3% (w/w) increased the acidity of the catalyst by adding Bronsted acid sites. Meanwhile, another report revealed that the catalyst acidity decreased by 5-30% (w/w) with copper impregnation due to pore blockage by excess Cu [8]. Pore blockage was also observed in this work, as seen in the decrease in micropore area (Table 3). However, Cu4/ANZ exhibited the most remarkable mean pore size, thus preventing pore blockage.

Morphology and mapping elements

Copper impregnation did not change the morphology of the catalysts. Meanwhile, small spherical particles (*i.e.*, copper particles) were found on the catalyst after impregnation. These copper particles are shiny and more clearly observed at a scale of 1 µm, as depicted in the inset image of Fig. 4. The number of copper particles was higher after copper impregnation at 8% (w/w) in comparison with that at 4% (w/w), which agrees with XRF results. Furthermore, the particles appear to be evenly distributed on the catalysts, indicating a good dispersion of copper.

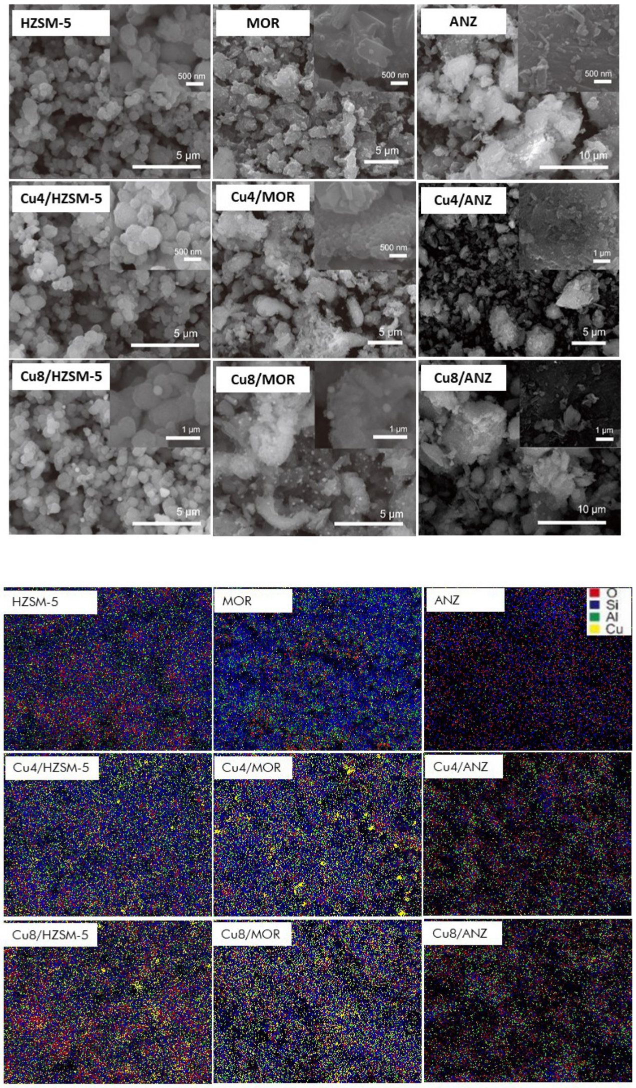


Figure 4. SEM image and mapping elements of the Cu/HZSM-5, Cu/MOR, and Cu/ANZ catalysts

Anisole hydrodeoxygenation

HZSM-5, MOR, and ANZ exhibited a higher conversion of anisole than the thermal reaction, indicating that the support can provide a catalytic effect, even without adding metals (Table 5). Zeolites have acid sites that can dissociate H_2 or absorb oxygenated compounds. Conversely, oxygen electrons can interact with the unsaturated orbitals of Lewis acid sites [5].

The conversion for anisole HDO is HZSM-5 > MOR > ANZ. Zeolites have different acidity and surface area, so they significantly affect HDO catalytic activity [8]. Higher catalyst acidity and surface area lead to higher HDO activity [8], [23]. The large surface area is an ideal adsorption site for reactants before they are activated and converted to products. Meanwhile, the acidity of the zeolite serves to dissociate H_2 or bind to the anisole's oxygen groups. The Bronsted acid site interacts with the oxygen atom because the π-electrons of the oxygen are more electronegative than the π-electrons of the aromatic ring [24]. On the other hand, the Lewis acid site can also adsorb oxygen atoms from the anisole through electron vacancy [5].

The HZSM-5 catalyst has the highest acidity and a large surface area, thus providing the best anisole HDO catalytic activity. HZSM-5 showed the highest conversion (41,09%) and the highest liquid yield (18,11%), as shown in Table 5. This excellent performance is probably due to its high surface area and acidity. The BTX product was successfully obtained from the anisole HDO reaction. The ANZ catalyst, on the other hand, has the lowest surface area, so it has poor catalytic activity. The results also demonstrated that the catalyst performed better as a powder than as a pellet in terms of anisole conversion and liquid products yield. The catalyst has a higher surface area in powder form, resulting in a better catalyst-anisole contact.

Table 5. Product of anisole HDO regarding the Cu/MOR, Cu/HZSM-5, and Cu/ANZ catalysts at 350 °C

Catalyst	**Conversion (%)[b]**	**Liquid product (mmol/g)[b]**				
		Yield (%)[b]			Selectivity (%)	
		Total	BTX[c]	Phs[d]	BTX[c]	Phs[d]
Thermal	10,31	0,68	0,37	0,14	53,95	21,05
HZSM-5[a]	36,55	12,00	0,51	6,37	4,28	53,12
HZSM-5	41,09	18,11	1,33	11,11	7,31	60,94
Cu4/HSZM-5	20,91	7,62	0,26	6,04	3,41	79,29
Cu8/HZSM-5	31,47	12,23	0,65	9,53	5,34	77,84
MOR	29,85	7,02	0,06	4,55	0,88	64,84
Cu4/MOR	8,04	4,96	0,11	3,07	2,15	61,91
Cu8/MOR	20,61	5,06	0,04	3,02	0,83	59,67
ANZ	23,77	3,23	0,03	2,45	0,98	75,68
Cu4/ANZ	21,17	4,32	0,05	3,26	1,15	75,38
Cu8/ANZ	24,43	1,60	0,000	1,25	0,00	78,26

[a] HZSM-5 catalyst in pellet form
[b] Calculated as weight percentage of reactant mass
[c] BTX
[d] Phs: phenolic compounds

According to [4], adding copper metal reduces anisole HDO conversion for HZSM-5, MOR, and ANZ. The loss of some zeolite acid sites due to metal loading indicates that it causes decreased catalytic activity. The metal is also believed to be less active at reaction temperatures below 400 °C. In addition, Cu loading also reduces the surface area of the HZSM-5, MOR, and ANZ micropores, which decreases the catalytic activity of anisole HDO [8], [23]. Nevertheless, due to the synergy effect between metal and zeolite, the Cu8/zeolite catalyst showed better conversion than Cu4/zeolite. The unique performance of the Cu4/MOR catalyst is thought to be due to the catalyst's metallic $Cu(OH)_2$ phase.

Coke is undesirable in HDO anisole reactions because it can deactivate the catalyst. However, it was discovered that, as the pore size of the catalyst decreased, the coke increased. The HZSM-5 catalyst produces the most coke due to its small pore size (136,8 nm). Metal additions to HZSM-5 and MOR increased the pore size of the catalyst, resulting in coke reduction. However, because the pore size was too large, the presence of metals did not affect coke formation in the Cu/ANZ catalyst. A certain amount of coke produced as a byproduct can clog the pores of the catalyst, rendering it inactive. As a result, the catalyst's large pore size is advantageous because it prevents coke formation and deactivation. On the other hand, too large a pore size can also have a harmful effect, as zeolites can lose their pore selectivity.

The total yield of the liquid product from the HDO anisole reaction using Cu/HZSM-5, Cu/MOR, and Cu/ANZ is presented in Fig. 5. Phenol and methyl anisole were the dominant liquid products obtained from HDO anisole. Based on kinetics studies, the anisole-HDO reaction occurred as a bimolecular reaction with a short contact time and high reactant concentrations [25]. The bimolecular reaction produced phenol and methyl anisole by combining two anisole compounds.

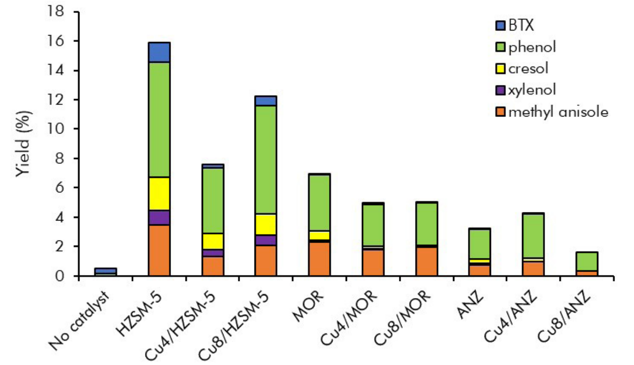


Figure 5. Liquid yield product of anisole HDO using the Cu/HZSM-5, Cu/MOR, and Cu/ANZ catalysts

Phenol (Phs) corresponds to each catalyst's main product of anisole HDO. The presence of phenolic products follows the theory that the anisole HDO reaction consists of two or more steps. One step is the transmethylation reaction between two anisoles, where the methoxy bond ($-O-CH_3$) from one anisole is hydrogenated and broken to produce Phs and methyl. The methyl group ($-CH_3$) is then bonded back to the phenol to produce methyl Phs. In addition, transmethylation reactions can also occur when the methyl group is released in the first step and reacts with another anisole, producing methyl anisole. The second bimolecular reaction produces cresol and xylenol with longer contact times. However, in this study, the contact time was less than one hour, so there was Phs dominance over cresol and xylenol.

The Cu/zeolite catalyst increased the Phs yield from 8 to 55 times compared to the thermal reaction. Except for MOR, the addition of Cu metal to zeolite increases the yield due to the synergistic effect. Copper metal plays a role in absorbing and separating hydrogen [26], and the dissociated hydrogen interacts with the methoxy group through spillover hydrogen, resulting in a hydrogenation reaction [24]. At the same time, zeolite provides Bronsted acid sites (Fig. 6).

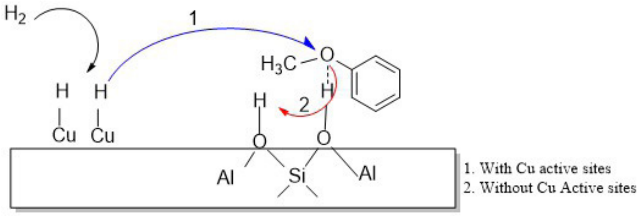


Figure 6. Proposed mechanism for the interaction of zeolite-embedded Cu catalysts with anisole in the modified HDO reaction [24]

BTX products indicate a successful anisole HDO reaction. BTX can be formed through one- and two-step deoxygenation routes [1]. Acidic catalysts, such as zeolites, have a tendency towards a two-step anisole HDO reaction [5]. This route was also found in this study, where BTX was the final product.

The maximum yield of the BTX product was 1,33% (w/w), which indicated that the HDO anisole reaction had not been completed. Contact time, bond energy, and metal activity can cause an unsuitable reaction. The short contact time caused the trans-alkylation reaction to dominate, and deoxygenation, the second step, did not occur. Meanwhile, the bond energy of RO-R (339 kJ/mol) was lower than that of Ar-OH (468 kJ/mol) [25]. Thereupon, anisole is more easily converted to phenol instead of forming benzene. Furthermore, metal-zeolite synergism is important in anisole HDO reactions such as transmethylation [27]. Unfortunately, metal activity is low at temperatures below 400 ˚C[28], so there is no BTX yield decrease caused by the synergy effect.

The effect of temperature on the catalyst activity was examined, finding that the conversion increased significantly (Table 6). The high temperature can provide a lot of energy, which allows the methoxy anisole groups to break and react with hydrogen to form BTX products easily.

At 500 ˚C, although HZSM-5 provided a higher liquid product yield and conversion, the Cu8/HZSM-5 catalyst produced the highest BTX yield and selectivity. This finding supports the notion that high temperatures are required for Cu metal to synergize with zeolite and increase BTX activity and selectivity.

Table 6. Effect of temperature at anisole HDO for the HZSM-5 and Cu8/HZSM-5 catalysts

Catalyst	Temperature (˚C)	Conversion (%)[1]	Liquid product (mmol/g)[2]			
			Yield (%)[a]		Selectivity (%)	
			BTX[b]	Phs[c]	BTX[b]	Phs[c]
HZSM-5	350	41,09	1,33	11,11	7,31	60,94
HZSM-5	500	61,87	12,15	16,38	37,55	50,62
Cu8/HSZM-5	350	31,47	0,65	9,53	5,34	77,84
Cu8/HZSM-5	500	53,28	13,06	14,79	42,45	48,05

aCalculated as the weight percentage of reactant mass
[b]BTX: benzene, toluene, and xylene
[c]Phs: phenolic compounds

Conclusion

This study was dedicated to the modification of copper attachment on HZSM-5, MOR, and ANZ, obtaining the best impregnation efficiency (94%), which corresponded to Cu4/MOR. The structures of HZSM-5, MOR, and ANZ were not damaged after impregnation with copper, which was confirmed via structural analysis using XRD and FTIR. Further analysis revealed copper as a metal phase of Cu Cu_2O, CuO, and $Cu(OH)_2$. Copper impregnation affects the surface area and pore volume of the parent zeolite. The internal surface area and micropore volume of the zeolite continued to decrease after impregnation, indicating that the copper adhered to the pores of the zeolite. This was supported by total acidity reduction data.

However, some copper particles adhered to the zeolite's outer surface, causing roughness and slightly increasing the BET surface area. Due to its high surface area and acidity, HZSM-5 provided the best anisole HDO conversion and BTX selectivity. At 350 ˚C, the contribution of Cu metal as the active site was less pronounced. However, Cu was more active at 500 ˚C, which significantly increased anisole HDO conversion (53,28%) and BTX selectivity (42,45%).

Acknowledgements

The project presented in this article was supported by SEBELAS MARET University under HIBAH MANDATORY (2021), with grant number 260/UN27.22/HK.07.00/2021.

CRediT author statement

All authors have participated in a) conception and design or data analysis and interpretation; (b) drafting the article or revising it critically for important intellectual content; and (c) approving of the final version of the manuscript.

N. K. D.: conceptualization and supervision; N. K. D., H. Y.: writing – review and editing; H. E., R. F., N. I. F.: formal analysis; S. A. I.: investigation.

Conflicts of interest

The authors confirm that this work is original and has not been published elsewhere nor is it currently under consideration for publication elsewhere. The authors declare that they have no conflict of interest.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

[1] A. R. K. K. Gollakota, M. Reddy, M. D. Subramanyam, and N. Kishore, "A review on the upgradation techniques of pyrolysis oil," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 1543-1568, 2016, https://doi.org/10.1016/j.rser.2015.12.180

[2] X. Li *et al.*, "Heterogeneous sulfur-free hydrodeoxygenation catalysts for selectively upgrading the renewable bio-oils to second generation biofuels," *Renew. Sust. Energy Rev.*, vol. 82, pp. 3762-3797, 2018, https://doi.org/10.1016/j.rser.2017.10.091

[3] X. Wang, B. Du, L. Pu, Y. Guo, H. Li, and J. Zhou, "Effect of particle size of HZSM-5 zeolite on the catalytic depolymerization of organosolv lignin to phenols," *J. Anal. Appl. Pyrolysis*, vol. 129, pp. 13-20, Jan. 2018, https://doi.org/10.1016/j.jaap.2017.12.011

[4] J. Zhang, B. Fidalgo, A. Kolios, D. Shen, and S. Gu, "Mechanism of deoxygenation in anisole decomposition over single-metal loaded HZSM-5: Experimental study," *Chem. Eng. J.*, vol. 336, pp. 211-222, Mar. 2018, https://doi.org/10.1016/j.cej.2017.11.128

[5] K. A. Rogers and Y. Zheng, "Selective deoxygenation of biomass-derived bio-oils within hydrogen-modest environments: A review and new insights," *ChemSusChem*, vol. 9, no. 14, pp. 1750-1772, 2016, https://doi.org/10.1002/cssc.201600144

[6] S. Popov and S. Kumar, "Rapid hydrothermal deoxygenation of oleic acid over activated carbon in a continuous flow process," *Energy and Fuels*, vol. 29, no. 5, pp. 3377–3384, 2015, https://doi.org/10.1021/acs.energyfuels.5b00308

[7] J. Zhang, B. Fidalgo, S. Wagland, D. Shen, X. Zhang, and S. Gu, "Deoxygenation in anisole decomposition over bimetallic catalysts supported on HZSM-5," *Fuel*, vol. 238, pp. 257-266, Feb. 2019, https://doi.org/10.1016/j.fuel.2018.10.129

[8] C. Li *et al.*, "Catalytic cracking of Swida wilsoniana oil for hydrocarbon biofuel over Cu-modified ZSM-5 zeolite," *Fuel*, vol. 218, no. January, pp. 59-66, Apr. 2018, https://doi.org/10.1016/j.fuel.2018.01.026

[9] E. Roduner, "Understanding catalysis," *Chem. Soc. Rev.*, vol. 43, no. 24, pp. 8226-8239, 2014, https://doi.org/10.1039/C4CS00210E

[10] R. Lippi *et al.*, "Unveiling the structural transitions during activation of a CO2 methanation catalyst Ru0/ZrO2 synthesised from a MOF precursor," *Catal. Today*, vol. 368, pp. 66-77, May 2021, https://doi.org/10.1016/j.cattod.2020.04.043

[11] J. Zhang, B. Fidalgo, A. Kolios, D. Shen, and S. Gu, "The mechanism of transmethylation in anisole decomposition over Brønsted acid sites: Density functional theory (DFT) study," *Sustain. Energy Fuels*, vol. 1, no. 8, pp. 1788-1794, 2017, https://doi.org/10.1039/C7SE00280G

[12] Y. Zheng *et al.*, "Efficient and stable Ni-Cu catalysts for ex situ catalytic pyrolysis vapor upgrading of oleic acid into hydrocarbon: Effect of catalyst support, process parameters and Ni-to-Cu mixed ratio," *Renew. Energy*, vol. 154, pp. 797-812, Jul. 2020, https://doi.org/10.1016/j.renene.2020.03.058

[13] Q. Che, W. Yi, Y. Liu, X. Wang, H. Yang, and H. Chen, "Effect of mesopores in ZSM-5 on the catalytic conversion of acetic acid, furfural, and guaiacol," *Energy and Fuels*, vol. 35, no. 7, pp. 6022-6029, 2021, https://doi.org/10.1021/acs.energyfuels.0c04415

[14] G. M. H. A. Ichsan, K. D. Nugrahaningtyas, D. M. Widjonarko, and F. Rahmawati, "Structure and morphology of the (Ni, Co) Mo/Indonesian natural zeolite," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 578, art. 012009, 2019, https://doi.org/10.1088/1757-899X/578/1/012009

[15] K. D. Nugrahaningtyas, R. S. R. Suharbiansah, W. W. Lestari, and F. Rahmawati, "Metal phase, electron density, textural properties, and catalytic activity of CoMo based catalyst applied in hydrodeoxygenation of oleic acid," *Evergreen*, vol. 9, no. 2, pp. 283-291, Apr. 2022, https://doi.org/10.5109/4793665

[16] K. D. Nugrahaningtyas, M. M. Putri, and T. E. Saraswati, "Metal phase and electron density of transition metal/HZSM-5," *AIP Conf. Proc.*, vol. 2237, art. 020003, 2020, https://doi.org/10.1063/5.0005561

[17] K. D. Nugrahaningtyas, M. F. Kurniawati, A. Masykur, N. 'Abidah Quratul'aini, and N. 'Abidah Quratul'aini, "Periodic trends in the character of first-row transition metals-based catalysts embedded on mordenite," *Moroccan J. Chem.*, vol. 10, no. 3, pp. 375-386, May 2022, https://revues.imist.ma/index.php/morjchem/article/view/32665/16923

[18] T. Barzetti, E. Selli, D. Moscotti, and L. Forni, "Pyridine and ammonia as probes for FTIR analysis of solid acid catalysts," *J. Chem. Soc. - Faraday Trans.*, vol. 92, no. 8, pp. 1401-1407, 1996, https://doi.org/10.1039/ft9969201401

[19] M. A. Klunk *et al.*, "Synthesis and characterization of mordenite zeolite from metakaolin and rice husk ash as a source of aluminium and silicon," *Chem. Pap.*, vol. 74, pp. 2481-2489, 2020, https://doi.org/10.1007/s11696-020-01095-4

[20] A. S. Al-Dughaither, H. De Lasa, and H. De Lasa, "HZSM-5 zeolites with different SiO2 Al2O3 ratios. Characterization and NH_3 desorption kinetics," *Ind. Eng. Chem. Res.*, vol. 53, no. 40, pp. 15303-15316, Oct. 2014, https://doi.org/10.1021/ie4039532

[21] A. I. Sabiilagusti, K. D. Nugrahaningtyas, and Y. Hidayat, "Synthesis and metal phases characterization of mordenite supported copper catalysts," *J. Phys. Conf. Ser.*, vol. 1912, art. 012032, 2021, https://doi.org/10.1088/1742-6596/1912/1/012032

[22] Y. Cudennec and A. Lecerf, "The transformation of Cu$(OH)_2$ into CuO, revisited," *Solid State Sci.*, vol. 5, no. 11-12, pp. 1471-1474, 2003, https://doi.org/10.1016/j.solidstatesciences.2003.09.009

[23] W. B. Widayatno et al., "Upgrading of bio-oil from biomass pyrolysis over Cu-modified β-zeolite catalyst with high selectivity and stability," *Appl. Catal. B Environ.*, vol. 186, pp. 166-172, 2016, https://doi.org/10.1016/j.apcatb.2016.01.006

[24] Z. He and X. Wang, "Hydrodeoxygenation of model compounds and catalytic systems for pyrolysis bio-oils upgrading," *Catal. Sustain. Energy, Versita*, vol. 1, no. 2013, pp. 28-52, Oct. 2012, https://doi.org/10.2478/cse-2012-0004

[25] X. Zhu, R. G. Mallinson, and D. E. Resasco, "Role of transalkylation reactions in the conversion of anisole over HZSM-5," *Appl. Catal. A Gen.*, vol. 379, no. 1-2, pp. 172-181, May 2010, https://doi.org/10.1016/j.apcata.2010.03.018

[26] J. Albo, M. Perfecto-Irigaray, G. Beobide, and A. Irabien, "Cu/Bi metal-organic framework-based systems for an enhanced electrochemical transformation of CO2 to alcohols," *J. CO2 Util.*, vol. 33, pp. 157-165, 2019, https://doi.org/10.1016/j.jcou.2019.05.025

[27] X. Zhu, L. L. Lobban, R. G. Mallinson, and D. E. Resasco, "Bifunctional transalkylation and hydrodeoxygenation of anisole over a Pt/HBeta catalyst," *J. Catal.*, vol. 281, no. 1, pp. 21-29, 2011, https://doi.org/10.1016/j.jcat.2011.03.030

[28] J. Zhang, B. Fidalgo, S. Wagland, D. Shen, X. Zhang, and S. Gu, "Deoxygenation in anisole decomposition over bimetallic catalysts supported on HZSM-5," *Fuel*, vol. 238, pp. 257-266, 2019, https://doi.org/10.1016/j.fuel.2018.10.129

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e101173)

Research Article/Chemical, Food, and Environmental Engineering https://doi.org/10.15446/ing.investig.101173

Trends in Electrochromic Materials: Industrial Perspective in Colombia

Tendencias en materiales electrocrómicos: perspectiva industrial en Colombia

Luis Felipe Hurtado-Palacios[1], Sandra Patricia Castro-Narváez[2], and Alonso Jaramillo-Aguirre[3]

ABSTRACT

Dissertations across various sectors (industrial, governmental, and research) advocate for the manufacture of products with innovative scientific and technological developments as a driving force to generate added value and, in a dual approach, respond to the expectations of domestic consumers and encourage exports in the country. This analytical and descriptive bibliometric review on electrochromism, a property of some materials that allows them to reversibly change their optical transmittance via the repeated and consecutive intercalation and extraction of electric charges under a small alternating polarity electric potential, contributes by describing the fundamentals, recent research advances, and the feasibility of national industrial application, given the existence of mineral resources, chemical inputs, companies, and qualified human talent at different levels of sectors related to chromic devices and energy storage.

Keywords: electrochromic devices, eco-energy technologies, resources assessment

RESUMEN

Disertaciones en diferentes sectores (industriales, gubernamentales y de investigación) abogan por la fabricación de productos con desarrollos científicos y tecnológicos innovadores como un motor para generar valor agregado y, en un enfoque dual, responder a las expectativas de los consumidores nacionales y fomentar las exportaciones en el país. Esta revisión bibliométrica de carácter analítico y descriptivo sobre el electrocromismo, una propiedad de algunos materiales que les permite cambiar de manera reversible su transmitancia óptica a través de la intercalación y extracción consecutiva y repetida de cargas eléctricas bajo un pequeño potencial eléctrico de polaridad alterna, hace aportes en la descripción de los fundamentos, los avances recientes en la investigación y la viabilidad de la aplicación industrial nacional, dada la existencia de recursos minerales, insumos químicos, empresas y talento humano calificado en diferentes niveles de sectores relacionados con dispositivos crómicos y almacenamiento de energía.

Palabras clave: dispositivos electrocrómicos, tecnologías eco-energéticas, evaluación de recursos

Received: February 21st, 2022
Accepted: May 13th, 2023

Introduction

Between 2000 and 2019, the increase in the *per capita* gross domestic product (GDP) of Colombia has reached 3,3%, the highest since 2014 (DANE, 2020a). This has not been enough to circumvent structural weaknesses at the international level in terms of technological change. The Colombian economy has a low complexity and is poorly diversified; although the Gini, a measure of a country's inequality, went from 0,56 in 2010 to 0,476 in 2017 (OECD, 2019), it is still one of the highest in Latin America. In 2019, the monetary poverty gap was 35,7%, and extreme monetary poverty was 9,6% in the national total (DANE, 2020b). Exports based mainly on unprocessed mining and total factor productivity (TFP) generated negligible growth between 2000 and 2019, which is associated with low leverage in research and development (R&D) and industrial interest in technology-based process innovation.

Reflections on improving the GDP converge to the interaction between science, technology, and innovation through the creation of products or services where public or private investment generates added value to the country's resources, with an important impact on social and environmental challenges (MINICIENCIAS, 2020). The introduction of technologies with chromic properties, in the form of materials with usable characteristics in the country's glass and mirror industry, could significantly contribute to the assertive engagement of research groups in the technological innovation required by this sector.

Chromic properties include the displacement of the absorption and emission spectra of certain dyes upon application of an electric field, an effect called *electrochromism*, which can analogously be referred to as *thermochromism* and *photochromism*, which describe the

---

[1] Chemist, Universidad Santiago de Cali. E-mail: luis.hurtado03@usc.edu.co

[2] Corresponding author. Associate professor, Universidad Santiago de Cali, Electrochemistry and Environment Research Group. Master in Basic Sciences, Universidad Del Valle. E-mail: sandracastro00@usc.edu.co

[3] PhD in Chemistry, Emory University, Atlanta, GA, USA. Electrochemistry and Environment Research Group, Universidad Santiago de Cali. Email: aloja12@yahoo.com

color changes produced by heat and light, respectively. In particular, an electrochromic material experiences persistent but reversible color change via an electrochemical reaction, as a result of different visible electronic absorption bands between redox states. The color change is caused by a transition of electrons in the region of the molecule (*i.e.*, the chromophore) to a higher energy orbital.

The first report of thin tungsten oxide films as an electrochromic material under potentiostatic or galvanostatic control dates to 1960 (Shi *et al.*, 2020a). This was achieved using three-electrode circuits, electrochemical techniques such as cyclic voltammetry, coulumbimetry, and chronoamperometry, and *in situ* spectroscopic measurements. In the 1960s, the American Cyanamide Corporation found analogous results for tungsten oxide films (Sorar *et al.*, 2019). Parallel developments in the Soviet Union showed similar results for niobium oxide (Bulja, 2017), and electrochromic materials started to attract research interest with a view to develop applied technologies.

Electrochromism was first considered for use in information displays, and, during the first half of the 1970s, several companies such as IBM, Philips, and Canon undertook research efforts (Yang *et al.*, 2016). However, this technology became less important in the 1970s due to the rise of liquid crystal in displays. In 1984, its use was proposed in energy-efficient buildings, and, since then, electrochromism research evidences a powerful momentum. The term *smart window* came into being between 1984 and 1985 and attracted immediate attention, not only from scientists but also from the media and the general public since smart optical variation can be used in saving energy (Bulja, 2017).

In the last two decades, scientific and technological interest in applications involving electrochromic materials, substrate, and binding electrolytes has been evident, given their potential applications in smart windows and conductive materials in the fields of energy conversion and nanotechnology (Varghese-Hansen *et al.*, 2018). Correspondingly, the question arises whether Colombia would be able to consolidate industries for the manufacture of chromic devices such as smart windows, solar cells, photometric mirrors, and thermochromic or gasochromic glasses, utilizing its mining resources for industrial infrastructure and national scientific agendas.

The analytical review presented herein has two purposes: on the one hand, to present the lines of research from the last five years which have focused on scientific advances regarding electrochromic materials and their components for manufacturing chromic devices; and, on the other hand, to define opportunities for establishing Colombian technology-based companies in this sector.

Methodology

Our descriptive and systematic review on electrochromism, including the definition of the concept, its applications, and research gaps, used primary sources (research articles) and secondary sources (review articles and technical reports) published between 2015-2022, which were retrieved from databases such as ScienceDirect, Springer, Taylor & Francis online, Redalyc, Ebsco, and SciELO. Scopus statistics associated with citations and other characteristics (*e.g.*, country of origin) were evaluated. The search strategy considered the following concepts in addition to electrochromism: photochromism, thermochromism, gasochromism, electrochromic material deposition techniques, and electrochromic material types (organic and inorganic). Since 2 634 articles were found, those with the highest number of citations were selected, for a total of 130 documents.

To determine the feasibility of implementing chromic technologies in the country, we analyzed the tertiary information detailed in technical reports and databases of governmental entities, such as the National Administrative Department of Statistics (DANE), the Ministry of Science, Technology, and Innovation (MinCiencias), the Ministry of National Education, the Ministry of Commerce, the Ministry of Mines and Energy, the Geological Service, and the National Mining Agency.

Results and discussion

Smart windows

An electrochromic device, also called a *smart window*, is a display unit that allows for electrochemically driven modulations of light transmission and reflections (Li *et al.*, 2019; Kumar *et al.*, 2016; Tsige *et al.*, 2020). The potential market for smart windows is very large and includes rear-view mirrors, sunroofs, and side and smart windows in the automotive sector. Other uses include large-area information displays for use in airports and railroad stations and electrochromic eyeglasses and sunglasses.

In smart windows, the entire system is in the optical path, which imposes some restrictions on the electrolyte used. This electrolyte must be transparent or electrochromic and complementary to the primary electrochromic electrode. These requirements are not easy to meet, and, therefore, most research is devoted to the characterization of materials capable of ensuring the desired switchable complementary optical function. Electrochromic windows can be combined with solar cells, so the energy required for electrochromic color switching can be obtained from them (Yun *et al.*, 2019).

The physicochemical principle of electrochromism in the manufacture of smart windows

The reversible change in the optical properties of a material, which is produced by oxidation-reduction induced by electrochemical processes, is called *electrochromism* (Shi *et al.*, 2020a). This process is exploited in the fabrication of electrodes by coating the surface of a transparent conductor with a thin film made of a conductive electrochromic material and by applying potentials that are sufficiently positive to induce oxidation (*i.e.*, the extraction of electrons from the electrochromic material to the transparent electrode) and sufficiently negative to generate reduction, where a color change is observed (Fletcher, 2015) (Figure 1). These oxidation and reduction processes give rise to variations in the electronic structure of the material, *i.e.*, a new distribution of occupied and unoccupied electronic levels, with exchanges in energy distances.

Electrochromic films can be elaborated using different methods and techniques. Figure 2 consolidates the reports associated with electrochromism and different chromic film deposition methods. Regarding the number of publications, the magnetron sputtering technique far outperforms the others, including sol gel chemical deposition (dip coating) (Wen-Cheun *et al.*, 2019). Other techniques with a smaller number of publications are electrochemical deposition (Kimura *et al.*, 2018), thermal evaporation (El-Nahass *et al.*, 2015), pulsed laser deposition (Farhad, 2021), spray pyrolysis, and spin coating (Paipitak *et al.*, 2011).

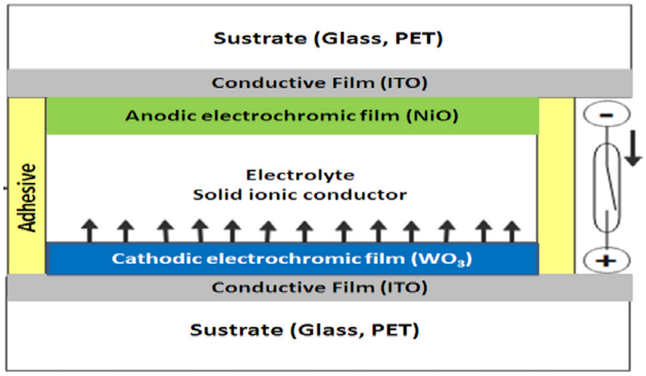


Figure 1. Electrochromic device design
Source: Authors

Electrochromic devices must include at least two electrodes and an electrolyte layer. The color change takes place when electrochromic species transition from a transparent to a colored state. In general, electrochromic devices can be classified into three categories: in the first one, both the colored and transparent (bleached) species are soluble in the electrolyte (Marchisio *et al.*, 2015); in the second one, the transparent state is soluble, while the colored state coats one of the electrolytes (Mendieta-Reyes *et al.*, 2020); and, in the third one, both the bleached and colored species coat one electrode (Hidalgo-Rodríguez, 2002). Some important aspects considered in the development of electrochromic devices include a high transmittance (>50) in the transparent state and a low transmittance in the colored state (<10%), color change times of a few minutes, good durability (5-20 years), costs, potentials of 1-3 V for small areas and 10-24 V for large areas, high open circuit durability (1-12 h). The open circuit must also be able to operate within a wide range of temperatures (between -20 and 85 °C).

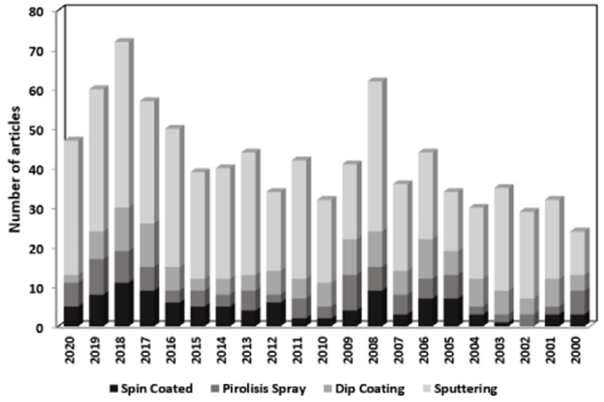


Figure 2. Number of publications on different types of chromic material deposition.
Source: Scopus data for "Electrochromism and deposition technique" as of November 2020

Multilayer devices, which consider sandwich structures, consist of inner active ion storage or counter electrode layers, an ion conductive layer (electrolyte), and an electrochromic film (working electrode), utilizing both organic and inorganic substances (Figure 3). Transparent electrode layers are designed while prioritizing a high transparency, low resistance, and easy deposition. Finally, the protective layers use substrates that can be rigid (*e.g.*, glass) or made up of polymers, allowing for flexible systems such as polyethylene terephthalate (TEP).

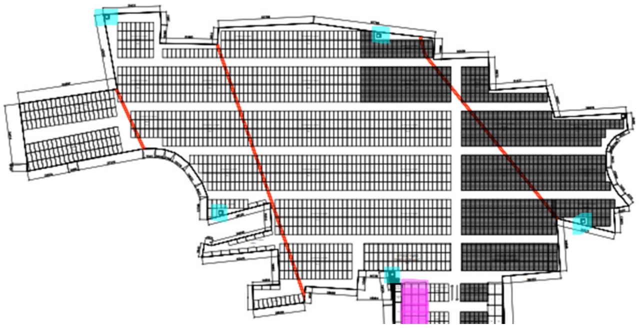


Figure 3. Electrochromic device with different conductive polymers
Source: Authors

The publications on electrochromism from 1960 to 2020 amount to 8 405, with a steeper slope in the last decade (Figure 4).

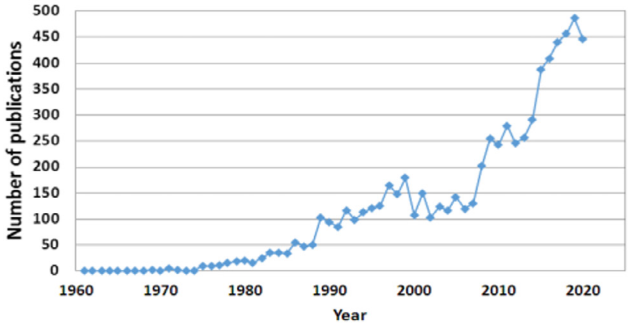


Figure 4. Evolution of publications on electrochromism
Source: Scopus data for "Electrochromism" as of November 2020

These figures reveal a more pronounced trend in Eastern countries, where technological application is in constant growth, followed by the United States. Figure 5 shows the countries reporting at least 200 related publications.

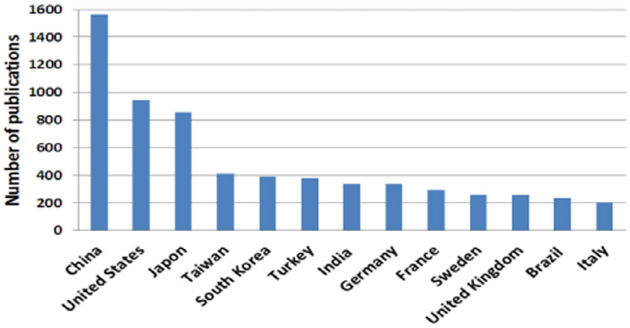


Figure 5. Publications on electrochromism by country
Source: Scopus data for "Electrochromism" as of November 2020

The Americas account for 16,84% of the publications. However, excluding the United States, only 473 contributions were considered (Figure 6). The particular case of Colombia exhibits a deficit in associated studies. According to Scopus, only five contributions feature Colombian researchers and three are linked to national universities (Chen P., 2020; Kalay *et al*., 2020), even though this country has mining resources such as W, Mo, Ni, Ti, and Zn, among others, which could be oriented towards the manufacture of electrochromic devices with technological added value (Zhang *et al*., 2020).

Electrochromic materials

Electrochromic materials fall into three categories: metal oxide films (De Ribamar-Martins-Neto *et al*., 2016), molecular dyes (Barile *et al*., 2017), and conducting polymers (Barus *et al*., 2018). Electrochromic layers switch their optical properties between their oxidized and reduced forms. In an electrochromic device based on cathodic and anodic layers, applying a voltage to transport ions between the two films in one direction causes both films to become colored, and transporting ions in the opposite direction decolorizes them. Thus, optical modulation and visual appearance can be extended by optimizing both electrochromic layers.

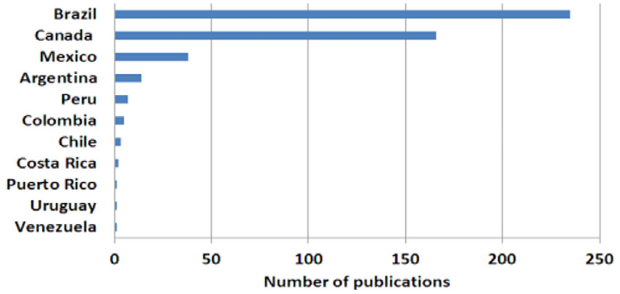


Figure 6. Publications in the Americas without the contribution of the United States
Source: Scopus data for "Electrochromism" as of November 2020

Table 1 exemplifies some electrochromic materials and their potential technological use: transition metal oxides, widely used in smart windows, thermochromics, satellite systems, and chemical paper; Prussian blue systems and metallic phthalocyanines that allow for different color ranges, which are used in monitors; viologen compounds, with applications in smart windows and smart mirrors (such as rear-view mirrors); and conductive polymers, which, in addition to mobile devices, are used in solar cells, artificial muscles, and smart windows.

Table 1. Electrochromic materials and potential technological applications

Electrochromic material	Examples	Applications	Reference
Transition metals (oxides)	WO_3, MoO_3, V_2O_5, Nb_2O_5, $Ir(OH)_3$, and NiO_xH_y	Intelligent windows, thermal satellite control	(Richardson, *et al*., 2018)
Prussian blue systems (PB)	Prussian blue, Prussian brown, Prussian green, and Prussian white	Monitors	(Pan, 2020)
Viologens	Salts of 1,1′-disubstituted -4,4′-bipyridine	Mirrors, monitors	(Zeng, 2021)
Conductive polymers	Poly(3,4-ethylenedioxy-thiophene), polypyrrole, polythiophene, Poly(styrene sulfonate)	Intelligent windows and screens	(Wang L., 2020),
Metallic phthalocyanines	$[Lu(Pc)_2]$ $(C_32H_{18}N_8)$	Monitors	(Ho *et al*., 2015)

Source: Authors

Metal oxides

Transition metal oxide films, *e.g.*, iridium (Sun, 2015; Jang and Lee, 2020), rhodium (Jeong, 2020), ruthenium (Wang Y. *et al*., 2019), tungsten (Atak *et al*., 2020; Pan, 2020), molybdenum (Hasani, 2017), possess electrochromic properties, with tungsten oxide and molybdenum oxide being the most widely studied. This type of electrochromic material can be classified into inorganic electrochromic materials that can be deposited through vacuum evaporation (Sahu *et al*., 2017; Gupta *et al*., 2022b), sputtering (Murphy *et al*., 2016; Kumar *et al*., 2022; Kumar *et al*., 2022a; Kumar *et al*., 2022b), electrodeposition (Dulgerbaki *et al*., 2018), tungsten metal electrochemical oxidation (Lin, 2020), and the sol-gel method (Zhang *et al*., 2019), among others. Metal oxide films can be electrochemically modulated to a non-stoichiometric redox state, which has an intense electrochromic absorption band due to optical range charge transfer.

Transition metal oxides with electrochromic properties can be colored cathodically or anodically, depending on the material that is colored under ion injection or extraction, *i.e.*, via a reduction or oxidation (Besnardiere, 2019). Cathodic coloration (blue) has been reported for Ti, Nb, Mo, Ta, and W oxides, with tungsten oxide (WO_3) being the first and most widely studied. Anodic coloration (red) is associated with Cr, Mn, Fe, Co, Ni, Rh, and Ir oxides, with nickel oxide (NiO) and iridium oxide (Ir_2O) being the most reported (Wang Z. *et al.*, 2020; Ding, 2021).

Table 2 presents some examples of coloration changes in both oxidized and reduced states, as well as the applied potential range of cathodic and anodic ionic insertion materials that are colored through electrochemical oxidation. The studies are aimed at obtaining the highest contrast between the oxidized and reduced state, as well as shades in the entire color range.

For cathodic oxides, the color variation can be established at any absorption level between states with maximum and minimum absorption. The mechanisms for optical absorption are still scarcely understood. However, they generally associated with charge transfer and polaron absorption (Sun, 2020; Szkoda *et al.*, 2020; Rathika, 2020).

Table 2. Electrochromic characteristics of some metal oxides

Cathode, c Anode, a	Oxidized state	Reduced state	E (V)	Area	Reference
MoO_3 c	Transparent	Blue	0,5	2,5 cm^2	(Szkoda *et al.*, 2020)
V_2O_5 c	Yellow	Black, blue	0,2	1,5 cm^2	(Rathika, 2020)
WO_3 c	Transparent	Blue	0,4	4 cm^2	(Bayrak *et al.*, 2021)
$Ir(OH)_3$ a	Blue black	Transparent	0,5	2,5 cm^2	(Jiang *et al.*, 2008)
$Ni(OH)_2$ a	Maroon	Transparent pale green	0,2	3 cm^2	(Morales-Torres *et al.*, 2009)

Source: Authors

For example, WO_3 exhibits a highly efficient blue coloration in amorphous films, which can reversibly change from transparent to dark blue by electrochemical redox reactions (Jittiarporn *et al.*, 2017; Gupta *et al.*, 2021; Gupta *et al.*, 2022a). Coloration characteristics depend on the constitution and composition of the film, as well as on the atomic structure and size of the nanoparticles, pores, and adsorbed substances. The insertion/extraction process can be represented via the following Equations:

$$WO_3 + x\left(M^+ + e^-\right)(Transp.)\ M_xWO_3\left(Blue\right) \quad (1)$$

$$WO_3 + x\left(Li^+ + e^-\right)(Transp.)\ Li_xWO_{x-1}W_x^{+4}WO_3^{+5}\left(Blue\right) \quad (2)$$

Tungsten trioxide, with all tungsten atoms in a W^{4+} oxidation state, is a transparent thin film. During electrochemical reduction, W^{5+} sites are generated to provide the electrochromic effect. The applications of electrochromism include manufacturing solid-state devices, where Li^+ is most often used. Applications have also been reported with Al^{+3} ions, given their affinity with tungsten trioxide (Li W., 2020; Tang *et al.*, 2020; Yao, 2020; Kumar *et al.*, 2022b).

The properties of WO_3 are combined with other electrochromic materials such as Nb_2O_5 (Tang, 2019), conductive polymers (Brooke *et al.*, 2018), and viologens (Yu, 2020) to improve contrast and color change efficiency. In commercial applications, as is the case of prototype alphanumeric displays, electrochromic mirrors are employed (Kim, 2016).

Although the W^{4+}/W^{5+} interval transition model mentioned above is the most accepted theory, which implies electron delocalization and is consistent with the conductivity enhancement that accompanies the insertion processes (Gesheva, 2016), other models involving non-localized electrons or band interaction (Meenakshi *et al.*, 2016) have been proposed to explain the electrochromic mechanism of tungsten oxide and inorganic ion insertion compounds in general.

The characteristics of the electrochromic process can be conventionally examined by directly comparing the optical and electrochemical response of the WO_3 electrode. The cyclic voltammetry and optical transmittance of a WO_3 electrode has been measured in a cell with a Li^+ ion-conducting electrolyte and a counter electrode (Dong, 2018).

Among the anodic oxides, nickel oxides with Li^+ electrolytes have been comprehensively studied. Nickel oxide can be bleached or colored by moving the Fermi level near the top of the valence band up or down. The electrochemical cycling of nickel oxide in different electrolytes takes place through the exchange of H^+ and OH^-. Given the nature of the film and the electrolyte, as well as the coloration mechanism of NiO, this is regarded as a surface phenomenon (Dong, 2020). After ion insertion and removal, the lattice constant of NiO exhibits changes in the grain unit cells (Kalanur *et al.*, 2020). The coloration of nickel oxide is based on the transition between Ni^{2+} and Ni^{3+} (Morales-Torres *et al.*, 2009)

$$NiOH_2 + Bleaching\ \ NiOOH\left(colored\right) + H^+ + e^- \quad (3)$$

Electrochromism has been studied under two possible cycling modes: the response between coloration and bleaching in steps of increasing potential (shift) and stepwise modulation towards coloration with increasing power and back to bleaching (modulation), observing that NiO films during oxidation exhibit a change in coloration

(Chen P. *et al.*, 2020). However, NiO films experience some issues such as inferior optical modulation and poor transparency in the bleached state. In this vein, doping with other transition metals such as Cu improves electrochromic performance with p-type conductivity for the films (Firat and Peksoz, 2019; Huang *et al.*, 2018). Likewise, the addition of surfactants such as dodecyl sulfate entails changes in the microstructure, morphology, and electrochromic performance of NiO films (Pham *et al.*, 2020).

Electrochromism and energy exchange functions can be achieved through redox reactions, indicating that the design of a bifunctional electrochromic energy storage device is feasible. In this context, NiO has been studied regarding its stability and compatibility with WO_3 electrodes (Thummavichai *et al.*, 2017). Interfacial charge transfer and ion transport generally cause a decrease in performance, which is due to irreversible redox reactions and ion accumulation or overpotential at the solution/electrode interface of a high-performance material. With the advent of multifunctional devices with electrochromic behavior and electrochemical energy storage capabilities, the complementary design of film structures using inorganic-organic materials has sparked great interest for applications in the field of optoelectronic research, as is the case of ambipolar electrochromic devices based on color-complementary WO_3 films (Zeng, 2021). In the technological architecture of hybrid systems in the construction of multifunctional tandem devices, replacing the conventional counter electrodes in a sensitized solar cell based on n-type titanium dioxide with a p-type nickel oxide photocathode yields energy conversion efficiencies close to 8,0%.

The hybrid cell design of WO_3 and NiO films obtained by sputtering reports a coloration efficiency of 360 cm^2C^{-1} at 640 nm, which depends on sputtering conditions (Koo *et al.*, 2020). The hexagonal electrode assembly of WO_3 and NiO nanoplates allows obtaining optical modulation values of 67,6% at 630 nm, a coloration efficiency of 109,6 cm^2C^{-1}, fast switching speeds (7,9 s), and a cycle life with a remaining ΔT of 84,1% after 4 000 cycles (Thummavichai *et al.*, 2017).

Another important factor affecting the electrochromic performance of NiO films is thickness; the optical transmission of the films decreases as their thickness increases. X-ray diffraction patterns reveal that polycrystalline phase and peak density increase with film thickness. The optical forbidden band decreases by about 0,2 eV with increasing film thickness, with values of 480 nm boasting a good performance (Atak and Coşkun, 2019). NiO nanocomposite films with gold for supercapacitor applications exhibit greatly enhanced pseudo-capacitive properties, 1,5 times higher than those of pure NiO (Jung, 2020).

Reports show more interest in NiO electrochromophores at the counter electrode, *i.e.*, as a redox reagent at the second electrode of an electrochromic cell. Here, there is a reactive primary electrochrome, such as WO_3 (Zheng, 2018) and conductive polymers such as poly(pyrrole), poly(thiophene), and poly(methylthiophene) (Liu Z., 2019). When it comes to the coloring mechanism, it should be noted that cathodic oxide (WO_3) and anodic oxide (NiO) films are different regarding their electrochemical performance. Ions can enter the lattice structure of cathodic oxide, whereas, in a Li^+-containing electrolyte, surface electrochemical processes dominate NiO films.

Viologens

The term *viologen* was coined by Michaelis and Hills for 4,4r-dipyridium compounds because they turn deep blue purple upon reduction. The ion may exhibit a two-step reduction, *i.e.*, one electron, or a two-state reduction (Monk *et al.*, 2015).

The most widely studied material in the viologen family is methyl viologen (MV). Elofson and Edsberg reported the first electron reduction of MV at 0,6 V vs. ECS. The coplanarity of the two heterocyclic nuclei facilitates a pH-independent, reversible reduction. The second electron reduction is not electrochemically reversible, but MV can be reoxidized by air. Reports state that the second reduction at 1,038 V vs. ECS is independent between 5,0 and 13,0 pH units and slightly dependent at a pH below 5,0 (Gugole, 2020).

Electrodes can be functionalized with electrochemically active polymeric materials derived from the hydrolytically unstable viologen, N, Nr-bis [-3-(trimethoxysilyl) propyl]-4,4r-bipyridinium dibromide, which shows potential for improved electrochromic efficiency in solid-state devices. Conventionally, electrochromic displays achieve contrast change by modulating light via the optical absorption of the electrodeposited film (Gugole, 2020).

The color change experienced by viologens is related to the functioning of the nitrogen substituent. For example, if the substituent is n-heptyl, a purple deposit is obtained; if it is p-cyano-phenyl, a green deposit is obtained. In addition, the electrical properties of the deposit can be tailored via the selection of the anion. Thus, for viologens of n-heptyl and bromide anions, the deposit is conductive, whereas the substitution of bromide with dihydrogen phosphate anions generates precipitates with insulating characteristics (Mishra, 2017). It has been reported that electrochromic devices based on viologens exhibit reversible multicolor changes at low voltages, with optical contrast values reaching 62,5% between 617 and 716 nm (Xu Z., 2020).

Conductive polymers

Conductive polymers have a leading role in the current challenges associated with efficient energy production. An example of this is their applications in fuel cells, with benefits for the economy and the versatility of shape and weight. Electroactive polymers show potential in technological applications such as electrodes in batteries (Wang L., 2020), microelectronics (Li S., 2020), electrochromic materials for sensor devices (Tsai *et al*., 2021), and catalyst support in fuel cells (Liu *et al*., 2019), among others. Conventionally, electrochromic displays achieve contrast change by modulating light via the optical absorption of the electrodeposited film, and they are good ionic and electronic conductors, with low processing costs and weight when compared to metallic materials used as electrodes.

Resistance to corrosive conditions in the cell and a porosity that enables permeability to ions (*e.g*., hydrogen ions) significantly increase the reaction area. The low thickness in comparison with metal oxides facilitates the serial assembly of several cells in reduced sizes, which has made electroactive polymers the most widely applied materials in fuel cells (Pandey, 2016). These are electrochromic materials that offer a high degree of color matching by coupling various polymer systems through functionalization and/or copolymerization of monomers, as well as through the use of blends, laminates, and composites. Complex colors are achieved by mixing two existing colors in a dual-polymer device. Electrochromic changes are induced by redox processes accompanied by ion insertion/expulsion, and the color exhibited by the polymer is closely related to the band gap and the doping ions. A focus area in the study of electrochromic polymeric materials is the control of their colors by modifying the structure of the main and pendant chain group. Polyheterocycles have proven to be of special interest in this regard due to their stability under ambient conditions (Gao *et al*., 2020). Table 3 provides a comparison of the electrochromic properties observed in the most widely studied conductive polymers.

Table 3. Electrochromic effect of some conductive polymers

Polymer	Anion	Oxidized state	Reduced state	Time (ms)
PPy	ClO_4^-	Brown	Yellow	20
Polythiophene	ClO_4^-	Brown	Green	45
Polymethylthiophene	BF_4^-	Blue	Red	12
PEDOT	ClO_4^-	Dark Blue	Blue	60
Poly-2,2r-bithiophene	SF_3^-, SO_3	Blue	Red	40
Octacyano phthalocyanine	H^+	Grey Green	Blue	1
PANI	Cl^-	Blue-green	Yellow	100

Source: Authors

Polypyrrole (*PPy*). The first PPy films were usually prepared via the electrochemical polymerization of pyrrole (Py) in an aqueous or organic solvent solution such as acetonitrile, together with a suitable electrolyte. This process involves the oxidation and dimerization of Py (Yin, 2020). The polymerization reaction involves two electrons per reacting Py molecule, and the resulting polymer is produced in an oxidized state with 0,25-0,33 cation centers per unit of Py depending on the electrolyte (Matysiak et al., 2020). The doped (oxidized) PPy film is blue-violet (λ max 670 nm), and electrochemical reduction produces an 'undoped' yellow-green form (λ max 420 nm) (Kazazi, 2019). PPy and polyindole (PIn)-based electrodes exhibit a high degree of flexibility, which makes them remarkably suitable and applicable for fabricating flexible supercapacitors. Py derivatives with methyl red (PMRPy) electropolymerized on glass electrodes improve chromatic contrast ($\Delta\%T=34,2\%$), switching times ($\tau=10$ s), and stability ($\Delta\%T=15\%$ in the 100th cycle) with respect to Ppy films. Their coloration changes (magenta in acidic media and yellow in basic ones) can be used in pH sensors (Choudhary et al., 2020).

Polythiophene. The electrochromic properties of polythiophene and polymers of various substituted thiophenes are a matter of great interest, as color states can be tuned via the appropriate choice of thiophene monomer. For example, the electrochromic properties of polymer films prepared with 3-methylthiophene-based oligomers strongly depend on the relative positions of the methyl groups in the polymer's main chain (Table 4), and multicolor systems can be constructed by combining different conducting polymer species. Polyethylenedioxythiophene (PEDOT) can prevent halogen ion poisoning in the reactive area in protic media (Dias et al., 2019). PPy thin films can be prepared via the electrochemical polymerization of 3,4-ethylenedioxythiophene. During the electropolymerization process, the anion in the electrolyte directly influences the uniformity, morphology, and electrochromic performance of PEDOT films (Xu D., 2020).

Table 4. Most common polytyophenes and their coloration

Polyphenols	Color	
Monomer	Oxidized state	Reduced state
Thiophene	Blue	Red
3-methylthiophene	Dark blue	Red
3,4-dimethylthiophene	Dark blue	Pale brown
2,2r-Bithiophene	Blue-gray	Orange

Source: Zhang and Feng (2020)

Electrochromic layers of PEDOT copolymerized with polystyrene sulfonate (PSS) and intercalated on a Ni lattice (Chen, C. 2020), Au/PEDOT/Pt microtubes on WO_3 films (Karaca *et al*., 2019), the imidazole/Fe(III) p-toluensulfonate weight ratio (Wang W., 2020), the adhesion to acidified nanostructured tubular nanostructured silicoaluminate clays, and Halloysite (HNT) (Hu *et al*., 2019) allow obtaining devices with short staining and decolorization times (less than 20 s), high optical contrast (close to 90%), staining efficiency (higher than 300 cm^2C^{-1}), and excellent mechanical reliability, which corresponds to optical contrast retention after 1 000 cycles (Zhang and Feng, 2020).

Polyaniline (*PANI*). The electrical and electrochromic properties of PANI depend on its oxidation state, its protonation state, the pH of the electrolyte used, the redox mechanisms involving protonation/deprotonation, and/or the input/output of anions, which allows it to exhibit reversible color changes at the potential: transparent leukoesmeraldine to yellowish green emeraldine to dark blue-black pernigraniline in the potential range of -0,2 to +1,0 V vs. SCE (Begum *et al*., 2013).

The use of donor-acceptor systems between PANI films and the porous spaces between them facilitates the diffusion of the ions involved in charge transfer reactions, exhibiting increased device performance, contrast parameters close to 38,5%, redox reaction efficiencies close to 200 cm^2C^{-1}, response speeds close to 1,0 s, and cyclic stability, with optical modulations of 98% for at least 5 000 cycles.

Some examples of multifunctional electrochromic devices that also exhibit energy storage capabilities and are based on organic/inorganic composites include the insertion of Cu(II) ions in PANI films (Deshmukh *et al*., 2017), nanostructured mesoporous Ni-PANI thin film electrodes (Inamdar *et al*., 2019), nanofibers of PANI entrapped in antimony-doped tin oxide, nanofilaments using aminophenyl diazonium cations as precursors of growth by electro-polymerization in silica mesochannels (Ullah *et al*., 2021), NiO nanoparticles functionalized with (3-aminopropyl) trimethoxysilane (3-APTS) subsequently attached to aniline monomers (Jamdegni and Kaur, 2020), $W_{18}O_{49}$/PANI hybrid films (Chang, 2018), and V_2O_5/ PANI composites obtained through in situ chemical oxidation polymerization (Zhang, 2018).

The versatility of polymers allows for couplings of different families, self-assemblies (e.g., polystyrene and PANI films using natural sedimentation and *in situ* polymerization of aniline monomers) (Wu, 2019), thin core/film structure formation through π-π stacking between aromatic rings of polydiphenylamine and monodisperse PANI via oxidative polymerization (Han *et al*., 2020), and semiconducting enzymatic polygalic acid (PGAL) templates with PEDOT and PANI in complexes that can be processed by sputter coating to produce flexible electrochromic devices (Díaz-Sánchez *et al*., 2019), among others, improving features such as the efficiency, contrast, and permutation cycles of smart windows.

The color change and conductivity of PANI films are applied in the design of other types of devices, such as multicolor immunoassay platforms for clinical diagnostics. With visible light excitation, Ag_2S NPs@ZnO NTs generate electron holes. The photoelectrons pass into the PB and the photoinduced holes oxidize the PANI, causing color changes from emerald green to purple-blue and black depending on the carcinoembryonic antigen's concentration (Wang Y. *et al*., 2020). Conductive emulsions of PANI and polyacrylate (PA) obtained via chemical polymerization in the presence of poly (maleic anhydride-co-co-acrylic acid-co-sodium p-styrenesulfonate) as a stabilizer exhibit the electrical conductivity and mechanical properties necessary for applications in antistatic fabric finishing and functional textile development (Wu L. *et al*., 2018). Likewise, the structural stabilization of PANI films obtained by coating conductive carbon nanotube networks provides multiple directions for electron transfer and, therefore, improves charge capacity by forming 3D networks and enhancing optical transmittance and capacity, which is relevant in the construction of smart supercapacitors (Xu K., 2020).

Electrolytes

In electrochromic devices, when a potential is applied, the ions in the electrolyte are transported between the layers, generating a change in the transparency of the chromic material. The initial electrolyte prototypes had aqueous characteristics, entailing conductivities that were nowhere near the expected values (around 1 $mS.cm^{-1}$) for industrial applications, which demand reliable, safe, and easily transportable and implementable architectures. Hence, there is a global search for electrolytes that allow manufacturing solid-state electrochromic devices.

Lithium-based electrolytes have gained great attention due to their environmental stability. They are made up of mixtures of $LiClO_4$ and one or more polymers such as polyvinyl chloride (PVC) (Hernández, 2020), propylene carbonate (PC) and ethylene carbonate (EC) (Dalavi *et al*., 2020), and aprotic salt with polymethylmethacrylate (PMMA) (Ganesh *et al*., 2015), which entails a transmittance higher than 90% and a durability greater than 5 000 cycles. However, there are other types of electrolytes, e.g., ionic liquids such as coPIL, which is composed of trioxyethylene and tetraoxyethylene spacers containing 1,2,3-triazolium and offers high degradation temperatures (230 °C) and an ionic conductivity lower than 9×10^{-5} Scm^{-1}. This has allowed for conversion times shorter than 12 s and stability after 3 000 cycles (Puguan and Kim, 2019). Gelatin-based solid electrolytes, cross-linked with formaldehyde, glycerol, acetic acid, and/or HCl have also been used, obtaining ionic conductivity values of $1,28\text{x}10^{-5}$ Scm^{-1} and a percent transmittance of 98% at wavelengths above 580 nm (Ramadan *et al*., 2017). Environmentally friendly hydrogels containing deionized water, sodium carboxymethylchitosan, and vitamin C are used as electrolytes of the electrochromophore bis (2-(2-(2-(2-hydroxyethoxy) ethoxy) ethyl) viologen bromide, with colorless to purple shifts between 0 and -1,2 V (Zhang Y. *et al*., 2021), among others. Some authors integrate ITO/VO_2/$LiTaO_3$/WO_3/ITO-type devices on glass substrates, where several chromic processes are evidenced: a thermochromic layer passively responding to ambient temperature, another electrochromic layer actively and simultaneously responding to external voltage, and reversible Li^+ migration within the lattice at low bias voltages (Jia, 2019).

Table 5 compiles some focus points regarding potential electrochromic devices for use in smart windows.

Possibilities for chromic materials innovation in the Colombian industry

Colombia exhibits a significant extraction of Fe, Co, Ni, Mo, C, W, and Ti, mainly in the northern part of the country, Cundinamarca, and the coffee axis region, which would allow consolidating industrial applications with tungsten and nickel oxide inputs. There are also tungsten deposits in the departments of Guaviare, Vichada, and Caquetá, where extraction activities were carried out between 2012 and 2016. Although the country lacks information on the extraction or production of tungsten as a mineral, information from DANE allows stating that the national industries' consumption is centered on tungsten plates. Imports of tungsten come in the form of sintered bars (industrial materials composed of tungsten) and not in the mineral's raw form. On the other hand, the country showed exports of 87 metric tons between 2012 and 2015 (Prieto *et al*., 2019). Likewise, in Colombia, there are six nickel deposits, three of which are in the Caribbean region, *i.e*., in the department of Córdoba: Cerro Matoso, Planeta Rica, and Uré. The remaining three are in the department of Antioquia: Ituango, Morro Pelón, and Medellín. Colombian nickel production remains above 40 000 tons since 2012 (DNP, 2017). Data as of 2018 establish Colombia as the first producer of nickel in South America and the third in Central America and the Caribbean, only after Cuba and the Dominican Republic (MinMinas, 2018).

A strategic union between academia and national companies could be the link that the national glass industry needs to become a beacon of electrochromic, gasochromic, or thermochromic technologies. Colombia has industries in the manufacture of flat glass whose inputs are national, such as Tecnoglass SA, Cristar SAS, Cristalería Peldar SA, AGP de Colombia SA, and Fiberglass Colombia SA, among others. The import of raw materials does not exceed 16%, and, regarding mirrors and fibers, it does not reach 19%. Plates and smooth sheets of cast or laminated, unassembled, colored, opacified, and plated glass, as well as glass with absorbent, reflective, or anti-reflective coating, are the most traded products abroad. The United States accounts for 60,25% of the exports, followed by Mexico with 7,76%, Ecuador with 7,65%, and Peru 7,15% (Legis-Comex, 2017).

The manufacture of chemical products in Colombia, which has all the necessary inputs to produce chromic technologies, grew with a compound annual rate of 8,5% between 2000 and 2018. The departments with the highest production of chemical products are Cundinamarca, Antioquia, Valle, and Bolívar. Between 2014 and 2016, the gross production of this sector increased by 21,1%. However, 8,4% of their production facilities closed. Projections for 2030 show a compound annual growth rate of 6,5%, with a market growth rate higher in proportion than that of production (4,7% by 2032) (MinCIT, 2019a). The expectations regarding innovation programs and the strengthening of the chemical supply chain, which include chromic devices, should provide added value, allowing the Colombian chemical industry to replace current imports (MinCIT, 2019b).

Industrial-level scientific initiatives on topics related to smart windows and mirrors can be covered by the lines of research that MINCIENCIAS has been promoting, which focus on converging technologies and the Industry 4.0, especially on nanotechnology, competitiveness, and productivity; sustainable energy, aiming for energy efficiency in heating systems, lighting in urban areas, direct and indirect heat, and/or efficient energy management; and geosciences, promoting the identification, characterization, quantification, and industrial exploitation of mineral deposits with added value (MinCiencias and Vicepresidencia de la República de Colombia, 2020).

The qualified human resources for the implementation and research of chromic devices or related technologies and materials can be supported by more than 50 research groups endorsed by MinCiencias in the 2021 call for proposals (MinCiencias, 2022). In addition, the higher education system has more than 180 professional technical and technological programs, technological specializations, and professional technical specializations, as well as more than 140 professional programs and more than 170 postgraduate programs in areas such as nanotechnology, materials, geology, electronics, clean technologies, energy resources, automation of industrial processes, mines, and chemistry, among others (SNIES, 2021). These programs can be oriented towards theoretical and experimental competencies in innovation in the implementation of industries associated with electrochromic, photochromic, and gasochromic technological devices, in addition to fuel cells, smart windows, and mirrors, among others.

Finally, as for the feasibility of implementing industries that apply chromic properties, the differential capabilities of the Colombian industry should be highlighted. The availability of natural resources; the country's location; free trade agreements with consumer and producer countries; the culture of cost reduction, quality, and compliance with requirements, regulations, and standards; the institutions that promote the culture of continuous improvement to the production process (Colombia Productiva, INNpulsa and Pro-Colombia); and customer service levels are all factors to be leveraged by companies. However, there is a need to streamline government processes, to train personnel in export policies, to promote internalization in productive clusters that project the acquisition of investors, to support new regulations, and to establish productive linkages in the chemical sector through dialogues aimed at collective development.

Table 5. Electrochromic effect of some conductive polymers

Substrate	Electrochromic material deposition method	η (cm^{2}C^{-1})	ΔT (%)	modulation potentials (V)	λ (nm)	Response time (s)	Cycles with retention >90% ΔT	Reference
Glass/ITO/ WO_3/ $LiClO_4$ in PC/WO_3-in/ITO/Glass	Cathodic sputtering			1,5 to 4,0, *vs.* Li / Li$^+$			500	(Shi *et al.*, 2020a)
Si-SiO_2/ WO_3/ $LiNbO_3$/NiO/C	Electron beam evaporation			0,7 to -0,9	500 700	15		(Buljam, 2017)
PET/ITO/ WO_3/ dmFc), PVDF-co-HFP, [EMI] [TFSI] [Li]/ ITO/PET	Spin-coating	57,4	85	−0,9 to 0,0		(tb) 15 (tc) 5	1 000	(Yun *et al.*, 2019)
PET/ITO/ WO_3 $LiClO_4$ in PMMA/ ITO/ PET	DC magnetron sputtering		64	-4 to +3	633	(tb) 3,75 (tc) 2,35	5 000	(Chen X., 2020)
Glass/ITO/ nanoplates of WO_3/ [C_3mpyr][BF4]/ ITO/Glass	Ultrasonic irradiation	52,4	70	1,0 to − 0,5	520	(tb) 23 (tc) 2,8	500	(De Ribamar Martins Neto *et al.*, 2016)
Glass/ITO/ Pt/ Cu-Pb gel in hydroxyethylcellulose/Cu-Ag	Electrodeposition		90	-0,6 to 0,8	600	10	5 500	(Barile *et al.*, 2017)
PC/ PEDOT:PSS, ethylene glycol, benzenesulfonic acid, methanesulfonic acid, p-toluenesulfonic acid/ PVA/ H_3PO_4/ PC	Spin-coating	22		0 to 1,8	630	20	9 000	(Barus *et al.*, 2018)
PET/ITO/ Nano sheets WO_3-H_2O/ $LiClO_4$ in PC/ ITO/PET	Spin-coating	120,9	48 67	±3	798 1 300	(tb) 5,1 (tc) 9,7	400	(Pan, 2020)
Glass/FTO/ TiO2 (TiPOT)/ ge of citric acid, ethanol, Li_2CO_3, and tetraethylorthosilicate/ FTO/Glass	Spin-coating	82,7	28	0,6 to -1,8	390 to 2 500	17,8	400	(Zhang, *et al*, 2019)
Glass/FTO/ h'-H_xWO_3/ H_2SO_4 0.1M/ Cu	Spin-coating	53	24	±0,8 *vs.* Ag/ AgCl	700	13	150	(Besnardiere, 2019),
Glass/ITO/NiO:(Li, Mg)/Ta_2O_5/1 M PC/$LiClO_4$/WO_3/ITO	DC magnetron sputtering	60	66	±1,5	500 to 600	(tb) 1,0 (tc) 2,5	1 500	(Dong, 2018)
Glass/ITO/ WO_3/$LiClO_4$ PC/ NiO/ITO/ Glass	Cathodic arc plasma	90	46	0,8 to -1,4	633	3,1	2 500	(Chen P. *et al.*, 2020)
Glass/ITO/EDBT; EDOT; thieno [3,4-b] pyrazine and benzo [1,2,3] triazole/PEDBT and PMTTP/ ITO/Glass	Electropolymerization		64	−0,2 to 1,0	590 1 500	2		(Xu Z., 2020)
Glass/FTO/PEDOT/KCl 3M, 0.1$LiClO_4$/Pt	Electropolymerization		15	±3	500-645	(tb) 0,15 (tc) 2,30	2 000	(Xu D., 2020)
Glass/ITO/Ni mesh / PEDOT:PSS/ $TABPF_6$/Ni mesh /ITO/Glass PET/ITO/Ni mesh / PEDOT:PSS/ $TABPF_6$/Ni mesh /ITO/PET	Spin-coating		15	3,5 to -2,5	644	1	1 000	(Chen C., 2020)
PET/ITO/ WO_3-Au/PEDOT/Pt/ $LiClO_4$ in PMMA, PC/ITO/PET	Electron beam evaporation: WO_3 Spin-coating: Au/ PEDOT/Pt		26	±2	750	(tb) 13,83 (tc) 19,25	1 000	(Karaca *et al.*, 2019)
Glass/FTO/PEDOT in (Im/Fe)/ $LiCLO_4$/ FTO7Glass	Vapor phase polymerization		18,5	0,4 to -0,2	550	10	1 500	(Wang, W., 2020)
Glass/ITO/HNT@PEDOT/0.1 M $LiClO_4$/DMSO/ITO/Glass	Oxidative chemical polymerization (HNT) Spin-coating (PEDOT)	174,3	59,3	−0,5 to 0,7	450 to 700	(tb) 2,8 (tc) 2,7		(Hu *et al.*, 2019)
Glass/ITO/(Cu(II)/PANI, H_2SO_4 0.2M/ ITO/Glass	Electrodeposition		10	-0,1 to 1,4 *vs.* Ag/AgCl	400-900	1,85		(Deshmukh *et al.*, 2017)

Glass/ITO/ Ni-PANI/0.5 M $LiClO_4$ +PC/ ITO/Glass	Electrodeposition	45,9	66	-0,5 to 0,7 *vs.* SCE	630			(Inamdar, *et al*, 2019)
Glass/ITO/ PANI-NiO/Pt	Electrodeposition	138	59 40 50	0,7 to -0,2	310 648 1 100	(tb) 2,0 (tc) 1,5	7 000	(Jamdegni and Kaur, 2020)
Glass/ITO/ $W_{18}O_{49}$/PANI/1 M $AlCl_3$/ ITO/Glass	Spin-coating: $W_{18}O_{49}$ Electrodeposition: PANI	45,68	25,8	±1	632	20		(Chang X., 2018)
Glass/ITO/ V_2O_5@PANI/1 M $LiClO_4$ in PC/ ITO/Glass	Polymerization by chemical oxidation		40	−0,6 to 0,7	665	(tb) 1,5 (tc) 2,3	1 000	(Zhang K., 2018)
Glass/ITO/ PS@PANI/ H_2SO_4 0.1 M/ Pt	Natural sedimentation (reflectance analysis)		4	0,2 to 1,0	278, 246, 224	10		(Wu, 2019)
PET/ITO/ PEDOT-PGAL/PANI-PGAL/ PMMA-PC-$LiClO_4$/ITO/PET	Spray coating	275	15	0,3 to -1,0 *vs.* Pt	590	3,5	1 000	(Díaz-Sánchez *et al*., 2019)
PET/ITO/ PEDOT-PSS/PANI-PGAL/ PMMA-PC-$LiClO_4$/ITO/PET	Spray coating	517	26	0,3 to -1,0 *vs.* Pt	635	(tb) 1,5 (tc) 2,1	1 000	
Glass/ITO/ PANI@CNT/1 M $LiClO_4$/ITO/Glass	Spin-coating: carbon nanotube (CNT) Electropolymerization PANI	120	46,6	±1 *vs.* Ag/AgCl	650	(tb) 5,8 (tc) 7,9	8 000	(Xu K., 2020),
Glass/FTO/ WO_3/V_4/0.5 M $LiClO_4$ + PC/CeO_2-V_2O_5 /FTO/Glass	Spray pyrolysis	76	8	± 0,9 *vs.* SCE	630			(Dalavi *et al*., 2020)
Glass/ITO/ PEDOT:PSS/ coPIL/ PEDOT:PSS/ITO/Glass	Spin-coating		24	0 to 1,8	630	(tb) 6,2 (tc) 11,8	3 000	(Puguan and Kim, 2019)
Glass/FTO/ WO_3/ gelatin with formaldehyde, glycerol, acetic acid or chlorhydric acid/NiO/FTO/Glass	Spray pyrolysis	38,1 43,1	67,7	± 2,5	600 1 100			(Ramadan *et al*., 2017)
Glass/ITO/ HEV^{2+} $2Br^-$/ carboxymethylchitosan sodium hydrogel, vitamin C /ITO/Glass		243 119	64,1	0 to 1,2	520 865	(tb) 7,8 (tc) 64,5 (tb) 9,8 (tc) 44,7	1 000	(Zhang Y. *et al*., 2021)
Glass/FTO/WO_3/ $LiTaO_3$/ VO_2/FTO/ Glass	DC magnetron sputtering		37,5	±3	580	(tb) 5 (tc) 15	1 000	(Jia, 2019)

Source: Authors

Conclusions

Studies on the physical and chemical properties of diverse organic and inorganic electrochromic materials show potential for the manufacture of electrochromic devices with industrial applications such as rear-view mirrors in the automotive sector, alphanumeric screens, monitors, and intelligent windows, among others. Among the inorganic materials used, NiO (anodic oxide) and WO_3 (cathodic oxide) stand out, which could be regarded as *classical electrochromic materials* since they boast a better electrochromic performance. Couplings have been made for them with organic materials such as conductive and viologen polymers in the construction of solid-state devices. In addition to electrochromic properties, studies on electrochromic devices have shown a tendency towards the modulation of radiation in the near-infrared, thermal-infrared, and microwave regions. 'Color' may represent the detectors' response to these wavelengths (*i.e.*, thermochromic, photochromic) as well as energy storage capacity in the manufacture of supercapacitors. Widespread applications of architectural electrochromic devices rely on reducing costs, increasing the useful life of the device, and overcoming the issue of degradation. Thus, the use of polymer electrolytes with lithium ions has been widely studied, as well as innovations with ionic liquids and polymers with a low environmental impact.

The trajectory of the Colombian industrial sector, together with the human talent in local research groups, and the training offer for personnel at all levels (technical, professional, and postgraduate) are inputs for the potential implementation of an industry with great social and economic impact. Colombia has mining resources and companies in the glass sector and in the manufacture of chemical products that, if directed towards industrial innovation clusters and leveraged by different sectors of the government and national and international Investors, could entail the establishing of companies with chromic applications in the medium term, especially regarding the use of inorganic inputs. In the long term, if the countr''s chemical industry, which works with petroleum derivatives, shows interest in obtaining inputs related to conductive polymer monomers and viologens applicable in chromic devices, the need to project the manufacture of organic and inorganic chromic materials could be resolved.

Acknowledgements

The authors would like to thank Dirección General de Investigaciones (DGI) of Universidad Santiago de Cali for funding the project 939-621119-477.

CRediT author statement

Conceptualization L.F.H.P. and S.P.C.N. Methodology: all authors. Validation and formal analysis: all authors. Writing (original draft, review): L.F.H.P. and S.P.C.N. Editing: S.P.C.N. and A.J.A. All authors have read and agreed to the published version of the manuscript

Conflicts of interest

The authors declare no conflicts of interest.

References

Atak, G., and Coşkun, Ö. D. (2019). Effects of anodic layer thickness on overall performance of all-solid-state electrochromic device. *Solid State Ionics, 341*, 115045. https://doi.org/10.1016/j.ssi.2019.115045

Atak, G., Bayrak Pehlivan, İ., Montero, J., Primetzhofer, D., Granqvist, C., G., and Niklasson, G. A. (2020). Electrochromism of nitrogen-doped tungsten oxide thin films. *Materials Today Proceedings, 33*(6), 2434-2439. https://doi.org/10.1016/j.matpr.2020.01.332

Barile, C. J. Slotcavage, D. J. Hou, J. Strand, M. T. Hernandez, T. S., and McGehee, M. D. (2017). Dynamic windows with neutral color, high contrast, and excellent durability using reversible metal electrodeposition. *Joule*, 1(1), 133-145, https://doi.org/10.1016/j.joule.2017.06.001

Barus, D. A., Sebayang, K., Ginting, J., and Ginting, R. T. (2018). Effect of chemical treatment on conducting polymer for flexible smart window application. *Journal of Physics: Conference Series, 1116*(3), 4-9. https://doi.org/10.1088/1742-6596/1116/3/032006

Bayrak Pehlivan, İ., Atak, G., Niklasson, G. A., Stolt, L., Edoff, M., and Edvinsson, T. (2021). Electrochromic solar water splitting using a cathodic WO3 electrocatalyst, *Nano Energy*, *81*, 105620. https://doi.org/10.1016/j.nanoen.2020.105620

Begum, A. N., Dhachanamoorthi, N., Saravanan, M. E. R., Jayamurugan, P., Manoharan, D., and Ponnuswamy, V. (2013). Influence of annealing effects on polyaniline for good microstructural modification. *Optik*, *124*(3), 238-242. https://doi.org/10.1016/j.ijleo.2011.11.096

Besnardiere J. (2019). Structure and electrochromism of two-dimensional octahedral molecular sieve h'-WO 3. *Nature Communications*, *10*(1), 1-9. https://doi.org/10.1038/s41467-018-07774-x

Brooke, R., Edberg, J., Iandolo, D., Berggren, M., Crispin, X., and Engquist, I. (2018). Controlling the electrochromic properties of conductive polymers using UV-light. *Journal of Materials Chemistry C, 6*, 4663-4670. https://doi.org/10.1039/c7tc05833k

Bulja, S. (2017). Tuneable dielectric and optical characteristics of tailor-made inorganic electro-chromic materials. *Nature Scientific Reports, 7*, 3-10. https://doi.org/10.1038/s41598-017-13941-9

Chang, X. (2018). Sunlight-charged electrochromic battery based on hybrid film of tungsten oxide and polyaniline. *Applied Surface Science*, *441*, 105-112. https://doi.org/10.1016/j.apsusc.2018.02.003

Chen, C. (2020). High-performance embedded nickel grid electrodes for fast-response and bendable all-solid PEDOT: PSS electrochromic devices. *Organic Electronics, 77*, 105506. https://doi.org/10.1016/j.orgel.2019.105506

Chen, X. (2020). Bio-inspired flexible vibration visualization sensor based on piezo-electrochromic effect. *Journal of Materiomics*, *6*(4), 643-650. https://doi.org/10.1016/j.jmat.2020.06.002

Chen, P. W., Te-Chang, C., Ko, T. F., Hsu, S. C., Li, K. D., and Wu, J. Y. (2020). Fast response of complementary electrochromic device based on WO3/NiO electrodes. *Nature Scientific Reports, 10*, 1-12, https://doi.org/10.1038/s41598-020-65191-x

Choudhary, R. B., Ansari, S., and Purty, B. (2020). Robust electrochemical performance of polypyrrole (PPy) and polyindole (PIn) based hybrid electrode materials for supercapacitor application: A review. *Journal of Energy Storage, 29*, 101302. https://doi.org/10.1016/j.est.2020.101302

Dalavi, D. S., Bhosale, A. K., Desai, R. S., and Patil, P. S. (2020). Energy efficient electrochromic smart windows based on highly stable CeO_2-V_2O_5 optically passive counter electrode. *Materials Today: Proceedings, 43*, 4, 2702-2706. https://doi.org/10.1016/j.matpr.2020.06.146

DANE (2020a). *Boletín técnico PIB IV trimestre 2019*. https://www.dane.gov.co/files/investigaciones/boletines/pib/bol_PIB_IVtrim19_producion_y_gasto.pdf

DANE (2020b). *Boletín técnico pobreza monetaria en Colombia año 2019*. https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/2019/Boletin-pobreza-monetaria_2019.pdf

De Ribamar-Martins-Neto, J., Torresi, R. M., and Cordoba De Torresi, S. I. (2016). Electrochromic behavior of WO_3 nanoplate thin films in acid aqueous solution and a protic ionic liquid. *Journal of Electroanalytical Chemistry, 765*, 111-117. https://doi.org/10.1016/j.jelechem.2015.08.032

Deshmukh, M. A., Gicevicius, M., Ramanaviciene, A., Shirsat, M. D., Viter, R., and Ramanavicius, A. (2017). Hybrid electrochemical/electrochromic Cu(II) ion sensor prototype based on PANI/ITO-electrode. *Sensors and Actuators B: Chemical, 248*(II), 527-535. https://doi.org/10.1016/j.snb.2017.03.167

Dias, O. A. T., Konar, S., Leão, A. L., and Sain, M. (2019). Flexible electrically conductive films based on nanofibrillated cellulose and polythiophene prepared via oxidative polymerization. *Carbohydrate Polymers, 220*, 79-85. https://doi.org/10.1016/j.carbpol.2019.05.057

Díaz-Sánchez, J., Roquero, P., Hernández-Alcántara, J. M., Rosas-Aburto, A., Vázquez-Torres, H., and Gimeno, M. (2019). Complementary electrochromic devices of PANI and PEDOT using the enzymatic poly(gallic acid). *Solar Energy Materials and Solar Cells*, *200*, 109973. https://doi.org/10.1016/j.solmat.2019.109973.

Ding, S. (2021). High-temperature flame spray pyrolysis induced stabilization of Pt single-atom catalysts. *Applied Catalysis B: Environmental*, *281*, 119471. https://doi.org/10.1016/j.apcatb.2020.119471

Departamento Nacional de Planeación (DNP) (2017). *Vidrio en Colombia*. https://www.revistaespacios.com/a02v23n01/02230122.htm

Dong D. (2018). Enhanced electrochromism in short wavelengths for NiO:(Li, Mg) films in full inorganic device ITO/NiO:(Li, Mg)/Ta_2O_5/WO_3/ITO. *Electrochimical Acta*, *263*, 277-285. https://doi.org/10.1016/j.electacta.2018.01.049

Dong, D. (2020). Electrochromic and colorimetric properties of anodic NiO thin films: Uncovering electrochromic mechanism of NiO. *Electrochimical Acta*, *335*, 135648. https://doi.org/10.1016/j.electacta.2020.135648

Dulgerbaki, C., Maslakci, N. N., Komur, A. I., and Oksuz, A. U. (2018) Electrochromic strategy for tungsten oxide/polypyrrole hybrid nanofiber materials. *European Polymer Journal, 107*, 173-180. https://doi.org/10.1016/j.eurpolymj.2018.07.050

El-Nahass, M. M., Saadeldin, M. M., Ali, H. A. M., and Zaghllol, M. (2015). Electrochromic properties of amorphous and crystalline WO_3 thin films prepared by thermal evaporation technique. *Materials Science in Semiconductor Processing, 29*, 201-205. https://doi.org/10.1016/j.mssp.2014.02.051

Farhad, S. F. U. (2021). The effect of substrate temperature and oxygen partial pressure on the properties of nanocrystalline copper oxide thin films grown by pulsed laser deposition. *Data in Brief*, *34*, 106644. https://doi.org/10.1016/j.dib.2020.106644

First, Y. E., and Peksoz, A. (2019). Efficiency enhancement of electrochromic performance in NiO thin film via Cu doping for energy-saving potential. *Electrochimica Acta, 295*, 645-654. https://doi.org/10.1016/j.electacta.2018.10.166

Fletcher, S. (2015). The definition of electrochromism. *Journal of Solid State Electrochemistry, 19*, 3305-3308. https://doi.org/10.1007/s10008-015-3039-9

Ganesh, G. P. T., Ravi, R., and Deb, B. (2015). A pragmatic approach to methyl methacrylate based solid polymer electrolyte processing: A case study for electrochromism. *Solar Energy Materials and Solar Cells*, *140*, 17-24. https://doi.org/10.1016/j.solmat.2015.03.022

Gao, S., Sun, F., Liu, N., Yang, H., and Cao, P.-F. (2020). Ionic conductive polymers as artificial solid electrolyte interphase films in Li metal batteries–A review, *Materials Today*, *40*, 140-159. https://doi.org/10.1016/j.mattod.2020.06.011

Gesheva, K. (2016). Optical, structural and electrochromic properties of sputter- deposited W-Mo oxide thin films. *Journal of Physics: Conference Series*, 764, 012010. https://doi.org/10.1088/1742-6596/764/1/012010

Gugole, M. (2020). High-contrast switching of plasmonic structural colors: Inorganic versus organic electrochromism. *ACS Photonics*, *7*(7), 1762-1772. https://doi.org/10.1021/acsphotonics.0c00394.

Gupta, J., Shaik, H., Kumar, K. N. (2021). A review on the prominence of porosity in tungsten oxide thin films for electrochromism. *Ionics, 27*(6), 2307-2334. https://doi.org/10.1007/s11581-021-04035-8

Gupta, J., Shaik, H., Kumar, K. N., and Sattar, S. A. (2022a). PVD techniques proffering avenues for fabrication of porous tungsten oxide (WO3) thin films: A review. *Materials Science in Semiconductor Processing, 143*, 106534. https://doi.org/10.1016/j.mssp.2022.106534

Gupta, J., Shaik, H., Kumar, K. N., Sattar, S. A., and Reddy, G. A. (2022). Optimization of deposition rate for E-beam fabricated tungsten oxide thin films towards profound electrochromic applications. *Applied Physics A*, *128*(6), 498. https://doi.org/10.1007/s00339-022-05609-7

Han, W. J., Lee, J. H., and Choi, H. J. (2020). Poly(diphenylamine)/polyaniline core/shell composite nanospheres synthesized using a reactive surfactant and their electrorheology. *Polymer, 188*, 122161. https://doi.org/10.1016/j.polymer.2020.122161

Hasani A. (2017). Facile solution synthesis of tungsten trioxide doped with nanocrystalline molybdenum trioxide for electrochromic devices. *Scientific Reports, 7*(1), 1-10. https://doi.org/10.1038/s41598-017-13341-z

Hernández, T. S. (2020). Electrolyte for improved durability of dynamic windows based on reversible metal electrodeposition. *Joule, 4*(7), 1501-1513. https://doi.org/10.1016/j.joule.2020.05.008

Hidalgo-Rodríguez., J. (2002). *Electrocromismo de los metales de transición*. [Undergraduate thesis, Universidad Nacional de Ingeniería]. http://repositorio.uni.edu.pe/bitstream/uni/9151/1/hidalgo_rj.pdf

Ho, C. H., Kuo, Y. M., Chan, C. H., and Ma, Y. R. (2015). Optical characterization of strong UV luminescence emitted from the excitonic edge of nickel oxide nanotowers. *Scientific Reports, 5*, 1-7. https://doi.org/10.1038/srep15856

Hu, F., Peng, H., Zhang, S., Gu, Y., Yan, B., and Chen, S. (2019). PEDOT nanoparticles fully covered on natural tubular clay for hierarchically porous electrochromic film. *Solar Energy Materials and Solar Cells, 199*, 59-65. https://doi.org/10.1016/j.solmat.2019.04.017

Huang, Q., Zhang, Q., Xiao, Y., He, Y., and Diao, X. (2018). Improved electrochromic performance of NiO-based thin films by lithium and tantalum co-doping. *Journal of Alloys and Compounds, 747*, 416-422. https://doi.org/10.1016/j.jallcom.2018.02.232

Inamdar, A. I., Chavan, H. S., Kim, H., and Im, H. (2019). Mesoporous Ni-PANI composite electrode for electrochromic energy storage applications. *Solar Energy Materials and Solar Cells, 201*, 110121. https://doi.org/10.1016/j.solmat.2019.110121

Jamdegni, M., and Kaur, A. (2020). Highly efficient dark to transparent electrochromic electrode with charge storing ability based on polyaniline and functionalized nickel oxide composite linked through a binding agent. *Electrochimica Acta*, *331*, 135359. https://doi.org/10.1016/j.electacta.2019.135359

Jang, H., and Lee, J. (2020). Iridium oxide fabrication and application: A review. *Journal of Energy Chemistry, 46*, 152-172. https://doi.org/10.1016/j.jechem.2019.10.026

Jeong, C. Y. (2020). Electrochromic properties of sputter-deposited rhodium oxide thin films of varying thickness. *Thin Solid Films, 709*, 138226. https://doi.org/10.1016/j.tsf.2020.138226

Jia, H. (2019). Dual-response and Li+-insertion induced phase transition of VO2-based smart windows for selective visible and near-infrared light transmittance modulation. *Solar Energy Materials and Solar Cells*, *200*, 110045. https://doi.org/10.1016/j.solmat.2019.110045

Jiang, F., Zheng, T., and Yang, Y. (2008). Preparation and electrochromic properties of tungsten oxide and iridium oxide porous films. *Journal of Non-Crystalline Solids, 354*(12-13), 1290-1293. https://doi.org/10.1016/j.jnoncrysol.2006.10.083

Jittiarporn, P., Badilescu, S., Al-Sawafta, M. N., Sikong, L., and Van-Truong, V. (2017). Electrochromic properties of sol-gel prepared hybrid transition metal oxides – A short review. *Journal of Science: Advanced Materials and Devices*, *2*(3), 286-300. https://doi.org/10.1016/j.jsamd.2017.08.005

Jung, H. K. (2020). Au-incorporated NiO nanocomposite thin films as electrochromic electrodes for supercapacitors. *Electrochimica Acta, 330*, 135203. https://doi.org/10.1016/j.electacta.2019.135203

Kalanur, S. S., Noh, Y. G., and Seo, H. (2020). Engineering band edge properties of WO3 with respect to photoelectrochemical water splitting potentials via a generalized doping protocol of first-row transition metal ions. *Applied Surface Science*, *509*, 145253. https://doi.org/10.1016/j.apsusc.2020.145253

Kalay, I., Yiğit, D., Güllü, M., Depci, T., Toppare, L., and Hacioglu, S. O. (2020). Enhancing electrochemical and electrochromic performances of carbazole comprising monomer via copolymerization with 3,4-ethylenedioxythiophene (EDOT). *Synthetic Metals, 267*, 116449. https://doi.org/10.1016/j.synthmet.2020.116449

Karaca, G. Y., Eren, E., Cogal, G. C., Uygun, E., Oksuz, L., and Uygun A. Oksuz (2019). Enhanced electrochromic characteristics induced by Au/PEDOT/Pt microtubes in WO3 based electrochromic devices. *Opical Materials,* 88, 472–478, https://doi.org/10.1016/j.optmat.2018.11.052

Kazazi, M. (2019). High-performance electrode based on electrochemical polymerization of polypyrrole film on electrophoretically deposited CNTs conductive framework for supercapacitors. *Solid State Ionics, 336*, 80-86. https://doi.org/10.1016/j.ssi.2019.03.021

Kim, H. N. (2016). Electrochromic mirror using viologen-anchored nanoparticles. *Materials Research Bulletin, 82*, 16-21. https://doi.org/10.1016/j.materresbull.2016.03.010

Kimura, R., Tsuboi, A., Nakamura, K., and Kobayashi, N. (2018). Effects of silver halide complexes on optical and electrochemical properties of silver deposition-based electrochromic device. *Solar Energy Materials and Solar Cells*, *177*, 128-133. https://doi.org/10.1016/j.solmat.2017.01.014

Koo, B. R., Jo, M. H., Kim, K. H., and Ahn, H. J. (2020). Multifunctional electrochromic energy storage devices by chemical cross-linking: Impact of a WO3.H2O nanoparticle-embedded chitosan thin film on amorphous WO3 films. *NPG Asia Materials, 12*(1), 10. https://doi.org/10.1038/s41427-019-0193-z

Kumar, N., Alam, F., Dwivedi, C., and Dutta, V. (2016). In-situ fabrication of metal-semiconductor (M-S) plasmonic thin films by a chemical spray pyrolysis technique: Optical properties. *Solar Energy Materials and Solar Cells*, *144*, 352-358. https://doi.org/10.1016/j.solmat.2015.09.030

Kumar, K. N., Shaik, H., Gupta, J., Sattar, S. A., Jafri, R. I., Pawar, A., Madhavi, H., Reddy G. V., A., and Nithya, G. (2022). Sputter deposited tungsten oxide thin films and nanopillars: electrochromic perspective. *Materials Chemistry and Physics*, *278*, 125706. https://doi.org/10.1016/j.matchemphys.2022.125706

Kumar, K. N., Nithya, G., Shaik, H., Hemanth, B., Chethana, M., Kishore, K., Madhavi, V., Imran Jafri, R., Sattar, S. A., Gupta, and Reddy, G. V. A. (2022a). Simulation and fabrication of tungsten oxide thin films for electrochromic applications. *Physica B: Condensed Matter*, *640*, 413932. https://doi.org/10.1016/j.physb.2022.413932

Kumar, K. N., Shaik, H., Pawar, A., Chandrashekar, L. N., Sattar, S. A., Nithya, G., Imran Jafri, R., Madhavi, V., Gupta, J., and Reddy, G. V. A. (2022b). Effect of annealing and oxygen partial pressure on the RF sputtered WO3 thin films for electrochromic applications. *Materials Today: Proceedings*, *59*, 339-344. https://doi.org/10.1016/j.matpr.2021.11.185

Kumar, K. N., Shaik, H., Chandrashekar, L. N., Aishwarya, P., Sattar, S. A., Nithya, G., Madhavi, V., Imran Jafri, R., Gupta, J., and Reddy, G. V. A. (2022c). On ion transport during the electrochemical reaction on plane and GLAD deposited WO3 thin films. *Materials Today: Proceedings*, *59*, 275-282. https://doi.org/10.1016/j.matpr.2021.11.113

Legis-Comex, Sistema de Inteligencia Comercial, Inteligencia de Mercados (2017). *Informe sectorial de vidrio en Colombia 2017*. https://www.legiscomex.com/BancoMedios/Documentos%20PDF/informe-sectorial-sector-vidrio-colombia-2017-completo-rci318.pdf

Li W. (2020). Preparation and performance of fast-response ITO/Li-NiO/Li-WO_3/ITO all-solid-state electrochromic devices by evaporation method. *Materials Letters, 265*, 127464. https://doi.org/10.1016/j.matlet.2020.127464

Li, B., Liu, J., Tian, S., Liu, B., Yang, X., Yu, Z., and Zhao, X. (2019). VO_2-ZnO composite films with enhanced thermochromic properties for smart windows. *Ceramics International, 46*(13), 2758-2763. https://doi.org/10.1016/j.ceramint.2019.09.264

Li, S. (2020). Grafting polymer from oxygen-vacancy-rich nanoparticles to enable protective layers for stable lithium metal anode. *Nano Energy, 76*, 105046. https://doi.org/10.1016/j.nanoen.2020.105046

Lin K. (2020). Star-shaped trithiophene and hexathiophene functionalized truxenes: Synthesis, electropolymerization, and electrochromism. *Reactive & Functional Polymers, 154*, 104. https://doi.org/10.1016/j.reactfunctpolym.2020.104674

Liu, Z. (2019). Three-dimensional ordered porous electrode materials for electrochemical energy storage. *NPG Asia Materials, 11*, 12. https://doi.org/10.1038/s41427-019-0112-3

Liu, Z., Shang, S., Chiu, K.-I., Jiang, S., and Dai, F. (2019). Fabrication of conductive and flame-retardant bifunctional cotton fabric by polymerizing pyrrole and doping phytic acid. *Polymer Degradation and Stability, 167*, 277-282. https://doi.org/10.1016/j.polymdegradstab.2019.06.029

Marchisio, S., Lerro, F., and Von Pamel, O. (2015). Empleo de un laboratorio remoto para promover aprendizajes significativos en la enseñanza de los dispositivos electrónicos. *Píxel-Bit. Revista de Medios y Educación, 38*, 129-139. http://www.redalyc.org/pdf/368/36816200010.pdf

Matysiak, W., Tański, T., Smok, W., Gołombek, K., and Schab-Balcerzak, E. (2020). Effect of conductive polymers on the optical properties of electrospun polyacrylonitryle nanofibers filled by polypyrrole, polythiophene and polyaniline. *Applied Surface Science, 509*, 145068. https://doi.org/10.1016/j.apsusc.2019.145068

Meenakshi, M., Sivakumar, R., Perumal, P., and Sanjeeviraja, C. (2016). Studies on electrochromic properties of RF sputtered vanadium oxide: Tungsten oxide thin films. *Materials Today: Proceedings,* 3, S30-S39. https://doi.org/10.1016/j.matpr.2016.01.005

Mendieta-Reyes, N. E., Cheuquepán, W., Rodes, A., and Gómez, R. (2020). Spectroelectrochemical study of CO_2 reduction on TiO_2 electrodes in acetonitrile. *ACS Catalysis, 10*(1), 103-113. https://doi.org/10.1021/acscatal.9b02932

Ministerio Colombiano de Ciencia, Tecnología e innovación (MinCiencias) and Vicepresidencia de la República de Colombia (2020). *Misión Internacional de sabios: Colombia y la nueva revolución industrial*. https://doi.org/10.17230/9789585135116vdyc

Ministerio Colombiano de Ciencia, Tecnología e innovación (MinCiencias) (2022). *Base de datos, grupos de investigación*. https://minciencias.gov.co/sites/default/files/convocatoria/listado_resultados_preliminares_-_convocatoria_894_de_2021_-_grupos.pdf

Ministerio Colombiano de Comercio (MinCIT) (2019a). *Plan de Negocios Sector de Plásticos a 2032*. https://www.colombiaproductiva.com/ptp-capacita/publicaciones/sectoriales/publicaciones-plasticos-y-pinturas/plan-de-negocio-industria-de-plasticos-2019-2032

Ministerio Colombiano de Comercio (MinCIT). *Plan de Negocios Sector de Químicos Visión a 2032*. https://www.colombiaproductiva.com/ptp-capacita/publicaciones/sectoriales/publicaciones-quimica-basica/plan-de-negocio-industria-quimica-basica-2019-2032

Ministerio Colombiano de Minas y Energía (MinMinas) (2018). *Níquel caracterización y análisis de mercado internacional de minerales en el corto, mediano, y largo plazo con vigencia al año 2035*. https://www1.upme.gov.co/simco/Cifras-Sectoriales/Datos/mercado inter/Producto2_Niquel_FINAL_12DIC2018.pdf#search=niquel

Mishra, S. (2017). Live spectroscopy to observe electrochromism in viologen based solid state device. *Solid State Communications, 261*, 17-20, https://doi.org/10.1016/j.ssc.2017.05.020

Monk, P. M. S., Rosseinsky, D. R., and Mortimer, R. J. (2015). Electrochromic materials and devices based on viologens. *Electrochromic Materials and Devices, 77*, 57-90. https://doi.org/10.1002/9783527679850.ch3

Morales-Torres, J.A., Olaya-Florez, J.-J., and Rojas, H. F. (2009). Evaluación de la capacidad protectora de recubrimientos Ni-SiC y Ni-Co-W depositados por proyección térmica. *Dyna, 76*(160) 195-206. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0012-73532009000400019

Murphy, N. R., Moreno-Tarango, A. J., Ramana, C. V., Sun, L., Jones, J. G., and Grant, J. T. (2016). Hybrid co-deposition of molybdenum doped niobium pentoxide (NbxMoyOz) thin films. *Journal of Alloys and Compounds, 681*, 350-358. https://doi.org/10.1016/j.jallcom.2016.04.233

OECD (2019) *Economic Surveys, 48*. http://www.oecd.org/economy/surveys/Overview_Colombia_ESP.pdf

Paipitak, K., Kahattha, C., Techitdheera, W., Porntheeraphat, S., and Pecharapa, W. (2011). Characterization of Sol-gel derived Ti-doped tungsten oxide electrochromic thin films. *Energy Procedia, 9*, 446-451. https://doi.org/10.1016/j.egypro.2011.09.050

Pan, J. (2020). A high-performance electrochromic device assembled with hexagonal WO3 and NiO/PB composite nanosheet electrodes towards energy storage smart window. *Solar Energy Materials and Solar Cells, 207*, 110337. https://doi.org/10.1016/j.solmat.2019.110337

Pandey, S. (2016). Highly sensitive and selective chemiresistor gas/vapor sensors based on polyaniline nanocomposite: A comprehensive review" *Journal of Science: Advanced Materials and Devices, 1*(4), 431-453. https://doi.org/10.1016/j.jsamd.2016.10.005

Pham, N. S., Seo, Y. H., Park, E., Nguyen, T. D. D., and Shin, I. S. (2020). Implementation of high-performance electrochromic device based on all-solution-fabricated Prussian blue and tungsten trioxide thin film. *Electrochimica Acta, 353*, 136446. https://doi.org/10.1016/j.electacta.2020.136446

Prieto R. G., Guatame, C. L., and Cárdenas, S. C. (Comps.) (2019). *Recursos minerales de Colombia*. Servicio Geológico Colombiano. https://www2.sgc.gov.co/Publicaciones/Cientificas/NoSeriadas/Documents/recursos-minerales-de--colombia-vol-2.pdf

Puguan, J. M. C., and Kim, H. (2019). Ionene copolymer electrolyte obtained from cyclo-addition of di-alkyne and di-azide monomers for solid-state smart glass windows. *Journal of Industrial and Engineering Chemistry, 74*, 1-6. https://doi.org/10.1016/j.jiec.2019.03.006

Ramadan, R., Elshorbagy, M. H., Kamal, H., Hashem, H. M., and Abdelhady, K. (2017). Preparation and characterization of protonic solid electrolyte applied to a smart window device with high optical modulation. *Optik, 135*, 85-97. https://doi.org/10.1016/j.ijleo.2017.01.056

Rathika, R. (2020). Tailoring the properties of spray deposited V2O5 thin films using swift heavy ion beam irradiation. *Nuclear Engineering and Technology, 52*(11), 2585-2593. https://doi.org/10.1016/j.net.2020.04.013

Richardson, T. J., Slack, J. L., and Rubin, M. D. (2018). Electrochromism in copper oxide. *Thin Films, 46*, 13-14. https://doi.org/10.1016/S0013-4686(01)00397-8

Sahu, D. R., Wu, T. J., Wang, S. C., and Huang, J. L. (2017). Electrochromic behavior of NiO film prepared by e-beam evaporation. *Journal of Science: Advanced Materials and Devices, 2*, 225-232. https://doi.org/10.1016/j.jsamd.2017.05.001

Shi, Y., Chen, Q., Zheng, J., and Xu, C. (2020a). Electrochromism of substituted phthalate derivatives and outstanding performance of corresponding multicolor electrochromic devices. *Electrochimica Acta*, *341*, 136023. https://doi.org/10.1016/j.electacta.2020.136023

Shi, Y., Wang, G., Chen, Q., Zheng, J., and Xu, C. (2020b). Electrochromism and electrochromic devices of new extended viologen derivatives with various substituent benzene. *Solar Energy Materials and Solar Cells, 208*, 110413. https://doi.org/10.1016/j.solmat.2020.110413

Sistema Nacional de Información para la Educación Superior en Colombia (SNIES) (2021). Base de datos programas académicos en Colombia. https://hecaa.mineducacion.gov.co/consultaspublicas/programas.

Sorar., I., Pehlivan, B., Granqvist, C. G., and Niklasson, G. A. (2019). Electrochromism of W–In oxide thin films: Implications for cycling durability. *Thin Solid Films, 697*(1) 137830. https://doi.org/10.1016/j.tsf.2020.137830

Sun H. (2015). Smart responsive phosphorescent materials for data recording and security protection. *Nature Communications, 5*, 2-10. https://doi.org/10.1038/ncomms4601

Sun, S. (2020). Flexible and rechargeable electrochromic aluminium-ion battery based on tungsten oxide film electrode. *Solar Energy Materials and Solar Cells, 207*, 110. https://doi.org/10.1016/j.solmat.2019.110332

Szkoda, M., Trzciński, K., Nowak, A. P., Gazda, M., Sawczak, M., and Lisowska-Oleksiak, A. (2020). The effect of morphology and crystalline structure of Mo/MoO$_3$ layers on photocatalytic degradation of water organic pollutants. *Materials Chemistry and Physics, 248*, 122908. https://doi.org/10.1016/j.matchemphys.2020.122908

Tang, C. J. (2019). An all-solid-state electrochromic device based on WO3-Nb2O5 composite films prepared by fast-alternating bipolar-pulsed reactive magnetron sputtering. *Coatings, 9*(1), 9. https://doi.org/10.3390/coatings9010009

Tang, X., Chen, G., Liao, H., Li, Z., Zhang, J., and Luo, J. (2020). Unveiling mechanical degradation for a monolithic electrochromic device: Glass/ITO/WO3/LiClO4 (PEO)/TiO2/ITO/glass. *Electrochimica. Acta, 329*, 135182. https://doi.org/10.1016/j.electacta.2019.135182

Thummavichai, K., Xia, Y., and Zhu, Y. (2017). Recent progress in chromogenic research of tungsten oxides towards energy-related applications. *Progress Materials Science, 88*, 281-324. https://doi.org/10.1016/j.pmatsci.2017.04.003

Tsai, H., Ceretti, E., Rizzi, D., Ginestra, P., Kao, T., and Leu, M. C. (2021). Laser induced metallization on flexible polymer coating: Analysis and application. *Progress in Materials Science, 290*, 116986. https://doi.org/10.1016/j.jmatprotec.2020.116986

Tsige, A., Ganesh, T., Mensur, D., and Tesfaye, D. (2020). Thermal studies on chemical bath deposited thermochromic VO$_2$ thin film for energy efficient glass windows. *Materials Today: Proceedings,* 45(Part 7), 6171-6175. https://doi.org/10.1016/j.matpr.2020.10.480

Ullah, W., Herzog, G., Vilà, N., and Walcarius, A. (2021). Electrografting and electropolymerization of nanoarrays of PANI filaments through silica mesochannels. *Electrochemistry Communications, 122*, 106896. https://doi.org/10.1016/j.elecom.2020.106896

Varghese-Hansen, R., Yang, J., and Zheng., L. (2018). Flexible electrochromic materials based on CNT/PDA hybrids. *Advances in Colloid and Interface Science, 258*, 21-35. https://doi.org/10.1016/j.cis.2018.07.003

Vicepresidencia de la República de Colombia and Ministerio de Ciencia, Tecnología e Innovación (MINCIENCIAS) (2020). *Misión Internacional de Sabios: Colombia hacia una sociedad del conocimiento* https://minciencias.gov.co/sites/default/files/upload/paginas/ebook_colombia_hacia_una_sociedad_del_conocimiento.pdf

Wang, L. (2020). Polymer of intrinsic microporosity (PIM) films and membranes in electrochemical energy storage and conversion: A mini-review. *Electrochemistry Communications, 118*, 106798. https://doi.org/10.1016/j.elecom.2020.106798

Wang, W. (2020). Controllable vapor phase polymerization of PEDOT films using imidazole as an inhibitor and their electrical and electrochromic properties. *Synthetic Metals, 269*, https://doi.org/10.1016/j.synthmet.2020.116523.

Wang, Y. C., Lu, H. C., Hsiao, L. Y., Lu, Y. A, and Ho, K. C. (2019). A complementary electrochromic device composed of nanoparticulated ruthenium purple and Fe(II)-based metallo-supramolecular polymer. *Solar Energy Materials and Solar Cells, 200*, 10. https://doi.org/10.1016/j.solmat.2019.109929

Wang, Y., Lei, Q., Dong, W., Mo, X., and Li, H. (2020). Photoelectric effect driving PANI/PB multicolor visualized detection of CEA based on Ag2S NPs@ZnO NTs. *Analytica Chimica Acta, 1108*, 61-69. https://doi.org/10.1016/j.aca.2020.02.053

Wang, Z., Wang, X., Cong, S., Geng, F., and Zhao, Z. (2020). Fusing electrochromic technology with other advanced technologies: A new roadmap for future development. *Materials Science & Engineering R: Reports, 140*, 54. https://doi.org/10.1016/j.mser.2019.100524

Wen-Cheun Au, B., Chan, K. Y., and Knipp, D. (2019). Effect of film thickness on electrochromic performance of sol-gel deposited tungsten oxide (WO$_3$). *Optical Materials, 94*, 387-392. https://doi.org/10.1016/j.optmat.2019.05.051

Wu, H. (2019). Electrically responsive structural colors from colloidal crystal arrays of PS@PANI core–shell nanoparticles. *Colloids and Surfaces A: Physicochemical and Engineering Aspects, 577*, 75-83. https://doi.org/10.1016/j.colsurfa.2019.05.049

Wu, L., Ge, Y., Zhang, L., Yu, D., Wu, M., and Ni, H. (2018). Enhanced electrical conductivity and competent mechanical properties of polyaniline/polyacrylate (PANI/PA) composites for antistatic finishing prepared at the aid of polymeric stabilizer. *Progress in Organic Coatings, 125*, 99-108. https://doi.org/10.1016/j.porgcoat.2018.09.002

Xu, D. (2020). Effect of counter anion on the uniformity, morphology and electrochromic properties of electrodeposited poly(3,4-ethylenedioxythiophene) film. *Journal of Electroanalytical Chemistry, 861*, 113833. https://doi.org/10.1016/j.jelechem.2020.113833

Xu, K. (2020). Integrated electrochromic supercapacitors with visual energy levels boosted by coating onto carbon nanotube conductive networks. *Solar Energy Materials and Solar Cells, 206*, 110330. https://doi.org/10.1016/j.solmat.2019.110330

Xu, Z. (2020). Color tuning for black-to-transmissive conjugated copolymer with excellent electrochromic properties via electrochemical copolymerization of two donor-acceptor type monomers. *Materials and Design, 194*, 108903. https://doi.org/10.1016/j.matdes.2020.108903

Yao, Y., Zhao, Q., Wei, W., Chen, Z., Zhu, Y., Zhang, P., and Gao, Y. (2020). WO_3 quantum-dots electrochromism. *Nano Energy*, *68*, 104350. https://doi.org/10.1016/j.nanoen.2019.104350

Yang, P., Sun, P., and Mai, W. (2016). Electrochromic energy storage devices. *Materials Today, 19*, 394-401. https://doi.org/10.1016/j.mattod.2015.11.007

Yin, J., Liu, Q., Zhou, J., Zhang, L., Zhang, Q., Rao, R., Liu, S., and Jiao, T. (2020). Self-assembled functional components-doped conductive polypyrrole composite hydrogels with enhanced electrochemical performances. *RSC Advances, 10*(18), 10546-10551. https://doi.org/10.1039/d0ra00102c

Yu, H., Guo, J., Wang, C., Zhang, J., Liu, J., Dong, G., Zhong, X., and Diao, X. (2020). Essential role of oxygen vacancy in electrochromic performance and stability for WO_3-y films induced by atmosphere annealing. *Electrochimica Acta, 332*, 135504. https://doi.org/10.1016/j.electacta.2019.135504.

Yun, T. Y., Li, X., Bae, J., Kim, S. H., and Moon, H. C. (2019). Non-volatile, Li-doped ion gel electrolytes for flexible WO_3-based electrochromic devices. *Materials and Design*, *162*, 45-51. https://doi.org/10.1016/j.matdes.2018.11.016

Zeng, J., Yang, H., Zhong, C., Rajan, K., Rehman Sagar, R., Qi, X., Deng, Y., Jiang, H., Liu, P., Liang, T. (2021). Colorless-to-black electrochromic devices based on ambipolar electrochromic system consisting of cross-linked poly(4-vinyltriphenylamine) and tungsten trioxide with high optical contrast in visible and near-infrared regions. *Chemical Engineering Journal, 404*, 126402. https://doi.org/10.1016/j.cej.2020.126402

Zhang, B., Xu, C., Xu, G., Tan, S., and Zhang, J. (2019). Amorphous titanium dioxide film with improved electrochromism in near-infrared region. *Optical Materials, 89*, 191-196. https://doi.org/10.1016/j.optmat.2019.01.034

Zhang, F., Feng, Y., and Feng, W. (2020). Three-dimensional interconnected networks for thermally conductive polymer composites: Design, preparation, properties, and mechanisms. *Materials Science and Engineering: R: Reports, 142*, 100580. https://doi.org/10.1016/j.mser.2020.100580

Zhang, H., Xu, H., Endres, F., and Li, Y. (2020). Multi-color poly(3-methylthiophene) films prepared by a novel pre-nucleation electrodeposition grown method for enhancing electrochromic stability. *Electrochimica Acta, 362*, 137103. https://doi.org/10.1016/j.electacta.2020.137103

Zhang, Y., Shi, X., Xiao, S., and Xiao, D. (2021). Visible and infrared electrochromism of bis(2-(2-(2-hydroxyethoxy)ethoxy)ethyl) viologen with sodium carboxymethyl chitosan-based hydrogel electrolytes. *Dyes and Pigments, 185*(PA), 108893. https://doi.org/10.1016/j.dyepig.2020.108893

Zheng, M., Xiao, X., Li, L., Gu, P., Dai, X., Tang, H., Hu, Q., Xue, H., and Pang, H. (2018). Hierarchically nanostructured transition metal oxides for supercapacitors. *Science China Materials, 61*(2), 185-209. https://doi.org/10.1007/s40843-017-9095-4

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e106558)

Research Article/Systems and Computer Engineering http://doi.org/10.15446/ing.investig.106558

Development and Validation of a Dry Electrode Array for sEMG Recording and Hand Movement Recognition

Desarrollo y validación de un arreglo de electrodos secos para la adquisición de señales sEMG y el reconocimiento de los movimientos de la mano

Cinthya L. Toledo-Peral[1], Ana I. Martín-Vignon-Whaley[2], Jorge A. Mercado-Gutiérrez[3],
Arturo Vera-Hernández[4], Lorenzo Leija-Salas[5], and Josefina Gutiérrez-Martínez[6]

ABSTRACT

Surface electromyography (sEMG) signals are an indirect measurement of muscle activity, and their applications range from biomechanics to control and rehabilitation. Hand movement recognition is a very difficult endeavor due to forearm anatomy. Hence, a multichannel approach for signal acquisition and processing is required. Conventional electrodes can limit the ease-of-use and repeatability of multi-channel sEMG recordings. New techniques have been proposed in this regard, with dry electrodes being one of them. Dry electrode technology has enabled the design of better donning and doffing procedures for multichannel sEMG recording, particularly for rehabilitation and prosthetic applications. However, there is a debate about the quality of the signals recorded with them and their usefulness for the recognition of multiple hand movements. To mitigate these quality issues, this work proposes an array of reusable stainless steel dry electrodes for multichannel sEMG recording with a design that facilitates its positioning on the forearm. The dry electrodes were characterized through electrical impedance measures and a Bland-Altman test. They were found to have similar characteristics to standard, disposable sEMG pre-gelled electrodes. For placement repeatability and application feasibility, an anatomy-based electrode positioning protocol was implemented with 17 healthy subjects and six hand movements. To evaluate the application feasibility of the electrode array, a feed-forward artificial neural network was trained to classify signals from the six movements, with a 97,86±0,58% accuracy. The amplitude of the sEMG signals for two antagonist movements was compared, finding a 24,81% variation. The dry electrode array showed feasibility in acquiring and classifying sEMG signals of hand movements with high accuracy.

Keywords: dry electrodes, Bland-Altman, anatomical positioning array, hand movement classification

RESUMEN

Las señales de electromiografía de superficie (sEMG) son una medida indirecta de la actividad muscular, y sus aplicaciones van desde biomecánica hasta control y rehabilitación. La identificación de movimientos de la mano es una tarea muy complicada debido a la anatomía del antebrazo, por lo que se requiere un enfoque multicanal para adquisición y procesamiento de señales. Los electrodos convencionales pueden limitar la facilidad de uso y la repetibilidad de los registros sEMG multicanal. Se han propuesto nuevas técnicas para ello, entre ellas los electrodos secos. La tecnología de electrodos secos ha permitido el diseño de mejores procedimientos de colocación y remoción para registro sEMG multicanal, particularmente en aplicaciones de rehabilitación y prótesis. Sin embargo, existe un debate sobre la calidad de las señales registradas con ellos y su utilidad para el reconocimiento de múltiples movimientos de la mano. Para mitigar estos problemas de calidad, se propone un arreglo de electrodos secos reutilizables de acero inoxidable para registro sEMG multicanal con un diseño que facilita su posicionamiento en el antebrazo. Estos electrodos se caracterizaron mediante mediciones de impedancia eléctrica y una prueba Bland-Altman. Se encontró que tienen características similares a los electrodos pregelados desechables estándar para sEMG. Para la repetibilidad de la colocación y su viabilidad de aplicación, se implementó un protocolo de colocación de electrodos basado en la anatomía con 17 sujetos sanos y seis movimientos

---

[1] Bachelor's in Electronics, Benemérita Universidad Autónoma de Puebla, Mexico. MSc in Electrical Engineer in Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. Affiliation: PhD Student in Electrical Engineer – Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. Email: phd.toledo@outlook.com

[2] Bachelor's in biomedical engineering, Universidad La Salle, Mexico. E-mail: anaimvw@gmail.com

[3] Bachelor in Electronics, Benemérita Universidad Autónoma de Puebla, Mexico. MSc in Biomedical Engineering, Universidad Autónoma Metropolitana, Mexico. PhD in Electrical Engineering, Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. Affiliation: Medical science researcher, Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, Mexico. E-mail: jorgea.mercadog@gmail.com

[4] Biomedical engineer, Instituto Politécnico Nacional, Mexico. MSc in Electrical Engineering, Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. PhD in Engineering, Institut Polytechinique de Lorraine, France. Affiliation: Researcher 3C, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. E-mail: arvera@cinvestav.mx

[5] Industrial engineer, Instituto Tecnológico de San Luis Potosí, Mexico. MSc in Electrical Engineering, Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. PhD in Engineering, Université de Nancy, France. Affiliation: Researcher 3D, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. E-mail: lleija@cinvestav.mx

[6] Bachelor's in biomedical engineering, Universidad Iberoamericana, Mexico. MSc in Electrical Engineering, Bioelectronics, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Mexico. PhD in Engineering, Universidad Nacional Autónoma de México. Affiliation: Head of the medical engineering research division, Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, Mexico. E-mail: josefina_gutierrez@hotmail.com

de la mano. Finalmente se entrenó una red neuronal artificial prealimentada para clasificar señales de los seis movimientos, con una precisión del 97,86±0,58 %. Se comparó la amplitud de las señales sEMG para dos movimientos antagonistas, encontrando una variación del 24,81 %. El arreglo de electrodos secos mostró viabilidad para adquirir y clasificar registros sEMG de los movimientos de la mano con gran precisión.

Palabras clave: electrodos secos, Bland-Altman, arreglo de posicionamiento anatómico, clasificación de movimientos de la mano

Received: December 28th, 2022
Accepted: July 18th, 2023

Introduction

Hand movements are complex to perform, acquire, and emulate. Many structures are involved in hand movements, especially forearm muscles. In order to analyze them, they can be divided into three regions: anterior, posterior, and external. The anterior region has several muscles divided into four planes that include the round pronator and anterior ulnar, the superficial common long flexor of the fingers, the flexor of the thumb and deep common flexor of the fingers, and the pronator square. The external region is composed of the long supinator, the first external radial, the second external radial, and the short supinator. Finally, the posterior region holds the flexor muscles, among others (Hall, 2011). Around half of these muscles are deep enough to be challenging to access from the surface of the skin (Mitchell and Whited, 2019), and most of them are either layered or braided with adjacent muscles. Moreover, from a bio-signal acquisition perspective, it is difficult to place enough electrodes in a such small area to match muscle anatomy and acquire signals, thus allowing to distinguish between hand movements.

Muscle fibers contract while being controlled by electrical signals that are transmitted through the nervous system and generated by units called *motor neurons*. These electrical signals are recorded on the surface of the skin, which is known as *surface electromyography* (sEMG). To record and interpret these signals, bipolar electrodes are placed at the belly of the muscle where the transversal area of the motor units is larger. This position is optimal for obtaining the highest amplitude signal, which also contains more information on muscle contraction. According to the recommendations provided by *The Surface Electromyography for the Non-Invasive Assessment of Muscles* (SENIAM) and the *Consensus for Experimental Design in Electromyography* (CEDE) projects, passive bipolar electrodes should have a separation of approximately 2,5 cm, the electrodes should be aligned relative to the fascicle direction, a reference should be placed at a dielectric site, and there has to be a contact area on the skin of at least 10 mm in diameter (Besomi *et al.*, 2019; Biga *et al.*, n.d.; Hermens *et al.*, 2000; Merletti and Farina, 2016; "Recommendations for the Practice of Clinical Neurophysiology: Guidelines of the International Federation of Clinical Neurophysiology", 1999).

In most cases, an sEMG signal from one muscle does not yield enough information, so the use of several acquisition channels has become common practice. Nevertheless, the number of channels needed for different applications varies depending on the aim and purpose (Kapelner *et al.*, 2016; Liu, 2015; Peng *et al.*, 2016).

For some applications (*e.g.*, a therapy session), placing more than two or three pairs of electrodes can be time-consuming. Moreover, placing too many electrodes in a limited time span could rush the process and lead to mistakes and variability. The use of disposable electrodes also represents higher costs and waste.

Commercially available acquisition systems come with pre-gelled electrodes that are ready to use, but most applications require placement in muscle targets, allowing for only minimal variations (Al-Ayyad *et al.*, 2023). When going through the literature, one may notice that the clinical applications of wearable devices tend to be unified to match conventional therapy, which is the gold standard. These systems may benefit from a ready-to-wear electrode array, useful for repetitive measurements.

Another example that requires the placement of many electrodes corresponds to prosthetic control applications, such as upper limb prostheses based on pattern recognition techniques, which may include up to eight pairs of electrodes to focus on biceps and forearm muscles. Most of these applications do not use arrays but personalized positioning (Liu, 2015; Phinyomark *et al.*, 2014). The Myo armband is an acquisition system embedded in an eight-channel geometrical dry electrode array. All channels are equidistantly placed around the forearm in the proximal region. This armband also has a three-axes gyroscope, a three-axes accelerometer, a three-axes magnetometer, and a Bluetooth module (Myo Armband, Thalmic Labs Inc.) (Visconti *et al.*, 2018). Other systems designed for upper limb sEMG signal acquisition that use dry electrodes include those from Biometrics Ltd. (Newport, UK) and the silver electrodes of Trigno Avanti (Delsys Inc., Massachusetts, USA) (Al-Ayyad *et al.*, 2023).

The electrode arrays found in the literature have a geometrical matrix design, and the signals acquired can be treated as one-dimensional vectors, 2D images, or 3D heat maps (Barbero *et al.*, 2012b; Dai *et al.*, 2019; Guzmán *et al.*, 2011; Moin *et al.*, 2018; Peng *et al.*, 2016; Reategui and Callupe, 2017; Rojas-Martínez *et al.*, 2013; Rui Ma *et al.*, 2010; Ruvalcaba *et al.*, 2017; Topalović *et al.*, 2019). These types of arrays are advantageous when one desires to acquire a larger number of channels, and they are mainly used to evaluate areas of muscle contribution during upper limb isometric contractions. In some cases, subjects arms are constrained to avoid undesired movements (Dai *et al.*, 2019; Moin *et al.*, 2018; Peng *et al.*, 2016; Rojas-Martínez *et al.*, 2013; Topalović *et al.*, 2019; Urbanek and van der Smagt, 2016). Furthermore, geometrical matrices do not require previous knowledge of the site's anatomy or customization for placement on each patient. Still, they are not recommended for evaluating movements (isotonic contraction) where the transient part of the signal is relevant; the subject must be able to move as normally as possible. Almost all arrays found in the literature say they use dry electrodes, but several have been found to need conductive gel to properly function (Dai *et al*., 2019; Guzmán *et al*., 2011; Peng *et al*., 2016; Rodrigues *et al*., 2020; Rojas-Martínez *et al*., 2013; Topalović *et al*., 2019; Urbanek and van der Smagt, 2016).

The only type of pre-positioned dry electrodes for sEMG signal acquisition is used in myoelectric prosthesis, where movements are activated through direct control, generally using two acquisition channels placed in antagonist muscles. In addition, the electrodes are normally integrated into the socket of the prosthesis (three electrodes: two larger bipolar ones and a small one for the reference in the middle), so they are used in large muscles, *e.g.*, the biceps and triceps, where electrodes can withstand a 2 cm location shift without considerable differences in the outcome signal (Losier *et al*., 2011; Young *et al*., 2011).

Dry electrodes play an essential role in the repeated usability of positioning electrode arrays. As a main advantage, dry electrodes do not need conductive electrolyte gel to function correctly; the user just needs to wait for a period of 10 to 15 min before recording a signal (Meziane *et al*., 2013), which falls within the preparation time for an sEMG recording session. Another benefit is that they do not need intensive skin cleaning, are reusable (thus reducing medical waste), and have long-term use applications, unlike gelled electrodes that dry out over time (Srinivasa and Pandian, 2017).

Dry electrodes are built using different materials, such as stainless steel, conductive textiles, and flexible polymers. They are used in research and are commercially available (Xie *et al*., 2013; Niu *et al*., 2021). The SENIAM and CEDE projects recommend metal materials such as stainless steel, platinum, and Ag/AgCl for building dry electrodes (Besomi *et al*., 2019; Hermens *et al*., 2000).

Moreover, there is a small recurrent set of materials used to fabricate dry electrodes and electrode arrays. Textile electrodes are widely used, but they are not suitable for dry applications in small areas (Ruvalcaba *et al*., 2017; Vojtech *et al*., 2013) because they have a higher sensibility to motion and pressure, which can lead to higher skin-contact impedance and noisy signals (Chi *et al*., 2010; Xie *et al*., 2013; Meziane *et al*., 2013; Puurtinen *et al*., 2006; Srinivasa and Pandian, 2017). In Kusche *et al*. (2019), five materials (gold, stainless-steel, rubber with/without texture, and textile) are compared when used for dry electrode fabrication and against Ag/AgCl gelled electrodes. Regarding frequency response, dry electrodes show lower impedance values. After 15 min of contact with the skin, the impedance of stainless-steel electrodes stabilizes and maintains its value for up to 30 min. Of the six materials, stainless-steel has a higher impedance than gold and Ag/AgCl, albeit lower than that of the other materials.

Some metallic options have been tested for sEMG signal acquisition, such as brass dry electrodes (Ruvalcaba *et al*., 2016) under 10 mm in diameter which yield impedance values like those of pre-gelled commercial electrodes (Ghoshdastider *et al*., 2012; Vojtech *et al*., 2013). Nevertheless, stainless-steel electrodes have two main advantages: they have a simple manufacturing process, and they can be reused because of their corrosion resistance (Albulbul, 2016; Gan *et al*., 2019). Therefore, this material is a good option for manufacturing small surface electrodes, wherein contact issues will not be a concern (Guo *et al*., 2020).

Stainless steel electrodes perform as well as commercial dry electrodes, and their manufacturing process is simple ("Recommendations for the Practice of Clinical Neurophysiology: Guidelines of the International Federation of Clinical Neurophysiology", 1999). Stainless steel electrodes are antibacterial, do not rust, and can be used for long-term measures. Furthermore, stainless steel is among the metals less likely to cause an allergic skin reaction, and they are recommended by the SENIAM (Hermens *et al*., 2000).

Even though there are different proposals regarding materials, configurations, and embedded acquisition systems, to the best of our knowledge, and after a thorough literature analysis, there is no evidence of anatomical positioning arrays using dry electrodes for sEMG signal acquisition (Palumbo *et al*., 2021).

To be able to simultaneously acquire sEMG signals from several muscles and easily find the best location for multiple electrodes within the forearm, in order to record the patterns of hand movements, this work proposes an array of 17 dry electrodes positioned on anatomical sites at the forearm, which are fixed using a textile sleeve with a design based on anthropometric measures. This design allows recording eight sEMG channels from the forearm muscles at their correct anatomical sites. We hypothesize that assessing the functionality of the forearm dry electrode anatomical array for the acquisition of sEMG signals regarding six movements (rest, hand opening, power grasp, fine pinch, supination, and pronation) will result in the acquisition of usable signals for pattern recognition.

Methodology

A total of 22 abled-body volunteers participated in two types of evaluations. 17 subjects participated in anthropometrical measurements, and five subjects participated in signal acquisition for the digital processing of the data. All subjects signed an informed consent. The procedures carried out in this work were aimed at performing a proof-of-concept to evaluate the functionality of an array of dry electrodes with anatomy-based positioning. All procedures were approved by the Institutional Research and Ethics Boards, as part of protocol No. 38/16.

To design a dry electrode array for sEMG signal acquisition based on anatomical positioning, the first step was to locate the muscles of interest in the forearm, which is a small area with multiple muscles, most of them narrow and short. For each hand movement, there is a set of known muscles involved (Barbero *et al*., 2012a; Hall, 2011). These can be seen in Table 1.

Table 1. Muscles involved in the target hand movements

Movement	Muscles Involved
Open Hand	Finger extensors, fifth finger extensor, index extensor, long extensor of the thumb, short extensor of the thumb, long abductor of the thumb
Power Grasp	Long palmar, superficial flexor of the fingers, deep flexor of the fingers of the hand, long flexor of the thumb, brachioradialis
Fine Pinch	Long palmar, superficial flexor of the fingers, long flexor of the thumb, brachioradialis
Supination	Long supinator, short supinator
Pronation	Pronator square, pronator teres

Source: Barbero *et al*. (2012a); Hall (2011b)

Regarding the first evaluation, 17 volunteer abled-body subjects participated in the protocol for the anatomical positioning of electrodes (eight females 22,8 ± 3,14 years of age, with a height of 1,72 ± 16,2 m and a weight of 72,12 ± 16,20 kg; and nine males 23 ± 3,1 years of age, with a height of 1,73 ± 0,12 m and a weight of 71,75 ± 16,10 kg). To trace the best position at the forearm for sEMG signal acquisition, eight muscle sites were located by identifying the insertions of the tendons and the ventral part of the muscles of interest. Then, the subjects were asked to perform five repetitions of six hand movements, *i.e.*, hand opening, power grasp, fine pinch, supination, pronation, and rest, while the muscles were palpated.

Afterwards, the position for the bipolar acquisition channel was marked, centered on top of the ventral area of the muscles of interest, with a separation of 2,5 cm. This process started at the brachioradialis muscle and was repeated towards the lateral region of the forearm (externally). A personalized map of the anatomical positioning of sEMG electrodes was obtained for each subject, which was then transferred to an acetate sheet to create *individual templates*. The initial marks for the electrodes were positioned in the brachioradialis, so this muscle was used as a common point to overlap the *individual templates*. The resulting positions were used to propose the *general template* of the anatomical positioning array (APA). These steps are exemplified in Figure 1.

The design of a set of dry electrodes was proposed and carried out. To place the dry electrodes, a textile sleeve design was developed, which was based on the anthropometric measures of the participants, as follows:

- Wrist circumference
- Length between the elbow and the wrist
- Half of the elbow-wrist length
- Half of the half of the elbow-wrist length
- Forearm circumference at the half of the half of the elbow-wrist length

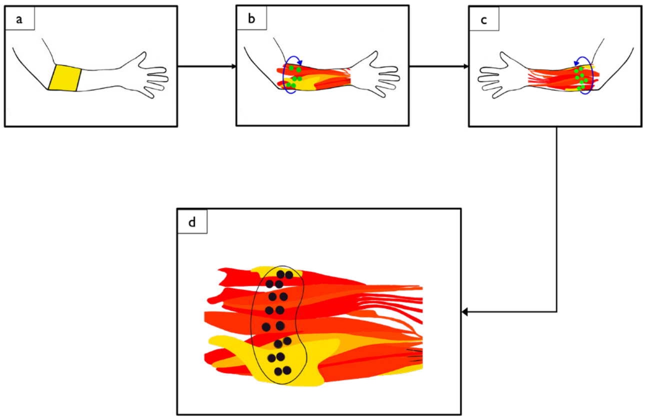


Figure 1. Schematic model of the protocol to map the anatomical sites for electrode positioning. a) Approximate area where the ventral site of the target muscles at the forearm are located. b) Location of the ventral zone of the forearm muscles from the lateral view. c) Location of the ventral zone of muscles from the medial view. d) Forearm template with the location of eight main sites of muscle activation during hand movements. This diagram is an unfolded representation of the muscles found in both lateral and medial views.
Source: Authors

In essence, the forearm length was divided in four segments: proximal, medial-proximal, medial-distal, and distal. The forearm circumference was measured at the boundary between the proximal and medial proximal segments. The latter site concurs with the area of the forearm where most electrodes are located in the *general template*. This area is shown in Figure 1a.

A common reference dry electrode was placed at the olecranon (a dielectric site) using a piece of elastic fabric that is separated from the textile sleeve. The fabric used to manufacture the sleeve was made of 92% polyester and 8% elastane, and it can withstand a temperature of 160 °C.

The APA was mounted on a textile sleeve to be placed over the forearm in order to acquire eight bipolar channels of sEMG signals from hand movements. This **F**orearm dry **EL**ec**T**rodes anatomical array was named FELT, an evolution from a first array proposed in Toledo-Peral *et al*. (2018).

The designed dry electrodes and the positioning array were subjected to the following functional evaluations:

- Dry electrodes evaluation 1: electrode-skin electrical impedance characterization
- Dry electrodes evaluation 2: Bland-Altman agreement assessment
- FELT functionality evaluation 1: anatomical array
- FELT functionality evaluation 2: signal acquisition and data processing

Dry electrodes design and manufacture

The electrodes were designed using CAD software (Dassault Systèmes SolidWorks Corporation, USA) and manufactured in a Computer Numeric Control (CNC) machine. The electrode was lathed as one piece using a ½ inch 304 stainless steel bar. It was designed to be 12 mm in diameter. On one side, the area for skin contact was flat and polished, and, on the other side, it had a 10 mm thread to connect the lead cable. An acorn nut was placed on the thread of the electrode to fix it to the sleeve and the connection cable.

Dry electrodes evaluations

The evaluations performed on the designed dry electrodes (sEMG-DRY) were electrode-skin impedance characterization and a Bland-Altman agreement assessment. To compare the electrodes against the standard reference, pre-gelled standard commercial disposable electrodes (sEMG-GEL) were used (Kendall, COVIDIEN).

Dry electrodes evaluation 1: electrode-skin electrical impedance characterization

According to Martin *et al.* (2018), the circuit used for impedance measurements was powered by an 8 Vpp sinusoidal signal provided by a function generator (AFG320, Sony Tektronix), which acted as a power supply. It was connected in series with a 100 kΩ resistor, and then to an electrode in contact with the skin. The second electrode in contact with the skin was connected to a 100 Ω resistor and then to the generator ground.

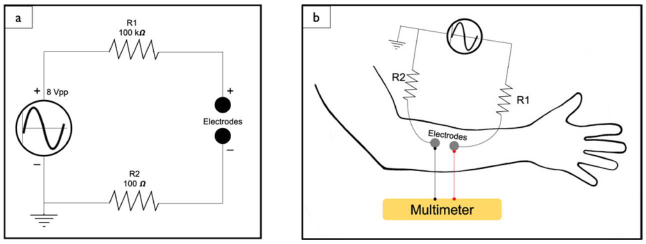


Figure 2. Measurements at the skin-dry electrode interface. a) Electronic circuit used to power the dry electrodes. b) Configuration used to measure the resulting electrode-skin impedance
Source: Authors

The sEMG-DRY electrodes were placed and left to settle for 15 min before any signals were acquired. The electrical impedance response of the electrode was measured for a frequency range between 20 and 500 Hz, with a frequency sweep and increments of 20 Hz (20-100 and 400-500 Hz) and 50 Hz (100-400 Hz). The peak-to-peak voltage variations were measured with a multimeter (FLUKE, model 189). The electric current was calculated through Equation (1), and the electrical impedance values were calculated with Equation (2) (Martin *et al.*, 2018):

$$I = \frac{V_{ab}}{R_1} \tag{1}$$

$$Z = \frac{V_{bc}}{I} \tag{2}$$

where I is the current; V_{ab} is the measured voltage that passes through R_1, which has a resistance of 100 kΩ; V_{bc} denotes the measured voltage that passes through the electrodes characterized; and Z represents the calculated electrical impedance, which was measured for each frequency and repeated for 10 sessions with a 23-year-old female abled-body subject for both the sEMG-DRY and the sEMG-GEL electrodes.

Dry electrodes evaluation 2: Bland-Altman agreement assessment

The Bland-Altman test was used to assess the agreement between the signals acquired by the sEMG-DRY and the sEMG-GEL electrodes. The first step was to choose the parameter to be extracted in order to perform the evaluation. In this case, we selected the mean temporal amplitude and calculated it using the area under the curve A of the signal for the period N regarding the signals obtained with the electrodes. This was calculated using Equation (3).

$$A = \int_0^N f(t)\,dt \tag{3}$$

where $f(t)$ is the signal of interest for a length N.

The Bland-Altman or difference plot is a statistical graphical method used to analyze the agreement between the measurements of two instruments, or, in this case, the acquired sEMG signals using two types of electrodes. This plot compares the sEMG-DRY signals (new measurement technique) against the sEMG-GEL ones (the reference/gold standard). This plot helps to assess the magnitude of disagreement (or bias) and determine if there is any trend. It is expected that, if the differences are normally distributed (Gaussian), 95% of them will lie between the limits of agreement (mean of differences ± 1,96 standard deviation, or SD). Therefore, if the bias (as estimated by the mean difference and the SD of the differences) is small, the two methods can be used interchangeably, and the new technique can replace the established one.

The precision of the estimates was determined by calculating the limits of agreement, the standard error of the mean (SEM), and the standard error of the limits of agreement (SELA). These limits were calculated using Equations (4) and (5), respectively.

$$SEM = \sqrt{SD^2 / n} \tag{4}$$

$$SELA = \sqrt{3 \cdot SD^2 / n} \tag{5}$$

where n is the sample size, and SD denotes the standard deviation of the differences.

A 95% confidence interval (CI) was calculated for the bias, as well as for the lower limit (LL) of agreement and for the upper limit (UL) of agreement. These values were calculated using Equations (6), (7), and (8), respectively.

$$CI\left(95\%\right) = -\bar{d} \pm \left(t \cdot SEM\right) \tag{6}$$

$$CI\left(95\%\right)LL = -LL \pm \left(t \cdot SELA\right) \tag{7}$$

$$CI\left(95\%\right)UL = -UL \pm \left(t \cdot SELA\right) \tag{8}$$

where $\bar{d}$ is the mean difference or bias of the measurements obtained from the sEMG-DRY and the sEMG-GEL electrodes, and t takes a value of 2,0010 (Bland and Altman, 1986).

The sEMG signals of four 20 to 22 years old abled-body female subjects who performed power grasp isotonic contractions using both types of electrodes were acquired for the Bland-Altman agreement assessment. These signals were recorded at the brachioradialis muscle of the right arm (dominant side). Three repetitions lasting 3 s were performed, with 10 s rest. The acquired signals were divided into 0,5 s segments for the statistical analysis, which is also the window length selected for the processing algorithm. In our experience, the first 3 s of the contraction contain the transient information of the hand movement. We found no need to record or use the posterior stationary segment; the 10 s of rest ensure that the subject has enough time to relax their muscles and prepare for the next repetition (Toledo *et al.*, 2018).

FELT functionality evaluation 1: anatomical array

The FELT was used to acquire sEMG signals from one 35-year-old abled-body male. FELT placement was performed following the *FELT positioning protocol* described below. This is shown in Figure 3.

1. Place the forearm in supine position
2. Place the FELT next to it with the seams looking upwards
3. Slide the hand through the FELT, maintaining the position for sections of the FELT (upper and lower)
4. Align the seams of the FELT with the line formed by the middle point of the flexor retinaculum and the ulnar fossa
5. Verify that the FELT has a tight fit over the forearm
6. Remove the bottom of the FELT
7. Place the reference electrode on the olecranon using a piece of elastic fabric

The electrodes placed in contact with the skin were left to settle for 15 min before any signals were acquired.

FELT functionality evaluation 2: signal acquisition and data processing

sEMG signals were acquired with the amplifier Shimmer3 EMG unit (Shimmer), connected via Bluetooth, at a 1 024 Hz sampling frequency, a fixed gain of 12 for the amplifier, and a bipolar configuration for eight channels. Signal processing was performed using MATLAB 2021b. For this evaluation, a 34-years-old abled-body male volunteer underwent five sessions of sEMG signal acquisition using the FELT.

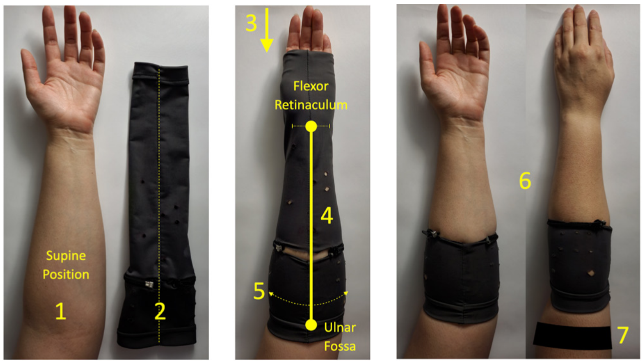


Figure 3. Positioning protocol for the FELT. The numbers match the steps designed to ensure a proper anatomical placement.
Source: Authors

In the case of the sEMG-GEL electrodes, the skin had to be previously cleaned to improve contact and for the adhesive to hold on longer. Then, the electrodes were positioned using the *general template*, and the reference electrode was placed at the olecranon.

An sEMG signal acquisition protocol was implemented, and five trials were performed, each one consisting of three contractions with a 3 s duration and 10 s rests. The algorithm acquired the sEMG signal using Shimmer3 and sent it directly to MATLAB. The hand movement protocol is shown in Figure 4.

Offline processing was performed based on a previously developed algorithm (Toledo-Peral *et al.*, 2018), with a few adjustments since a different acquisition system was being used. A time-frequency approach using the discrete wavelet transform (DWT) discriminated the signal content from artefacts and components not in the range of the frequencies of interest.

For either algorithm, the signal went through a Notch filter to eliminate the 60 Hz line interference. Then, a Daubechies wavelet 'db9' eight-level decomposition was performed for baseline elimination. Later, a wavelet Daubechies 'db4' two-level decomposition was performed to obtain the frequency range of interest for the sEMG components. The absolute value of the signal was calculated to finally go through a wavelet Haar nine-level decomposition to find the envelope. Using the last signal, a threshold was established to segment the active portion of the sEMG signals. This is depicted in Figure 5. From the segmented signal, several parameters were calculated. The parameters calculated using Equations (9) to (15) (Phinyomark *et al.*, 2013) are shown in Table 2.

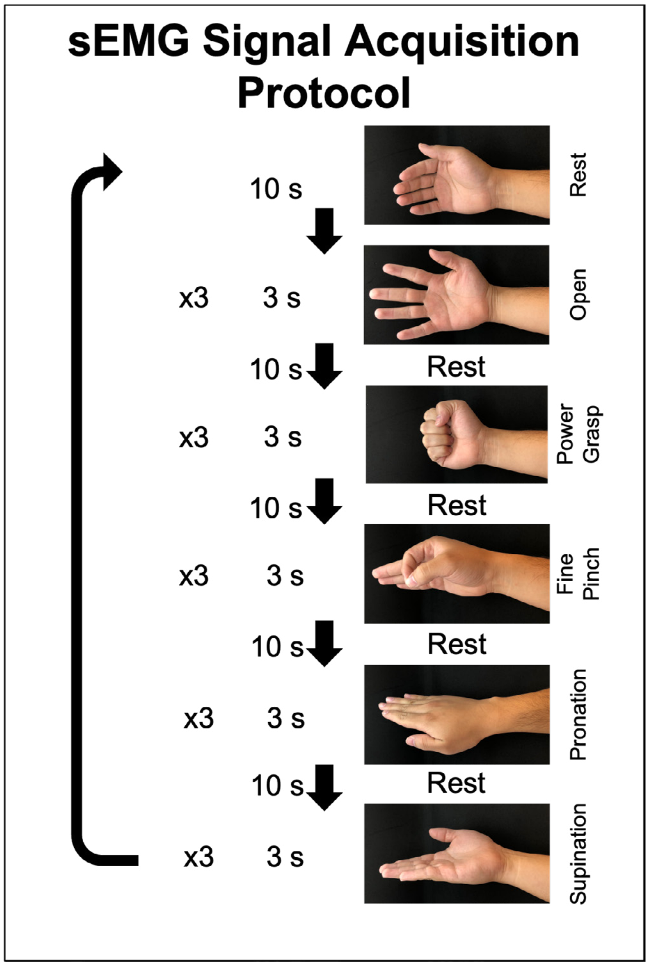


Figure 4. sEMG signal acquisition protocol for six hand movements. Each movement was performed three times for 3 s, with 10 s rest between movements.
Source: Authors

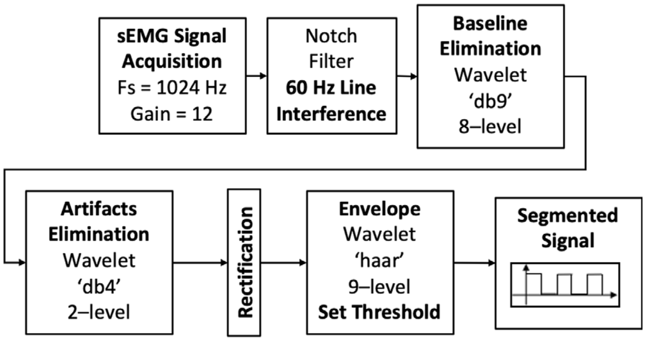


Figure 5. Signal processing algorithm based on DWT decomposition. This algorithm was used for signals acquired with either sEMG-DRY or sEMG-GEL electrodes.
Source: Authors

Table 2. Equations for seven features that characterize sEMG signals

Feature	Equation	Equation Number
Mean absolute value	$MAV = \frac{1}{n}\sum_{i=1}^{n}\lvert x_i \rvert$	(9)
Integral of the absolute value	$IAV = \sum_{i=1}^{n} x_i$	(10)
Root mean square	$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i)^2}$	(11)
Variance	$V = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$	(12)
Skewness	$s = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3}$	(13)
Slope signal change	$SSC = \begin{cases} 1 & x_i > x_{i+1} \ \& \ x_i > x_{i-1} \\ 1 & x_i < x_{i+1} \ \& \ x_i < x_{i-1} \\ 0 & else \end{cases}$	(14)
Wave-length	$WL = \sum_{i=1}^{n}\lvert x_i - x_{i-1} \rvert$	(15)

Source: Phinyomark *et al.* (2013)

Here, x denotes the sEMG signal, x_i is a sample of the signal, $\bar{x}$ represents the mean of x, σ is the standard deviation, and n denotes the length of the signal.

For each of the six hand movements, the seven parameters (Phinyomark *et al.*, 2013) were calculated and fed to an Artificial Neural Network (ANN).

Results and Discussion

The functionality of the proposed dry electrode APA for sEMG signal acquisition at the forearm while performing six hand movements is evaluated in this section. The 17 sEMG-DRY electrodes are composed of flat, round, stainless steel cylindrical pieces, each 12 mm in diameter, with a threaded section used to affix them to the textile sleeve and the connection cable using an acorn nut. Compared to previous designed prototypes (Martin *et al.*, 2018; Toledo-Peral *et al.*, 2018), this one yields a signal with fewer noise artefacts such as line interference. The contact area of the sEMG-DRY electrode ensured uniform contact with the skin. All electrodes were identically built.

Then, the designed sEMG-DRY electrodes were evaluated regarding electrode-skin electrical impedance and compared to the sEMG-GEL ones using the Bland-Altman agreement assessment test.

Dry electrodes evaluation 1: electrode-skin electrical impedance characterization

The electrode-skin electrical impedance (mean ± standard deviation) was measured for a frequency range of 20 to 550 Hz for both the sEMG-DRY and the sEMG-GEL electrodes. The corresponding electrical impedance plots are shown in Figure 6. The electrical impedance response had a similar behavior in both electrode types regarding the attenuation of their amplitude along the frequency range. The impedance values were higher for lower frequencies in sEMG-DRY electrodes. The literature states that dry electrodes have higher impedance values than pre-gelled standard electrodes (Garcia *et al.*, 2007; Xie *et al.*, 2013).

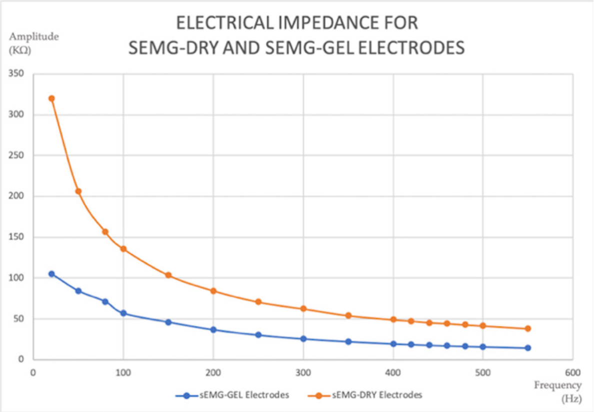


Figure 6. Electrical impedance response for a frequency range of 20 to 550 Hz for both sEMG-DRY and sEMG-GEL electrodes
Source: Authors

sEMG-DRY electrodes have some advantages over sEMG-GEL ones. Firstly, their placement is faster. Over time, the conductive gel degrades and dries, yielding low quality signals, while sEMG-DRY electrodes have improved contact because they take advantage of the properties of the skin. sEMG-DRY electrodes allow for long-term measurements, can be reused endless times with proper care, do not generate waste, are antibacterial, and pose a lower risk of skin rashes (Srinivasa and Pandian, 2017).

Dry electrodes evaluation 2: Bland-Altman agreement assessment

To assess whether the measurements provided by a new instrument are valid, they must be compared with those obtained via a gold standard instrument. To this effect, a Bland-Altman agreement assessment was performed. We wanted to know if sEMG-DRY electrodes could be used as sEMG-GEL ones, so we acquired sEMG signals using both types from four abled-body female volunteers. The parameter chosen for the test was the mean temporal amplitude, and a total 60 records of sEMG signals were involved in the calculations. The data were obtained using Equation (3) and are shown in Table 3. The Bland-Altman plot is shown in Figure 7.

Table 3. Mean temporal amplitude (*A*) of sEMG signals acquired using sEMG-DRY and sEMG-GEL electrodes

Reps	sEMG–DRY Electrodes Mean temporal amplitude [mV]				sEMG–GEL Electrodes Mean temporal amplitude [mV]			
	Subject				*Subject*			
	1	2	3	4	1	2	3	4
1	34,08	8035	29,84	3566	50,28	37,23	57,97	65,83
2	50,71	100,16	45,20	29,48	65,55	50,09	60,18	69,47
3	63,62	85,05	39,77	30,40	50,45	41,14	60,09	59,08
4	56,38	66,48	42,90	29,31	48,80	44,99	55,66	49,87
5	52,21	46,00	35,57	27,17	41,77	45,46	48,76	37,03
6	40,39	93,62	37,06	22,55	36,01	27,03	27,63	44,99
7	45,19	85,02	25,89	25,50	40,55	48,07	47,19	74,51
8	59,60	79,76	31,23	48,86	44,72	50,94	43,87	51,41
9	54,84	65,69	37,71	51,56	33,07	45,78	48,57	40,32
10	61,46	55,55	36,56	25,61	31,17	42,58	50,55	40,91
11	56,70	85,59	26,79	36,92	36,20	41,73	38,93	44,23
12	54,37	87,13	26,58	26,91	32,45	43,51	43,20	70,20
13	57,30	79,17	34,02	54,40	39,12	39,36	43,91	64,18
14	51,14	77,06	33,80	46,94	30,54	39,29	53,28	48,93
15	44,63	63,27	36,90	26,82	33,57	42,69	49,12	42,23

Source: Authors

The Bland-Altman method was used to plot the differences in the average mean temporal amplitude values of the acquired sEMG signal using both types of electrodes (Figure 7). The limits of agreement (-6,7164 mV, 49,2496 mV) exhibited 95% (57/60) of difference scores. The mean difference (bias) in the measurements of the sEMG-DRY and the sEMG-GEL electrodes was $\overline{d} = 21,2666\,\mathrm{mV}$. The SD of the differences was 13,9915 mV, and the width of 95% of the limits of agreement was 55,966 mV.

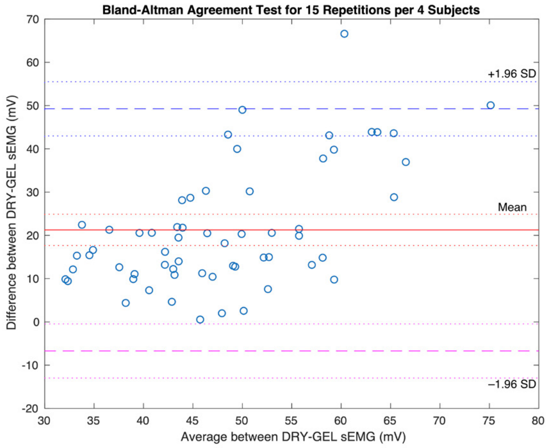


Figure 7. Bland-Altman plot of agreement. The differences between the sEMG-DRY and the sEMG-GEL electrodes are drawn against the mean of the paired measurements. For a normal distribution, at least 95% of the data are expected to land within the limits of agreement.
Source: Authors

The values of difference were normally distributed (Gaussian), showing approximately 73% of the differences (44/60) within one SD of the mean. 95% of the differences landed within two SDs (57/60) in the plot, and 98,3% were within three SDs (59/60).

The 95% CI was calculated by finding the point of t-distribution, with n-1 degrees of freedom, adding/subtracting the t standard error. For the amplitude data corresponding to Equations (4) and (5), the SD was 13,9915 mV, the SEM was 1,8063 mV, and the SELA was 3,1286 mV. For the 95% CI in Equation (6), there were 59 degrees of freedom and t = 2,0010. Hence, the 95% CI for the bias ranged from 17,6522 to 24,8810 mV. The 95% CI for the lower limit (LL) of agreement in Equation (7) was from -12,9767 to -0,4560 mV. For the upper limit (UL) of agreement, as per Equation (8), the 95% CI ranged from 42,9893 to 55,5099 mV.

FELT evaluation 1: anatomical array

The FELT anatomical array was designed and manufactured using dry electrodes and an anthropometrical textile sleeve. The FELT is a wearable APA using a sleeve placed on the forearm. It houses eight bipolar channels of sEMG electrodes to acquire hand movement signals with the Shimmer3 system. The average positions of the *individual templates* were used to determine the final anatomical positioning, which resulted in the proposal of a *general template*. This development translated into an easy-to-place APA fixed on a textile sleeve, able to hold eight pairs of sEMG-DRY. The original design, along with several views of the manufactured FELT, is shown in Figure 8.

There is an elastic band at the top of the sleeve, which contains the reference electrode that is placed at the olecranon (Figure 5a). The FELT was designed using anthropometric measurements taken from 17 abled-body volunteers with heights ranging from 1,50 to 1,90 m and weights from 50 to 110 kg.

The fabric used is soft, antibacterial, stretchable, and breathable, providing numerous benefits. These advantages make the sleeve comfortable to wear and help to secure the electrodes in place on the forearm. According to Kusche *et al.* (2019), dry electrode measurements should ensure that the electrode has enough contact force and that it works within the desired frequency range. The less the electrodes move, the fewer artifacts there will be.

The FELT anatomical array is designed to acquire sEMG signals closer to the main source for each movement. To this effect, it incorporates eight bipolar channels of dry electrodes and a reference. When a hand movement is performed, the FELT moves with the forearm and the hand, without restricting the range of movement. Although the FELT is wired, the use of Shimmer3 connected via Bluetooth to a computer transforms the FELT into a wearable, portable device that can have applications not only for control, but also for evaluations during various actions.

FELT evaluation 2: signal acquisition and data processing

The proper placement of the FELT is the difference between useful and useless signals. A FELT positioning protocol was designed and followed each time the sleeve was used in a trial. Placing the FELT was a straightforward process, which can be seen in Figure 3. Once the sleeve is on the forearm, the most important step is to ensure proper positioning by aligning the seams of the sleeve with two anatomical points: the ulnar fossa (the middle point of the elbow crease) and the middle point of the flexor retinaculum (a transversal ligament located across the wrist).

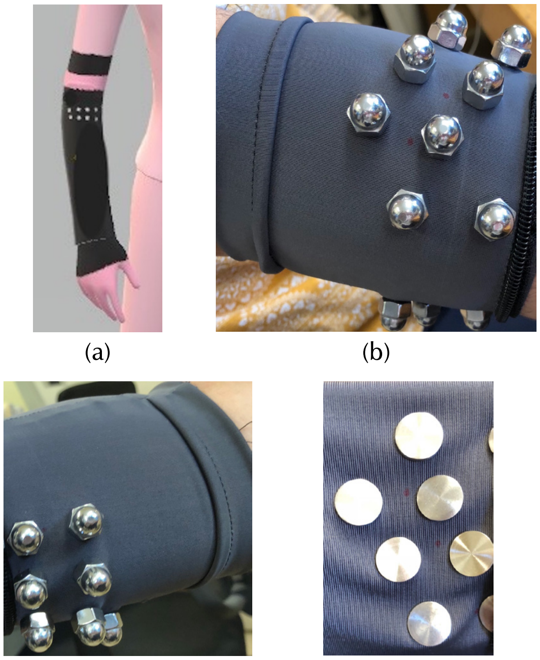

(a) (b)
(c) (d)

Figure 8. FELT Prototype. A) Simulation of the FELT, b) top view of the FELT prototype, c) bottom view of the FELT prototype, and d) internal view of FELT prototype, showing the inner side of the sEMG-DRY electrodes in contact with the skin.
Source: Authors

After placement, the sEMG signals were acquired. Two bipolar channels were recorded for hand opening and power grasp movements. Hand opening and power grasping are antagonist movements and activate antagonist muscles (brachioradialis and finger flexors *vs.* fingers extensors, which correspond to channels 1 and 5 of the FELT).

Over ten sessions, the main contribution of hand opening movements can be seen in channel 1 (mean absolute value or MAV of 80,42 ± 6,70 mV), while Channel 5 shows lower amplitudes (MAV of 18,69 ± 0,79 mV). For power grasping, the behavior is quite different, as channels 1 and 5 activate, albeit with lower amplitude values (channel 1: MAV of 39,93 ± 4,29 mV; channel 2: MAV of 32,38 ± 3,48 mV). The variations observed over ten sessions are measured for each channel and movement. These results are shown in Table 4. The maximum amplitude variation, occurring in channel 1, showed a 24,81% change in the MAV of the sEMG signal between sessions 8 and 9 for hand opening movements.

Five sessions were conducted, and sEMG signals were acquired using the FELT and Shimmer3 with eight recording channels for six hand movements: hand opening, power grasp, fine pinch, supination, pronation, and rest. Then, the parameters were calculated for each movement and were used to build the features vectors to feed the ANN. To this effect, the Equations in Table 2 were used.

Table 4. Results obtained for the amplitude variations over 10 sessions of sEMG signal acquisition using the FELT for six hand movements. Channel 1 and 5 amplitude variations between contiguous sessions are presented, featuring the maximum variation value.

Session No.	Movement			
	Hand Opening		Power Grasp	
	Channel 1 Amplitude Variation [%]	Channel 5 Amplitude Variation [%]	Channel 1 Amplitude Variation [%]	Channel 5 Amplitude Variation [%]
1	-	-	-	-
2	11,17	8,55	1,69	3,04
3	3,54	6,29	1,94	9,18
4	14,73	2,78	13,73	11,59
5	7,17	8,36	5,92	6,56
6	4,91	3,88	2,56	0,31
7	7,89	6,62	1,54	4,04
8	1,37	2,03	5,46	2,64
9	24,81	10,22	17,92	14,09
10	5,24	6,71	9,37	9,44
MAXIMUM VARIATION VALUE (%)	**24,81**	**10,22**	**17,92**	**14,09**

Source: Authors

A six-class feed-forward ANN was used, with Bayesian regularization as the transfer function ('trainbr'). Bayesian regularization is a technique that introduces prior knowledge into the learning process to prevent overfitting and improve the network's generalization ability. This type of model is used for pattern recognition when there is a low number of repetitions and large features vectors. The ANN had six outputs and 56 inputs set up according to the previously calculated parameters. Data division was randomized. Performance was calculated through the sum squared error function. The vectors of parameters used for validation were not shared with the training and testing process.

sEMG signals are stochastic and their behavior depends on several factors such as the number of repetitions in a session, the lactate released the day before, and fatigue during the session or from days before the trial. A way to customize the algorithms to fit every individual is to build a new classifier in each session. For the five sessions conducted, the average accuracy was 97,86 ± 0,58%. This six-class classifier has a sensitivity of 0,9787, and a specificity of 0,9955. These values are shown in Table 5.

Table 5. Results obtained regarding classifier accuracy over five sEMG signal acquisition sessions

Session No.	Classifier Efficiency (%)	Classifier Error (%)
1	97,8	2,2
2	98,5	1,5
3	98,4	1,6
4	97,4	2,6
5	97,2	2,8
Mean ± SD	**97,86 ± 0,58**	**2,14 ± 0,58**

Source: Authors

Classifier efficiency values showed that the FELT allowed for the acquisition of sEMG signals containing the characteristic patterns of muscles to differentiate between the six hand movements. These are useful evaluations, since the algorithms adjust to the characteristic features of the trial subject during the session. This type of flexibility works in favor of the subject because they do not have to force-fit the system; the system fits them. Thus, if the session is longer, the system may readjust, or the subjects could take a break, take out the FELT, and come back later, position the FELT, re-calibrate the classifier, and have a second session. No extra electrodes are wasted, and no skin irritation caused by taking the electrodes off and on will occur.

Regarding other works that classify hand movements based on sEMG signals and pattern recognition using neural network approaches, Kang *et al*. (2023) present a hand gesture recognition system using a binarized neural network, reporting a 95,4% classification accuracy for four gestures from a single sEMG channel. Lee *et al*. (2023) report 97% accuracy while using a stretchable array of eight pairs of fabric-based sEMG electrodes, and Tepe and Demir (2022) report mean accuracies of around 95,83% while using the commercial Myo band sEMG acquisition system.

The Myo armband is probably the most used device for prosthetic control in research (Abduo and Galster, 2015; Chen *et al*., 2020; Morais *et al*., 2016; Visconti *et al*., 2018; Yakob *et al*., 2021). It was originally designed to be used by the public as an everyday-life control device. It has a very sturdy design, and the circumference adapts to different arm thicknesses. It is calibrated at the beginning of each use, meaning that it adapts the algorithm to the sEMG signals acquired, so it can be re-calibrated as needed when the sEMG signals change. The embedded acquisition system has a limited frequency spectrum (0-100 Hz), but the full sEMG spectrum goes up to 500 Hz (Pizzolato *et al*., 2017). Moreover, the battery time of this device is limited.

Other acquisition configurations have been tested for recording sEMG signals, such as the Otto Bock 13 E200 and the Delsys Trigno setups (Pizzolato *et al*., 2017). Both systems have pre-positioned dry electrodes. In the case of the Delsys Trigno, the placing sites must be determined by the user. The Otto Bock system has an equidistant array of electrodes around the circumference of the forearm.

The proposed FELT array has an anatomical design that positions the dry electrodes closer to the source of different sEMG signal patterns. After signal processing, this allows an ANN to classify the signal with high accuracy.

Our device can be used with any acquisition system, and one can use as many electrodes as needed. Some drawbacks of the FELT array include issues such as the fact that it is wired to the acquisition system and, although it can be detached, processing must be conducted in order to enhance the sEMG signal, as it has a noisier-than-desirable raw signal.

As described by Al-Ayyad *et al*. (2023), the use of sEMG signals is migrating from prosthetic control towards use in rehabilitation training and physiological tracking. sEMG signal applications are moving towards anticipating trajectories for one arm based on the movement of the other.

Future work regarding this development will focus on the implementation of a small detachable acquisition system aimed at transforming the FELT into a wearable that can be used in different applications. An important one has to do with virtual/augmented reality (VR/AR) interfaces and serious games (Toledo-Peral *et al*., 2022). We believe that the proposal of arrays such as the one in this article may provide a better starting point for these developments, and that VR/AR applications will certainly benefit from sEMG signals as a control for present and future movements, in addition to being an evaluation stage for previous ones, thus closing the biofeedback loop.

The FELT anatomical array is a proposal aimed at embracing an organic design that is focused on the user and tries to harness bio-signals with less intervention during movements, hoping for better results in clinical applications, mainly in the field of rehabilitation.

As with most developments, this design is subject to limitations, which include the low number of test subjects. In addition, not all subjects may fit the textile sleeve; a larger or smaller one may be needed. This also limits the variety of sEMG signals acquired. It is necessary to include a larger and more varied demographic for both sleeve design and signal acquisition, which might imply changes in the classifier algorithm. Moreover, there is always room for improvement in the design of the electrodes.

Conclusions

An anatomical positioning forearm dry electrode array fixed on a textile sleeve (called *FELT*) was designed and built to acquire eight bipolar channels of sEMG signals from hand movements. The fabric has antibacterial and breathability properties that allow subjects to be comfortable while ensuring that the dry electrodes have a tight fit to the skin in order to avoid movement artefacts, which is crucial in dynamic measurements. The anatomical design is based on anthropometrical measures, which allows the FELT to pick up sEMG signals closer to the ventral area of the muscles which corresponds to the movement performed, thus improving signal quality as well as the sleeve's fit around the forearm.

Even though the sEMG-DRY electrodes, as any dry electrode, yield a noisier signal than that of sEMG-GEL electrodes, the digital signal processing performed, which is based on the DWT technique, allows making the signals comparable to those acquired with the pre-gelled electrodes. We proved this by means of the Bland-Altman agreement assessment, which showed a 95% agreement between both signals. This means that the signals from the sEMG-DRY electrodes can be used with the same confidence as those from the others.

Still, this might sound like a contradiction. Why should we substitute pre-gelled electrodes for noisier dry electrodes and have to go through a different type of processing to get a similar result? The answer is simple. It is complicated to place that many electrodes repeatedly for long periods of time without skin irritation, l*et al*one place them always at the same location, and it is even more difficult to find the correct anatomical positions without prior knowledge. These results show that the FELT is usable for hand movement signal acquisition.

Another important aspect to consider is placement. A positioning protocol was developed to serve as a guideline to properly place the FELT on the forearm, and, even though the protocol is based on anatomical markers, the procedure is straightforward. This is seen in Figure 3. It is also important to ensure repeatability, and, in this case, the difficulty was to find a parameter to evaluate and help calibrate the changes between sessions. To this effect, we calculated the MAV of sEMG signals from the antagonist muscles for hand opening and power grasp movements (brachioradialis and finger flexors *vs.* finger extensors, corresponding to channels 1 and 5). We then calculated the variations between each session, and found the maximum value to be 24,81%. The variations in the positioning of the FELT could thus be quantified using the same signals acquired in the trials. On top of that, as shown in the *Results and discussion* section (*FELT evaluation 2: signal acquisition and data processing*), we built a classifier that recalibrates for each session and proved to have an efficiency of over 97,86 ± 0,58% in differentiating between six hand movements.

Acknowledgements

Funding was received by the National Science and Technology Council (CONACYT) of Mexico, through the project CONACYT-SALUD-2016-01-272983. We also appreciate the funding for the development of this work through ERAnet-EMHE 200022, CYTED-DITECROD-218RT0545, and Proyecto IV-8 call Amexcid-Auci 2018-2020.

The authors would like to thank Gerardo Hernández-Nava, José Antonio Mejía-Licona, and Othón Benítez-Keller for their technical support during the experiments performed, as well as Dr. Jaime H. Guadarrama-Becerril for his input in helping to establish the anthropometrical bases to develop the textile sleeve.

CRediT author statement

Cinthya Lourdes Toledo-Peral: conceptualization, visualization, methodology, validation, formal analysis, investigation, writing – original draft, review, and editing. Ana Isabel Martín-Vignon-Whaley: methodology, formal analysis, investigation, writing – original draft, review, and editing. Jorge A. Mercado-Gutiérrez: methodology, writing – review & editing. Arturo Vera-Hernández: supervision, funding acquisition, writing – review and editing. Lorenzo Leija-Salas: supervision, project administration, funding acquisition, writing – review and editing. Josefina Gutiérrez-Martinez: supervision, project administration, writing – review and editing.

Conflicts of interest

The authors declare that they have no conflict of interest.

Data availability

The data are protected by the Institutional Ethics Board, and, in case of need, they should be requested through them.

CRediT author statement

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

References

Abduo, M., and Galster, M. (2015). *Myo gesture control armband for medical* applications [Bachelor thesis, University of Canterbury]. http://hdl.handle.net/10092/14449

Al-Ayyad, M., Owida, H. A., De Fazio, R., Al-Naami, B., and Visconti, P. (2023). Electromyography monitoring systems in rehabilitation: A review of clinical applications, wearable devices and signal acquisition methodologies, *Electronics (Switzerland), 12*(7), 1520. https://doi.org/10.3390/electronics12071520

Albulbul, A. (2016). Evaluating major electrode types for idle biological signal measurements for modern medical technology. *Bioengineering*, *3*(3), 20. https://doi.org/10.3390/bioengineering3030020

Barbero, M., Merletti, R., and Rainoldi, A. (2012a). *Atlas of muscle innervation zones*. Springer. https://doi.org/10.1007/978-88-470-2463-2

Barbero, M., Merletti, R., and Rainoldi, A. (2012b). EMG imaging: Geometry and anatomy of the electrode-muscle system. In Springer (Eds.), *Atlas of Muscle Innervation Zones* (pp. 39-47). Springer. https://doi.org/10.1007/978-88-470-2463-2_4

Besomi, M., Hodges, P. W., van Dieën, J., Carson, R. G., Clancy, E. A., Disselhorst-Klug, C., Holobar, A., Hug, F., Kiernan, M. C., Lowery, M., McGill, K., Merletti, R., Perreault, E., Søgaard, K., Tucker, K., Besier, T., Enoka, R., Falla, D., Farina, D., ... Wrigley, T. (2019). Consensus for experimental design in electromyography (CEDE) project: Electrode selection matrix. *Journal of Electromyography and Kinesiology*, *48*, 128-144. https://doi.org/10.1016/j.jelekin.2019.07.008

Biga, L. M., Dawson, S., Harwell, A., Hopkins, R., Kaufmann, J., LeMaster, M., Matern, P., Morrison-Graham, K., Quick, D., and Runyeon, J. (n.d.). *Explain the organization of muscle fascicles and their role in generating force*. https://open.oregonstate.education/aandp/chapter/11-2-explain-the-organization-of-muscle-fascicles-and-their-role-in-generating-force/#:~:text=When%20a%20group%20of%20muscle,the%20muscle's%20range%20of%20motion

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307-310.

Chen, L., Fu, J., Wu, Y., Li, H., and Zheng, B. (2020). Hand gesture recognition using compact CNN via surface electromyography signals. *Sensors*, *20*(3), 672. https://doi.org/10.3390/s20030672

Chi, Y. M., Jung, T.-P., and Cauwenberghs, G. (2010). Dry-contact and noncontact biopotential electrodes: Methodological review. *IEEE Reviews in Biomedical Engineering*, *3*, 106-119. https://doi.org/10.1109/RBME.2010.2084078

Dai, C., Zhu, Z., Martinez-Luna, C., Hunt, T. R., Farrell, T. R., and Clancy, E. A. (2019). Two degrees of freedom, dynamic, hand-wrist EMG-force using a minimum number of electrodes. *Journal of Electromyography and Kinesiology*, *47*, 10-18. https://doi.org/10.1016/j.jelekin.2019.04.003

Gan, Y., Vauche, R., Pons, J. F., and Rahajandraibe, W. (2019). *Dry electrode materials for electrocardiographic monitoring* [Conference presentation]. 2018 25th IEEE International Conference on Electronics Circuits and Systems, Bourdeaux, France. https://doi.org/10.1109/ICECS.2018.8617992

Garcia, G. A., Zaccone, F., Ruff, R., Micera, S., Hoffmann, K.-P., and Dario, P. (2007). *Characterization of a new type of dry electrodes for long-term recordings of surface-electromyogram* [Conference presentation]. 2007 IEEE 10th International Conference on Rehabilitation Robotics, Noordwijk, Netherlands. https://doi.org/10.1109/ICORR.2007.4428523

Ghoshdastider, U., Lange, C., Viga, R., and Grabmaier, A. (2012). *A modular and wireless exg signal acquisition system with a dense array of dry electrodes* [Conference presentation]. IEEE Sensors 2012, Taipei, Taiwan. https://doi.org/10.1109/ICSENS.2012.6411473

Guo, L., Sandsjö, L., Ortiz-Catalan, M., and Skrifvars, M. (2020). Systematic review of textile-based electrodes for long-term and continuous surface electromyography recording. *Textile Research Journal*, *90*(2), 227-244. https://doi.org/10.1177/0040517519858768

Guzmán, R. A., Silvestre, R. A., and Arriagada, D. A. (2011). Biceps brachii muscle innervation zone location in healthy subjects using high-density surface electromyography. *International Journal of Morphology*, *29*(2), 347-352. https://doi.org/10.4067/s0717-95022011000200007

Hall, J. E. (2011). *Guyton & Hall: tratado de fisiología medica*. Elsevier.

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., and Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology*, *10*(5), 361-374. https://doi.org/10.1016/s1050-6411(00)00027-4

Kang, S., Kim, H., Park, C., Sim, Y., Lee, S., and Jung, Y. (2023). sEMG-based hand gesture recognition using binarized neural network. *Sensors*, *23*(3), 1436. https://doi.org/10.3390/s23031436

Kapelner, T., Jiang, N., Holobar, A., Vujaklija, I., Roche, A. D., Farina, D., and Aszmann, O. C. (2016). Motor unit characteristics after targeted muscle reinnervation. *PLoS ONE*, *11*(2), 0149772. https://doi.org/10.1371/journal.pone.0149772

Kusche, R., Kaufmann, S., and Ryschka, M. (2019). Dry electrodes for bioimpedance measurements - Design, characterization and comparison. *Biomedical Physics and Engineering Express*, *5*(1), 015001. https://doi.org/10.1088/2057-1976/aaea59

Lee, H., Lee, S., Kim, J., Jung, H., Yoon, K. J., Gandla, S., Park, H., and Kim, S. (2023). Stretchable array electromyography sensor with graph neural network for static and dynamic gestures recognition system. *Npj Flexible Electronics*, *7*(1), 20. https://doi.org/10.1038/s41528-023-00246-3

Liu, J. (2015). Adaptive myoelectric pattern recognition toward improved multifunctional prosthesis control. *Medical Engineering and Physics*, *37*(4), 424-430. https://doi.org/10.1016/j.medengphy.2015.02.005

Losier, Y., Clawson, W., Scheme, E., Englehart, K., Kyberd, P., and Hudgins, B. (2011). *An overview of the UNB hand system* [Conference presentation]. MyoElectric Controls/Powered Prosthetics Symposium Fredericton, NB, Canada.

Martin, A. I., Toledo, C., Mercado, J. A., Vera, A., Leija, L., and Gutierrez, J. (2018). *Evaluation of dry electrodes for sEMG recording* [Conference presentation]. 2018 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE), Porto, Portugal. https://doi.org/10.1109/GMEPE-PAHCE.2018.8400758

Merletti, R., and Farina, D. (Eds.) (2016). *Surface electromyography physiology, engineering, and applications*. John Wiley & Sons.

Meziane, N., Webster, J. G., Attari, M., and Nimunkar, A. J. (2013). Dry electrodes for electrocardiography. *Physiological Measurement*, *34*(9), R47-R69. https://doi.org/10.1088/0967-3334/34/9/R47

Mitchell, B., & Whited, L. (2019). *Anatomy, shoulder and upper limb, forearm muscles*. https://www.ncbi.nlm.nih.gov/books/NBK536975/

Moin, A., Zhou, A., Rahimi, A., Benatti, S., Menon, A., Tamakloe, S., Ting, J., Yamamoto, N., Khan, Y., Burghardt, F., Benini, L., Arias, A. C., and Rabaey, J. M. (2018). *An EMG gesture recognition system with flexible high-density sensors and brain-inspired high-dimensional classifier* [Conference presentation]. IEEE International Symposium on Circuits and Systems, Florence, Italy. https://doi.org/10.1109/ISCAS.2018.8351613

Morais, G. D., Neves, L. C., Masiero, A. A., and Castro, M. C. F. (2016). *Application of Myo Armband System to control a robot interface* [Conference presentation]. BIOSIGNALS 2016 - 9th International Conference on Bio-Inspired Systems and Signal Processing, Rome, Italy. https://doi.org/10.5220/0005706302270231

Niu, X., Gao, X., Liu, Y., and Liu, H. (2021). Surface bioelectric dry Electrodes: A review. *Measurement: Journal of the International Measurement Confederation*, *183*, 109774. https://doi.org/10.1016/j.measurement.2021.109774

Palumbo, A., Vizza, P., Calabrese, B., and Ielpo, N. (2021). Biopotential signal monitoring systems in rehabilitation: A review. *Sensors*, *21*(21), 7172. https://doi.org/10.3390/s21217172

Peng, Y., He, J., Yao, B., Li, S., Zhou, P., and Zhang, Y. (2016). Motor unit number estimation based on high-density surface electromyography decomposition. *Clinical Neurophysiology*, *127*(9), 3059-3065. https://doi.org/10.1016/j.clinph.2016.06.014

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., and Laurillau, Y. (2013). EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with Applications*, *40*(12), 4832-4840. https://doi.org/10.1016/j.eswa.2013.02.023

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., and Laurillau, Y. (2014). Feature extraction of the first difference of EMG time series for EMG pattern recognition. *Computer Methods and Programs in Biomedicine*, *117*(2), 247-256. https://doi.org/10.1016/j.cmpb.2014.06.013

Pizzolato, S., Tagliapietra, L., Cognolato, M., Reggiani, M., Müller, H., and Atzori, M. (2017). Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLoS ONE*, *12*(10), 0186132. https://doi.org/10.1371/journal.pone.0186132

Puurtinen, M. M., Komulainen, S. M., Kauppinen, P. K., Malmivuo, J. A. V., and Hyttinen, J. A. K. (2006). *Measurement of noise and impedance of dry and wet textile electrodes, and textile electrodes with hydrogel* [Conference presentation]. 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA. https://doi.org/10.1109/IEMBS.2006.260155

Reategui, J., and Callupe, R. (2017). *Surface EMG multichannel array using active dry sensors for forearm signal extraction* [Conference presentation]. 2017 IEEE 24th International Congress on Electronics, Electrical Engineering and Computing, Cusco, Peru. https://doi.org/10.1109/INTERCON.2017.8079699

"Recommendations for the practice of clinical neurophysiology: Guidelines of the International Federation of Clinical Neurophysiology" (1999). *Electroencephalography and Clinical Neurophysiology*, *52*(supp.), 1-304.

Rodrigues, M. S., Fiedler, P., Küchler, N., Domingues, R. P., Lopes, C., Borges, J., Haueisen, J., and Vaz, F. (2020). Dry electrodes for surface electromyography based on architectured titanium thin films. *Materials*, *13*(9), 2135. https://doi.org/10.3390/ma13092135

Rojas-Martínez, M., Mañanas, M. A., Alonso, J. F., and Merletti, R. (2013). Identification of isometric contractions based on High Density EMG maps. *Journal of Electromyography and Kinesiology*, *23*(1), 33-42. https://doi.org/10.1016/j.jelekin.2012.06.009

Rui Ma, Dae-Hyeong Kim, McCormick, M., Coleman, T., and Rogers, J. (2010). *A stretchable electrode array for non-invasive, skin-mounted measurement of electrocardiography (ECG), electromyography (EMG) and electroencephalography (EEG)* [Conference presentation]. 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina. https://doi.org/10.1109/IEMBS.2010.5627315

Ruvalcaba, A., Altamirano, A., Toledo, C., Munoz, R., Vera, A., and Leija, L. (2016). Design and measurement of the standards of a miniaturized sEMG acquisition system with dry electrodes integrated [Conference presentation]. 2015 International Conference on Mechatronics, Electronics, and Automotive Engineering, Cuernavaca, Mexico. https://doi.org/10.1109/ICMEAE.2015.34

Ruvalcaba, A., Muñoz, R., Vera, A., and Leija, L. (2017a). Design and test of a dry electrode array implemented on wearable sEMG acquisition sleeve for long term monitoring [Conference presentation]. Pan American Health Care Exchanges, *PAHCE*, *2017-March*. https://doi.org/10.1109/GMEPE-PAHCE.2017.7972111

Srinivasa, M. G., and Pandian, P. S. (2017). *Dry electrodes for bio-potential measurement in wearable systems* [Conference presentation]. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India. https://doi.org/10.1109/RTEICT.2017.8256600

Tepe, C., and Demir, M. C. (2022). Real-time classification of EMG Myo Armband data using support vector machine. *IRBM*, *43*(4), 300-308. https://doi.org/10.1016/j.irbm.2022.06.001

Toledo, C., Flores, E., Mercado, J., Castellanos, P., and Gutiérrez, J. (2018, September 26). *Multiclass sEMG signal processing and classification for upper-limb FES-NP control* [Conference presentation]. Annual Conference of the German Society for Biomedical Engineering, Aachen, Germany.

Toledo-Peral, C. L., Gutiérrez-Martínez, J., Mercado-Gutiérrez, J. A., Martín-Vignon-Whaley, A. I., Vera-Hernández, A., and Leija-Salas, L. (2018). sEMG Signal Acquisition Strategy towards Hand FES Control. *Journal of Healthcare Engineering*, *2018*, 2350834. https://doi.org/10.1155/2018/2350834

Toledo-Peral, C. L., Vega-Martínez, G., Mercado-Gutiérrez, J. A., Rodríguez-Reyes, G., Vera-Hernández, A., Leija-Salas, L., and Gutiérrez-Martínez, J. (2022). Virtual/augmented reality for rehabilitation applications using electromyography as control/biofeedback: Systematic literature review. *Electronics*, *11*(14), 2271. https://doi.org/10.3390/electronics11142271

Topalović, I., Graovac, S., and Popović, D. B. (2019). EMG map image processing for recognition of fingers movement. *Journal of Electromyography and Kinesiology*, *49*, 102364. https://doi.org/10.1016/j.jelekin.2019.102364

Urbanek, H., and van der Smagt, P. (2016). IEMG: Imaging electromyography. *Journal of Electromyography and Kinesiology*, *27*, 1-9. https://doi.org/10.1016/j.jelekin.2016.01.001

Visconti, P., Gaetani, F., Zappatore, G. A., and Primiceri, P. (2018). Technical features and functionalities of Myo armband: An overview on related literature and advanced applications of myoelectric armbands mainly focused on arm prostheses. *International Journal on Smart Sensing and Intelligent Systems*, *11*(1), 1-25. https://doi.org/10.21307/ijssis-2018-005

Vojtech, L., Bortel, R., Neruda, M., and Kozak, M. (2013). Wearable textile electrodes for ECG measurement. *Advances in Electrical and Electronic Engineering*, *11*(5), 410-414. https://doi.org/10.15598/aeee.v11i5.889

Xie, L. Yang, G., Xu, L., Seoane, F., Chen, Q., and Zheng, L. (2013). *Characterization of dry biopotential electrodes* [Conference presentation]. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan. https://doi.org/10.1109/EMBC.2013.6609791

Yakob, M. Y. bin, Baharuddin, M. Z. bin, Khairudin, A. R. M., and Karim, M. H. B. A. (2021). Telecontrol of prosthetic robot hand using Myo Armband [Conference presentation]. 2021 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2021, Shah Alam, Malaysia. https://doi.org/10.1109/I2CACIS52118.2021.9495919

Young, A. J., Hargrove, L. J., and Kuiken, T. A. (2011). The Effects of Electrode Size and Orientation on the Sensitivity of Myoelectric Pattern Recognition Systems to Electrode Shift. *IEEE Transactions on Biomedical Engineering*, *58*(9), 2537-2544. https://doi.org/10.1109/TBME.2011.2159216

http://doi.org/10.15446/ing.investig.106657

Reduction of Large Scale Linear Dynamic MIMO Systems Using ACO-PID Controller

Reducción de sistemas MIMO dinámicos lineales a gran escala mediante el controlador ACO-PID

Jafaar M. Daif-Alkhasraji[1], Salam W. Shneen[2], and Mohammed Q. Sulttan[3]

ABSTRACT

The MIMO technique is an essential element in the standards of communication systems (IEEE 802.11n, IEEE 802.11ac, WiMAX, and LTE) because it helps to increase their capacity. This paper employs a model order reduction (MOR) technique with a PID controller, an ACO algorithm, and the ITAE fitness function to reduce the large-scale linearity of the MIMO technique. The numerator and denominator parameters are set by minimizing the ITAE fitness function between the transient responses of the original and the reduced model. The objectives are achieved with the PID controller and the ACO algorithm for the unit step input. The simulation results show a good system performance. The controller performance is presented with regard to the dynamic response in terms of rising time, settling time, and overshoot/undershoot. Moreover, the results of the proposed method are compared with four literature reports for validation purposes. Evaluating the parameters within the time frame and the error values with and without the PID controller and ACO algorithm allowed validating the functioning of the proposed method. Furthermore, the simulation results revealed that the proposed scheme exhibited sufficient robustness and demonstrated a reduction in the time-domain response and error values.

Keywords: LS-MIMO, model order reduction (MOR), PID controller, ant colony optimization

RESUMEN

La técnica MIMO es un elemento esencial en los estándares de sistemas de comunicación (IEEE 802.11n, IEEE 802.11ac, WiMAX y LTE) porque ayuda a aumentar la capacidad del sistema. En este trabajo se utiliza una técnica de reducción del orden del modelo (MOR) con un controlador PID, un algoritmo ACO y la función fitness de ITAE para reducir la linealidad de gran escala de la técnica MIMO. Los parámetros del numerador y denominador se establecen minimizando la función de aptitud de ITAE entre las respuestas transitorias del modelo original y el modelo reducido. Se alcanzan los objetivos con el controlador PID y el algoritmo ACO para la entrada de paso unitario. Los resultados de la simulación muestran un buen rendimiento del sistema. El rendimiento del controlador se presenta con base en la respuesta dinámica en términos de tiempo de subida, tiempo de asentamiento y sobreimpulso y subimpulso. Además, los resultados del método propuesto se comparan con cuatro reportes de la literatura para su validación. La evaluación de los parámetros en el marco temporal y de los valores de error con y sin el controlador PID y el algoritmo ACO permitió comprobar el funcionamiento del método propuesto. Asimismo, los resultados de la simulación revelaron que el esquema propuesto presentaba suficiente robustez y demostraba una reducción de la respuesta en el dominio temporal y de los valores de error.

Palabras clave: LS-MIMO, reducción del orden del modelo (MOR), controlador PID, optimización por colonias de hormigas

Received: January 6th 2023
Accepted: July 24th 2023

Introduction

Enhancing spectral efficiency is essential, especially in 5G wireless communication systems. In this regard, large-scale systems with multiple inputs and multiple outputs (LS-MIMO) are a promising technology that supports a broad notch of the data for multiple concurrent users, along with the ability to increase spectral efficiency. Furthermore, high-order MIMO can considerably improve the dependability and security of communication systems, obtaining extraordinary locative multiplexing gains, unlike restricted range systems (Ouyang & Yang, 2018; Sulttan, 2019; Han et al., 2020; X. Yang et al., 2020; Dang et al., 2021). Moreover, by increasing the physical antenna size, many intrinsic advantages can be exploited, *e.g.*, channel hardening and increased cell coverage, area, and throughput. This is allowed by high-order MIMO in integration with massive infrastructures such as airport roofs, the walls of theaters and stadiums, and large indoor shopping centers, also providing an extremely reliable technology for 6G systems (Khudhair et al., 2016; Sulttan, 2016). Despite the benefits of LS-MIMO, these advantages come at the price of increasing the complexity of the hardware, which is most notable in signal processing operation for transceivers. Therefore, there is still a need to seek an additional pragmatic, low-complexity reducing algorithm for large-scale systems (Rusek et al., 2012). It is noteworthy that, for point-to-point configurations, the receiver aspect has additional complexity when compared to the transmitter side. For example, the more antennas are employed for the transmitter aspect, the higher the system complexity, which rises exponentially with a huge number of antennas.

---

[1] Electrical engineer, University of Technology (UOT), Iraq. PhD in Optical Wireless Communication Engineering, Newcastle University, United Kingdom. Affiliation: Lecturer, University of Technology, Electromechanical Engineering Department, Baghdad, Iraq. E-mail: jaafar.m.dhaif@uotechnology.edu.iq
[2] Electrical engineer, University of Technology (UOT), Iraq. PhD in Power Electronics, Huazhong University of Science and Technology (HUST), China. Affiliation: Energy and Renewable Energies Technology Center, University of Technology, Baghdad, Iraq. E-mail: salam.w.shneen@uotechnology.edu.iq
[3] Electrical engineer, University of Technology (UOT), Iraq. PhD in Information and Communications Engineering, Huazhong University of Science and Technology (HUST), China. Affiliation: Assistant professor, University of Technology, Electromechanical Engineering Department, Baghdad, Iraq. E-mail: mohammed.q.mohammed@uotechnoloy.edu.iq

On the other hand, extra attention is paid to the receiver aspect of the MIMO model, as the transmitted data must be provided to several users (Jaldén & Ottersten, 2005; Larsson, 2009). In terms of applications, MIMO has been considered to be one of the crucial technologies in third, fourth, and fifth-generation (3G, 4G, and 5G) wireless communication systems since it offers considerable increases in link range and data rate without additional bandwidth or transmit power increases. Such technologies play an essential role in many applications in communication systems, such as GSM, Bluetooth, WiMAX, WLAN, wireless sensors, portable devices, handheld gadgets, automobiles, smartphones, and many more (Sharma et al., 2022; Bizaki, 2011). Furthermore, in order to provide an acceptable service, the number of antennae must be increased for multi-user scenarios. In this vein, large arrays come with a crucial challenge regarding reliability and computationally effective detection (Lyu et al., 2019). As for analysis or controller design, complicated expressions to simulate physical techniques, which imply complex handling of high-order differential formulas, have inevitably led to low-order designs. Such mathematical formulas could be adopted to provide a sufficient performance. Hence, comprehensive knowledge is required, along with decreasing complexity with simple system designs (Vasu et al., 2012). A considerable amount of literature on designing PID controllers for SISO and MIMO prototypes has been published (Lengare et al., 2012; Labibi et al., 2009; Rajinikanth & Latha, 2012; Hu et al., 2010). One of the critical missions of LS-MIMO technology is to reduce the complexity of dynamic designs. Given that dynamic systems contain many variables, notably many mathematical models, it is hard to adopt a complex mathematical form, since the system works in real time. Additionally, memory capacity requirements must be fulfilled, and quick and reliable results must be obtained based on system design (Juneja & Nagar, 2015; Al-doraiee & Al-Qaraawi, 2013; Muhsen & Raafat, 2021). Therefore, over the past few years, heuristic techniques have been widely employed in the literature to achieve optimal performance regarding the obstacles posed by MIMO controller configurations (Hassanzadeh & Mobayen, 2011; Lones, 2014). The difficulty of controlling the MIMO scheme arises from the disruption occurring in one branch, which entails changes in the other branches. Compared to the traditional methods, such as the Zeigler and Nichols and Cohen and Coon tuning techniques, the heuristic process is free of any mathematical calculations. Much of the current literature on MIMO proposes several heuristic approaches to obtain an optimum performance via tuning approaches, *e.g.*, Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Bacterial Foraging Optimization (BFO), the Firefly Algorithm (FA), and Ant-Colony Optimization (ACO) (Saroja et al., 2017; Sivakumar et al., 2015; X.-S. Yang, 2009; Daraji & Hale, 2014; AL-Suhail & Miry, 2015; Aziz et al., 2022). It is common knowledge that organizing and executing a controller-based centralized method is a complicated issue when compared to the decentralized method. It is worth adding that, according to the criteria for designing MIMO systems, several methods must be adopted, *i.e.*, the centralized method (Vu et al., 2007), the decentralized method (Lengare et al., 2012; Lones, 2014; Labibi et al., 2009), and optimal controllers with decouplers (Dhanraj & Nanjundappan, 2014; X.-S. Yang, 2009; Liu et al., 2003; Sivakumar et al., 2015).

Suman and Kumar (2020) suggested an approach to reducing the order of large-scale linear dynamical systems based on the Balanced Singular Perturbation Approximation method. This method relies on retaining the dominant system modes and comparatively eliminating the less significant eigenvalues. Furthermore, through the proposed approach, the time and frequency response of the system has shown an excellent match when compared to that of other approaches in the literature. Kumari and Vishwakarma (2019) proposed an algorithm that employs the renovated pole clustering technique and numerator coefficients computed by factor division in order to reduce the order of large-scale linear dynamic MIMO systems. This method is also applicable to linear MIMO systems. Abdullah (2016) proposed a new modified PSO technique to reduce the order of large-scale linear SISO and MIMO systems based on the linear descending inertia weight strategy. The suggested modification of the PSO algorithm provided better characteristics in terms of global search capabilities and fast convergence speed, thus making it effective and efficient. Abu-Al-Nadi et al. (2013) proposed a model order reduction using an invasive weed optimization technique in a linear multivariable system. This technique was applied with the combined advantages of retaining the dominant poles and error minimization. The proposed algorithm showed good results, reducing the 10th order of the MIMO linear model in a practical power system to the 3rd order. Vishwakarma and Prasad (2009) suggested a new mixed method to reduce the order of large-scale linear dynamic MIMO systems. This method uses modified pole clustering to synthesize the common denominator polynomial of the reduced-order transfer function matrix while a GA computes the coefficients of the numerator elements by reducing the integral square error between the time responses and the system elements. The proposed algorithm was illustrated and compared against other reduction techniques, showing good stability when the original high-order system is stable.

In this manuscript, the MIMO PID control is optimized while adopting the ACO algorithm, and it is integrated with ITAE as a fitness function. Two scenarios are analyzed, *i.e.*, open and closed loops with and without the PID controller, aiming to achieve system robustness and reduce the order of the large-scale MIMO system.

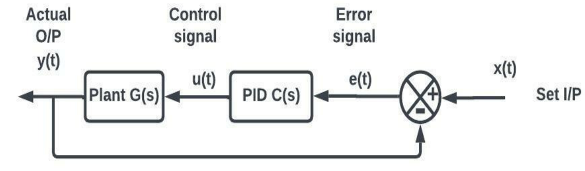
PID controller

The most straightforward and attractive solution in industrial control design implementation corresponds to Proportional-Integral-Derivative (PID) controllers, which do not have a complicated configuration and are very efficient. This, in addition to their competitive pricing and low maintenance expenses. Furthermore, through adequate parameter tuning, the dynamic response of a system can be improved by reducing the overshoot while reducing and/or eliminating the steady-state error, consequently enhancing the system's stability (Hanifah et al., 2013; Nagaraj & Murugananth, 2010; Xu, 2010).

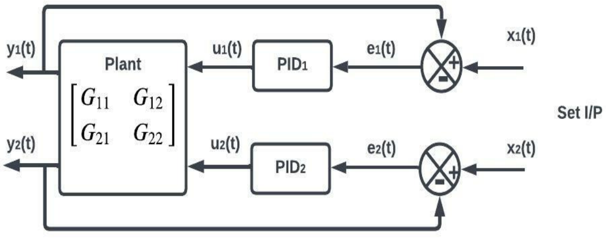
The purpose of tuning a PID controller is to ensure that its proportional, integral, and derivative factors are appropriate to the design, aiming to guarantee a closed-loop scheme implementation and a strong arrangement of the control loop over a broad spectrum of functions (Xu, 2010). It is worth mentioning that it is hard to adjust the PID controller without expertise. A manual optimum adjustment could be easy for experts because of the long times needed to obtain ideal operation results. In this regard, every factor in the PID controller could yield different outcomes. For instance, when adjusting to obtain a more satisfactory transient design response, the model response might be delayed under disruption circumstances (Hanifah et al., 2013; Nagaraj & Murugananth, 2010). Such a controller for SISO (Figure 1a) can be expressed mathematically in the time domain as in Prainetr et al. (2019); Oudah et al. (2021).

$$u(t) = k_p\, e(t) + k_i \int_0^t (t)dt + k_d \frac{d}{dt} e(t), \tag{1}$$

where k_p, k_i, and k_d represent the proportion gain factor, the integral gain factor, and the derivative gain, respectively, while e stands for the error and t denotes simultaneous time.



(a) Closed loop structure of SISO system



(b) Closed loop structure of MIMO system

Figure 1. Block diagram of SISO and MIMO controllers
Source: (Prainetr et al., 2019; Oudah et al., 2021)

In the s-domain, the transfer function of Equation (1) takes the form of Equation (2) (Saad et al., 2012).

$$C(s) = \frac{U(s)}{E(s)} = k_p + \frac{k_i}{s} + k_d s, \tag{2}$$

where $E(s)$, $U(s)$ represent the error and control signals applied to the plant model, respectively. For convenience, it is said that using two SISO schemes to construct a MIMO system is more applicable. Hence, this research adopts a heuristic strategy integrated with PID controller design and the ACO method for considering MIMO approaches. Figure 1b shows a schematic diagram configuration that includes the pair of controllers PID_1 and PID_2. Such a scheme yields the following expressions (Suresh et al., 2017):

$$\text{PID}_1 = k_{p_1} + \frac{k_{i_1}}{s} + k_{d_1} s, \tag{3a}$$

$$\text{PID}_2 = k_{p_2} + \frac{k_{i_2}}{s} + k_{d_2} s, \tag{3b}$$

Ant Colony Optimization (ACO)

Scientists have studied the social behavior of insects and animals and simulated them using algorithms. One such technique in the swarm intelligence family is Ant Colony Optimization (ACO) (Hanifah et al., 2013; Nagaraj & Murugananth, 2010; Xu, 2010). This approach has different applications, notably in industrial systems. It was developed by (1991, 1996; ?) and inspired by an empirical implementation conducted by Goss et al. (1989) to characterize ant behavior when it comes to using the shortest trajectories to travel over a binary bridge, which was designed so that one path was longer than the other. The insects were allowed to randomly select the path, with identical ant numbers in each branch. In his experiment, as time went by, it was found that the number of ants using the shortest trajectory was greater than those using the long lane. The reason for this is that, in order to return to the nest earlier, the ants must follow the short route. This, in turn, leads to an extra focus of the pheromone on the fastest time track (Goss et al., 1989).

Figure 2 (Hanifah et al., 2013) depicts the proposed PID controller with ACO. It is noteworthy that all the quantities of each factor were allocated to three different vectors. Moreover, it was assumed that the ants used all vectors as trajectories between the nests by selecting a lane between the terminations (Start-End).

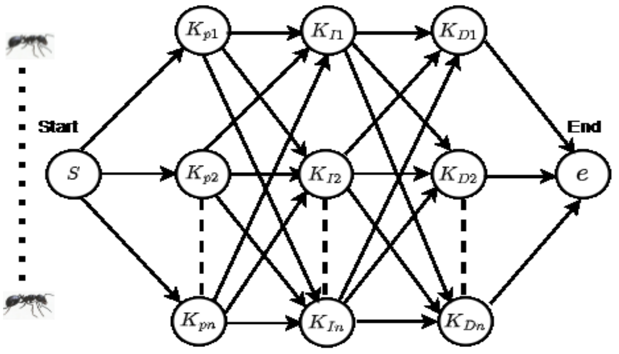


Figure 2. Pictorial description of PID tuning employing ACO for a three-parameter controller
Source: (Hanifah et al., 2013)

When using the ant algorithm, the pheromone is regularly updated at the start of each trajectory between the three nests, aiming to achieve the fastest round with the lowest expense. This work adopted ACO as the tuning approach for the PID controller. In this context, an arbitrary initial value is set as a pheromone for each junction point. The ants then arbitrarily decide the route they want to follow. Additionally, a number of ants are positioned at the start node and at every single iteration in order to determine their solution to the problem, which is known as a *trip*. Using pseudo-random proportional action choice rules, the probability P_{ij} in Equation (4) is computed as the estimated value between two nodes i, j according to two parameters: heuristic and metaheuristic (Prainetr et al., 2019).

$$P_{ij}^{k} = \frac{\tau_{ij}^{\alpha}(t).\,\eta_{ij}^{\beta}}{\sum\limits_{I\in\mathscr{R}_i^k} \tau_i^{\alpha}(I).\,\eta_{iI}^{\beta}}, \tag{4}$$

where τ_{ij} stands for the pheromone initial value and is updated regularly according to two parameters: α and β; and η represents the reciprocal of the span between the two nodes. The value ($\eta_{ij} = 1/d_{ij}$) directs the allocation of ants to the nodes, preventing them from being excessively far away. The impact of pheromone density on a certain route refers to the direction that the ant adopts, denoted by α, which is a positive value in Equation (4). It is noteworthy that, if $\alpha = 0$, the closest node will be selected every time, whereas, if $\beta = 0$, only the chemical factor (pheromone) tracks recreate on the selection (Prainetr et al., 2019; Chebli et al., 2018).

ITAE

ITAE stands for integrating the multiplication of time by an absolute error. The main feature of this approach is that it has acceptable damped oscillation with a fairly low overshoot value. The expression (5) defines such a method, while Figure 3 exhibits its block diagram.

$$\text{ITAE} = \int_0^\infty t\,|e(t)|\,dt, \tag{5}$$

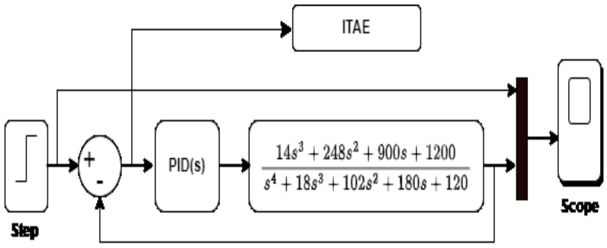


Figure 3. Simulation model showing ITAE with a PID controller
Source: (Prainetr et al., 2019; Chebli et al., 2018)

Methodology

In this section, the procedures of the ACO algorithm are presented in some detail with the aid of the flowchart shown in Figure 4 (Prainetr et al., 2019; Chebli et al., 2018).

According to this Figure, the first step is to assign the initial values of the number of ants, chemical factors, and control parameter constants, such as K_p, K_i, and K_d, using a random variable uniformly distributed over [0, 1]. Then, several ants are positioned at the node, and the values related to the multi-goals is calculated. Afterwards, the consecutive node is selected, and its probability is estimated via Equation (6).

$$P_{(rob)}{}^{q}_{ij}(t) = \begin{cases} \dfrac{\tau^{\alpha}_{ij}(t).\,\eta^{\beta}_{ij}}{\sum_{I\in\mathscr{R}^q_i}\tau^{\alpha}_{i}(I).\,\eta^{\beta}_{iI}}, & \text{if } i \in \mathscr{R}^q \\ 0, & \text{if } i \notin \mathscr{R}^q \end{cases} \tag{6}$$

where $j = [p,\ i,\ d]$ is investigatory, and $\mathscr{R}^q$ denotes path impacts of ant q within a specific interval. In the next step, the value of chemical factor $\Delta\tau^{q}_{ij}(t)$ for any iteration t on each route is determined via Equation (7).

$$\Delta\tau^{q}_{ij}(t) = \begin{cases} \frac{L^{min}}{L^{q}(t)}, & \text{if } i, j \in \mathscr{R}^q \\ 0, & \text{else,} \end{cases} \tag{7}$$

where $L^q(t)$ signifies the total length traveled by the ant q, and L^{min} represents the optimization obtained by a group of ants up to the current iteration.

Note that a high value of α might lead the formula to quickly converge to the sub-optimal route by increasing the

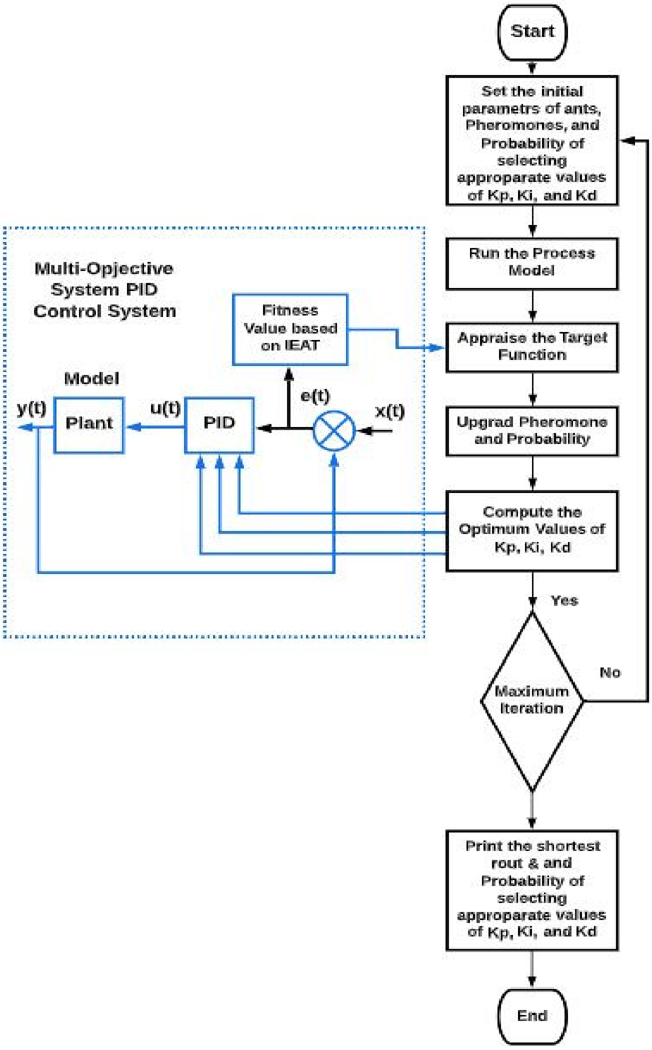


Figure 4. Flowchart for tuning the PID algorithm via ACO for SISO as a single branch of a MIMO scheme
Source: Authors

chemical factor, particularly regarding the initial arbitrary value. For the sake of convenience, the chemical factor concentration was allowed to evaporate in order to prevent early convergence. This, in order to avoid an unlimited increment in chemical factor trails. In other words, a reasonable trade-off between these two factors is required. Thus, the chemical factor is given by Equation (8).

$$\tau_{ij}(t) = \rho\tau_{ij}(t-1) + \sum_{k=1}^{m} \Delta\tau^{k}_{ij}(t), \tag{8}$$

where m denotes the amount of ants, t denotes the iteration, and ρ indicates the rate of chemical factor vaporization in the range of $0 < \rho \leq 1$. Afterwards, as soon as the steady-state zone of the PID controller is solved, the optimal quantities must be explored over a limited zone according to the multi-parameters.

It is well known that closed-loop schemes are commonly evaluated in terms of resilience, preciseness, reliability, and steady-state transition response, to name a few. In this vein, modern technologies demand additional advanced performance criteria, so the time at which errors occur must be considered. This issue has been widely explained in the literature.

Concerning the system's dynamic, this work considers a remarkable performance criterion known as the *Integrated Time Absolute Error standard* (ITAE). This index is widely used since it has higher selectivity when compared to other methods. It is less concentrated on the initial error and prioritizes the system error.

It is worth mentioning that this work was conducted using a Simulink model (MATLAB 2018b). This model entrusted with the the transfer function (TF), the ACO algorithm, and the remarkable performance criterion (ITAE) using codes of the M-file type. The PID controller was also implemented using the Simulink model.

Simulation results and discussions

This section presents and discusses the simulation results. This is done in two directions. The first one is to offer and discuss the results of the open-loop system, comparing them against to those of a closed-loop system with and without control (PID controller and ACO-PID controller). Figure 5 shows the step response of the system in an open loop and in a closed loop.

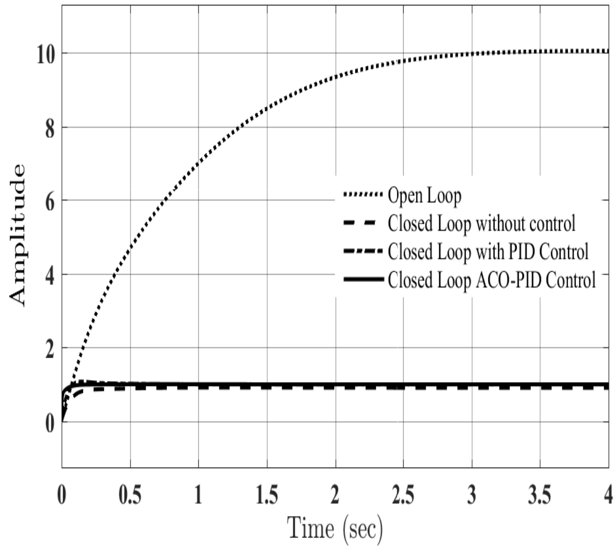


Figure 5. Step response of the system in open and closed loops (with and without control)
Source: Authors

The second direction concerns the results obtained for the system in a closed loop without control and in a closed loop with control (PID controller and ACO-PID controller) Figure 6 shows the step response of the system in a closed loop (with and without control). Of course, all types of the system referred to earlier include the cost function ITAE actions which help to reduce the large scale of a MIMO system.

According to the results shown in Figure 5, there is a large performance gap between the open- and closed-loop systems, as there is no optimization in the former, which is also less reliable. On the contrary, the closed-loop system is optimized and more reliable.

The results of the system without control (Figure 6) show the following: rising time = 0,147 s, overshoot = 0,446%, and undershoot = 1,98%. The parameters of the system with a closed-loop PID controller have K_p = 2, K_i = 12, and K_d = 0,001. Thus, the results are as follows: rising time = 0,0572 s, settling time = 0,418 s, overshoot = 7,27%, and undershoot = -0,003%. Furthermore, the results obtained for the closed-loop system with the ACO-PID controller, which works with K_p = 4,8, K_i = 9,5, and K_d = 0,11, are as follows: rising time = 0,0441 s, settling time = 0,115 s, and overshoot = 0,132%.

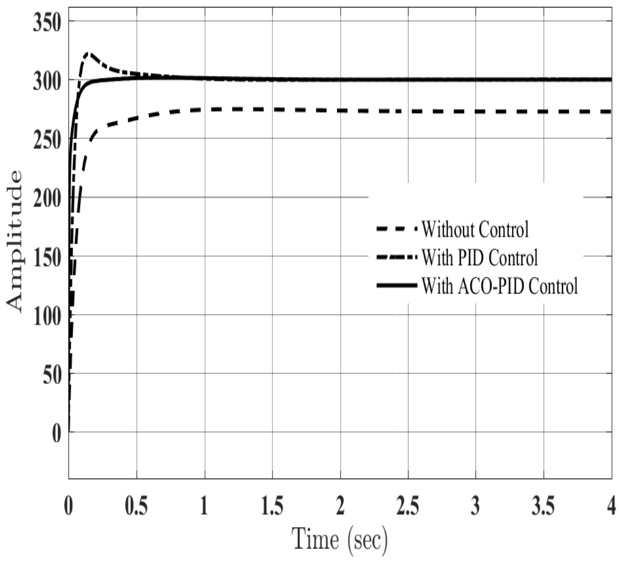


Figure 6. Step response of the closed-loop system with and without control (PID controller and ACO-PID controller)
Source: Authors

In light of the above, the closed-loop system with the ACO-PID controller responds quickly and is relatively stable (*i.e.*, there is a reduction in the large scale of the MIMO system) in comparison with the system with the PID controller and that without control.

Table 1 compares the response parameters of the proposed method based on step input regarding rising time (s), settling time (s), and overshoot (%) in four ways that have already been described in the literature (Goyal et al., 2019; Saraswat et al., 2015; Parmar et al., 2007; Prasad & Pal, 1991). It can be seen that the combination of ITAE and the ACO-PID controller reduces the response time, effectively reducing the large scale of MIMO techniques.

Conclusion

This work proposes a control method based on the combination of ITAE as a fitness function and ACO as an optimization algorithm with a PID controller. Using MATLAB 2018b, the m.file was dedicated to the transfer function (TF), the ACO algorithm, and the ITAE. Simulink was used to concstruct the PID controller. The desired response model was obtained by changing the three PID parameters (K_p, K_i, and K_d values). The method was validated by measuring time domain parameters such as rising time, settling time, and overshoot/undershoot. According to the simulation results for two scenarios (with and without control), it was concluded that improvements of about 30 and 29,5% can be achieved for response speed and relative stability, respectively. By comparing the results with those in the literature, it can be observed that our proposed approach can provide a remarkable and superior dynamic response. Therefore, it is inferred that this enhanced ratio can reduce the time response of large-scale MIMO systems.

Table 1. Comparison between the proposed method and four techniques introduced in the literature based on step input response parameters

Models	**Rising time (s)**	**Settling time (s)**	**Overshoot (%)**
4^{th} order (proposed method)	0,044 1	0,115	0,132
4^{th} order by Goyal *et al.*	2,26	3,93	0
2^{nd} order by Saraswat *et al.*	2,34	4,09	0
2^{nd} order by Parmar *et al.*	2,19	3,22	1,3
2^{nd} order by Prasad *et al.*	15,4	27,4	0

Source: Authors

CRediT author statement

All of the authors have significantly contributed to the research. Jafaar M. Daif-Alkhasraji was involved in the background research, method development, simulation results, drafting of the paper, and results discussion. Salam W. Shneen contributed to the method proposal, the simulation results, and the data collection and tabulation. Mohammed Q. Sulttan provided guidance during the entire research and the writing process. He supervised the research and provided critical feedback.

References

Abdullah, H. N. (2016). Reduction of large scale linear dynamic siso and mimo systems using modified particle swarm optimization algorithm. In *2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)* (pp. 166–171).

Abu-Al-Nadi, D. I., Alsmadi, O. M., Abo-Hammour, Z. S., Hawa, M. F., & Rahhal, J. S. (2013). Invasive weed optimization for model order reduction of linear mimo systems. *Applied Mathematical Modelling*, *37*(6), 4570–4577.

Al-doraiee, F., & Al-Qaraawi, H. M. S. H. J., Salih. (2013). Design and implementation of adaptive wavelet network pid controller for aqm in the tcp network. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, *13*(1), 9–17. https://ijccce.uotechnology.edu.iq/article_78386.html.

AL-Suhail, G. A., & Miry, M. H. (2015). Combining genetic algorithm and direction of arrival for mimo wireless communication system. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, *15*(2. https://ijccce.uotechnology.edu.iq/article_106699.html).

Aziz, G. A., Jaber, M. H., Sulttan, M. Q., & Shneen, S. W. (2022). Simulation model of enhancing performance of tcp/aqm networks by using matlab. *Journal of Engineering and Technological Sciences*, *54*(4. https://doi.org/10.5614/j.eng.technol.sci.2022.54.4.4).

Bizaki, H. K. (2011). *Mimo systems: Theory and applications.* BoD–Books on Demand. https://doi.org/10.5772/610.

Chebli, S., Elakkary, A., & Sefiani, N. (2018). Ant colony optimization based on pareto optimality: application to a congested router controlled by pid regulation. *Systems Science and Control Engineering*, *6*(1), 360–369. https://doi.org/10.1080/21642583.2018.1509395.

Dang, H. N., Nguyen, H. T., & Nguyen, T. V. (2021). Joint detection and decoding of mixed-adc large-scale mimo communication systems with protograph ldpc codes. *IEEE Access*, *9*, 101013–101029. https://doi.org/10.1109/ACCESS.2021.3097444.

Daraji, A., & Hale, J. (2014). Reduction of structural weight, costs and complexity of a control system in the active vibration reduction of flexible structures. *Smart Materials and Structures*, *23*(9), 095013. https://doi.org/10.1088/0964-1726/23/9/095013.

Dhanraj, A. V., & Nanjundappan, D. (2014). Design of optimized pi controller with ideal decoupler for a non linear multivariable system using particle swarm optimization technique. *International Journal of Innovative Computing, Information and Control*, *10*(1), 341–355. https://api.semanticscholar.org/CorpusID:16599826/.

Dorigo, M., Maniezzo, V., & Colorni, A. (1991). Positive feedback as a search strategy.

Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *26*(1), 29–41.

Goss, S., Aron, S., Deneubourg, J., & Pasteels, J. (1989). Self-organized shortcuts in the argentine ant. *Naturwissenschaften*, *76*(12), 579–581. https://doi.org/10.1007/BF00462870.

Goyal, R., Parmar, G., & Sikander, A. (2019). A new approach for simplification and control of linear time invariant systems. *Microsystem Technologies*, *25*, 599–607. https://doi.org/10.1007/s00542-018-4004-1.

Han, Y., Jin, S., Wen, C.-K., & Ma, X. (2020). Channel estimation for extremely large-scale massive mimo systems. *IEEE Wireless Communications Letters*, *9*(5), 633–637. https://doi.org/10.1109/LWC.2019.2963877.

Hanifah, R., Toha, S., & Ahmad, S. (2013). Pid-ant colony optimization (aco) control for electric power assist steering system for electric vehicle. In *2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)* (pp. 1–5. https://doi.org/10.1109/ICSIMA.2013.6717979).

Hassanzadeh, I., & Mobayen, S. (2011). Controller design for rotary inverted pendulum system using evolutionary algorithms. *Mathematical Problems in Engineering*, *2011*, 1-17. https://doi.org/10.1155/2011/572424.

Hu, W., Cai, W.-J., & Xiao, G. (2010). Decentralized control system design for mimo processes with integrators/differentiators. *Industrial and Engineering Chemistry Research*, *49*(24), 12521–12528.

Jaldén, J., & Ottersten, B. (2005). On the complexity of sphere decoding in digital communications. *IEEE Transactions on Signal Processing*, *53*(4), 1474–1484.

Juneja, M., & Nagar, S. (2015). Comparative study of model order reduction using combination of pso with conventional reduction techniques. In *2015 International Conference on Industrial Instrumentation and Control (ICIC)* (pp. 406–411).

Khudhair, A. A., Jabbar, S. Q., Sulttan, M. Q., & Wang, D. (2016). Wireless indoor localization systems and techniques: survey and comparative study. *Indonesian Journal of Electrical Engineering and Computer Science*, *3*(2), 392–409.

Kumari, A., & Vishwakarma, C. (2019). Order reduction of dynamic systems by using renovated pole clustering technique. In *2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC)* (pp. 532–537).

Labibi, B., Marquez, H. J., & Chen, T. (2009). Decentralized robust pi controller design for an industrial boiler. *Journal of Process Control*, *19*(2), 216–230.

Larsson, E. G. (2009). Mimo detection methods: How they work [lecture notes]. *IEEE Signal Processing Magazine*, *26*(3), 91–95.

Lengare, M., Chile, R., & Waghmare, L. M. (2012). Design of decentralized controllers for mimo processes. *Computers and Electrical Engineering*, *38*(1), 140–147.

Liu, G., Yang, J., & Whidborne, J. (2003). Multiobjective optimization and control. *Baldock, Hertfordshire,, England: Research Studies Press Limited*, *15*(2), 77–78.

Lones, M. A. (2014). Metaheuristics in nature-inspired algorithms. In *Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation* (pp. 1419–1422).

Lyu, S., Wen, J., Weng, J., & Ling, C. (2019). On low-complexity lattice reduction algorithms for large-scale mimo detection: The blessing of sequential reduction. *IEEE Transactions on Signal Processing*, *68*, 257–269.

Muhsen, A. N., & Raafat, S. M. (2021). Optimized pid control of quadrotor system using extremum seeking algorithm. *Engineering and Technology Journal*, *39*(6), 996–1010.

Nagaraj, B., & Murugananth, N. (2010). A comparative study of pid controller tuning using ga, ep, pso and aco. In *2010 international conference on communication control and computing technologies* (pp. 305–313).

Oudah, M. K., Sulttan, M. Q., & Shneen, S. W. (2021). Fuzzy type 1 pid controllers design for tcp/aqm wireless networks. *Indonesian Journal of Electrical Engineering and Computer Science*, *21*(1), 118–127.

Ouyang, C., & Yang, H. (2018). *Massive MIMO antenna selection: Asymptotic upper capacity bound and partial CSI.* https://doi.org/10.48550/arXiv.1812.06595.

Parmar, G., Prasad, R., & Mukherjee, S. (2007). Order reduction of linear dynamic systems using stability equation method and ga. *International Journal of Electrical and Computer Engineering*, *1*(2), 244–250.

Prainetr, S., Phurahong, T., Janprom, K., & Prainetr, N. (2019). Design tuning pid controller for temperature control using ant colony optimization. In *2019 ieee 2nd international conference on power and energy applications (icpea)* (pp. 124–127).

Prasad, R., & Pal, J. (1991). Stable reduction of linear systems by continued fractions. *Journal-Institution of Engineers India, Part El, Electrical Engineering Division*, *72*, 113–113.

Rajinikanth, V., & Latha, K. (2012). Setpoint weighted pid controller tuning for unstable system using heuristic algorithm. *Archives of Control Sciences*, *22*(4), 481–505.

Rusek, F., Persson, D., Lau, B. K., Larsson, E. G., Marzetta, T. L., Edfors, O., & Tufvesson, F. (2012). Scaling up mimo: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, *30*(1), 40–60.

Saad, M. S., Jamaluddin, H., & Darus, I. Z. (2012). Pid controller tuning using evolutionary algorithms. *Wseas Transactions on Systems and Control*, *7*(4), 139–149.

Saraswat, P., Parmar, G., & Bhatt, R. (2015). Application of differential evolution in order reduction of large scale systems. In *4th international conference,icatetr-2015, bkit kota* (pp. 124–127).

Saroja, K., Stefka Sharon, R., Meena, S., & Chitra, K. (2017). Multi-loop pid controller design for distillation column using firefly algorithm. *International Journal of Engineering and Technology (IJET)*, *9*(2), 1404–1410.

Sharma, P., Tiwari, R. N., Singh, P., Kumar, P., & Kanaujia, B. K. (2022). Mimo antennas: Design approaches, techniques and applications. *Sensors*, *22*(20), 7813.

Sivakumar, R., Rajinikanth, V., & Sankaran, D. (2015). Multi-loop pi controller design for tito system: An analysis with ba fa pso and bfo. *Australian Journal of Basic and Applied Sciences*, *9*(16), 249–254.

Sulttan, M. Q. (2016). A new approach of detection algorithm for reducing computation complexity of mimo systems. *Indonesian Journal of Electrical Engineering and Computer Science*, *1*(1), 159–167.

Sulttan, M. Q. (2019). Enhancement of k-best sphere detection algorithm performance in mimo systems. In *IOP Conference Series: Materials Science and Engineering* (Vol. 518, p. 052007).

Suman, S. K., & Kumar, A. (2020). Reduction of large-scale dynamical systems by extended balanced singular perturbation approximation. *International Journal of Mathematical, Engineering and Management Sciences*, *5*(5), 939.

Suresh, A., Meena, S., & Chitra, K. (2017). Controller design for mimo process using optimization algorithm. *International Journal of Pure and Applied Mathematics*, *117*, 163–170.

Vasu, G., Santosh, K., & Sandeep, G. (2012). Reduction of large scale linear dynamic siso and mimo systems using differential evolution optimization algorithm. In *2012 IEEE Students' Conference on Electrical, Electronics and Computer Science* (pp. 1–6).

Vishwakarma, C., & Prasad, R. (2009). Mimo system reduction using modified pole clustering and genetic algorithm. *Modelling and Simulation in Engineering*, *2009*, 1–5.

Vu, T. N. L., Lee, J.-T., & Lee, M.-Y. (2007). Design of multi-loop pid controllers based on the generalized imc-pid method with mp criterion. *International Journal of Control, Automation, and Systems*, *5*(2), 212–217.

Xu, Z. (2010). A novel robust pid controller design method. In *2010 international conference on computer application and system modeling (iccasm 2010)* (Vol. 6, pp. V6–332).

Yang, X., Cao, F., Matthaiou, M., & Jin, S. (2020). On the uplink transmission of extra-large scale massive mimo systems. *IEEE Transactions on Vehicular Technology*, *69*(12), 15229–15243.

Yang, X.-S. (2009). Firefly algorithms for multimodal optimization. In *International Symposium on Stochastic Algorithms* (pp. 169–178).

http://doi.org/10.15446/ing.investig.103210

Insulation system diagnosis in power transformers using DGA analysis and Megger DC tests

Diagnóstico de sistemas de aislamiento en transformadores de potencia utilizando análisis DGA y pruebas Megger DC

Juan S. Juris[1], Iván C. Durán-Tovar[2], Josimar Tello-Maita[3], and Agustín Marulanda-Guerra[4]

ABSTRACT

Dissolved gas chromatography (DGA) analysis and Megger DC insulation tests are performed to diagnose the condition of a power transformer. In this context, the objective of this study is to assess whether there is a relationship between both types of tests. To this effect, a database with DGA and Megger DC test protocols is analyzed, using the theory of variable correlation as well as current DGA techniques. The results allow stating that there is indeed a relationship between both tests under specific conditions. Therefore, the DC isolation status of a transformer can be estimated via DGA tests.

Keywords: insulation, correlation, chromatography, maintenance

RESUMEN

Los analisis por cromatografıa de gases disueltos (DGA) y las pruebas de aislamiento Megger DC son realizados para diagnosticar el estado de un transformador de potencia. En este contexto, el objetivo de este estudio es evaluar si existe una relacion entre ambos tipos de pruebas. Para ello, se analiza una base de datos con protocolos de pruebas DGA y Megger DC, utilizando la teorıa de correlacion de variables y las tecnicas actuales de DGA. Los resultados permiten afirmar que sı existe una relacion entre ambas pruebas bajo condiciones especıficas. Por lo tanto, se puede estimar el estado de aislamiento DC de un transformador a partir de pruebas DGA.

Palabras clave: aislamiento, correlacion, cromatografıa, mantenimiento

Received: January 16th 2022
Accepted: January 28th 2024

Introduction

Power systems, as well as industrial sectors, are made up of a variety of electrical equipment that enables the flow of energy at adequate voltage levels for various applications. There is an indispensable piece of equipment responsible for supplying said voltage levels: the power transformer. Ensuring the continuous operation of this equipment is essential for the reliability of any electrical system, so timely maintenance allows foreseeing any failures with a negative impact in both technical and economic aspects.

The insulation of a transformer is exposed to the climatic conditions of its work environment, as well as to its operating characteristics. In general, insulation is a key component that deteriorates throughout the useful life of the equipment. This is where insulation evaluations like the Megger test take on importance, but they require a de-energized transformer to be executed. In the Megger DC test, a DC voltage is applied for 10 minutes to determine the static resistance of the insulation. On the other hand, dissolved gas chromatography (DGA) tests have become important in establishing the dielectric condition of the oil and detecting failures in a transformer from the combustible gases present in it, with the main advantage that they can be performed with energized equipment.

Problem statement

A power transformer's likelihood of failure depends on the behavior of its subsystems, which include the insulation system (IEEE Standards Association, 2017). The insulation of a transformer can be diagnosed through a set of electric tests, among them the Megger DC resistance test (Torkaman and Karimi, 2015), the dissipation factor tan-δ test (Malpure and Baburao, 2008), the power factor tip-up test (IEEE Standards Association, 2000), the dielectric frequency response test (IEC, 2006), and frequency response analysis (Sardar *et al.*, 2017). These tests require the transformer to be off-line, which leads to economic impacts for the owner.

---

[1] Electrical engineer, Universidad Escuela Colombiana de Ingeniería Julio Garavito, Colombia. MSc in Electrical Engineering, Universidad Escuela Colombiana de Ingeniería Julio Garavito, Colombia. E-mail: juan.juris@mail.escuelaing.edu.co

[2] Electrical engineer, Universidad Nacional de Colombia, Colombia. MSc in Electrical Engineering, Universidad de los Andes, Colombia. PhD in Electrical Engineering, Universidad Nacional de Colombia, Colombia.
Affiliation: Associate professor, Universidad Escuela Colombiana de Ingeniería Julio Garavito, Colombia. E-mail: ivan.duran@escuelaing.edu.co

[3] Electrical engineer, Universidad del Zulia, Venezuela. MSc in Applied Mathematics, Universidad del Zulia, Venezuela. PhD in Engineering, Universidad del Zulia, Venezuela. Affiliation: Assistant professsor, Universidad Escuela Colombiana de Ingeniería Julio Garavito, Colombia. E-mail: josimar.tello@escuelaing.edu.co

[4] Electrical engineer, Universidad del Zulia, Venezuela. PhD in Electrical Engineering, Universidad de Sevilla, España. Affiliation: Associate professsor, Universidad Escuela Colombiana de Ingeniería Julio Garavito, Colombia. E-mail: agustin.marulanda@escuelaing.edu.co

On the other hand, a DGA test can be performed with an online transformer and allows detecting the presence of different gases associated with its condition (IEEE Standards Association, 2019; IEC, 2015). Thus, this test might allow identifying the condition of a transformer's insulation based on statistical inference from a set of samples taken from a group of power transformers. This can help reduce the need to frequently disconnect transformers in order to perform insulation tests, as well as its economic impact.

Contributions

The main contribution of this paper lies in the existence of a correlation between two states of the DGA test and the Megger DC resistance test (regular and bad states). This finding was obtained from a statistical treatment applied to a set of real data on these two types of tests from 337 industrial power transformers. This encourages the application of the DGA test to predict potential insulation damages and reduce the need for off-line tests and their economic impact.

State of the art

In the area of dielectric oils chromatography, the interpretation of the gases present in a sample allows diagnosing faults in power transformers, such as the degradation of their insulation system (Díaz and Schmidt, 2014; Dhini *et al.*, 2020). The analysis consists of taking a dielectric oil sample and testing it in an accredited laboratory with a calibrated chromatograph, in order to obtain quantitative results regarding its gas composition, which are expressed in parts per million (ppm) and as a percentage (%).

Currently, the focus of DGA analysis consists of characterizing the behavior of combustible gases in the transformer with the aim of estimating the possible trends throughout its lifespan (Dukarm, 2019), in addition to estimating any internal failures of the equipment which may be derived from the aging of its insulation system (oil and paper) due to both thermal and electrical factors (Mahmoudi *et al.*, 2019). Carrying out DGA analysis on a regular basis allows observing the evolution of gases in the transformer, but there is no clear relationship between electrical failures and gas trends. Failures are usually determined via preventive electrical tests that directly evaluate transformer components (Fofana and Hadjadj, 2016) but require the disconnection of the equipment. However, there are four main DGA analysis techniques that are complementary and have the potential to detect failures in equipment components without affecting operation continuity (Juris *et al.*, 2020).

Recent works have combined current DGA analysis techniques with more complex modeling and probabilistic tools to improve status monitoring in power transformers. Some works use tools such as fuzzy logic and artificial neural networks (Prasojo *et al.*, 2020; Aciu *et al.*, 2021; Patekar and Chaudhry, 2019; Saravanan *et al.*, 2020), which allow for fault prediction and a better interpretation of the states of the transformer based on gas levels. On the other hand, the work by Aizpurua *et al.* (2019) focuses on the combination of probabilistic methods and soft computing to improve health monitoring in transformers under conditions of uncertainty.

Document structure

This article is organized as follows: the *Introduction* section presents the problem, a review of the state of the art and the contributions of our developed work; the section titled *DGA techniques* describes the theory associated with current DGA techniques for dielectric oil in transformers; the *Methodology* section presents the two models applied, the first of which is of the statistical type and focuses on the behavior of trends in fuel gases in addition to applying the theory of correlation of variables, while the second is a technical model to characterize the state of dielectric oil in transformers by applying the current techniques of DGA tests; in the *Results* section, the results obtained with the proposed models in their corresponding case studies are presents; and the *Conclusions* section draws the main conclusions of this study.

DGA techniques

The measurement of dissolved gases must be accompanied by appropriate interpretation, as provided by the DGA techniques, which aim to determine the relationship between the concentration of combustible gases and the presence of thermal or electrical faults in a power transformer. This section describes the current DGA techniques that are also recommended as fault identification methods in IEEE Std. C57.104-2019 (IEEE Standards Association, 2019).

Doernenburg analysis

Doernenburg analysis, or the Doernenburg ratios method, is employed to determine whether the origin of a transformer failure is thermal or electrical in nature, upon the basis of a group of relationships between the concentrations of different gases present in the transformer oil. This technique considers the concentration of Methane (CH_4), Hydrogen (H_2), Acetylene (C_2H_2), Ethane (C_2H_6), and Ethylene (C_2H_4). The Doernenburg ratios are calculated as follows:

$$R_1 = \frac{CH_4}{H_2} \tag{1}$$

$$R_2 = \frac{C_2H_2}{C_2H_4} \tag{2}$$

$$R_3 = \frac{C_2H_2}{CH_4} \tag{3}$$

$$R_4 = \frac{C_2H_6}{C_2H_2} \tag{4}$$

According to this interpretation method, the failures can be as follows: (a) thermal failures, which can occur when the transformer load exceeds the rated capacity over long periods of time, accelerating the natural rate of oil decomposition, generally at temperatures between 125 and 600 ^oC; (b) Corona effect failures, which is a low-energy partial discharge that occurs when dielectric oil is ionized, likely evolving into discharges in the cellulose of the insulating paper; and (c) internal arc failures, which is a short-duration, high-intensity discharge that can compromise the solid insulation of the transformer, where the oil has started its deterioration process (Piegar *et al.*, 2015). Table 1 shows the conditions of the Doernenburg ratios for each type of fault.

Table 1. Doernenburg fault diagnosis

Possible fault	R1	R2	R3	R4
Thermal decomposition (a)	>0,1	<0,75	<0,3	>0,4
Corona effect (b)	<0,1	–	<0,3	>0,4
Internal arcing (c)	0,1 to 1,0	>0,75	>0,3	<0,4

Source: (IEEE Standards Association, 2019)

Key gas analysis

This technique analyzes the individual concentration of gases. According to the temperature variations to which the oil is subjected, it is possible to find higher concentrations of certain gases. For instance, when there is an electrical failure, the temperature rises inside the transformer tank, and high concentrations of combustible gases are liberated. The key gas technique yields an estimate of gas generation for different temperature ranges in the transformer tank (IEEE Standards Association, 2008; Wannapring *et al.*, 2016).

According to Rogers (1978) and IEEE Standards Association (2008), the key gases analyzed with this method and their possible fault diagnoses are as follows:

Ethylene. When the concentration of this gas exceeds 63%, failure due to overheating in the oil is suspected, as it has lost its cooling and insulating properties (*i.e.*, thermal mineral oil failure). Typically, this type of failure is associated with a predominant concentration of ethylene and smaller proportions of ethane, methane, and hydrogen.

Carbon monoxide. If the concentration of this gas exceeds 92%, overheating is suspected in the cellulose due to fluctuations in the transformer's operating temperature, which is derived from load changes (*i.e.*, thermal mineral oil and cellulose fault).

Hydrogen. If the hydrogen concentration exceeds 86%, failure by corona effect is suspected. Low-energy partial discharges predominantly generate hydrogen, with small amounts of methane and traces of ethylene and ethane.

Hydrogen and acetylene. If the predominant gases are hydrogen and acetylene, exceeding concentrations of 60 and 30%, respectively, the fault is most likely of the electric arc type. In this case, the oil undergoes a breakdown in its structure and tends to evaporate in small quantities.

Rogers analysis

Rogers analysis, or the Rogers ratios method, evaluates gas concentration ratios to diagnose faults while considering the transformer tank's operation temperature. This technique evaluates ratios from Equations (1) and (2). Moreover, it incorporates the ratio between the concentrations of ethylene and acetylene (5) (Syafruddin and Nugroho, 2020).

$$R_5 = \frac{C_2H_4}{C_2H_6} \tag{5}$$

This fault identification technique that the temperature failures in the transformer tank are independent of electrical failures such as partial discharges (Wang, 2003). Thus, it allows identifying five types of failure, namely low-temperature thermal failure, which can be caused by the inrush current or high charging periods; thermal failure below 700 °C, typical of equipment with long periods of operation; thermal failure over 700 °C, which can be caused by the loss of the oil's cooling capacity; low-energy density arcing; partial discharges (Corona effect); and high-energy arcing, which can be caused by high energy discharges.

Table 2 presents the Rogers ratios for each type of fault. The limitation of this technique is that there can be many cases where the ratios do not meet the conditions established by the method, so the failure cannot be identified.

Table 2. Rogers fault diagnosis

Possible fault	R1	R2	R5
No failure	<0,1	0,1 to 1	>1
Low-temperature thermal	<0,1	0,1 to 1	1 to 3
Thermal fault < 700^oC	<0,1	>1	1 to 3
Thermal fault > 700^oC	<0,1	>1	<1
Corona effect	<0,1	<0,1	<1
Electrical arcing	0,1 to 3	0,1 to 1	>3

Source: (Sarria-Arias *et al.*, 2014)

Duval analysis

The Duval triangle analysis may be the most complete fault interpretation technique, as it allows identifying the six basic types of faults established in IEEE Standards Association (2019), plus mixtures of electrical and thermal faults. The Duval equations shown in (6)-(8) determine the percentages of methane (X), ethylene (Y), and acetylene (Z) over the sum of the three concentrations. These gases correspond to the increasing energy content or temperature faults: methane for low energy/temperature faults, ethylene for high energy/temperature faults, and acetylene for very high/temperature/arcing faults.

$$\%C_2H_2 = \frac{X}{X+Y+Z} \times 100 \tag{6}$$

$$\%CH_4 = \frac{Y}{X+Y+Z} \times 100 \tag{7}$$

$$\%C_2H_4 = \frac{Z}{X+Y+Z} \times 100 \tag{8}$$

Graphically, the concentrations of the gases shown above correspond to the coordinate axes of an equilateral triangle, and the intersection of lines parallel to the axes allows finding the zone of possible failure in a transformer. Figure 1 shows the Duval triangle and its different fault zones, and Table 3 presents their corresponding diagnoses.

The mixed fault zone (DT) allows classifying both thermal and electrical failures. This is a factor related to equipment that operates for long periods of time, as characteristic gases will necessarily be generated for both types of failure.

DGA tests

Dissolved gas chromatography testing in a power transformer consists of systematically calculating the concentrations of the combustible gases present in a sample of dielectric oil. Its results are quantitative and are expressed in parts per million (ppm) and as a percentage (%).

The results obtained for each of the combustible gases are compared with the ranges stipulated by Table 1 of the IEEE

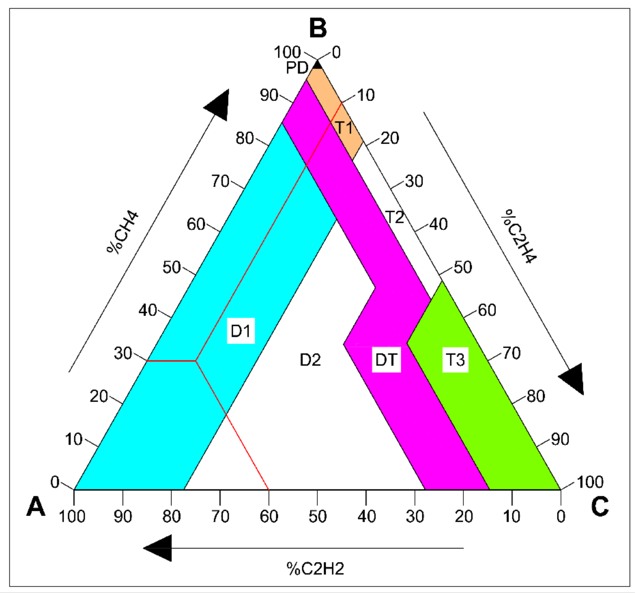


Figure 1. Duval triangle for fault diagnosis
Source: (Sarria-Arias *et al.*, 2014)

Table 3. Duval fault diagnosis

Zone	Possible fault	Gas limit values
PD	Partial discharge	$CH_4 = 98$
D1	High-energy discharge	$C_2H_4 = 36$; $C_2H_2 = 13$
D2	Low-energy discharge	$C_2H_4 = 23; 40$ $C_2H_2 = 13; 29$
T1	Thermal fault $< 300^oC$	$CH_4 = 98$; $C_2H_4 = 20$ $C_2H_2 = 4$
T2	Thermal fault 300^oC to 700^oC	$C_2H_4 = 20$; 50; $C_2H_2 = 4$
T3	Thermal failure $> 700^oC$	$C_2H_4 = 50$; $C_2H_2 = 15$
DT	Mix of thermal/electrical faults	$C_2H_4 = 13$; 4; 29; $C_2H_2 = 40$; 50

Source: (Sarria-Arias *et al.*, 2014)

Std. C57.104-2008 (*Dissolved gas concentrations*) (IEEE Standards Association, 2019). These ranges in ppm have been classified and are presented in Table 4.

Table 4. Transformer conditions and ppm gas ranges for the separate gases method

Gas	1 (Normal)	2 (Regular)	3 (Bad)	4 (Dangerous)
H_2	100	101 – 700	701–1 800	>1 800
CH_4	120	121 – 400	401–1 000	>1 000
CO	350	351 – 570	571–1 400	>1 400
C_2H_4	50	51 – 100	101 – 200	>200
C_2H_6	65	66 – 100	101 – 150	>150
C_2H_2	1	2 – 9	10 – 35	> 35

Source: (IEEE Standards Association, 2019)

According to Table 4 and the criteria established by the IEEE, condition 1 represents a transformer in a normal operating state, condition 2 denotes an equipment that must be analyzed using at least one of the DGA techniques, condition 3 implies failures in the oil (all the DGA techniques must be applied), and condition 4 represents a high probability that the oil is undergoing decomposition, so the transformer must be intervened by the manufacturer and be temporarily or permanently withdrawn from operation.

Megger DC resistance test

In these tests, DC voltage is applied under different configurations to the primary and secondary windings of the transformer. This, in order to observe the behavior of the insulation for 10 min. Good insulation should increase as time increases. The settings for the test are primary winding *vs.* ground, primary winding *vs.* secondary winding + ground, and secondary winding *vs.* ground.

The most important parameter in this test is the polarization index (*PI*), which is calculated using Equation (9). The PI is a dimensionless qualifying factor of the condition of the insulation system for the conditions of temperature and humidity under which a transformer operates.

$$PI = \frac{R_{insulation-10minutes}}{R_{insulation-1minute}} \tag{9}$$

Section 7.2.13.4 of IEEE Std. C57.152-2013 (*Polarization index test*) (IEEE Standards Association, 2013) qualifies PI conditions according to Table 5.

Table 5. Evaluation of the polarization index

IP	Condition	Criteria
≥ 2	1 (Excellent)	New transformer
$1,25 \leq IP < 2$	2 (Good)	Normal operation
$1,1 \leq IP < 1,25$	3 (Regular)	Under surveillance
$1 < IP < 1,1$	4 (Bad)	Corrective maintenance
< 1	5 (Dangerous)	Factory repair or decommissioning

Source: (IEEE Standards Association, 2013)

Methodology

The methodology consisted of applying two analyses. The first one was a comparative statistical analysis using the correlation between two conditions of the DGA test and the Megger DC resistance test, and the second one compared the DGA techniques in order to identify advantages and disadvantages for the detection of thermal and electrical faults. These analyses were applied to 337 mineral oil-filled industrial power transformers, whose main characteristics are shown in Table 6.

Table 6. Characteristics of the studied transformers

Power range [kVA]	Primary voltage range [kV]	Secondary voltage range [kV]	Manufacturing year range
75 - 22 000	0,480 - 115	0,215 - 110	1996 - 2016

Source: Authors

Figure 2 describes the methodology from the acquisition of the database to the results. Initially, the information in the database was organized according to the type of test (DGA or Megger DC resistance) and its results. Once the data were classified, they were separated for use in the proposed study cases.

At this stage, it was verified whether each transformer had a DGA test and a Megger DC test. The tests were classified according to the conditions established in the IEEE standards C57.104-2008 (IEEE Standards Association, 2019) and C57.152-2013 (IEEE Standards Association, 2013). The transformers with DGA tests in condition 4 and Megger DC tests in condition 5 were discarded, as these conditions imply a high risk of failure and are dangerous for the operation.

For the correlation analysis, two cases were studied. The first case (Case I) aimed at determining the correlation between conditions 2 (regular) and 3 (bad) in the DGA analysis and conditions 3 (regular) and 4 (bad) in the Megger DC tests. This, in order to determine the relationships between the states that can lead to future failures in the transformers, *i.e.*, the probability that a transformer with regular or bad conditions in the DGA analysis has the same conditions in the Megger DC tests.

The second case (Case II) sought to find the correlation between the tests exhibiting condition 1 (normal) in the DGA analysis and conditions 3 and 4 in the Megger DC test, aiming to evaluate the probability that one transformer with a normal DGA condition could be in a regular or bad state according to the Megger DC test.

To determine the correlation, Pearson's coefficient (r) was employed, which is a quantifying value of the linear relationship between two variables. This coefficient can take values from -1 to 1, with the following characteristics: a value of -1 indicates an inverse (negative) proportional relationship, a value of 1 indicates a direct (positive) proportional relationship, and a value of 0 indicates that there is no relationship between the variables.

The Pearson coefficients obtained for Case I were subjected to a hypothesis test in order to verify their significance. The theory of hypothesis testing for small samples was used to determine whether the coefficient was significantly different from 0. This, by means of the t-student distribution (Walpole *et al.*, 2017; Devore, 2016). This test confirms whether the DGA states and the Megger DC conditions are actually related or only show a relationship because of chance (Devore, 2016). The selected confidence level was 95%.

Finally, considering the 337 power transformers, a technical comparison between the DGA techniques was performed. This analysis encompassed the application of the aforementioned DGA techniques to each transformer in order to explore different types of failures as well as matches between techniques. Furthermore, the results were contrasted with the correlations obtained in previous study cases.

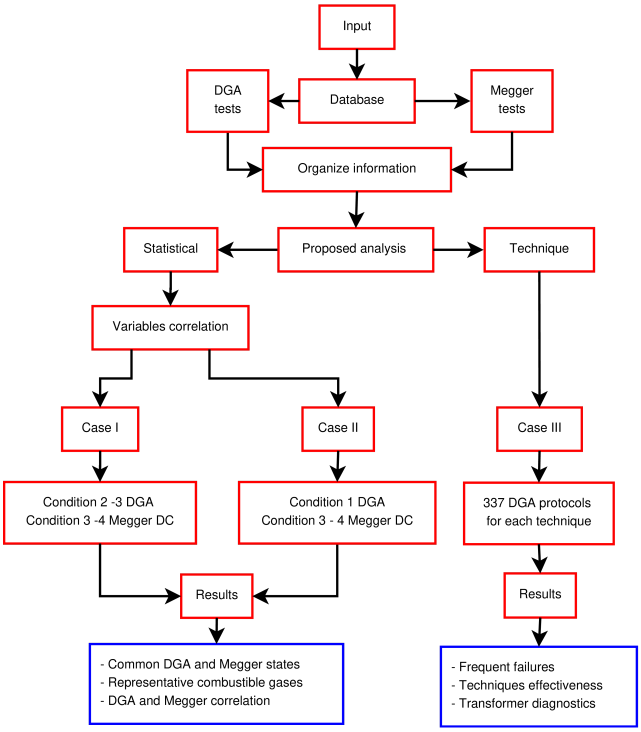


Figure 2. Methodology flow chart
Source: Authors

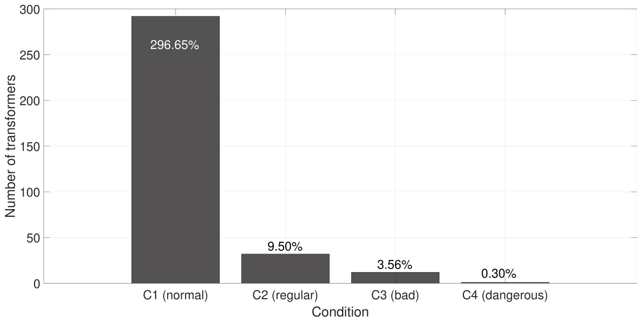


Figure 3. DGA tests classification
Source: Authors

Results

According to the classifications in the standards for the DGA and Megger tests, the results in the DGA tests place 86,65% of the transformers in normal conditions, 9,50% in regular conditions, 3,56% in a bad condition, and only 0,30% in dangerous conditions. These data can be seen in Figure 3, which shows the distributions of the DGA tests.

As for the Megger DC resistance tests, 10,68% of the transformers were found to be in excellent conditions, 63,80% in normal conditions, 10,39% in regular conditions, 2,97% in bad conditions, and 0% in dangerous conditions. 12,17% of the transformers fall into the category dubbed *No protocol*, which refers to the transformers that did not undergo a Megger DC resistance test, since it was not possible to de-energize them. These transformers were excluded from all study cases. Figure 4 shows the Megger DC resistance test results according to the conditions detected.

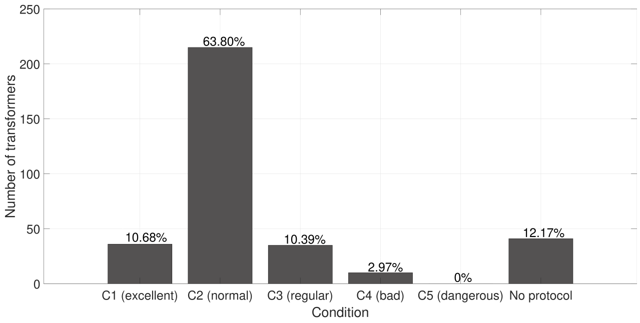


Figure 4. Megger DC test classification
Source: Authors

Analysis of Case I

For this case, out of the 337 transformers, a total of 44 DGA tests reported conditions 2 (regular) or 3 (bad), while, regarding the Megger DC tests, there was a total of 45 transformers with conditions 3 (regular) or 4 (bad). Out of these, there were only 38 transformers that had one of the DGA conditions and one of the Megger DC test states in common. Table 7 shows the conditions of these 38 power transformers.

To calculate the correlation coefficient, Equations (10) and (11) were used, where X represents the transformer's DGA test condition, Y is its Megger DC test state, and $\overline{X}$ and $\overline{Y}$ denote the averages of each condition, which were $2,2894$ and $3,2632$ respectively. Table 8 presents the results of each term in Equation (10).

$$covariance = \frac{\sum((X - \overline{X}) \times (Y - \overline{Y}))}{n-1} \tag{10}$$

$$r = \frac{covariance}{S_x \times S_Y} \tag{11}$$

Table 7. DGA and Megger DC condition values for Case I

Transformer ID	DGA (X)	Megger DC (Y)	Transformer ID	DGA (X)	Megger DC (Y)
007	3	4	217	2	3
023	2	3	218	2	3
031	2	3	219	3	4
035	2	3	220	3	4
049	2	3	224	3	4
069	2	3	237	3	4
070	2	3	238	2	3
101	2	3	239	2	3
110	2	3	242	2	3
115	2	3	243	2	3
137	2	3	244	3	3
152	3	4	245	3	3
156	3	4	270	2	3
186	2	3	276	2	3
190	2	4	287	2	3
203	3	4	297	2	3
204	3	4	304	2	3
211	2	3	332	2	3
212	2	3	333	2	3

Source: Authors

The results obtained show a covariance of $0,1650$ and a correlation coefficient of $0,8045$. The latter indicates that, if a transformer is in condition 2 (regular) or 3 (bad) in the DGA tests, then it can be said that its insulation is at least in condition 3 (regular) or in condition 4 (bad) of the Megger DC test, with a correlation of 80,45%.

The correlation coefficient obtained was subjected to the hypothesis test described in Table 9, with H_0 being the null hypothesis and H_1 the alternative hypothesis. A t-test based on Student's t-distribution with $N-2$ degrees of freedom and a standard deviation calculated with Equation (12) was applied to evaluate the hypotheses by means of the t_0 statistic (Equation (13)).

$$S_r = \sqrt{\frac{1-r^2}{N-2}} \tag{12}$$

Table 8. Transformer correlation coefficient for Case I

Transformer ID	$X-\overline{X}$	$Y-\overline{Y}$	$(X-\overline{X}) \times (Y-\overline{Y})$
007	0,7105	0,7368	0,5235
023	-0,2895	-0,2632	0,0762
031	-0,2895	-0,2632	0,0762
035	-0,2895	-0,2632	0,0762
049	-0,2895	-0,2632	0,0762
069	-0,2895	-0,2632	0,0762
070	-0,2895	-0,2632	0,0762
101	-0,2895	-0,2632	0,0762
110	-0,2895	-0,2632	0,0762
115	-0,2895	-0,2632	0,0762
137	-0,2895	-0,2632	0,0762
152	0,7105	0,7368	0,5235
156	60,7105	0,7368	0,5235
186	-0,2895	-0,2632	0,0762
190	-0,2895	-0,2632	0,0762
203	0,7105	0,7368	0,5235
204	0,7105	0,7368	0,5235
211	-0,2895	-0,2632	0,0762
212	-0,2895	-0,2632	0,0762
217	-0,2895	-0,2632	0,0762
218	-0,2895	-0,2632	0,0762
219	0,7105	0,7368	0,5235
220	0,7105	0,7368	0,5235
224	0,7105	0,7368	0,5235
237	0,7105	0,7368	0,5235
238	-0,2895	-0,2632	0,0762
239	-0,2895	-0,2632	0,0762
242	-0,2895	-0,2632	0,0762
243	-0,2895	-0,2632	0,0762
244	0,7105	0,7368	0,5235
245	0,7105	0,7368	0,5235
270	-0,2895	-0,2632	0,0762
276	-0,2895	-0,2632	0,0762
287	-0,2895	-0,2632	0,0762
297	-0,2895	-0,2632	0,0762
304	-0,2895	-0,2632	0,0762
332	-0,2895	-0,2632	0,0762
333	-0,2895	-0,2632	0,0762

Source: Authors

Table 9. Hypotheses raised to asses the significance of the correlation coefficient

Hypotheses raised	
$H_0 : r = 0$ indicates that the coefficient obtained comes from a population whose correlation is 0	$H_1 : r \neq 0$ indicates that the coefficient obtained comes from a population whose correlation is different from 0

Source: Authors

$$t_0 = \frac{r}{\sqrt{\frac{1-r^2}{N-2}}} \tag{13}$$

The t-test states that, if $t_0 > t_{(\alpha/2,N-2)}$, the null hypothesis is rejected, which means that the correlation coefficient of the population is different from 0, *i.e.*, there is a relationship between the variables. On the other hand, if $t_0 \leq t_{(\alpha/2,N-2)}$, the null hypothesis is accepted, meaning that the correlation of the population is equal to 0.

In the current study case, $t_0 = 8,1269$, and the selected significance level was $\alpha = 0,05$, which gives $t_{(\alpha/2,N-2)} = 2,0281$. These results lead to the rejection of the null hypothesis, which means that the sample comes from a population where the correlation is not equal to 0, with a confidence level of 95%.

Analysis of Case II

In this case, the number of transformers with a DGA test condition 1 (normal) was 292, and the transformers with the Megger DC resistance test conditions 3 and 4 were the same 45 of Case I. When comparing the conditions for these transformers, it was found that only seven of them share the conditions described above.

In these transformers, the PI meets two out of the three configurations of the Megger DC test, and the one that did not comply with the standard exhibits a value close to that required. In these protocols, the behavior of Case I was not observed; three or two configurations of the Megger DC test were not met, and their results were far from the required limit.

By applying correlation theory to the data in Table 10, it is observed that $\overline{X} = 1$ and $\overline{Y} = 1$. The covariance, as well as the standard deviations S_X and S_Y, have values equal to 0, so the correlation coefficient is an indeterminacy.

Table 10. DGA and Megger DC condition values for Case II

Transformer ID	DGA (X)	Megger DC (Y)
135	1	3
148	1	3
158	1	3
197	1	3
235	1	3
236	1	3
323	1	3

Source: Authors

Table 11. Total combustible gases (TDCG) conditions for Case II

Transformer ID	TDCG (ppm)	Limit allowed by the standard
135	193	< 720 ppm
148	367	
158	205	
197	65	
235	384	
236	377	
323	134	

Source: Authors

When a total combustible gases (TDCG) analysis was conducted in these seven transformers, it was found that the highest concentration was 384 ppm and the lowest one was 65 ppm (Table 11). Both values match the normal conditions of the IEEE Std. C57.104-2008 (IEEE Standards Association, 2019), with 0-720 ppm being the allowed range.

The irregular insulation results can be attributed to climatic conditions when the Megger DC tests were run, as a high humidity directly affects the measurement. For this case, it is not possible to apply variable correlation theory, since the DGA and Megger states are the same for all tests. A larger sample would be needed to identify new relationships between the states analyzed.

Analysis of Case III

When the aforementioned DGA techniques were applied to the transformers, it was found that most of them showed no identifiable faults. The majority of the faults found via the Doernenburg, key gas, and Rogers methods were thermal in nature, whereas the Duval technique reported a higher percentage of low-energy partial discharge electrical failures.

In the Doernenburg analysis, 92,88% of the transformers showed no faults. Only 21 transformers reported failures, which were classified according to the Doernenburg ratios (Table 1). 11 of these could not be identified, since at least one of the ratios was indeterminate, implying the need to rerun the tests. As for the rest, eight faults were identified as thermal and two as internal arc failures (Figure 5).

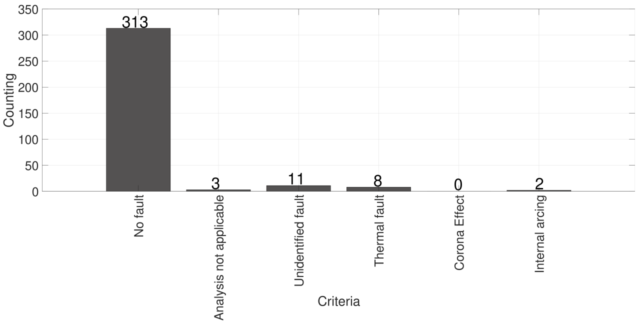


Figure 5. Results for the transformers analyzed via the Doernenburg technique
Source: Authors

When comparing the Doernenburg analysis results against those of the correlation analysis, it was noted that there were more transformers in regular and bad conditions than transformers with identifiable faults according to the Doernenburg analysis.

The key gas technique also classified the majority of the transformers as being in a normal state. The most representative failure with this method was cellulose overheating, with carbon monoxide being the representative gas (Figure 6).

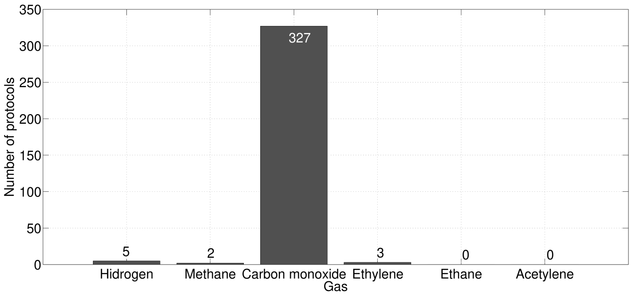


Figure 6. Results regarding relative gas distribution in the 337 power transformers
Source: Authors

As for the Rogers ratios method, 31,45% of the transformers were not suitable for analysis due to indeterminacies in the ratios, 20,47% did not have faults, 15,43% reported faults at low temperatures, and the rest exhibited thermal faults below 700 ^{o}C. No electrical faults were found with this method, which supports its ability to better predict thermal faults.

According to the Duval triangle, 81,31% of the results correspond to lower-intensity partial discharge electrical faults, which are typical of transformers with long periods of operation. Even though this is a minor fault type, it is likely that it will evolve into more critical faults in the triangle. This is the only method that identified more faulty transformers in comparison with the correlation analysis, given that Duval analysis evaluates failures in early stages, while correlation requires more advanced failure conditions.

Conclusions

The results obtained for the study cases with real data from 337 transformers allow concluding that there is a correlation between DGA and Megger DC tests. Therefore, the insulation status of a transformer can be diagnosed based on the results of gas concentration analysis, with a correlation of 80,45% when the DGA condition is regular or bad. Estimating the insulation status of equipment without the need to disconnect it constitutes an advantage, *i.e.*, it minimizes its downtime and increases its availability.

Based on the theory and the development of the exposed models, it was observed that DGA tests and the available analysis techniques are robust tools for diagnosing the state of the insulation system in a power transformer. The analyses carried out made it possible to distinguish between thermal and electrical faults according to the presence of specific combustible gases. The gases with the highest likelihood in a DGA test were hydrogen and carbon monoxide, which are inherent to the operation of a power transformer. In contrast, it was less likely to find acetylene in high concentrations, since it is linked to dangerous electrical fault conditions.

Acknowledgements

The authors wish to thank Ingeniería Especializada en Sistematización de Mantenimiento (INESSMAN LTDA.) for their invaluable support in the development of this document. They would also like to thank the support of Universidad Escuela Colombiana de Ingeniería Julio Garavito and the MEEP research group.

CRediT author statement

All authors contributed equally to the research.

References

Aciu, A. M., Nicola, C. I., Nicola, M, and Nitu M. C. (2021). Complementary Analysis for DGA Based on Duval Methods and Furan Compounds Using Artificial Neural Networks. *Energies, 14*(3), 588. `https://doi.org/10.3390/en14030588`

Aizpurua, J. I., Stewart, B. G., McArthur, S. D. J., Lambert, B., Cross, J. G., and Catterson, V. M. (2019). Improved power transformer condition monitoring under uncertainty through soft computing and probabilistic health index. *Applied Soft Computing, 85*, 1-15. `https://doi.org/10.1016/j.asoc.2019.105530`

Cerón, A., Orduña, I., Aponte G., and Romero, A. (2014) Panorama de la Gestión de Activos para transformadores de Potencia. *Información Tecnológica, 26*(4), 99-110. `http://dx.doi.org/10.4067/S0718-07642015000300014`

Dawson, S. B., and Trapp, R. (2005). *Bioestadística médica* (4th ed.). Editorial Manual Moderno.

Devore, J. L. (2016). *Probability & statistics for engineers and the Sciences* (9th ed.). Editorial Cengage Learning.

Dhini, A., Faqih, A., Kusumoputro, B., Surjandari, I., and Kusiak, A., (2020). Data-driven fault diagnosis of power transformers using dissolved gas analysis (DGA). *International Journal of Technology, 11*(2), 388-399. `https://doi.org/10.14716/ijtech.v11i2.3625`

Díaz, M. S., and Schmidt, E. (2014). Diagnóstico del estado de Transformadores de Potencia mediante análisis DGA. *TECNOTRANS, 2*, 62-69. `https://doi.org/10.30972/eitt.20281`

Dukarm, J. (2019). Análisis de gas disuelto basado en fiabilidad: un nuevo enfoque para DGA. *Revista CIDET, 20*, 22-24. `http://revista.cidet.org.co/revistas/revista-20/analisis-de-gas-disuelto-basado-en-fiabilidad-un-nuevo-enfoque-para-dga/`

Fofana, I., and Hadjadj, Y. (2016). Electrical-based diagnostic techniques for assessing insulation condition in aged transformers. *Energies, 9*(9), 2-26. `https://doi.org/10.3390/en9090679`

International Electrotechnical Commission (IEC) (2006). *IEC 60060-3:2006. High-voltage test techniques Part 3: Definitions and requirements for on-site testing*. IEC.

International Electrotechnical Commission (IEC) (2015). *IEC 60599:2015. Mineral oil-filled electrical equipment in service { Guidance on the interpretation of dissolved and free gases analysis*. IEC.

IEEE Standards Association (2000) *IEEE recommended practice for measurement of power factor tip-up of electric machinery stator coil insulation, IEEE Std 286-2000*. IEEE. `https://doi.org/10.1109/IEEESTD.2000.8685849`

IEEE Standards Association (2008). *IEEE guide for the interpretation of gases generated in oil-immersed iransformer, IEEE Std C57.104-2008* (revision of *IEEE Std C57.104-1991*). IEEE. `https://doi.org/10.1109/IEEESTD.2009.4776518`

IEEE Standards Association (2013) *IEEE guide for diagnostic field testing of fluid-filled power transformers, regulators, and reactors, IEEE Std C57.152-2013*. IEEE. `https://doi.org/10.1109/IEEESTD.2013.6544533`

IEEE Standards Association (2017). *IEEE guide for evaluation and reconditioning of liquid immersed power transformers, IEEE Std C57.140-2017* (revision of *IEEE Std C57.140-2006*). IEEE. `https://doi.org/10.1109/IEEESTD.2017.8106924`

IEEE Standards Association (2019). *IEEE guide for the interpretation of gases generated in mineral oil-immersed transformer, IEEE Std C57.104-2019* (revision of *IEEE Std C57.104-1991*). IEEE. `https://doi.org/10.1109/IEEESTD.2019.8890040`

Juris, J. S., Duran, I. C., and Marulanda, A. R. (2020). *Correlation between insulation resistance and dissolved gas analysis tests in power transformers* [Conference presentation]. 2020 IEEE PES Transmission & Distribution Conference and Exhibition - Latin America (T&D LA), Montevideo, Uruguay. `https://doi.org/10.1109/TDLA47668.2020.9326248`

Kruger, M., and Engelen, C. (2018). *Diagnostic testing of cast resin transformers*. Omicron Electronics Gmbh.

Mahmoudi, N., Samimi, M. H., and Mohseni, H. (2019). *Experiences with transformer diagnosis by DGA: Case studies*. *IET Generation, Transmission & Distribution, 13*(23), 5431-5439. `https://doi.org/10.1049/iet-gtd.2019.1056`

Malpure, B. D., and Baburao, K. (2008). *Failure analysis & diagnostics of power transformer using dielectric dissipation factor* [Conference presentation]. 2008 International Conference on Condition Monitoring and Diagnosis, Beijing, China, `https://doi.org/10.1109/CMD.2008.4580334`

Márquez, M., and Molina., G. (2016). *Evaluación de la vida útil del aislamiento en trasformadores de potencia a partir del ciclo térmico del calentamiento del transformador* [Doctoral thesis, Universidad de El Salvador]. `https://ri.ues.edu.sv/id/eprint/10341/`

Patekar, K. D., and Chaudhry, B. (2019). *DGA analysis of transformer using artificial neutral network to improve reliability in power transformers* [Conference presentation]. 2019 IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), Chennai, India. `https://doi.org/10.1109/CATCON47128.2019.PID6178475`

Piegar, J. S., Álvarez, R. E., and Catalano L. J. (2015). *Herramienta de gestión y análisis para los ensayos DGA en transformadores de potencia inmersos en aceite mineral* [Conference presentation]. XVI Encuentro Regional Iberoamericano de CIGRÉ, Puerto Iguazú, Argentina. `https://sedici.unlp.edu.ar/handle/10915/51477`

Prasojo, R. A., Gumilang, H., Suwarno, Maulidevi, N. U., and Soedjarno, B. A. (2020). A fuzzy logic model for power transformer faults' severity determination based on gas level, gas rate, and dissolved gas analysis interpretation. *Energies, 13*(4), 1009. `https://doi.org/10.3390/en13041009`

Rogers, R. R. (1978). *IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis*. `https://doi.org/10.1109/TEI.1978.298141`

Saravanan, D., Hasan, A., Singh, A., Mansoor, H., and Shaw, R. N. (2020). *Fault prediction of transformer using machine learning and DGA* [Conference presentation]. 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India. `https://doi.org/10.1109/GUCON48875.2020.9231086`

Sardar, S., Kumar, A., Chatterjee, B., and Dalai, S. (2017). *Application of statistical interpretation technique for frequency response analysis and detection of axial displacement in transformer winding* [Conference presentation]. 2017 IEEE Calcutta Conference (CALCON), Kolkata, India. `https://doi.org/10.1109/CALCON.2017.8280776`

Sarria-Arias, J., Guerrero-Bello, N. and Rivas-Trujillo, E. (2014) Estado del arte del análisis de gases disueltos en transformadores de potencia. *Revista Facultad de Ingeniería, 23*(36), 105-122. `https://doi.org/10.19053/01211129.2716`

Syafruddin, H., and Nugroho, H. P. (2020). *Dissolved gas analysis (DGA) for diagnosis of fault in oil-immersed power transformers: A case study* [Conference presentation]. 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), Medan, Indonesia. `https://doi.org/10.1109/ELTICOM50775.2020.9230491`

Torkaman, H., and Karimi, F. (2015). Measurement variations of insulation resistance/polarization index during utilizing time in HV electrical machines A survey. *Measurement, 59*, 21-29. `https://doi.org/10.1016/j.measurement.2014.09.034`

Walpole, R. E., Myers, R. H., and Ye, K. (2017). *Probability & statistics for engineers & scientists* (9th ed.). Pearson editors.

Wang, M., Vandermaar, A. J., and Srivastava, K. D. (2002). Review of condition assessment of power transformers in service. *IEEE Electrical Insulation Magazine, 18*(6), 12-25. `https://doi.org/10.1109/MEI.2002.1161455`

Wang, H. (2009). A novel extension method for transformer fault diagnosis. *IEEE Transactions on Power Delivery, 18*(1), 164-169. `https://doi.org/10.1109/TPWRD.2002.803838`

Wannapring, E., Suwanasri, C., and Suwanasri, T. (2016). *Dissolved gas analysis methods for distribution transformers* [Conference presentation]. 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand. `https://doi.org/10.1109/ECTICon.2016.7561320`

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e102345)

Review Article/Mechanical Engineering, Mechatronics, and Materials Science http://doi.org/10.15446/ing.investig.102345

Surface Roughness Value Recommended for the Manufacture of Antibacterial Metal Surfaces: A Review

Rugosidad superficial recomendada en la manufactura de superficies metálicas antibacterianas: una revisión

Martha Patricia Calvo-Correa[1], Carlos Julio Cortés-Rodríguez[2], and Julián R. Camargo-López[3]

ABSTRACT

The manufacturing of antibacterial metal surfaces has been widely studied in the elaboration dental and orthopedic implants. Surface characteristics such as wettability, chemistry, electrostatics, and roughness have been described as factors for avoiding bacterial adhesion. However, surface roughness is still debated among authors regarding its effect on antibacterial surfaces. This paper reviews the existing literature to identify the recommended surface roughness values for metal implants to avoid bacterial adhesion, and it evaluates the different roughness parameters used in this regard. This compilation found no agreement when it comes to the exact roughness that a metal implant's surface should have to avoid bacterial adhesion and the subsequent formation of biofilms. In general, different authors recommend manufacturing smooth surfaces with a nanoscale roughness, smaller than the size of the target bacterium.

Keywords: bacterial adhesion, metal implant surface, surface roughness, bacterial biofilm

RESUMEN

La manufactura de superficies metálicas antibacterianas ha sido ampliamente estudiada en la elaboración de implantes ortopédicos y dentales. Ciertas características de las superficies, tales como mojabilidad, química, electrostática y rugosidad, han sido señaladas como factores para prevenir la adhesión bacteriana. Sin embargo, la rugosidad superficial aún se encuentra en debate en cuanto a su efecto en las superficies antibacterianas. Este artículo realiza una revisión de la literatura existente para identificar los valores de rugosidad superficial recomendados para evitar la adhesión bacteriana en implantes metálicos, y evalúa los distintos parámetros de rugosidad utilizados en este contexto. Esta compilación no encontró un consenso en términos de la rugosidad exacta que la superficie de un implante metálico debería tener para evitar la adhesión bacteriana y la posterior formación de *biofilms*. En términos generales, diferentes autores recomiendan manufacturar superficies con rugosidades de escala nanométrica, menores que el tamaño de la bacteria objetivo.

Palabras clave: adhesión bacteriana, superficie de implantes metálicos, rugosidad superficial, *biofilm* bacteriano

Received: April 28th, 2022
Accepted: August 1st, 2023

Introduction

Bacterial surface adhesion has been widely studied for many years, seeking to understand and prevent possible future infections [1], [2]. Nosocomial infection, also called *hospital-acquired infection*, is "an infection acquired in hospital by a patient who was admitted for a reason other than that infection. An infection occurring in a patient in a hospital or other health care facility in whom the infection was not present..." [3, p. 1].

Every year, 1 to 5% of orthopedic implants end up with an infection [4]. The number one reason for persisting infections has been the formation of bacterial biofilms [5]. According to the reports, biofilms have been present in 60 to 80% of chronic infections [6], [7].

Treating this type of infection increases the cost of healthcare as well as the recovery times [6], [8], [9]. Among other factors, roughness, wettability, chemistry, and electrostatic charge influence the adhesion of bacteria to a surface. The values of these factors differ if metal, polymer, or ceramic is used. Knowing the exact values of these factors helps to manufacture an antibacterial surface that reduces the patient's probability of having a nosocomial infection [2], [10].

This paper aims to review one of the factors that influence bacterial adhesion in metallic surfaces: surface roughness. It aims to summarize the recommended values for avoiding bacterial adhesion to orthopedic metal implants according to different authors.

---

[1] BS, Universidad Pedagógica Nacional, Colombia. MSc, Universidad Nacional de Colombia, Colombia. Affiliation: PhD student, Universidad Nacional de Colombia, Colombia. Email: mpcalvoc@unal.edu.co

[2] Eng, Universidad Nacional de Colombia, Colombia. PhD, Universität Kassel, Germany. Affiliation: Full professor, Universidad Nacional de Colombia, Colombia. Email: cjcortesr@unal.edu.co

[3] Eng, Universidad Distrital Francisco José de Caldas, Colombia. MSc, Universidad Distrital Francisco José de Caldas, Colombia. Affiliation: Full professor, Universidad Distrital Francisco José de Caldas, Colombia. Email: jcamargo@udistrital.edu.co

This article is divided into four sections. *What is biofilm?* elucidates the process of bacterial adhesion to the implant surface, the difficulty of eradicating it, and the role of surface roughness in it. *Surface roughness parameters used to measure bacterial adhesion* outlines various roughness parameters utilized for measuring surface roughness, and it discusses why specific parameters are more suitable than Sa to measure bacterial adhesion, as per the ISO 25178 – 2 standard [33]. *How do nano- and micro-roughness surfaces affect adhesion?* compares the findings of different authors regarding the surface roughness ranges that prevent bacterial adhesion. In this section, the lack of consensus on these values is addressed. The *Conclusions* section provides a comprehensive overview based on this review, offering a general recommendation for manufacturing antibacterial metal surfaces. By synthesizing the findings from various sections, valuable insights for creating effective antibacterial metal surfaces are presented.

What is biofilm?

Peri-implant infections are a particular type of infection produced by a biomaterial in the body. The microorganisms in these infections take advantage of the weak interface between the antibodies and the surface of the implant [11]-[13].

Because planktonic bacteria are associated with orthopedic implants, these infections create multicellular communities on the implant's surface, which are challenging to eradicate. Bacterial adhesion, in general terms, has five stages [5], [14]-[16]:

- *Stage one*: Planktonic bacteria adhere reversibly to the surface implant. This adhesion depends on characteristics such as the surface material, hydrodynamics, electrostatic interactions, surface roughness, *etc.* [1].
- *Stage two*: The adhered bacteria send information to the nearby bacteria and secrete proteins with adhesive components, which makes the adhesion irreversible in a few hours [1].
- *Stage three*: The bacterial community colonizes the surface and creates a biofilm matrix that will encapsulate the bacteria. This results in a complex structure which will cover the colony [1].
- *Stage four*: The bacterial colony starts to secrete extracellular polymeric substances (EPS) and develops a complex microorganism known as a *mature biofilm*. This biofilm consists mainly of bacteria, polysaccharides, proteins, and DNA. In this last stage, some ducts are created to spare more bacteria, spreading the infection to different body parts [1].

Preventing bacterial adhesion and anchorage to the implant's surface is crucial for preventing further infections [17]. Biofilms protect bacteria from the immune system and antibiotics, and, once they have created a biofilm, treating the infection with antibiotics becomes harder [1]. It has been found that the dose of antibiotics needed to treat some bacteria in biofilms is 1 000 times higher than that needed in the first adhesion stage [14], [18]. Many authors such as [19] and [20] have stated that modifying the implant surface, either chemically or physically, may lead to antibacterial surfaces.

Regarding the physical modification of the surface, micro- and nano-topographies are essential in creating antibacterial surfaces. However, the exact values are still unknown [21]-[23].

Surface roughness parameters used to measure bacterial adhesion

Surface roughness is the way in which the topography of a surface can be described in terms of how high the peaks are and how deep the valleys are on a surface after manufacturing. Surface roughness is just one of the many characteristics that influence bacterial adhesion [24]-[28].

Surface roughness can be measured with different parameters: profile roughness parameters (Ra, Rq...) calculated in a profile (line), and area roughness parameters (Sa, Sq...) calculated in a completed area. Profile parameters, especially Ra, are the most commonly reported, but they cannot describe the surface thoroughly, providing a reduced portion of surface information [26]-[28].

Area roughness parameters better describe the surface as a whole. Some parameters can provide specific information about the surface, and their use depends on the application [29]-[32]. The surface roughness parameters described below are reported in the ISO 25178 – 2 standard and are used to measure surface roughness for antibacterial adhesion [33].

Sa – Arithmetic mean deviation: This parameter is the arithmetic mean of an absolute value corresponding to the height of the sampling area. This parameter does not adequately describe the surface topography because different surfaces with different peaks and valleys can have the same Sa.

Sq – Squared mean height: This parameter is similar to Sa and is related to the surface energy.

Ssk – Asymmetry: This parameter describes the distribution of peaks of the topography. The height difference is distributed uniformly if a surface has a Ssk = 0.

Sku – Kurtosis: This parameter describes the spikiness of the surfaces if the peaks are uniformly distributed.

Sdr – Surface area ratio: This is one of the most important parameters to describe an antibacterial surface. A surface with a Sdr = 0 is ideally 'smooth'. This parameter can also identify the differences between two surfaces with the same Sa.

The measurement of these parameters is affected by the equipment used. Some studies describe how contact and no-contact equipment affect measurements, so this criterion should be considered in future research [32], [34].

How do nano- and micro-roughness surfaces affect adhesion?

Surface roughness has been widely studied with the aim of preventing bacterial adhesion to metal surfaces. Many authors have described the roughness values found to be the most appropriate in this regard. However, they disagree on each type of topography [13], [35].

Some authors have stated that some surface feature patterns potentially determine the bacteria's early location and adhesion. In this vein, nanoscale topographies with an organized pattern reduce bacterial adhesion by 40-95% [36], [37].

In a study on bacterial adhesion to titanium nitride, [38] found that surfaces with a roughness lower than 0,066 μm had better anti-adhesion properties. Other studies such as [39] argue that the adhesion of bacteria is minimal if the surface has a roughness value or Ra of 160 nm. This study evaluated stainless steel surfaces with Ra between 30 and 890 nm. A different study [41] evaluated five types of surfaces with Ra between 172,5 and 45,2 nm, finding that bacterial adhesion was restrained on the nanoscale surface.

While investigating different polishing surfaces, [41] found that bacteria had less adhesion in metal surfaces with higher porosity. The authors found increased bacterial adhesion in surfaces with a Ra smaller than 0,2 μm. In another study, no relation between superficial roughness and bacterial adhesion could be found [42].

In contrast, the authors of [43] argue that there is a negative correlation between bacterial adhesion and metal surfaces with a roughness of less than 6 nm. They found that bacteria prefer a smoother surface under static culture conditions if the roughness is smaller than 0,23-6,13 nm. When the roughness is more significant than 6-30 nm, bacteria prefer surfaces with higher roughness values. The authors also found that a higher roughness does not necessarily benefit adhesion; it will increase the production of EPS to create a biofilm. It will help protect the bacteria because deep valleys trap them and provide cover from shear force [16].

On the other hand, some authors such as [44], [45], and [46], among others, have stated that nano-topography has an antibacterial effect, as peaks with sizes similar to those of the bacteria exert strain forces on the bacterial membrane, causing it to break. This phenomenon is known as the *cicada wings effect*.

[43] suggests that increasing the surface roughness more significantly than 1,24 μm will increase bacterial adhesion. [47] support this idea, stating that surfaces with a roughness higher than 0,4 μm increase the risk of peri-implant diseases. They support the conclusions of [48], *i.e.*, metal surfaces of 130-360 nm are desired in avoiding bacterial adhesion. A similar idea is shared by [49]: surfaces with flaws such as groves, gaps, or cracks provide a favorable environment for bacteria [39]. These authors suggest that bacteria adhere to surfaces that correspond to their size since this maximizes their contact area [50]-[52].

[20] suggest that the adhesion of the bacterium *Staphylococcus aureus* decreases by using surfaces with a roughness value under 1,51 nm and a unidirectional surface texture. Some related results were obtained [53]. In this study, the adhesion of *Escherichia coli* and *Staphylococcus aureus* was reduced by 55,6 and 40,5%, respectively, by using a titanium surface smaller than 6 nm (surface peak density). The study by [15] with the same bacteria found similar results using pattern surfaces of 1 μm. A similar result with the same value of roughness was found by [54].

As is the case of [55], some authors agree that, despite the vast amount of works, there is not enough information about the adequate topography. However, in general, the best approximation would be a nanoscale surface roughness with a structure similar structure to that of bone tissue [56].

In [57], by analyzing the way in which bacterial adhesion is evaluated, the authors found that using parameters such as Ra may not be the most accurate approach. Ra is an arithmetic mean or the absolute value of the vertical deviation from the mean line of the profile [31], [58], [59]. In other words, Ra is described as the arithmetic mean deviation of an assessed profile, which is the average of the peaks and valleys of the surface with regard to a centerline [35]. According to this, there could be different surfaces with the same Ra and significant differences in peaks and valleys [60]. Therefore, other superficial roughness parameters should be considered in future research, such as Rq, Rsk, or even area roughness parameters such as Sq, Sq, Ssk, and Sku [61].

Other works [61], [66] have used Sq, Ssk, and Sku to measure surface roughness, finding, for instance, that *Staphylococcus epidermis* creates biofilms on surfaces with high Ssk values and that smoother surfaces have the best antibacterial properties. Despite not providing an exact value for the surface parameters to create antibacterial surfaces, the findings are consistent with the idea that a surface with low peaks and valleys hinders the adhesion of bacteria.

Moreover, bacteria will avoid surfaces with a roughness smaller than their size, as this increases their contact area, which can damage the bacterial membrane [46], [66].

Conclusions

According to this review, there are two things to consider while creating antibacterial surfaces.

Table I. Different roughness values suggested by different authors to create antibacterial surfaces

Authors	Title	Results
Songze Wu *et al.* [18]	Influence of surface topography on bacterial adhesion	The height of the peaks on the topography should not be greater than the flagella size. The adequate roughness should be Ra > 172,5-45,2 nm or Rq > 217,9-56,6 nm.
Kurup *et al.* [38]	Surface modification techniques of titanium and titanium alloys for biomedical, dental applications	Surfaces with a roughness lower than 0,066 µm had better anti-adhesion properties than treated ones.
Vadillo-Rodríguez *et al.* [37, p. 347]	Bacterial response to spatially organized microtopographic surface patterns with nanometers scale roughness	"Surfaces with shorter or less deep features than 21,1 nm inhibited the growth of the bacteria..."
Costa *et al.* [63,p. 4]	Fitting pieces into the puzzle: The impact of titanium-based dental implant surface modifications on bacterial accumulation and polymicrobial infections	"Nanostructured surfaces have been demonstrated to be more efficient at reducing bacterial attachment..." Roughness values of Ra = 0,51-1,36 µm or Sa = 0,66-2,91 µm have no antibacterial effects, although they induce osteoblast adhesion.
Yuan *et al.* [14, p. 53]	Bacterial anti-adhesion surface design: surface patterning, roughness, and wettability: A review	A Ra smaller than 6 nm has a negative correlation with bacterial adhesion. If the roughness is low (Ra: 0,23-6,13 nm), "... bacteria prefer to adhere on the smoother surface under static culture conditions and the increased production of EPS to create a biofilm...". When the roughness has a value between 6 and 30 nm, bacteria prefer to adhere to the rougher surface.
Dantas *et al.* [47, p. 2]	Bacteria co-culture adhesion on different texturized zirconia surfaces	"Roughness surfaces higher than 0,4 µm increase the affinity of microorganisms and the risk of peri-implant diseases..." Surfaces between 130 and 360 nm are more likely to be antibacterial surfaces.
Lu *et al.* [20]	Effects of surface roughness and texture on the bacterial adhesion on the bearing surface of bio-ceramic joint implants: An *in vitro* study	The Ra to decrease the number of adherent bacteria (*Staphylococcus aureus*) should be under 1,51 nm.
Filipović *et al.* [49, p. 5]	Bacterial adhesion on orthopedic implants	"A tiny increase in surface roughness (Ra = 0,04–1,24 µm) resulted in a pronounced increase of bacteria..." Bacteria adhere to surfaces similar in size since this allows them to maximize the contact area.
Annunziata *et al.* [64]	Bacterial adhesion to direct laser metal formed and mildly acid-etched implants surfaces	Roughness should be less than 0,2 µm. The texture should be Sa < 0,5 µm.
Lüdecke *et al.* [53]	Nanorough titanium surface adhesion of *Escherichia coli* and *Staphylococcus aureus* via nano adhesion points	*Escherichia coli* and *Staphylococcus aureus* adhesion was reduced by 55,6 and 40,5%, respectively, by using a titanium surface smaller than 6 nm (surface peak density).
Yang *et al.* [15]	Control of bacterial adhesion and growth on honeycomb-like patterned surfaces	*Escherichia coli* and *Staphylococcus aureus*, in patterns of Ra between 1 and 10 µm, showed bacterial growth inhibited to a greater extend than on a flat surface.
Lüdecke *et al.* [54 p. 587]	Physical vapor-deposited titanium thin films for biomedical applications: Reproducibility of nanoscale surface roughness and microbial adhesion properties	"Escherichia Coli adhesion decreased from approximately 23 % to 19 % with increasing surface nano roughness from 2nm to 6 nm..."

- Bacteria have fewer probabilities of initial adhesion and biofilm formation in surfaces with nanoscale roughness.
- Bacteria have fewer odds of adhesion in surfaces with a roughness value less than their size.

In conclusion, antibacterial surfaces should be smooth, avoiding high peaks and valleys where the bacteria could be protected from shear forces. The surface roughness should be smaller than the bacteria's size. This parameter depends on the type of bacteria studied, but *Escherichia coli* and *Staphylococcus aureus* can serve as a reference since they are the most common in nosocomial infections. The sizes of these two bacteria are 1-5 and 1 µm in diameter, respectively.

Acknowledgments

The authors would like to thank Universidad Nacional de Colombia, the GIBM research group, and the Colombian Ministry of Sciences (MinCiencias) for their funding through the Bicentennial Doctoral Excellence Scholarships program.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

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

References

[1] J. Palmer, S. Flint, and J. Brooks, "Bacterial cell attachment, the beginning of a bio-film," *J. Ind. Microbiol. Biotechnol*, vol. 34, no. 9, pp. 577-588, 2007, https://doi.org/10.1007/s10295-007-0234-4

[2] B. Bhushan and Y. C. Jung, "Natural and biomimetic artificial surfaces for superhy-drophobicity, self-cleaning, low adhesion, and drag reduction," *Prog. Mater. Sci.*, vol. 56, no. 1. pp. 1-108, 2011. https://doi.org/10.1016/j.pmatsci.2010.04.003

[3] World Health Organization, "Prevention of hospital-acquired infections: A practical guide," 2002. [Online]. Available: https://iris.who.int/bitstream/handle/10665/67350/WHO_CDS_CSR_EPH_2002.12.pdf?sequence=1&isAllowed=y

[4] P. Ginestra *et al.*, "Post processing of 3D printed metal scaffolds: A preliminary study of antimicrobial efficiency," *Procedia Manuf.*, vol. ESAFORM 20, pp. 1106-1112, 2020. https://doi.org/10.1016/j.promfg.2020.04.126

[5] S. Kumar, D. N. Roy, and V. Dey, "A comprehensive review on techniques to create the anti-microbial surface of biomaterials to intervene in biofouling," *Colloids Interface Sci.*

Com., vol. 43, art. 100464, 2021. https://doi.org/10.1016/j.colcom.2021.100464

[6] R. A. Mendoza, J.-C. Hsieh, and R. D. Galiano, "The impact of biofilm formation on wound healing," in *Wound Healing – Current Perspectives*, K. Hakan Dogan, Ed., London, UK: IntechOpen, 2019, pp. 3-17. https://doi.org/10.5772/intechopen.85020

[7] Instituto Nacional de Salud, "Boletín epidemiológico. Semana epidemiológica 09," 2021. https://doi.org/10.33610/23576189.2021.09

[8] M. Malone *et al.*, "The prevalence of biofilms in chronic wounds: A systematic review and meta-analysis of published data," *J. Wound Care*, vol. 26, no. 1, pp. 20-25, 2017. https://doi.org/10.12968/jowc.2017.26.1.20

[9] Boletín Epidemiógico Semanal, "Infecciones asociadas a procedimientos médicos quirúrguicos," 2017. [Online]. Available: https://www.ins.gov.co/BibliotecaDigital/Boletin-epidemiologico-semana-25-2019.pdf

[10] L. M. Pandey, "Design of biocompatible and self-antibacterial titanium surfaces for biomedical applications," *Curr. Opinion Biomedical Eng.*, vol. 25, art. 100423, 2022. https://doi.org/10.1016/j.cobme.2022.100423

[11] D. Campoccia, L. Montanaro, and C. R. Arciola, "The significance of infection related to orthopedic devices and issues of antibiotic resistance," *Biomater.*, vol. 27, no. 11, pp. 2331-2339, 2006. https://doi.org/10.1016/j.biomaterials.2005.11.044

[12] I. B. Beech, J. A. Sunner, C. R. Arciola, and P. Cristiani, "Microbially-influenced corrosion: Damage to prostheses, delight for bacteria," *Int. J. Artif. Organs*, vol. 29, no. 4, pp. 443-452, 2006. https://doi.org/10.1177/039139880602900415

[13] R. Jia, T. Unsal, D. Xu, Y. Lekbach, and T. Gu, "Microbiologically influenced corrosion and current mitigation strategies: A state of the art review," *I. Biodeterioration Bio-degradation*, vol. 137. pp. 42-58, 2019. https://doi.org/10.1016/j.ibiod.2018.11.007

[14] Z. Yuan, Y. He, C. Lin, P. Liu, and K. Cai, "Antibacterial surface design of biomedical titanium materials for orthopedic applications," *J. Mater Sci. Tech.*, vol. 78. pp. 51-67, 2021. https://doi.org/10.1016/j.jmst.2020.10.066

[15] M. Yang, Y. Ding, X. Ge, and Y. Leng, "Control of bacterial adhesion and growth on honeycomb-like patterned surfaces," *Colloids Surf. B Biointerfaces*, vol. 135, pp. 549-555, 2015. https://doi.org/10.1016/j.colsurfb.2015.08.010

[16] K. B.-C. Justyna Mazurek-Popczyk, L. Palka, K. Arkusz, and B. Dalewski, "Personalized, 3D- printed fracture fixation plates versus commonly used orthopaedic implant materials – Biomaterials characteristics and bacterial biofilm formation," *Injury*, vol. 53, no. 3, pp. 938-946, 2022. https://doi.org/10.1016/j.injury.2021.12.020

[17] M. Lorenzetti *et al.*, "The relationship between the nanostructure of titanium surfaces and bacterial attachment," *ACS Nano*, vol. 31, no. 4, pp. 706-713, 2010. https://doi.org/10.1016/j.biomaterials.2009.09.081

[18] S. Wu, B. Zhang, Y. Liu, X. Suo, and H. Li, "Influence of surface topography on bacterial adhesion: A review," *Biointerphases*, vol. 13, no. 6, art. 060801. https://doi.org/10.1116/1.5054057

[19] S. B. Chinnaraj *et al.*, "Modelling the combined effect of surface roughness and topography on bacterial attachment," *J. Mater. Sci. Tech.*, vol. 81, pp. 151-161, Jan. 2021. https://doi.org/10.1016/j.jmst.2021.01.011

[20] A. Lu, Y. Gao, T. Jin, X. Luo, Q. Zeng, and Z. Shang, "Effects of surface roughness and texture on the bacterial adhesion on the bearing surface of bio-ceramic joint implants: An in vitro study," *Ceram. Int.*, vol. 46, no. 5, pp. 6550-6559, 2020. https://doi.org/10.1016/j.ceramint.2019.11.139

[21] Y. Ammar, D. Swailes, B. Bridgens, and J. Chen, "Influence of surface roughness on the initial formation of biofilm," *Surf. Coat. Tech.*, vol. 284, pp. 410-416, 2015. https://doi.org/10.1016/j.surfcoat.2015.07.062

[22] Y. S. Huang and H. H. Huang, "Effects of clinical dental implant abutment materials and their surface characteristics on initial bacterial adhesion," *Rare Metals*, vol. 38, no. 6, pp. 512-519, 2019. https://doi.org/10.1007/s12598-019-01219-0

[23] N. Mitik-Dineva, J. Wang, R. C. Mocanasu, P. R. Stoddart, R. J. Crawford, and E. P. Ivanova, "Impact of nano-topography on bacterial attachment," *Biotech. J.*, vol. 3, no. 4, pp. 536-544, 2008. https://doi.org/10.1002/biot.200700244

[24] R. Krishna Alla, K. Ginjupalli, N. Upadhya, M. Shammas, R. Krishna Ravi, and R. Sekhar, "Surface roughness of implants: A review," *Trends Biomater. Artif. Organs*, vol. 25, no. 3, pp. 112-118, 2011. [Online]. Available: https://brnskll.com/wp-content/uploads/2019/02/taat11i3p112.pdf

[25] G. R. M. Matos, "Surface roughness of dental implant and osseointegration," *J. Maxillofac. Oral Surg.*, vol. 20, no. 1, pp. 1-4, 2021. https://doi.org/10.1007/s12663-020-01437-5

[26] B. Azarhoushang and A. Daneshi, "Work-piece surface roughness," in *Tribology and Fundamentals of Abrasive Machining* Processes, 3rd ed., B. Azarhoushang, I. D. Marinescu, W. B. Rowe, B. Dimitrov, and H. Ohmori, Eds., Amsterdam, The Netherlands: Elsevier, 2022, pp. 575-590. https://doi.org/10.1016/B978-0-12-823777-9.00015-X

[27] V. S. Lukyanov, "Surface roughness and parameters," *Precis. Eng.*, vol. 5, no. 3, pp. 99-100, 1983. https://doi.org/10.1016/0141-6359(83)90001-6

[28] W. Group, "3D Roughness Metrology," 2020[Online]. Available:. https://www.ptb.de/cms/en/ptb/fachabteilungen/abt5/fb-51/ag-514.html

[29] "Roughness parameter," in *CIRP Encyclopedia of Production Engineering*, Berlin, Heidelberg, Germany: Springer, 2019, p. 1497. https://doi.org/10.1007/978-3-662-53120-4_300588

[30] T. Jeyapoovan and M. Murugan, "Surface roughness classification using image processing," *Measurement*, vol. 46, no. 7, pp. 2065-2072, Apr. 2013. https://doi.org/10.1016/j.measurement.2013.03.014

[31] B. Bhushan, *Mordern Tribology Handbook*, 1st ed., Oxfordshire, UK: Taylor & Francis Group, 2000. https://doi.org/10.1201/9780849377877

[32] C. J. Cortés-Rodríguez, F. C. Herreño Cuestas, and I. Z. Areque-Salazar, *Medición de Rugosidad Superficial 3D*, 1st ed., Kassel, Germany: Kassel University Press, 2019.

[33] *Geometrical product specification (GPS). Surface texture: Areal. Part 606: Nominal characteristics of non-contact (focus variation) instruments*, ISO 25178-606, International Organization for Standardization, Switzerland, Jun. 2015.

[34] P. Podulka, "Selection of methods of surface texture characterisation for reduction of the frequency-based errors in the measurement and data analysis processes," *Sensors*, vol. 22, no. 3, art. 791, 2022. https://doi.org/10.3390/s22030791

[35] H. Johnson, "Surface roughness," in *Optical Properties of Surfaces*, J. Vileger and D. Bedeaux, Singapore: World Scientific, 2001, pp. 401-429. https://doi.org/10.1142/9781860945434_0014

[36] N. Encinas *et al.*, "Submicrometer-sized roughness suppresses bacteria adhesion," *ACS Appl. Mater. Interfaces*, vol. 12, no. 19, pp. 21192-21200, 2020. https://doi.org/10.1021/acsami.9b22621

[37] V. Vadillo-Rodríguez *et al.*, "Bacterial response to spatially organized microtopographic surface patterns with nanometer scale roughness," *Colloids Surf. B Biointerfaces*, vol. 169, pp. 340-347, 2018. https://doi.org/10.1016/j.colsurfb.2018.05.038

[38] A. Kurup, P. Dhatrak, and N. Khasnis, "Surface modification techniques of titanium and titanium alloys for biomedical dental applications: A review," *Mater. Today Proc.*, vol. 39, pp. 84-90, 2020. https://doi.org/10.1016/j.matpr.2020.06.163

[39] E. Medilanski, K. Kaufmann, L. Y. Wick, O. Wanner, and H. Harms, "Influence of the surface topography of stainless steel on bacterial adhesion," *Biofouling*, vol. 18, no. 3, pp. 193-203, 2002. https://doi.org/10.1080/08927010290011370

[40] S. Wu, S. Altenried, A. Zogg, F. Zuber, K. Maniura-Weber, and Q. Ren, "Role of the surface nanoscale roughness of stainless steel on bacterial adhesion and microcolony formation," *ACS Omega*, vol. 3, no. 6, pp. 6456-6464, 2018. https://doi.org/10.1021/acsomega.8b00769

[41] D. H. Kang, H. Choi, Y. J. Yoo, J. H. Kim, Y. B. Park, and H. S. Moon, "Effect of polishing method on surface roughness and bacterial adhesion of zirconia-porcelain veneer," *Ceram. Int.*, vol. 43, no. 7, pp. 5382-5387, 2017. https://doi.org/10.1016/j.ceramint.2016.11.036

[42] M. Annunziata *et al.*, "Bacterial adhesion to direct laser metal formed and mildly acid etched implant surfaces," *Surf. Coat. Tech.*, vol. 328, pp. 390-397, 2017. https://doi.org/10.1016/j.surfcoat.2017.09.011

[43] K. Yang *et al.*, "Bacterial anti-adhesion surface design: Surface patterning, roughness and wettability: A review," *J. Mater. Sci. Tech.*, vol. 99, pp. 82-100, 2022. https://doi.org/10.1016/j.jmst.2021.05.028

[44] T. Wuirk, "Insect wings shred bacteria to pieces," *Nature News*, Mar. 04, 2013. https://doi.org/10.1038/nature.2013.12533

[45] G. Lazzini, A. H. A. Lutey, L. Romoli, and F. Fuso, "Molecular dynamics model for the antibactericity of textured surfaces," *Colloids Surf. B Biointerfaces*, vol. 199, art. 111504, 2021. https://doi.org/10.1016/j.colsurfb.2020.111504

[46] A. Elbourne, R. J. Crawford, and E. P. Ivanova, "Nano-structured antimicrobial surfaces: From nature to synthetic analogues," *J. Colloid Interface Sci.*, vol. 508, pp. 603-616, 2017. https://doi.org/10.1016/j.jcis.2017.07.021

[47] T. Dantas *et al.*, "Bacteria co-culture adhesion on different texturized zirconia surfaces," *J. Mech. Behav. Biomed. Mater.*, vol. 123, art. 104786, 2021. https://doi.org/10.1016/j.jmbbm.2021.104786

[48] L. Yin, Y. Nakanishi, A. R. Alao, X. F. Song, J. Abduo, and Y. Zhang, "A review of engineered zirconia surfaces in biomedical applications," *Procedia CIRP*, vol. 65, pp. 284-290. https://doi.org/10.1016/j.procir.2017.04.057

[49] U. Filipović, R. G. Dahmane, S. Ghannouchi, A. Zore, and K. Bohinc, "Bacterial adhesion on orthopedic implants," *Adv. Colloid Interface Sci.*, vol. 283, art. 102228, 2020. https://doi.org/10.1016/j.cis.2020.102228

[50] N. J. Bassous, C. L. Jones, and T. J. Web-ster, "3D printed Ti-6Al-4V scaffolds for supporting osteoblast and restricting bacterial functions without using drugs: Predictive equations and experiments," *Acta Biomater.*, vol. 96, pp. 662-673, 2019. https://doi.org/10.1016/j.actbio.2019.06.055

[51] H. L. Huang, Y. Y. Chang, M. C. Lai, C. R. Lin, C. H. Lai, and T. M. Shieh, "Antibacterial TaN-Ag coatings on titanium dental implants," *Surf. Coat. Tech.*, vol. 205, no. 5, pp. 1636-1641, 2010. https://doi.org/10.1016/j.surfcoat.2010.07.096

[52] L. C. D. M. Dantas, J. P. Da Silva-Neto, T. S. Dantas, L. Z. Naves, F. D. Das Neves, and A. S. Da Mota, "Bacterial adhesion and surface roughness for different clinical techniques for acrylic polymethyl methacrylate," *Int. J. Dent.*, vol. 2016, art. 8685796, 2016. https://doi.org/10.1155/2016/8685796

[53] C. Lüdecke, M. Roth, W. Yu, U. Horn, J. Bossert, and K. D. Jandt, "Nanorough titanium surfaces reduce adhesion of *Escherichia coli* and *Staphylococcus aureus* via nano adhesion points," *Colloids Surf. B Biointerfaces*, vol. 145, pp. 617-625, 2016. https://doi.org/10.1016/j.colsurfb.2016.05.049

[54] C. Lüdecke, J. Bossert, M. Roth, and K. D. Jandt, "Physical vapor deposited titanium thin films for biomedical applications: Re-producibility of nanoscale surface roughness and microbial adhesion properties," *Appl. Surf. Sci.*, vol. 280, pp. 578-589, 2013. https://doi.org/10.1016/j.apsusc.2013.05.030

[55] K. Harawaza, B. Cousins, P. Roach, and A. Fernandez, "Modification of the surface nanotopography of implant devices: A translational perspective," *Mater. Today Bio*, vol. 12, no. 12, art. 100152, 2021. https://doi.org/10.1016/j.mtbio.2021.100152

[56] J. Alipal *et al.*, "An updated review on surface functionalisation of titanium and its alloys for implants applications," *Mater. Today Proc.*, vol. 42, pp. 270-282, 2019. https://doi.org/10.1016/j.matpr.2021.01.499

[57] S. Al-Amshawee, M. Y. B. M. Yunus, J. G. Lynam, W. H. Lee, F. Dai, and I. H. Dakhil, "Roughness and wettability of biofilm carriers: A systematic review," *Environ. Tech. Innov.*, vol. 21, art. 101233, 2021. https://doi.org/10.1016/j.eti.2020.101233

[58] R. J. Crawford, H. K. Webb, V. K. Truong, J. Hasan, and E. P. Ivanova, "Surface topographical factors influencing bacterial attachment," *Adv. Colloid. Interface Sci.*, vol. 179, no. 182, pp. 142-149, 2012. https://doi.org/10.1016/j.cis.2012.06.015

[59] K. Koyama, H. Abe, S. Kawamura, and S. Koseki, "Stochastic simulation for death probability of bacterial population considering variability in individual cell inactivation time

and initial number of cells," *Int. J. Food Microbiol.*, vol. 290, pp. 125-131, 2019. https://doi.org/10.1016/j.ijfoodmicro.2018.10.009

[60] L. Vepsäläinen, P. Stenberg, P. Pääkkönen, M. Kuittinen, M. Suvanto, and T. A. Pakkanen, "Roughness analysis for textured surfaces over several orders of magnitudes," *Appl. Surf. Sci.*, vol. 284, pp. 222-228, 2013. https://doi.org/10.1016/j.apsusc.2013.07.085

[61] M. F. Kunrath, "Customized dental implants: Manufacturing processes, topography, osseointegration and future perspectives of 3D fabricated implants," *Bioprinting*, vol. 20, art. e00107, 2020. https://doi.org/10.1016/j.bprint.2020.e00107

[62] M. F. Kunrath, M. S. G. Monteiro, S. Gupta, R. Hubler, and S. D. de Oliveira, "Influence of titanium and zirconia modified surfaces for rapid healing on adhesion and biofilm formation of Staphylococcus epidermidis," *Arch. Oral Biol.*, vol. 117, no. 117, art. 104824, 2020. https://doi.org/10.1016/j.archoralbio.2020.104824

[63] R. C. Costa *et al.*, "Fitting pieces into the puzzle: The impact of titanium-based dental implant surface modifications on bacterial accumulation and polymicrobial infections," *Adv. Colloid. Interface Sci.*, vol. 298, art. 102551, 2021. https://doi.org/10.1016/j.cis.2021.102551

[64] M. Annunziata *et al.*, "Bacterial adhesion to direct laser metal formed and mildly acid etched implant surfaces," *Surf. Coat. Tech.*, vol. 328, pp. 390-397, 2017. https://doi.org/10.1016/j.surfcoat.2017.09.011

[65] V. K. Truong *et al.*, "The influence of nano-scale surface roughness on bacterial adhesion to ultrafine-grained titanium," *Biomaterials*, vol. 31, no. 13, pp. 3674-3683, 2010. https://doi.org/10.1016/j.biomaterials.2010.01.071

[66] A. Jain, N. Kumari, S. Jagadevan, and V. Bajpai, "Surface properties and bacterial behavior of micro conical dimple textured Ti6Al4V surface through micro-milling," *Surf. Interfaces*, vol. 21, no. 21, art. 100714, 2020. https://doi.org/10.1016/j.surfin.2020.100714

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e105266)

Research Article/Mechanical, Mechatronics, and Materials Engineering http://doi.org/10.15446/ing.investig.105266

Experimental Study of Innovative FRC Dome-Shaped Structures with Industrial, Recycled, and Alternative Reinforcing under Compressive Load

Estudio experimental de estructuras innovadoras en forma de cúpula de FRC con refuerzo industrial, reciclado y alternativo bajo cargas de compresión

Alejandro Meza-de Luna[1], Elia Mercedes Alonso-Guzman[2], and Adrián Bonilla-Petriciolet[3]

ABSTRACT

Arc concrete structures are aesthetic and suitable for buildings with large spans. Within the framework of this research, fiber-reinforced concrete (FRC) dome-shaped structures were studied in a lab, analyzing mechanical performance via strength, deformation, and failure mode. The studied FRC was elaborated with industrial, recycled, and alternative fibers. The mold used for producing the dome specimens was made up of two parts in order to favor extraction. This study considered 21 dome samples reinforced with fibers and one control, with replication for those with atypical behavior. The results show that the load-displacement behavior of dome-shaped elements increases with the incorporation of fibers, which depends on the fiber dose and the material. Moreover, the efficiency analysis proved that specimens with low fiber content (20 and 2 kg/m^3) have the best strength-reinforcement relation. The recycled and alternative fibers exhibit good mechanical performance and ductility. The main contribution of this research is a study of the scope and limitations of different types of FRC as the only reinforcement in arched structures.

Keywords: fiber-reinforced concrete, thick-walled shell, dome-shaped structures, experimental research, mechanical behavior, alternative and recycled materials

RESUMEN

Las estructuras de concreto en arco son estéticas y adecuadas para edificios de grandes claros. En el marco de esta investigación se estudiaron las estructuras en forma de domo de concreto reforzado con fibra (FRC) en un laboratorio, analizando el rendimiento mecánico a través de la resistencia, la deformación y el modo de falla. Los FRC estudiados fueron elaborados con fibras industriales, recicladas y alternativas. El molde utilizado para elaborar los especímenes de cúpula estuvo compuesto de dos secciones para favorecer la extracción. Este estudio consideró 21 muestras de domos reforzados con fibras y un control, con replicación sobre aquellas que presentaron un comportamiento atípico. Los resultados muestran que el comportamiento carga-desplazamiento de los elementos en forma de cúpula aumenta con la incorporación de fibras, lo cual depende de la dosis de fibra y el material. Además, el análisis de eficiencia demostró que las muestras con bajo contenido de fibra (20 y 2 kg/m^3) tienen la mejor relación resistencia-refuerzo. Las fibras recicladas y alternativas presentan buen desempeño mecánico y ductilidad. La principal contribución de esta investigación es el estudio de los alcances y limitaciones de diferentes tipos de FRC como único refuerzo en estructuras arqueadas.

Palabras clave: concreto reforzado con fibra, cascarón de paredes gruesas, estructuras en forma de cúpula, investigación experimental, comportamiento mecánico, materiales alternativos y reciclados

Received: October 13th, 2022
Accepted: August 17th, 2023

Introduction

Concrete is a material with high compressive strength. However, it is a ceramic material, so it exhibits brittle failure with low resistance under tensile load, which produces cracking and structural failures (Abhishek and Arabinda, 2020; Abbood *et al.*, 2020). This composite material has been the basis for complex structures, complying with the architectural ideas of different designers, such as the Felix Candelas hyperbolic-paraboloid structures (Pereira, 2018). Researchers have proven that curved structures have a high compressive strength. Usually, this shape is seen in sports stadiums, exhibition centers, transport terminals, airports, and religious precincts, highlighting their elegance and strength-weight relationship (Wenfeng *et al.*, 2019; Cocchetti *et al.*, 2011; Gomes *et al.*, 2018; Tomás *et al.*, 2010; Caycedo *et al.*, 2019).

---

[1] Mechanical engineer, PhD, Affiliation: Professor, Tecnológico Nacional de México/IT de Aguascalientes, México. E-mail: alejandro.meza@mail.ita.mx

[2] Civil engineer, PhD, Affiliation: Professor, FIC, Universidad Michoacana de San Nicolas de Hidalgo, México. E-mail: elia.alonso@umich.mx

[3] Chemical engineer, PhD, Affiliation: Tecnológico Nacional de México/IT Aguascalientes, México. E-mail: petriciolet@hotmail.com

A shell element is a curved structure with a wall thickness smaller than the diameter and height. It is classified into two branches: thin-shell and thick-shell (Berger *et al.*, 2020a; Berger *et al.*, 2020b). These structures are generally studied using two strategies: analysis by computer and full-scale experimentation. Caycedo *et al.* (2019), Verwimp *et al.* (2016), and Zingoni *et al.* (2013) evaluated the capacity of thin-walled shell elements using finite element analysis. Their results proved that these structural concrete elements have great plastic strain and buckling load, where the mechanical behavior depends on the structure geometry. Based on experimentation on thin shells with a span of 3 m, Du *et al.* (2019) highlighted the importance of the arc on the mechanical performance of concrete shell elements. Sharei *et al.* (2017) studied the structural performance of concrete shells reinforced with textile material, finding a good agreement between experimental and analytical studies. In addition, analytical studies have been conducted to understand the mechanical performance of concrete shell elements. Tamayo *et al.* (2013) and Chun (2005) analyzed the strength of concrete shell structures, and their results showed an approximation between the proposed mathematical model and the experimental data.

On the other hand, the search for light concrete structures with quick reinforcing has promoted the development of fiber-reinforced concrete (FRC). FRC is a composite material of cement, water, and aggregates reinforced with randomly distributed discrete elements. Fibers are used to enhance the mechanical performance of the concrete matrix, as they have a superior tensile strength to the base material (Meza *et al.*, 2021; Daud *et al.*, 2020; Meza and Shaikh, 2020; Meza and Siddique, 2019; Chiaia *et al.*, 2009; Cadoni *et al.*, 2009; Stähli *et al.*, 2008; Chiaia *et al.*, 2007). An example of the application of FRC with steel fibers is the hyperbolic paraboloid Oceanographic Museum in Valencia, Spain. Nevertheless, studies have reported that FRC has certain limitations, such as a scarce contribution of the fibers under compression load (Meza *et al.*, 2021; Abhishek and Arabinda, 2020; Daud *et al.*, 2020), low workability due to the incorporation of the fibers (Meza *et al.*, 2021; Abhishek and Arabinda, 2020; Daud *et al.*, 2020), and the high cost of the reinforcing elements (Meza *et al.*, 2021; Meza and Siddique, 2019; Emon *et al.*, 2017). As a solution to the high price of FRC, researchers have proposed using alternative and recycled fibers, aiming to reduce the cost of FRC and promote green engineering. Signorini and Volpini (2021) studied the mechanical performance of fiber-reinforced cement composites with fully recycled plastic elements. Their results showed that recycled fibers could sustain the economy of construction fields. Malek *et al.* (2020) analyzed the characteristics of recycled polypropylene fibers and reported variability in the mechanical responses of FRC using said elements. Ming *et al.* (2021) investigated the compressive behavior of recycled waste fibers in cement-based composites, concluding that steel waste fibers can provide similar mechanical properties to industrial ones.

This literature review indicates that the materials used in curve structures are plain concrete, concrete with a textile reinforcement, and concrete with conventional reinforcements. These structures are usually analyzed with finite element methods or full-scale experimentation. It is worth noting that the experimental study of FRC applied to dome-shaped structures is still a developing field that deals with topics such as the capacities of FRC in arc structures.

This research experimentally analyzed FRC in thick dome-shaped structures using seven different reinforcing elements including industrial, recycled, and alternative materials, varying the fiber doses in a total of 21 tests. This study considered 22 samples and replicated those with atypical behavior. The dome-shaped samples were tested under a central load until failure. The main objectives of this study were a) to analyze the mechanical capacities of FRC in dome-shaped structures, b) to study the effect of fiber dosage on mechanical properties, and c) to explore recycled and alternative materials as reinforcing elements in these structures. Based on these objectives, the authors aim to contribute to a better understanding of the scope and limitations of FRC with different reinforcing elements. In this study, the specimens had fibers as the only reinforcement. This technique is typical; structures typically use fibers as reinforcement elements without incorporating re-bars (*i.e.*, Oceanografic Valencia by Felix Candelas), thus reducing building times.

Material and methods

Fibers used as reinforcing elements

This study considered seven reinforcing elements, including two materials (steel and polypropylene), with diverse origins: industrial (those manufactured under a registered trademark), recycled (those obtained from discarded materials), and alternative (those generated from materials obtained in local stores). Furthermore, the fibers had different shapes, *i.e.*, end-hooked, straight, and crimped. Table 1 indicates the fiber characteristics of the reinforcing elements used.

Industrial fibers were selected because they have a higher mechanical capacity than other industrial reinforcing elements (Meza *et al.*, 2014). On the other hand, recycled PET fibers were produced using waste bottles, which were cut manually to form a plastic sheet and sectioned with a guillotine to make the reinforcing elements. The PET fibers had the following dimensions: a width of 3 mm, a thickness of 0,3, and a length of 23,5 mm. The methodology followed to produce the PET fibers was similar to that of a previous study (Meza and Siddique, 2019). Moreover, alternative fibers were made from commercial wire steel (annealed and galvanized). The annealed wire had a diameter of 1,78 mm, with a composition of 0,1% carbon, 0,04% phosphorus, 0,30-0,50% magnesium, and 0,05% sulfur. The galvanized wire had a diameter of 1,2 mm and was composed of 0,07%

carbon, 0,03-0,04% phosphorus, 0,30-0,31% magnesium, 0,035-0,045% sulfur, and 0,075-0,085% silicon. The wire had a 250 kg/m^2 zinc layer.

The fibers (alternative annealed, alternative galvanized, and recycled PET) were crimped with the device shown in Figure 1, which was developed by A. Meza and the research group. The crimped device has a guide, two gears, and a crank. The two gears have a 2,5 module, with 22 teeth 3,75 mm in height and a crank to activate the crimped gears in the rear.

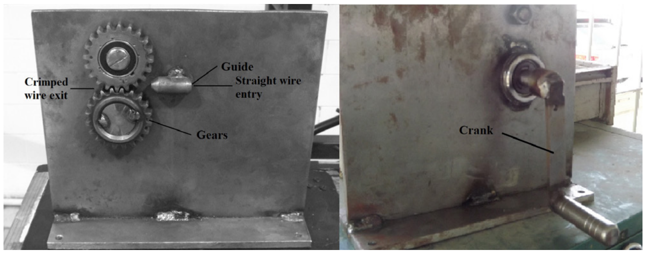


Figure 1. Front and rear view of the crimping device
Source: Authors

Table 1. Fibers used

Fiber series	Origin	Shape	Material	Aspect ratio (mm/mm)	Tensile strength (MPa)
A1	Alternative	Crimped	Annealed steel	22	340
A2	Alternative	Crimped	Galvanized steel	22	479
R1	Recycled	Straight	PET	22	250
R2	Recycled	Crimped	PET	22	250
D1	Dramix	Hooked-end	Steel	66,6	1 160
E1	Master-fiber	Straight-microfiber	Polypropylene	NA	300
E2	Tuf-strand-sf-Euclid	Straight-macrofiber	Polypropylene	74	625

The tensile strength was extracted from Meza and Shaikh (2020), Meza and Siddique (2019), and Meza *et al.* (2014).

Source: Authors

Experimental design

The concrete design consisted of ordinary Portland cement CPC 30R, in compliance with the NMX-C-414-ONNCE Mexican norm (2017) and two aggregates (fine and coarse river sand). The characteristics of these materials are as follows: i) the cement had a density of 3,15 g/cm^3; ii) the fine aggregate had a density of 2,5 g/cm^3, a fineness modulus of 2,8, and a water absorption of 3,3%; and iii) the coarse aggregate had a density of 2,68 g/cm^3, a maximum size of 20 mm (3/4 in), and a water absorption of 1,1%.

The experiments included a control and 21 fiber-reinforced concrete samples, employing seven reinforcement types and three dose levels. Samples with atypical behavior were repeated. Atypical values were evaluated using both the trends of the tested fiber-reinforced concrete and literature reports via a normalized test. Changes in the fiber aspect ratio were not evaluated because the fiber dose had a greater impact on mechanical characteristics (Meza *et al.*, 2021; Abhishek and Arabinda, 2020; Daud *et al.*, 2020; Meza and Shaikh, 2020; Chiaia *et al.*, 2019; Meza and Siddique, 2019; Meza *et al.*, 2014). There were four dome-shaped specimens with atypical behavior due to issues related to preparation or testing. Furthermore, this performance could be attributed to the random distribution of the fibers in the concrete matrix. It is essential to indicate no trend was observed in the atypical samples. Table 2 shows the proportions of the 22 different mixes used. An asterisk denotes the samples that showed atypical behavior and were repeated.

Table 2. Description of the proportions used in the experiments of this study

Specimen series	Fiber series	Fiber dose, kg/m^3	w/c ratio	Cement, kg/m^3	Fine aggregate, kg/m^3	Coarse aggregate, kg/m^3	Additive, kg/m^3	Slump, mm
C.0	---	---	0,5	383	672	1 100	10	90
A1.20	A1	20	0,5	383	672	1 100	10	50
A1.30*	A1	30	0,5	383	672	1 100	10	35
A1.40	A1	40	0,5	383	672	1 100	10	20
A2.20	A2	20	0,5	383	672	1 100	10	70
A2.30	A2	30	0,5	383	672	1 100	10	62
A2.40	A2	40	0,5	383	672	1 100	10	55
R1.2	R1	2	0,5	383	672	1 100	10	79
R1.4	R1	4	0,5	383	672	1 100	10	80
R1.6*	R1	6	0,5	383	672	1 100	10	75
R2.2	R2	2	0,5	383	672	1 100	10	65
R2.4	R2	4	0,5	383	672	1 100	10	63
R2.6	R2	6	0,5	383	672	1 100	10	60
D1.20*	D1	20	0,5	383	672	1 100	10	75
D1.30	D1	30	0,5	383	672	1 100	10	68
D1.40	D1	40	0,5	383	672	1 100	10	59
E1.2	E1	2	0,5	383	672	1 100	10	85
E1.4	E1	4	0,5	383	672	1 100	10	80
E1.6	E1	6	0,5	383	672	1 100	10	83
E2.2*	E2	2	0,5	383	672	1 100	10	80
E2.4	E2	4	0,5	383	672	1 100	10	72
E2.6	E2	6	0,5	383	672	1 100	10	68

Source: Authors

Specimen mixing, casting, and curing

First, the materials (cement, aggregates, fibers, additive, and water) were weighed under lab conditions. Second, the materials were manually stirred for up to 3 min without water. The water was added, and the materials were mixed for another 3 min to achieve homogeneity. Finally, the specimens were cast in steel molds, demolded and cured for 28 days, in accordance with ASTM C192 (ASTM International, 2017).

Mold design and testing

The process of elaborating the dome-shaped concrete specimens was based on female and male molds. The female mold controlled the outer shape of the dome sample, and the male mold controlled the inner one. Figure 2 shows the sketch of the molds and the concrete dome element.

The demolding of the dome sample was simulated using SolidWorks software to ensure a correct extraction (Figure 3). The analysis suggested a corner rounding for the two sections (in black) located in the bottom fillets of the dome (inner and outer edges). Grey sections indicate negative and positive draft angles, implying an adequate extraction of the hardened sample. Figure 3 shows the demolding analysis results.

The mold used to create the dome-shaped samples was made of stainless steel and had a thickness of 1,21 mm (this material was workable and allowed shaping the dome mold). The mold had two parts (male and female). The female mold was sectioned into two parts, aiming to facilitate the removal of the dome specimen when it hardened. The mold also had two spacers that ensured a uniform thickness. The male mold was a single element (without a section) with the same dimensions as the inner shape of the specimen. At the bottom of the male mold was a handle that allowed extracting the hardened sample. Figure 4 shows the male and female molds.

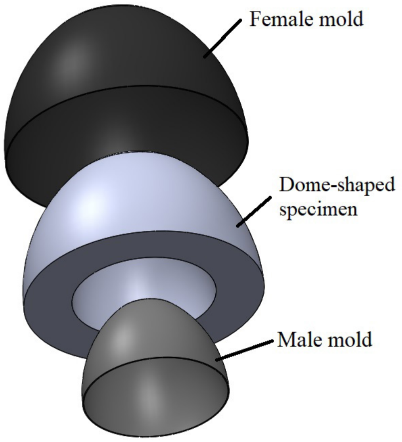


Figure 2. Sketch of the molds proposed to produce the dome-shaped specimen
Source: Authors

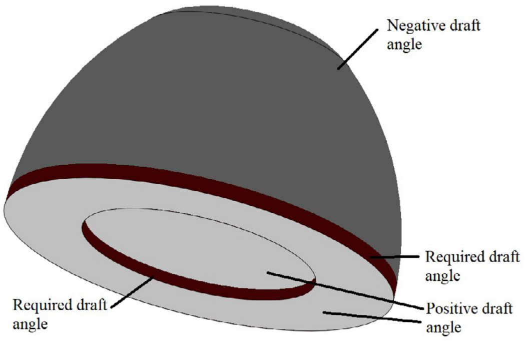


Figure 3. Dome extraction analysis using SolidWorks
Source: Authors

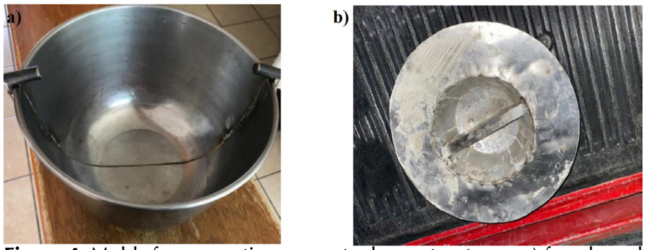


Figure 4. Molds for generating concrete dome structures: a) female and b) male
Source: Authors

The male and female molds were assembled and placed on a flat surface. The open section was facing up, and the closed one was facing down to take advantage of the force of gravity and prevent the dispersion of the material. Then, the mixture was settled into the mold, filling the third part of the sample's total volume, and it was then compacted to avoid air encapsulation. The process was repeated until the concrete covered all the mold. Compaction was performed with a flat rod 16 mm in diameter and 24 mm in length and rounded tamping ends. The rod of each layer was uniformly distributed over the cross-section of each layer with 25 strokes. Finally, the specimen was covered with a plastic sheet to reduce moisture loss. The dome-shaped sample was demolded after 48 h and cured according to the ASTM C192 (ASTM International, 2017). Figure 5 shows the procedure followed to elaborate the specimens.

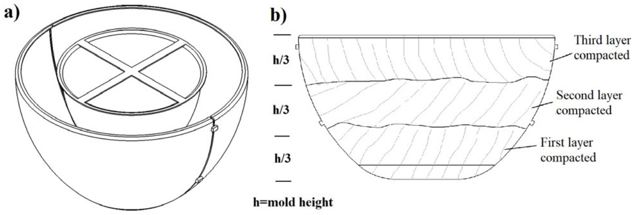


Figure 5. a) Assembly of male and female molds; b) compaction procedure
Source: Authors

Figure 6 shows the detailed geometric dimensions of the dome specimens. The demolding issues observed in the analysis were solved using 3 mm fillets. The dome-shaped elements have the following dimensions: 50 mm thickness, 150 mm height, 250 mm outer diameter, and a ratio (ratio/ thickness) of 2,5. A thick shell was considered (Gere, 2009). The upper section of the dome specimen has a flat 58 mm part, where the load was applied.

A 59 kN Galdabini hydraulic machine was used during sample testing, which was controlled at 5 mm/min. The possible irregularities in the contact surface between the top of the dome specimen and the load were adjusted using a neoprene pad. The neoprene pad had a diameter of 100 mm and a thickness of 25 mm. Specimen deformation was measured via a plunger dial indicator with a measuring range of 25 mm, a resolution of 0,25 mm, and an accuracy of ±0,25 μm. The load was on the top of the structure, simulating a dead punctual load. Figure 7 shows the test setup of the dome specimen.

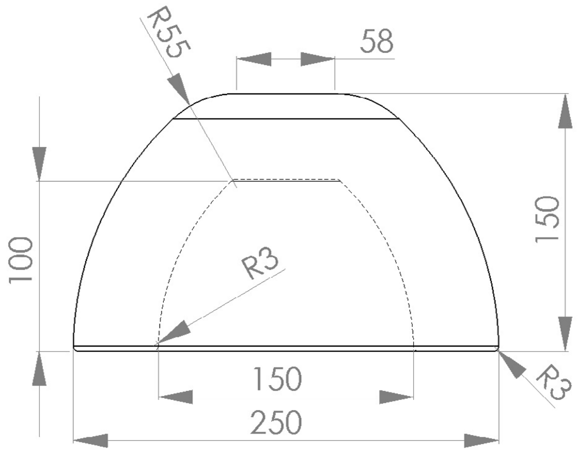


Figure 6. Geometric dimensions of the dome-shaped structure (dimensions in mm)
Source: Authors

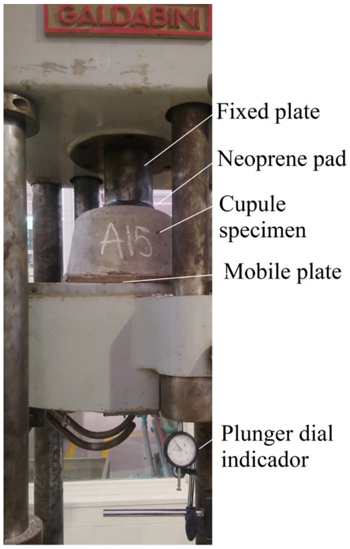


Figure 7. Test setup of the cupule specimen
Source: Authors

Results and discussion

Slump

The properties of the concrete were evaluated in the fresh state while following the standard test method for the flow of FRC through an inverted slump cone, according to ASTM C995 (ASTM International, 2018). Table 2 shows the results of the slump tests for batches with different fibers and without fibers (control). The data show a reduced workability due to the addition of the fibers, which depends on fiber material and dosage, where the minimum difference was 6% and the maximum one was 78% with respect to the control sample. The reduction in the workability of FRC is a normal phenomenon that has been reported in other studies (Meza *et al*., 2021; Meza and Siddique, 2019). The maximum and minimum slumps were found on samples with polymeric microfibers and annealed crimped fibers. The highest workability corresponded to polymeric microfibers and is attributed to the flexibility of the reinforcing elements, an effect that has also been previously reported (Meza *et al*., 2021).

On the other hand, the lowest slump, observed in concrete with annealed steel fibers, was related to the ability to trap water on the wire surface. Figure 8 shows the surface conditions of the galvanized and annealed steel wires with a 300X zoom lens. There was a higher slump on the annealed surface than on the galvanized one, with hints of corrosion pitting due to the ability to trap the humidity of the environment. This surface condition on annealed wire absorbed the moisture of the concrete mix, thus reducing workability. In contrast, the zinc layer on the surface of the galvanized wire entailed a smooth surface. Moreover, the comparison between the industrial, recycled, and alternative fibers indicated a superior slump performance for industrial vs. recycled or alternative elements. In some cases, the reduction was scarce, i.e., concrete with galvanized reinforcing elements showed a mean workability decrease of 7% with respect to Dramix elements. However, other samples exhibited a considerable slump change: the difference between Dramix and annealed fibers was 48%. Finally, the batches with polymer fibers (industrial and recycled) were similar, with a difference lower than 24%.

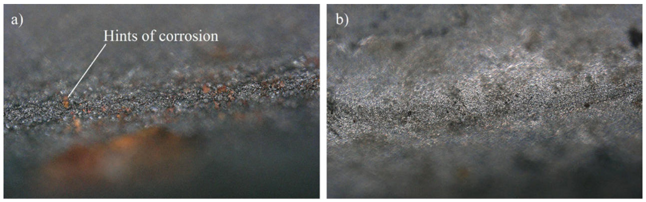


Figure 8. Surface images of a) annealed and b) galvanized steel
Source: Authors

Dome-shaped samples

Load deflection

Figure 9 presents the load displacement curves for the tested samples. These curves show that the dome-shaped structures, with or without fibers, started with an elastoplastic behavior, followed by a nonlinear load-displacement relation, with different strength and deformation ranges, attributed to the incorporation of the reinforcing elements, the dose, and efficiency. The following sections analyze aspects such as stiffness, peak load, the ductility index, and toughness.

Stiffness

Equation (1) shows the relation used to calculate the relative stiffness, which is based on the NMX-C-128-ONNCCE norm (2013) and was adapted to this study. Table 3 shows the stiffness results, with the data indicating that all the fiber-reinforced dome-shaped structures had a higher initial stiffness than the control specimen. However, it must be pointed out that the change in fiber dose had a low contribution to stiffness, with a maximum difference of 22%.

The highest stiffness values were found for dome-shaped steel FRC. Annealed samples had the best performance, which can be attributed to their rough surface and larger cross-section in comparison with the other steel fibers tested. Regarding stiffness, the samples with annealed fibers were followed by those with industrial and galvanized steel ones. On the other hand, the samples with polypropylene fibers showed a lower stiffness than those with steel fibers, which can be attributed to their high flexibility and ductility (Meza *et al*., 2021). The lowest values were reported by samples with straight PET fibers, exhibiting a higher stiffness (7%) with respect to the control.

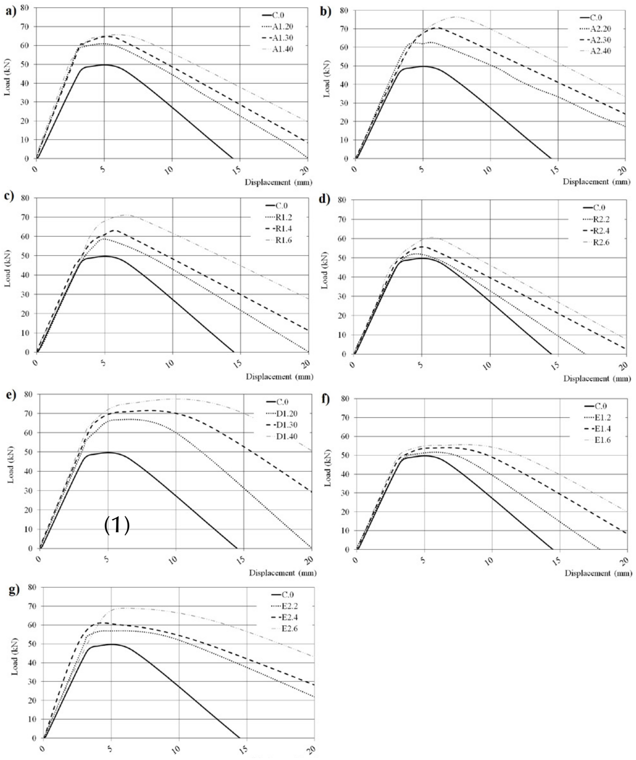


Figure 9. Load-deflection curves of dome-shaped concrete samples reinforced with a) annealed crimped steel fibers, b) galvanized crimped steel fibers, c) straight PET fibers, d) crimped PET fibers, e) industrial hooked-end steel fibers, f) straight polymeric microfibers, and g) straight polymeric macro fibers
Source: Authors

$$S_o = \frac{L_o^{40\%} - L_o^{0,00005}}{\delta_o^{40\%} - \delta_o^{0,00005}} \qquad (1)$$

where S_o is the stiffness, $L_o^{40\%}$ is the load corresponding to 40% of the ultimate value, $L_o^{0,00005}$ is the load corresponding to 0,00005 of the ultimate value, $\delta_o^{40\%}$ is the displacement corresponding to 40% of the ultimate value, and $\delta_o^{0,00005}$ is the displacement corresponding to 0,00005 of the ultimate value

Peak load

Table 3 shows the peak load results for samples with and without fibers. The data indicate that the peak load capacity increases by incorporating reinforcing elements into the concrete. In fact, all the specimens with fibers exhibited a superior performance when compared to the control. The highest difference was 54% for the sample with galvanized crimped steel fibers and a fiber dose of 40 kg/m^3, while the lowest difference was reported by a dose of 2 kg/m^3 of industrial polymeric microfibers, with a maximum difference of 4% regarding the control sample. On the other hand, the mean responses of the samples with fibers indicated the superior performance of dome-shaped specimens with steel elements, followed by those with polymer reinforcements. Moreover, changing the fiber dose had a low effect on the maximum load (lower than 23%).

Table 3. Mechanical behavior of the dome-shaped specimens

ID	Stiffness, MN/m	Peak load, kN	Ductility index, mm/mm	Toughness, J
C.0	14,13	48,8	1,2	411,2
A1.20	19,06	60,3	1,8	710,4
A1.30	18,5	63,8	2,3	787,5
A1.40	20,15	65,0	2,8	952,3
A2.20	16,06	62,4	2,1	841,7
A2.30	15,02	70,7	2,6	990,8
A2.40	14,99	76,5	3,2	1 218,8
R1.2	14,2	58,6	1,6	635,8
R1.4	16,11	62,7	2,0	774,5
R1.6	14,99	70,0	2,1	1 066,2
R2.2	14,91	52,0	1,5	493,2
R2.4	15,54	55,1	1,8	609,7
R2.6	15,6	60,5	2,2	731,1
D1.20	16,3	66,7	1,7	795,2
D1.30	17,06	70,8	1,9	1 121,7
D1.40	17,98	75,0	2,0	1 486,3
E1.2	14,41	50,5	1,3	546,4
E1.4	15,75	53,9	1,9	723,5
E1.6	16,46	55,3	2,1	869,9
E2.2	16,71	56,9	1,5	889,2
E2.4	18,78	60,0	2,0	1 026,5
E2.6	14,95	68,9	2,3	1 389,5

Source: Authors

Ductility index

Nehdi *et al*. (2015) proposed calculating the ductility index by dividing the displacement at peak load by the corresponding cracking load. The researcher uses the peak load as a reference because concrete's mechanical performance resembles that of a fragile material. Thus, it is more feasible to localize the peak load than the yield displacement. Table 3 shows the results of this ratio applied to FRC dome-shaped samples. The data indicate that the FRC specimens have a higher ductility index than the control. The maximum and minimum differences are found for crimped steel fibers and industrial polymeric microfibers, with doses of 40 and 2 kg/m^3, respectively. The differences reach 2,2 and 1,4 times the control sample's capacity. These results also prove that increasing the fiber dose enhances the ductility index of dome-shaped structures by up to 61%. The mean data show that polymeric fibers have a better ductility index than steel ones. This behavior was contrary to that of the stiffness and peak load. In comparison with the data reported by Nehdi *et al.* (2015), the range of the ductility index in FRC dome-shaped structures was similar to that of FRC tunnel structures.

Toughness

Table 3 shows the toughness of the dome-shaped samples, which was measured through the energy calculated based on the area under the load-displacement curve until failure. The data indicate that the toughness of the dome-shaped samples depends on the characteristics of the reinforcement, i.e., all the fiber-reinforced specimens exhibited a better energy absorption than the control, with a difference that increased with the fiber amount. The highest toughness was reported by the specimen with industrial polymeric macrofibers at a dose of 6 kg/m^3, with a difference of 3,6 times the capacity of the control sample. In contrast, the lowest difference corresponded to the specimen with crimped PET fibers at a dose of 2 kg/m^3, with a superior toughness of 20% regarding the control. Moreover, the data reported the best toughness values for dome-shaped elements with industrial fibers (steel and polymeric macrofibers), followed by samples with alternative and recycled ones. Nehdi *et al.* (2015) reported that the increase in the toughness of concrete tunnel specimens reinforced with steel fibers amounts to 95% compared to control samples. In this study, some FRC dome-shaped samples exceed the records presented by Nehdi *et al.* (2015), thus proving the contribution of fibers to dome elements.

FEM study

The failure mode of the dome structure was studied using the finite element method (FEM). The aim was to study the stress gradients and the area where the cracks couldstart up. Only linear properties were considered. The FEM study was static, using a first-order tetra mesh with four Jacobian points, a maximum element size of 17,7 mm, and a total of nodes and elements of 11 367 and 7 223, respectively. The concrete properties were isotropic linear elastic, with a modulus of elasticity of 13 GPa, a Poisson ratio of 0,2, and a density of 2 400 kg/m^3 (based on the data indicated by Meza *et al.* in 2021). The movement of the dome structure was constrained in its low section in the z-axis (perpendicular to the load direction). The other axes (x-y) were not constrained, simulating the conditions of the tests. The load was applied in a perpendicular direction (15 kN) onto the superior face of the dome structure, with a diameter of 58 mm. Figure 10 shows the FEM results, where the maximum stress was observed in the upper and lower corners of the dome sections, corresponding to the load application and holder areas. During the test, the failure started in the top section, where the stress was higher, and moved to the opposite zone. Moreover, the dome samples with high fiber contents demonstrated a higher formation of microcracks than those without fibers or with low proportions. However, the fibers delayed crack propagation in the dome specimens, enhancing the rate of load recovery after cracking. Bhosale *et al.* (2020) and Adel *et al.* (2021) studied the crack behavior of concrete structures, indicating a similar performance (the fibers controlled crack propagation but were more susceptible to a high number of micro-cracks). This phenomenon was attributed to the presence of air in the interface of the concrete fiber.

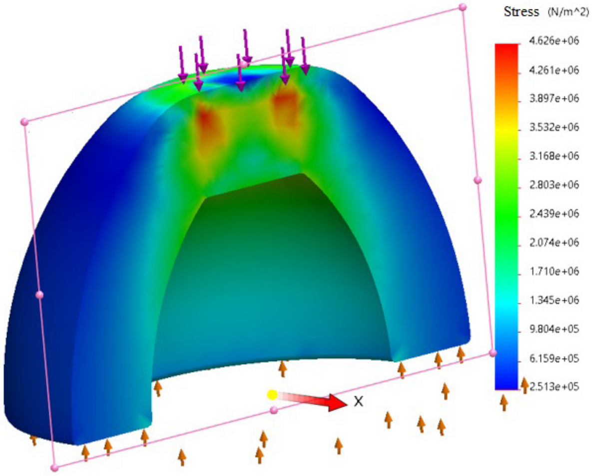


Figure 10. a) FEM study of the dome-shaped sample; b) cracking of the dome-shaped sample
Source: Authors

Cracking behavior

All the dome-shaped samples exhibited a progressive failure, including those without fibers. This effect was attributed to the arc shape of the specimens, which maintained the load after the first crack and gradually declined. Figure 11 shows the procedure observed during the cracking of the dome-shaped specimens, which had three stages. Stage I was characterized by an elastoplastic behavior, observed with a semi-linear behavior in the load-displacement graph. Stage II showed a nonlinear load-displacement relation, reaching the maximum load zone. Stage III corresponded to a decreasing segment in the load-defection curve, which depended on the ductility behavior of the dome structure until the failure point. The range of strength and deformation is different in each part of the graphs, which can attributed to the incorporation of fibers, their dose, and their efficiency.

Fiber efficiency

Equation (2) shows the relation used to calculate the fiber efficiency, which compares the capacity of the FRC dome-shaped structure against that of the control sample. Moreover, the Equation considers the fiber dose incorporated into the mix. Figure 12 shows the results obtained for the fiber efficiency, in correspondence with the peak load, ductility index, and toughness. Since the fiber dose depends on the density of the fiber materials, the analysis is depicted in two groups (steel and polymeric FRC samples).

$$\eta_f = \frac{\varphi_f - \varphi_c}{\varphi_f D_f} \tag{2}$$

where η_f is the fiber efficiency, φ_f is the capacity of the dome-shaped structure reinforced with fibers, φ_c is the capacity of the control dome-shaped structure, and D_f is the fiber dose.

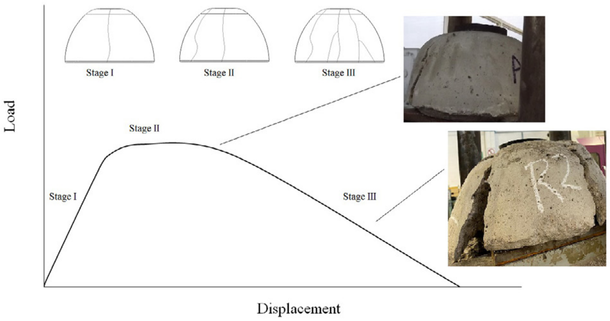
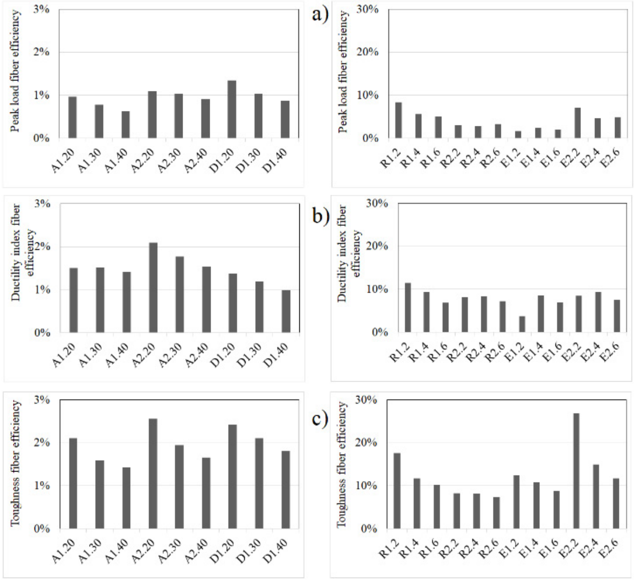


Figure 11. Cracking process of the dome-shaped samples, general performance
Source: Authors

Figure 12. Fiber efficiency of dome-shaped samples: a) peak load, b) ductility index, and c) toughness



Source: Authors

Fiber efficiency and peak load

Figure 12a shows the results obtained for the fiber efficiency regarding the peak load of dome samples. The data indicate that steel FRC with annealed, galvanized, and industrial elements and a low fiber dose is more effective than that with a high proportion of reinforcing elements. Furthermore, the dome-shaped samples with straight PET fibers and industrial polymeric macrofibers showed a similar trend to elements with steel fibers (the specimens with a low fiber content had high efficiency). This phenomenon was attributed to i) the nonlinear increase in the peak load, which was not proportional to the fiber dose; and ii) the fact that the mechanical strength of the fiber was higher than its adherence to the concrete. This result agreed with other studies, which indicate that the fiber material has a greater influence than the adherence between the fiber and the concrete (Meza and Shaikh, 2020; Meza *et al.*, 2004; Poveda *et al.*, 2020; Qi *et al.*, 2018; Breitenbücher *et al.*, 2014). On the other hand, the data show that dome-shaped specimens with industrial reinforcements have the best efficiency, followed by those with galvanized and annealed ones. In contrast, the samples with polymeric fibers show the best behavior with recycled, straight elements, followed by those with industrial polymeric macrofibers, recycled crimped PET, and industrial polypropylene microfibers.

Fiber efficiency and ductility

Figure 12b shows the results regarding fiber efficiency and ductility. The overall results indicate that dome-shaped elements reinforced with straight PET and with crimped annealed fibers perform better than the other fibers tested. Moreover, the data prove that samples reinforced with high steel fiber contents have reduced efficiency, i.e., the lowest efficiency was reported by samples A1.40, A2.40, and D1.40. In contrast, samples with polymeric reinforcement show no trend with respect to fiber efficiency. Meanwhile, the elements with straight PET fibers exhibit a similar behavior to those with steel ones (the best efficiency reported for samples with low fiber contents), and the other dome samples show different performances. Specimens with crimped PET fibers, industrial polymeric macrofibers, and polymeric microfibers show the best efficiency at a central dose (4 kg/m^3).

Fiber efficiency in the toughness

The data show better efficiencies in dome-shaped samples with a low proportion of reinforcing elements. The global results indicate that the specimens with alternative annealed and industrial fibers perform better than those with galvanized ones. This was attributed to the mechanical properties of the fibers, i.e., the deformation capacity of annealed steel and the strength of industrial steel fibers (Ming *et al.*, 2021; Meza *et al.*, 2014). On the other hand, the specimens with polymeric industrial macrofibers showed the best efficiency, followed by those with recycled straight PET fibers, industrial microfibers, and recycled crimped PET fibers. These results are similar to those of previous studies, which point out that macrofibers and recycled PET have adequate adherence to concrete and tensional strength (Ming *et al.*, 2021; Meza *et al.*, 2021). Figure 12c shows the fiber efficiency results regarding the dome samples' toughness.

Conclusions

The experimental results obtained for the dome-shaped samples show the load-displacement curves to have different ranges of strength and deformation, which is attributed to the incorporation of fibers, their dose, and their efficiency, with higher mechanical performances than the control. Each fiber type showed both advantages and limitations in the different stages of this research, as is presented below.

I: The data showed a reduced workability due to the addition of fibers, which was higher on FRC with annealed elements (78% regarding the control) and is attributed to the ability to trap water on its surface.

II: The FEM study located the maximum stress at the edges of the load and the holder sections of the dome-shaped specimen, where the cracks start up and grow.

III: The samples with high fiber contents showed a larger number of microcracks than the control and those with a low fiber proportion. However, the reinforcing elements controlled the cracks and increased the ductility of the structural elements.

IV: The fiber efficiency study pointed out that samples with a low fiber content (20 and 2 kg/m^3) had a higher mechanical efficiency than those with a high proportion (40 and 6 kg/m^3).

V: The dome shape maintained the load after the first crack and declined gradually according to the reinforcement capacity, highlighting the mechanical performance of industrial hooked steel fibers, alternative crimped annealed fibers, alternative crimped galvanized fibers, industrial polymeric macrofibers, and recycled straight PET fibers.

Acknowledgments

The authors wish to acknowledge Tecnológico Nacional de México and IT Aguascalientes for the support provided to this research and the access granted to the laboratory for testing composite materials. Moreover, the authors wish to thank ME Rodolfo B. Sierra Ortiz, Eng. Fernando Chavez Valdivia, and Eng. Rafael Reyes Cortés for participating in this study.

CRediT author statement

Alejandro Meza de Luna: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing (original draft, review, and editing).

Elia M. Alonso and Adrián Bonilla: conceptualization, investigation, methodology, supervision, validation, visualization, writing (review and editing).

Conflicts of interest

The authors declare that no known financial interest or relationship could have influenced the work reported in this paper.

References

Abbood, I. S., Odda, S. A., Hasan, K. F., and Jasim, M. A. (2020). Properties evaluation of fiber reinforced polymers and their constituent materials used in structures – A review. *Materials Today: Proceedings, 43*(2), 1003-1008.

Abhishek, J., and Arabinda, S. (2020). Experimental study on the properties of steel fibre reinforced concrete. *Indian Journal of Engineering, 17*(47), 151-162.

Adel, M., Matsumoto, K., and Nagai, K. (2021). Crack-bridging degradation and evolution in SFRC structural beams under variable amplitude flexural cyclic loading. *Computers and Structures, 272*, 114176. https://doi.org/10.1016/j.compstruct.2021.114176

ASTM International (2017). *ASTM C192: Standard practice for making and curing concrete test specimens in the laboratory*. ASTM International.

ASTM International (2018). *ASTM C39/C39M: Standard test method for compressive strength of cylindrical concrete specimens*. ASTM International.

ASTM International (2018). ASTM C78/C78M Standard test method for flexural strength of concrete (using simple beam with third-point loading). ASTM International.

ASTM International (2018). *ASTM C995-01: Standard test method for time of flow of fiber-reinforced concrete through inverted slump cone*. ASTM International.

Berger, J., Gericke, O., Feix, J., and Sobek, W. (2020). Actively bent concrete shells. *Structural Concrete, 1*(11), 2282-2292. https://doi.org/10.1002/suco.201900505

Berger, J., Gericke, O., Feix, J., and Sobek, W. (2020). Experimental investigations on actively bent concrete shells. *Structural Concrete, 21*(6), 2268-2281. https://doi.org/10.1002/suco.202000045

Bhosale, A. B., and Prakash, S. S. (2020). Crack propagation analysis of synthetic *vs.* steel *vs.* hybrid fibre-reinforced concrete beams using digital image correlation technique. *International Journal of Concrete Structures and Materials, 14*, 57. https://doi.org/10.1186/s40069-020-00427-8

Breitenbücher, R., Meschke, G., Song, F., and Zhan, Y. (2014). Experimental, analytical and numerical analysis of the pullout behaviour of steel fibres considering different fibre types, inclinations and concrete strengths. *Structural Concrete, 15*(2), 126-135. https://doi.org/10.1002/suco.201300058

Cadoni, E., Meda, A., and Plizzari, G.A. (2009). Tensile behaviour of FRC under high strain-rate. *Materials and Structures, 42*, 1283-1294. https://doi.org/10.1617/s11527-009-9527-6

Caycedo, M. S., Siqueira, G. H., Vieira, L. C. M., and Vizotto, I. (2019). Evaluation of structural capacity of triangular and hexagonal reinforced concrete free-form shells. *Engineering Structures, 188*, 519-537. https://doi.org/10.1016/j.engstruct.2019.03.044

Chiaia, B., Fantilli, A. P., and Vallini, P. (2007). Evaluation of minimum reinforcement ratio in FRC members and application to tunnel linings. *Materials and Structures, 40*, 593-604. https://doi.org/10.1617/s11527-006-9166-0

Chiaia, B., Fantilli, A. P., and Vallini P. (2009). Evaluation of crack width in FRC structures and application to tunnel linings. *Materials and Structures, 42*, 339. https://doi.org/10.1617/s11527-008-9385-7

Chun, N. H. (2005). Ultimate strength of large scale reinforced concrete thin shell structures. *Thin Walled Structures, 43*, 1418-144.

Cocchetti, G., Colasante, G., and Rizzi, E. (2011). On the analysis of minimum thickness in circular masonry arches. *Applied Mechanics Reviews, 64*(5), 050802. https://doi.org/10.1115/1.4007417

Daud, R. A., Daud, S. A., and Al-Azzawi, A. A. (2020). Tension stiffening evaluation of steel fibre concrete beams with smooth and deformed reinforcement. *Journal of King Saud University, 33*(3), 147-152. http://doi.org/10.1016/j.jksues.2020.03.002

Du, W., Liu, Q., Zhou, Z., Uddin, N. (2019). Experimental investigation of innovative composite folded thin cylindrical concrete shell structures. Thin-Walled Structures, 137 (2019), 224–230. https://doi.org/10.1016/j.tws.2019.01.014

Emon, M. A. B., Manzur, T., and Sharif, M. S. (2017). Suitability of locally manufactured galvanized iron (GI) wire fiber as reinforcing fiber in brick chip concrete. *Case Studies in Construction Materials, 7*, 217-227. https://doi.org/10.1016/j.cscm.2017.08.003

Gere, M. J. and Barry, J. G. (2009). *Mechanics of materials* (7th ed.). Cengage Learing.

Gomes, C., Parente, M., Azenha, M., and Lino, J.C. (2018). An integrated framework for multi-criteria optimization of thin concrete shells at early design stages. Advanced Engineering Informatics, 38, 330–342. https://doi.org/10.1016/j.aei.2018.08.003

Małek, M., Jackowski, M., Łasica, W., and Kadela, M. (2020). Characteristics of recycled polypropylene fibers as an addition to concrete fabrication based on Portland cement. *Materials, 13*(8), 1827. https://doi.org/10.3390/ma13081827

Meza, A., Ortiz, J. A., Peralta, L., Pacheco, J., Soto, J. J., Rangel, S. H., Padilla, R., and Alvarado, J. (2014). Experimental mechanical characterization of steel and polypropylene fiber reinforced concrete. *Revista Técnica Facultad de Ingeniería Universidad de Zulia, 37*(2), 106–115.

Meza, A., Pujadas, P., Meza, L. M., Pardo-Bosch, F., and López-Carreño, R. D. (2021). mechanical optimization of concrete with recycled PET fibres based on a statistical-experimental study. *Materials, 14*(2), 240. https://doi.org/10.3390/ma14020240

Meza, A. and Shaikh, F. U. A. (2020). Anisotropy and bond behaviour of recycled polyethylene terephthalate (PET) fibre as concrete reinforcement. *Construction and Building Materials, 265*, 120331. https://doi.org/10.1016/j.conbuildmat.2020.120331

Meza, A., and Siddique, S. (2019). Effect of aspect ratio and dosage on the flexural response of FRC with recycled fiber. *Construction and Building Materials, 213*, 286-291. https://doi.org/10.1016/j.conbuildmat.2019.04.081

Ming, Y., Chen, P., Li, L., Gan, G., and Pan, G. (2021). A comprehensive review on the utilization of recycled waste fibers in cement-based composites. *Materials, 14*(13), 3643. https://doi.org/doi.org/10.3390/ma14133643

Nehdi, M. L., Abbas, S., and Soliman, A. M. (2015). Exploratory study of ultra-high performance fiber reinforced concrete tunnel lining segments with varying steel fiber lengths and dosages. *Engineering Structures, 101*, 733-742. https://doi.org/10.1016/j.engstruct.2015.07.012

NMX-C-128-ONNCCE (2013). *Determination of modulus of elasticity and Poisson ratio*. https://www.dof.gob.mx/nota_detalle.php?codigo=5577342&fecha=31/10/2019#gsc.tab=0

NMX-C-414-ONNCCE (2017). *Construction industry-hydraulic cementants-specifications and test methods*. https://www.onncce.org.mx/es/venta-normas/fichas-tecnicas?view=item&id=1879

Pereira, M. (2022, September 10). Cáscaras de hormigón: principios de diseño y ejemplos construidos. *Arch Daily*. https://www.archdaily.mx/mx/895405/cascaras-de-hormigon-principios-de-diseno-y-ejemplos-construidos

Poveda, E., Yu, R. C., Tarifa, M., Ruiz, G., Cunha, V. M. C. F., and Barros, J. A. O. (2020). Rate effect in inclined fibre pullout for smooth and hooked-end fibres: A numerical study. *International Journal of Fracture, 223*, 135-149. https://doi.org/10.1007/s10704-019-00404-7

Qi, J., Wu, Z., Zhongguo, J.M., and Wang, J. (2018). Pullout behavior of straight and hooked-end steel fibers in UHPC matrix with various embedded angles. *Construction and Building Materials, 191*, 764-774. https://doi.org/10.1016/j.conbuildmat.2018.10.067

Sharei, E., Scholzen, A., Hegger, J., and Chudoba, R. (2017). Structural behavior of a lightweight, textile-reinforced concrete barrel vault shell. *Composite Structures, 171*, 505-514. https://doi.org/10.1016/j.compstruct.2017.03.069

Signorini, C., and Volpini, V.D. (2021). Mechanical performance of fiber reinforced cement composites including fully-recycled plastic fibers. *Fibers, 9*(3), 16. https://doi.org/10.3390/fib9030016

Stähli, P., Custer, R., and Van-Mier, J. G. M. (2008). On flow properties, fibre distribution, fibre orientation and flexural behaviour of FRC. *Materials and Structures, 41*, 189-196. https://doi.org/10.1617/s11527-007-9229-x

Tamayo, J. L. P., Morsch, I. B., and Awruch, A. M. (2013). Static and dynamic analysis of reinforced concrete shells. *Latin American Journal of Solids and Structures, 10*, 1109-1134.

Tomás, A., and Martí, P. (2010). Shape and size optimisation of concrete shells. *Engineering Structures, 32*(6), 1650-1658. https://doi.org/10.1016/j.engstruct.2010.02.013

Verwimp, E., Tysmans, T., Mollaert, M., and Wozniak, M. (2016). Prediction of the buckling behavior of thin cement composite shells: Parameter study. *Thin Walled Structures, 108*, 20-29. https://doi.org/10.1016/j.tws.2016.07.011

Wenfeng, D., Qi, L., Zhiyong, Z., and Nasim, U. (2019). Experimental investigation of innovative composite folded thin cylindrical concrete shell structures. *Thin Walled Structures, 137*, 224-230. https://doi.org/10.1016/j.tws.2019.01.014

Zingoni, A., Mudenda, K., French, V., and Mokhothu, B. (2013). Buckling strength of thin-shell concrete arch dams. *Thin Walled Structures, 64*, 94-102. https://doi.org/10.1016/j.tws.2012.12.001

INGENIERÍA E INVESTIGACIÓN VOL. 44 NO. 1, JANUARY - 2024 (e100573)

Research Article/Civil and Sanitary Engineering http://doi.org/10.15446/ing.investig.100573

Helicobacter pylori Removal through Gravel Filtration in a Water Treatment System of the Municipality of Popayán, Cauca

Remoción de *Helicobacter pylori* a través de sistemas de tratamiento por filtración en gravas en el municipio de Popayán, Cauca

Cristina Ledezma[1], Javier Fernández[2], Patricia Acosta[3], and Javier Leyton[4]

ABSTRACT

Multi-stage filtration technology (MSFT) is an alternative that reduces the risk of fecal contamination, allowing for reliable water purification in rural water supplies. MSFT is composed of two gravel filtration (GF) stages: one including dynamic gravel filters (DyGF) and up-flow gravel filters (UGF), and a final stage with slow sand filters (SSFs). However, with the purpose of reducing construction costs, this technology is partially implemented, leaving SSFs for a later construction stage and limiting its treatment potential. To evaluate the removal capabilities regarding fecal contamination (especially that by the pathogen *H. pylori*) of a two-stage GF system, the quality of raw and treated water and the hydraulic parameters of Los Llanos treatment system (municipality of Popayán, Cauca) were monitored for 15 weeks. This system is operated and maintained by the community. The results showed the removal efficiency regarding turbidity to be 16% ($p=0,045$) and 34% ($p=0,030$) for the DyGF and the UGF, respectively. The color removed by the DyGF reached 19% ($p=0,033$), and the UGF reported a value of 30% ($p=0,041$). The reduction of total coliforms was limited by the system's operation and maintenance, exhibiting a tendency towards increased concentrations at the outlet. The presence or absence of *H. pylori* was determined via the PCR molecular technique. A greater presence was evidenced in treated water than in raw one, which may be associated with a limited operation and a low maintenance frequency of the system. The implementation of MSFT, without the complement of SSFs, is not reliable in ensuring quality of water, particularly from a perspective of microbiological control and *H. pylori* control.

Keywords: *Helicobacter pylori*, gravel filtration, multi-stage filtration technology, rural water supply

RESUMEN

La tecnología de filtración en múltiples etapas (FIME) es una alternativa que reduce el riesgo por contaminación fecal, permitiendo la potabilización del agua de manera confiable en acueductos rurales. La tecnología FiME está compuesta por dos etapas de filtros en grava (FG): una que incluye filtros dinámicos (FGDi) y filtros gruesos ascendentes (FGAC), y una etapa final con filtros lentos en arena (FLA). Sin embargo, con el fin de reducir los costos de construcción, esta tecnología se implementa de manera parcial, dejando los FLA para una etapa posterior de construcción y limitando su potencial de tratamiento. Con el fin de evaluar la capacidad de remoción de la contaminación fecal (especialmente del patógeno *H. pylori*) en un sistema de FG de dos etapas, se monitoreó durante 15 semanas la calidad del agua cruda y tratada y los parámetros hidráulicos del sistema de tratamiento Los Llanos en el municipio de Popayán, Cauca, el cual es operado y mantenido por la comunidad. Los resultados mostraron que la eficiencia de remoción respecto a la turbiedad es de 16 % ($p=0,045$) y 34 % ($p=0,030$) para el FGDi y el FGAC respectivamente. El color removido por el FGDi alcanzó el 19 % ($p=0,033$), y el FGAC reportó un valor de 30 % ($p=0,041$). La reducción de los coliformes totales se vio limitada por la operación y mantenimiento del sistema, presentando una tendencia a incrementar su concentración en la salida. La presencia o ausencia del *H. pylori* se determinó con la técnica molecular PCR. Se evidenció una mayor presencia en agua tratada que en agua cruda, lo cual puede asociarse con una limitada operación y una baja frecuencia de mantenimiento del sistema. La implementación de la tecnología FiME, sin el complemento de FLA, es poco confiable para garantizar la calidad del agua, particularmente desde el punto de vista microbiológico y el control del *H. pylori*.

Palabras clave: *Helicobacter pylori*, filtración en gravas, filtración en múltiples etapas, acueducto rural

Received: January 19th, 2022
Accepted: July 24th, 2023

---

[1] Chemist. MSc in Continental Hydrobiological Resources, Universidad del Cauca. Affiliation: Associate professor, Department of Environmental and Sanitary Engineering, Environmental Engineering research group, Universidad del Cauca, Colombia. Email cledezma@unicauca.edu.co

[2] Sanitary engineer. PhD in Engineering, Universidad del Valle. Affiliation: Full professor, Department of Environmental and Sanitary Engineering, Environmental Engineering research group, Universidad del Cauca, Colombia.

[3] Biologist. PhD in Environmental Sciences, Universidad del Cauca. Affiliation: Full professor, Department of Physiological Sciences, Applied Human Genetics research group, Universidad del Cauca, Colombia.

[4] Environmental engineer. Specialist in Water Resources, Universidad del Cauca. Affiliation: Research assistant, Environmental Engineering and Applied Genetics research groups, Universidad del Cauca, Colombia.

Introduction

In a worldwide systematic review by Hooi *et al.* (2017), it was revealed that 60% of the population is infected with *Helicobacter Pylori* (*H. pylori*), with variations in distribution. The regions with developing countries that show the highest prevalence of this infection are Africa (79,1%), Latin America and the Caribbean (63,4%), and Asia (54,7%), contrasting with a lower prevalence in developed countries in North America (37,1%) and Oceania (24,4%).

H. pylori infection is acquired by the population before the age of ten, and its prevalence increases with age (Ramos and Sánchez, 2009). In Colombian children, it is close to 75%, and, in the population older than 20, it reaches 86% (Moncayo *et al.*, 2006). The incidence of this pathogen is related to gastric cancer figures. In Colombia, it represented 13,7% of all cancer deaths in the country, being the first cause in men and the third in women. The highest risk of mortality was reported in the departments of Quindío, Huila, and Cauca for men and in Cauca, Norte de Santander, and Huila for women (Pardo *et al.*, 2017; Adrada *et al.*, 2008).

Three routes of diffusion have been considered for this bacterium, oral-oral, gastro-oral, and fecal-oral (Aziz *et al.*, 2015; Bartram and Cairncross, 2010; Goh *et al.*, 2011). However, for the latter, water is considered as an intermediate vehicle for transmission, acting as an environmental reservoir (Hulten *et al.*, 1996; Engstrand, 2001). Its presence was reported in surface sources used in water supply systems in Venezuela (Adrada *et al.*, 2008; Fernández-Delgado *et al.*, 2008) and, more worryingly, in water from storage tanks and in water distribution systems for consumption in Peru (Hulten *et al.*, 1996), Portugal (Gião *et al.*, 2008), United States (Baker and Hegarty, 2001), Mexico (Premoli *et al.*, 2004), Costa Rica (Campos *et al.*, 2011; Montero-Campos, 2019), Iran, and Pakistan (Samra *et al.*, 2011; Al-Sulami *et al.*, 2010), thus confirming the presence of the pathogen in drinking water as an important source of transmission. Additionally, some studies provide evidence that surface waters may contain considerable levels of viable *H. pylori* cells, (Acosta *et al.*, 2018; Agustí *et al.*, 2010; Santiago, 2016).

In Colombia, *H. pylori* has been found in the water sources used in the public supply systems of cities such as Bogotá (Vesga, 2018), as well as in rural areas of the department of Cauca (Ordóñez, 2015). In addition, the purification treatments used have great operating limitations in small municipalities, particularly in rural areas, where the necessary infrastructure for water treatment is non-existent or has only been partially implemented. This entails negative impacts on the quality of the water, which was reflected on the Water Quality Risk Index (IRCA) of 2019. This index showed low risk values for urban areas (5,73%) and medium risk values for the rural sector (16,23%). These figures do not include self-sufficiency, since most rural water supplies have not been legalized, and 42,4% of these systems are classified as having high risk (Ministerio de Vivienda, Ciudad y Territorio, 2020), evidencing a great gap between urban and rural areas in terms of the quality of water for human consumption.

The use of MSFT depends on the water sources' degree of contamination. This method can comprise two or three stages of filtration, *i.e.*, dynamic gravel filters (DyGF), up-flow gravel filters (UGF), and slow sand filters (SSF) (Galvis, 1999). The first two stages make up the pretreatment phase, which allows reducing the concentration of suspended solids as well as microbiological contamination. Upon reaching the SSF, a more effective removal of the remaining microbiological contaminants is achieved via biological predation mechanisms. Finally, low-turbidity waters are produced, which are virtually free of enterobacteria, enterovirus, and protozoan cysts (S. Haig, 2014; Haig *et al.*, 2011; Sánchez *et al.*, 2007; Raman *et al.*, 1992). MSFT is considered to be a simple, reliable, and efficient technology that is suitable for rural communities due to its ease of operation and maintenance.

Considering a gradual and progressive approach, the implementation of this technology could be carried out in two phases. In the first phase, the pretreatment system would be built (DyGF and UGF), while, the SSFs would be constructed in the second one (Galvis *et al.*, 1992; Galvis, 1999; Galvis *et al.*, 1999; Sánchez *et al.*, 2007).

Numerous studies carried out in Colombia and different parts of the world have shown that the combination of pretreatment systems involving gravel and slow sand filters yields water with low sanitary risk for users (Mushila *et al.*, 2016; Vega, 2013; Nkwonta and Ochieng, 2009; Sánchez *et al.*, 2007; Fernández *et al.*, 2006; Ochieng *et al.*, 2004; Galvis, 1999; Wegelin, 1996; Galvis *et al.*, 1992). In all the aforementioned studies, pre-treatment systems have been able to achieve water qualities that would allow applying gradualness and progressiveness in rural treatment systems of water for human consumption, which could benefit the development of projects when there are no resources for a full implementation of MSFT.

MSFT is considered to be highly efficient in the elimination of pathogens, but there are no studies on the removal of *H. pylori* using this technology, much less in gravel prefiltration systems (DyGF and UGF), which, despite their high microbiological removal capabilities, may not be enough to achieve optimal levels of potability.

This research presents the results obtained by monitoring the removal of *H. pylori* in water pretreated with gravel filtration in Los Llanos water supply system in the municipality of Popayán, Cauca.

Methodology

This full-scale study was carried out at the drinking water treatment plant in the village of Los Llanos, which is part of Las Guacas water supply system and is located in the northeastern part of the municipality of Popayán.

The water treatment plant is composed of a DyGF, followed by a layered UGF (Figure 1). It does not have SSFs or employ a chlorination stage as a disinfection method. The raw water supply comes from the Velasquillo and San Isidro streams.

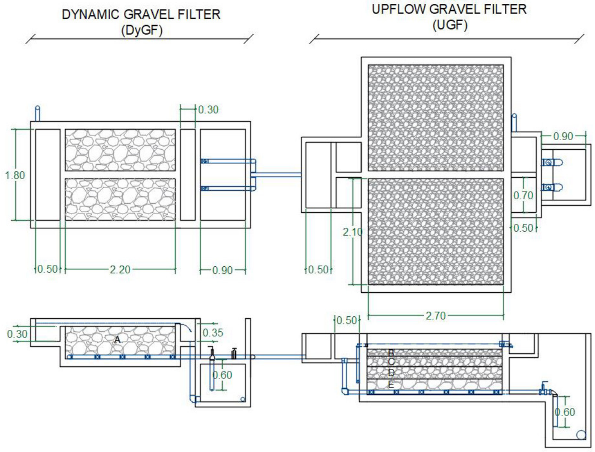


Figure 1. Scheme of treatment plant and gravel size
Source: Authors

Water quality monitoring

The physical-chemical and microbiological quality of the water was monitored by analyzing the parameters shown in Table 1.

Considering the configuration of the plant, the sampling points were as follows: i) raw water, ii) mixture of the effluent water from the two dynamic filters, and iii) outlet of the layered UGF. Physical and hydraulic parameters were measured twice a week, and microbiological aspects were evaluated once a week, for a total of 15 weeks.

Table 1. Measurement protocol for water quality parameters

Parameter	Unit of measurement	Equipment	Method
Turbidity	Nephelometric turbidity units (NTU)	Turbidity meter HACH 2100P (HACH, USA)	Nephelometric (SM 2130B)
Color	Platinum-cobalt units (PCU)	Spectrophotometer HACH DR2010 (HACH, USA)	Spectrophotometric (SM 2120C)
pH	Dimensionless	pH-meter (HACH, USA)	Potentiometric (SM 4500H+B),
Flow	L/s.	Volumetric material and stopwatch	Volumetric method (Ministry of the Environment, 2007)
Head loss	cm	Piezometers	Difference of levels
Total coliforms and *E. coli*	CFU/100ml	Membrane filtration equipment, laminar flow chamber, incubator	Membrane filtration, code: TP0314 IDEAM

Source: Authors

By means of a macro-drip system, 500 mL of water were collected for 30 min in amber glass containers and sterilized via a Gemmyco S232 autoclave at 121 °C for 15 min. They were transported, in portable refrigerators at 4 °C to the Applied Human Genetics laboratory of Universidad del Cauca, where they were processed for later *Helicobacter pylori* analysis. They were also transported to the SENA-Cauca's Environmental Analysis Laboratory, where the presence of total coliforms and *E. coli* was determined. Physical parameters were measured in the field. As operating control parameters of the treatment plant, the flow treated by the filtration units and the head losses were measured in the UGF.

Microbiological contamination indicators

The total coliforms and *E. coli* were measured by filtering 100 mL of water through sterile cellulose membranes with 0,45 μm pores and 47 mm in diameter (S/PAD, MS, MCE squared membrane, Sterile White). Chromocult Coliform Agar was used as the culture medium. As positive controls for *E. coli* and total coliforms, the *Escherichia coli* ATCC 25922 strain was used, in addition to *Pseudomona aeruginosa* ATCC 10145 strain as a negative control.

Helicobacter pylori detection

The water samples were concentrated by centrifugation in two stages. A first step at 4 000 rpm and 8 °C for 30 min (Heraeus Megafuge 1.0R centrifuge) was carried out to reach a concentration volume of 4,5 mL. Previously resuspended, this volume was transferred to 1,5 mL microcentrifuge tubes, which were in turn centrifuged at 12 000 rpm for 5 min, eliminating the supernatant until final volumes of 200 uL were obtained. This was done using an Eppendorf 5415C centrifuge.

Subsequently, DNA extraction was carried out. The concentrates were processed using the E.Z.N.A. Water DNA kit (Omega Bio-Tek, Doraville, USA) while following the manufacturer's recommendations. The isolated DNA was eluted in 200 μL of buffer solution and stored at -20 °C until amplification.

DNA amplification was carried out by means of polymerase chain reaction (PCR), using specific primers previously reported for *H. pylori:* VacA si/s2, VacA m1/m2, and CagA (Atherton et al., 1995; Erzin et al., 2006; Yamaoka et al., 1999). Additionally, PCR specificity was verified using DNA from three strains of *H. pylori* (TX30a, 11638, and 11637). The amplification mix contained 0,36 μL of each primer (Mm), 5 μL of Qiagen Multiplex PCR, prepared according to the manufacturer's recommendations, and 7 μL of genomic DNA. The PCR tubes were briefly centrifuged at low speed in order to bring the entire sample to the bottom of the tube. The samples were taken to the thermal cycler (MyCycler Thermal Cycler, BIORAD) under the conditions shown in Table 2.

Table 2. PCR conditions for amplification

Step	Process	T (˚C)	Time	Cycles
1	Initial denaturation	95	10 min	1
2	Denaturation Coupling Extension	94 57 70	30 s 90 s 45 s	15
3	Denaturation Coupling Extension	94 54 72	30 s 1 min 1 min	25
4	Final extension	72	5 min	1

Source: Authors

Finally, the amplified DNA was visualized in 1,5% agarose gels in 1X TAE buffer (Tris-Acetic-EDTA), using EZ-VISION THREE as a dye. Gel electrophoresis was performed at 80 V for 60 min. The presence of the amplified fragments was verified by visual inspection on the Gel DocTM XR + Imaging System (BIO-RAD, USA), taking the amplification bands of strain 11637 as reference.

The data were processed in the STATGRAPHICS Centurion XVI.I statistical software. The maximum efficiency of each filtration unit was determined through hypothesis tests, with a confidence level of 95% regarding the turbidity and color parameters. The reduction of microbiological characteristics (total coliforms and *E. coli*) was performed in logarithmic units.

Results and discussion

Raw water quality

The turbidity of the raw water ranged between 1,6 and 12,7 NTU, the average apparent color was 39,6 UPC, the total coliforms and the the average *E. coli* were 1 213 and 103 CFU/100 mL, respectively, the average temperature was 18,8 ˚C, and the pH ranged from 6,6 to 7,5. These quality conditions were recommended by Galvis *et al*. (1999) for the use of MSFT systems with DyGF, UGF, and SSF plus disinfection as a safety barrier. However, the treatment system only has the first two stages, which is usable within a scheme of progressivity for improving water quality if disinfection is performed collectively or individually in the households.

Treatment system operation

In general, some operating conditions limited the proper functioning of the different treatment structures. The DyGF did not exhibit a water level over the filter bed, which does not allow blocking the turbidity peaks that occur in raw water during the rainy season due to the dragging of solids in the basin. The flow treated by the plant during the study period was 3,05 L/s, which corresponded to a filtration rate of 2,94 m/h in the DyGF. Meanwhile, in the studied UGF, the treated flows per unit ranged between 0,84 and 2,04 L/s, with an average of 1,52 L/s, which represents a filtration rate range of 0,53-1,14 m/h, with an average value of 0,85 m/h. These rates are above those recommended (0,30 and 0,60 m/h) by Galvis *et al*. (1999), which may entail reduced removal efficiency and increased maintenance requirements due to a greater hydraulic head loss in the UGF.

It can be highlighted that there were periods of low maintenance frequencies in the treatment units, with an impact on hydraulic head losses. During the first five weeks, the UGF reached hydraulic head loss values of up to 4,7 cm, with large amounts of mud on the filter surface generated by a very low frequency of maintenance, which included only two bottom drainages without surface washing. Figure 2a shows substantial sludge buildup on the surface filter medium, while Figure 2b illustrates a clean filter medium.

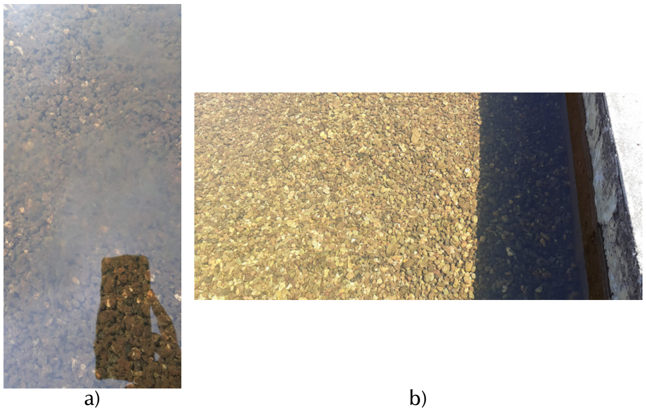

a) b)

Figure 2. Gravel filter with presence of mud on the surface. a) week 3; b) week 10.
Source: Authors

During the following five weeks, the maintenance frequency was increased to once a week, with deep washing and superficial washing every two weeks, showing reductions in hydraulic head loss (from 4,7 to 2,1). For the last four weeks of monitoring, the frequency of maintenance was reduced again (no bottom drainage or surface washing was evidenced), yielding a much greater increase in hydraulic head loss (between 2,6 and 5,8 cm).

These data, particularly those between the first five weeks and the last four, indicate that there was an accumulation of sludge within the filter medium of the UGF.

Treated water quality

The results obtained regarding turbidity, color, total coliforms, and *E. coli* are summarized in Table 3. None of these parameters reached the average quality value suggested by the World Health Organization (WHO, 2017) for drinking water, which shows low or no removal efficiency. These statistical results were obtained with a confidence level of 95%, through which the removal efficiency was estimated (Table 4).

As per the statistical analysis, with regard to turbidity and color, the DyGF exhibited removal values lower than 16 and 19%, respectively (Table 4). These results are consistent with the efficiency of other MSFTs treating similar raw water qualities. Galvis *et al*. (1999) reported a removal efficiency ranging from 24 to 38% for turbidity and between 17 and 21% for color. However, these values were obtained under appropriate DyGF operating and maintenance conditions.

Table 3. Quality of raw water and effluent of DyGF and UGF

Parameter	Statistical	Raw water	DyGF	UGF
Turbidity	Average	3,8	3,3	2,3
	SD	2,1	1,4	1,0
	Maximum	12,7	7,3	5,9
	Minimum	1,6	1,9	1,4
Apparent Color	Average	39,6	32,8	26,3
	SD	21,7	17,0	16,5
	Maximum	92,3	67,0	67,7
	Minimum	8,3	7,7	3,7
Total Coliforms	Average	1 213	1 343	1 184
	SD	732	958	736
	Maximum	3 300	4 100	3 069
	Minimum	483	154	392
E. coli	Average	103	83	57
	SD	71	63	39
	Maximum	263	239	148
	Minimum	5	20	8
H. pylori	Total samples	30	30	30
	Number of positives	3	6	6

Source: Authors

In the case of the UGF, the results indicate removal values of less than 34,0% for turbidity and 30% for color (Table 4), while the same study by Galvis *et al*. (1999) reported 31-36% for turbidity and 21-26% for color, with a filtration rate of 0,6 m/h, lower than that of our study. Posso (2012) reported 32 and 36% turbidity removals in the Golondrinas and Arroyohondo plants, respectively, with filtration rates between 0,45 and 0,57 m/h. This indicates that the UGF units under study exhibit a similar removal efficiency, despite the higher filtration rates used.

The results obtained regarding the removal of total coliforms and *E. coli* (Tables 3 and 4) show great variation and low performance in the filtration units. In 71% of the cases, there was a tendency towards an increased concentration of total coliforms at the outlet of the DyGF. In the remaining 29%, a reduction of 0,17 log was observed. The UGF exhibited a similar but less accentuated trend, with a reduction of 0,22 log for 64% of the cases (Table 4).

The performance of the DyGF in the removal of *E. coli* also exhibited limitations, as reductions were only achieved in 57% of the cases, reaching a value lower than 0,22 log (Table 4). Meanwhile, the UGF, managed to obtain reductions of less than 0,3 log for 79% of the cases. These results evidence the low performance of the system when compared to the reduction of microorganisms through gravel filters (Mushila *et al*., 2016; Vega, 2013; Nkwonta and Ochieng, 2009; Sánchez *et al*., 2007; Fernández *et al*., 2006; Galvis *et al*., 1999; Posso, 2012), where removals and/or reductions between 30 and 40% and between 0,6 and 2,5 log have been reported, The behavior of the studied units was favored by the environmental conditions caused by inadequate operation and maintenance,

Table 4. Efficiency of the filters in the removal and reduction of the parameters

Efficiency	Parameter	DyGF	UGF
% Removal, (p-value)	Turbidity	16 (0,045)	34 (0,030)
	Apparent Color	19 (0,033)	30 (0,041)
Reduction in Log, (p-value)	* Total Coliforms	0,17 (0,041)	0,22 (0,032)
	**E. coli*	0,21 (0,047)	0,30 (0,037)

* Includes only the values that caused reductions.
Source: Authors

The removal of particles in a gravel filter occurs by different mechanisms, with sedimentation being one of the most important. The efficiency of the filter depends on variables such as the filtration speed, the length and size of the filter medium, porosity, and the particle size to removed (Boller, 1993; Sánchez, 2017). The adherence of the particles to the surface of the filter medium is given by the interaction between van der Waals attraction forces and hydrodynamics. The accumulation of particles inside the pores reduces their size, which entails increased interstitial flow velocity, hydraulic head losses, and shear forces that allow for the detachment of the retained particles (Boller, 1993). Sánchez (2017) found that particles less than 5 µm in diameter can only be removed with efficiencies between 45 and 70% and a filtration rate of less than 0,5 m/h.

In this sense, the low or null efficiency found especially in the removal of coliforms and *E. coli* could be explained by the detachment of accumulated excess particles inside the gravel filters, which is caused by the low frequency of maintenance and high filter operation speeds (0,53-1,14 m/h), likely entailing shear forces large enough to generate drag in the retained particles and cause increases in both turbidity and color, as well as in the concentration of microorganisms, as described by Boller (1993).

Helicobacter pylori removal

The presence of *H. pylori* was detected in 15 of the 90 samples taken during monitoring. Three samples correspond to raw water, six to the outlet of the DyGF, and another six to the outlet of the UGF (Table 5). Considering that the bacteria can survive in water by employing several strategies, including the formation of biofilm (Campos *et al*., 2011; Moreno *et al*., 2007; Santiago, 2016; Chowdhury, 2012; Watson *et al*., 2004), conversion to coccoid form in a non-cultivable viable state (CVNC) (Azevedo *et al*., 2007; Oliver, 2005), or adherence to abiotic surfaces (Fernández *et al*., 2008), their greater presence at the outlet of the filters could be associated with the possible detachment of retained solids due to the high cutting forces generated by the accumulation of particles in the gravel. However, this finding should be studied in more detail.

As shown in Table 5, during weeks 1 to 5 and 12 to 15, in which maintenance was not adequately carried out, nor with the frequency required for this type of filtration units, the number of samples with presence of *H. pylori* was greater at the outlet of the gravel filters (DyGF and UGF). On the other hand, between weeks 6 and 11, during which maintenance activities were performed more frequently, no positive values were detected.

Table 5. Presence (+)/absence (-) of *H. pylori* per week

Sampling point	**Week**						**Total samples**
	*1 to 5 **		*6 to 11***		*12 to 15****		
	(+)	(-)	(+)	(-)	(+)	(-)	
Raw water	0	10	0	12	3	5	30
DyGF output	4	6	0	12	2	6	30
UGF output	2	8	0	12	4	4	30

Maintenance activities:
** Two bottom drainages without surface washing*
*** Weekly bottom drainages with surface washing*
**** Without bottom drainages or surface washing*
Source: Authors

Statistical analysis revealed a relationship between the lack of filter maintenance and the presence of *Helicobacter pylori*, with p-values of 0,036 and 0,018 for the DyGF and the UGF respectively. Meanwhile, no significant differences were found in terms of color, turbidity, *E. coli*, and total coliforms removal, with p-values greater than 0,27.

Conclusions

The results obtained for the studied system's operating conditions do not show significant microorganism removal (coliforms and *Helicobacter pylori*). On the contrary, there is a greater number of positive values at the outlet than at the inlet. It should be noted that the operation and maintenance of the treatment system was mostly inappropriate, which could explain the results. In general, when they are well operated and maintained, these systems are expected to improve quality of water. Thus, it is evident that the selection and implementation of treatment technologies must be thoroughly considered according to local conditions. In our case, the partial introduction of MSFT in Los Llanos water supply system, despite the fact that its operation and maintenance scheme is much simpler than that of other technologies, does not seem to be sufficient to guarantee a good quality of water for human consumption if interventions that facilitate the appropriation of technology by the community are not considered. Similarly, without the complement of SSFs, MSFT seems to be unreliable in ensuring water quality, particularly from a microbiological point of view.

These results also show that applying the concept of *graduality* to rural water treatment systems, which considers the construction of staged systems, must ensure quality of water for all users. This should also be complemented with disinfection techniques.

Although the findings regarding microbiological removal are linked to a low frequency of maintenance, a more detailed study is required to better explain the presence of *H. pylori* at the outlet of the filters.

Acknowledgments

The authors express their gratitude to the project titled *Validation of slow sand filters for the removal of the emerging pathogen Helicobacter pylori in rural water supply systems*, which is attached to the call for the funding of projects aimed at strengthening research groups classified in category C, which are recognized and registered by the Vice-Principalship for Research of Universidad del Cauca. The authors would also like to thank the community of Los Llanos, which is part of Las Guacas water supply system, for their valuable support in the execution of our research activities; as well as the Environmental Engineering and Human Genetics laboratory of Universidad del Cauca and the SENA-Cauca Environmental Analysis Laboratory.

CRediT author statement

All authors contributed to the conceptualization, methodology, validation, formal analysis, investigation, data curation, and writing of this work. They have approved the definitive version of this manuscript for publication.

References

Acosta, C. P., Codony, F., Fittipaldi, M., Sierra-Torres, C. H., and Morató, J, (2018). Monitoring levels of viable Helicobacter pylori in surface water by qPCR in Northeast Spain. *Journal of Water and Health*, *16*(5), 839-845. https://doi.org/10.2166/wh.2018.195

Adrada, J. C., Calambás, F. H., Díaz, J. E., Delgado, D. O., and Sierra, C. H. (2008). Características sociodemográficas y clínicas en una población con cáncer gástrico en el departmento de Cauca, Colombia. *Revista Colombiana de Gastroenterologia*, *23*, 309-314. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-99572008000400004

Agustí, G., Codony, F., Fittipaldi, M., Adrados, B., and Morató, J. (2010). Viability determination of *Helicobacter pylori* using propidium monoazide quantitative PCR. *Helicobacter*, *15*(5), 473-476. https://doi.org/10.1111/j.1523-5378.2010.00794.x

Al-Sulami, A. A., Al-Taee, A. M. R., and Juma'a, M. G. (2010). Isolation and identification of helicobacter pylori from drinking water in Basra governorate, Iraq. *Eastern Mediterranean Health Journal*, *16*(9), 920-925. https://doi.org/10.26719/2010.16.9.920

Atherton, J. C., Cao, P., Peek, R. M., Tummuru, M. K. R., Blaser, M. J., and Cover, T. L. (1995). Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*, Association of specific vacA types with cytotoxin production and peptic ulceration. *Journal of Biological Chemistry, 270*(30), 17771-17777. https://doi.org/10.1074/jbc.270.30.17771

Aziz, R. K., Khalifa, M. M., and Sharaf, R. R. (2015). Contaminated water as a source of *Helicobacter pylori* infection: A review. *Journal of Advanced Research*, *6*(4), 539-547. https://doi.org/10.1016/j.jare.2013.07.007

Baker, K. H., and Hegarty, J. P. (2001). Presence of *Helicobacter pylori* in drinking water is associated with clinical infection. *Scandinavian Journal of Infectious Diseases*, *33*(10), 744-746. https://doi.org/10.1080/003655401317074536

Bartram, J., and Cairncross, S. (2010). Hygiene, sanitation, and water: Forgotten foundations of health *PLoS Medicine*, *7*(11), e1000367. https://doi.org/10.1371/journal.pmed.1000367

Boller, M. (1993). Filter mechanisms in roughing filters. *Aqua*, *42*(3), 174-185.

Campos, M., Soto, H., Meléndez, M., Sandoval, C., Santamaría, G., Rojas, B., Cascante, L., Gutiérrez, O., and Montero Campos, V. (2011). Hallazgo de la bacteria Helicobacter pylori en agua de consumo humano y su relación con la incidencia de cáncer gástrico en Costa Rica. *Tecnologia En Marcha*, *24*(3), 3-14.

Chowdhury, S. (2012). Heterotrophic bacteria in drinking water distribution system: A review. *Environmental Monitoring and Assessment, 184,* 6087-6137. https://doi.org/10.1007/s10661-011-2407-x

Engstrand, L. (2001). Helicobacter in water and waterborne routes of transmission. *Symposium Series* (*Society for Applied Microbiology*), *30*, 80-84.

Erzin, Y., Koksal, V., Altun, S., Dobrucali, A., Aslan, M., Erdamar, S., Dirican, A., and Kocazeybek, B. (2006). Prevalence of *Helicobacter pylori vacA, cagA, cagE, iceA, babA2* genotypes and correlation with clinical outcome in Turkish patients with dyspepsia. *Helicobacter, 11*(6), 574-580. https://doi.org/10.1111/j.1523-5378.2006.00461.x

Fernández-Delgado, M., Contreras, M., García-Amado, M. A., Michelangeli, F., and Suárez, P. (2008). Evidencias de la transmisión acuática de *Helicobacter pylori*. *Interciencia, 33*(6), 412-417. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0378-18442008000600005

Fernández, J., Cruz, A., and Benavides, E. (2006). Remoción de materia orgánica por filtración en múltiples etapas. *Ingeniería Hoy*, *24*, 20-27.

Galvis, G., Teun Visscher, J., and Lloyd, B. (1992). Multi-stage surface water treatment for community water supply in Colombia. *Waterlines*, *10*(3), 26-29, https://doi,org/10.3362/0262-8104.1992.008

Galvis, G. (1999). *Development and evaluation of multistage filtration plants: An innovative, robust and efficient water treatment technology* [Doctoral thesis, Univertsity of Surrey]. https://openresearch.surrey.ac.uk/esploro/outputs/doctoral/Development-and-evaluation-of-multistage-filtration/99513881902346

Galvis, G., Latorre, J., and Visscher, J. T. (1999). *Filtración en múltiples etapas: tecnología innovativa para el tratamiento de agua*. Artes Gráficas de Univalle. https://www.ircwash.org/sites/default/files/255.9-99FI-17025.pdf

Gião, M. S., Azevedo, N. F., Wilks, S. A., Vieira, M. J., and Keevil, C. W. (2008). Persistence of *Helicobacter pylori* in heterotrophic drinking-water biofilms. *Applied and Environmental Microbiology*, *74*(19), 5898-5904. https://doi.org/10.1128%2FAEM.00827-08

Goh, K., Chan, W., Shiota, S., and Yamaoka, Y. (2011). Epidemiology of *Helicobacter pylori* infection and public health implications. *Helicobacter*, *16*, 1-9. https://doi.org/10.1111/j.1523-5378.2011.00874.x

Haig, S. (2014). *Characterising the functional ecology of slow sand filters through environmental genomics* [Doctoral thesis, University of Glasgow]. https://theses.gla.ac.uk/5523/1/2014HaigPhD.pdf

Haig, S. J., Collins, G., Davies, R. L., Dorea, C. C., and Quince, C. (2011). Biological aspects of slow sand filtration: Past, present and future. *Water Science and Technology: Water Supply, 11*(4), 468-472. https://doi.org/10.2166/ws.2011.076

Hooi, J. K. Y., Lai, W. Y., Ng, W. K., Suen, M. M. Y., Underwood, F. E., Tanyingoh, D., Malfertheiner, P., Graham, D. Y., Wong, V. W. S., Wu, J. C. Y., Chan, F. K. L., Sung, J. J. Y., Kaplan, G. G., and Ng, S. C. (2017). Global prevalence of *Helicobacter pylori* infection: Systematic review and meta-analysis. *Gastroenterology*, *153*(2), 420-429. https://doi.org/10.1053/j.gastro.2017.04.022

Hulten, K., Han, S. W., Enroth, H., Klein, P. D., Opekun, A. R., Gilman, R. H., Evans, D. G., Engstrand, L., Graham, D. Y., and El-Zaatari, F. A. K. (1996). *Helicobacter pylori* in the drinking water in Peru. *Gastroenterology*, *110*(4), 1031-1035. https://doi.org/10.1053/gast.1996.v110.pm8612990

Ministerio de Vivienda Ciudad y Territorio (2020). *Informe Nacional de Calidad del Agua para Consumo Humano INCA 2019*. https://www.minvivienda.gov.co/sites/default/files/documentos/informe-calidad-de-agua-2019.pdf

Moncayo, J. I., Santacruz, J. J., Álvarez, A. L., Franco, B., López, M. A., Ángel, A., Gallego, M. L., and Serrano, H. (2006). Comparación de métodos diagnósticos en la infección por *Helicobacter pylori* en Quindío, Colombia. *Colombia Medica*, *37*(3), 203-212. https://www.redalyc.org/articulo.oa?id=28337306

Montero-Campos, V. (2019). *Helicobacter pylori* en Costa Rica, más de una década de investigaciones. *Revista Tecnología En Marcha*, *32*, 94-103. https://doi.org/10.18845/tm.v32i9.4636

Moreno, Y., Piqueres, P., Alonso, J. L., Jiménez, A., González, A., F. M. (2007). Survival and viability of *Helicobacter pylori* after inoculation into chlorinated drinking water. *Water Research*, *41*(15), 3490-3496. https://doi.org/10.1016/j.watres.2007.05.020

Mushila, C. N., Ochieng, G. M., Otieno, F. A. O., Shitote, S. M., and Sitters, C. W. (2016). Hydraulic design to optimize the treatment capacity of multi-stage filtration units. *Physics and Chemistry of the Earth*, *92*, 85-91. https://doi.org/10.1016/j.pce.2015.10.015

Nkwonta, O., and Ochieng, G. (2009). Roughing filter for water pre-treatment technology in developing countries: A review. *International Journal of Physical Sciences, 4*(9), 455-463). https://www,researchgate,net/publication/237827490_Roughing_filter_for_water_pre-treatment_technology_in_developing_countries_A_review

Ochieng, G. M. M., Otieno, F. A. O., Ogada, T. P. M., Shitote, S. M., and Menzwa, D. M. (2004). Performance of multistage filtration using different filter media against conventional water treatment systems. *Water SA*, *30*(3), 361-367. https://doi.org/10.4314/wsa.v30i3.5085

Oliver, J. D. (2005). The viable but nonculturable state in bacteria. *Journal of Microbiology, 43*, 93-100.

Ordóñez, J. F. (2015). *Evaluación de la calidad microbiológica de las fuentes de abastecimiento del acueducto rural El Saladito – Timbío – Cauca* [Undergraduate thesis, Universidad del Cauca]. http://repositorio.unicauca.edu.co:8080/xmlui/handle/}123456789/6862

Pardo, C., De Vries, E., Buitrago, L., and Gamboa, Ó. (2017). *Atlas de mortalidad por cáncer en Colombia*. https://www.cancer.gov.co/ATLAS_de_Mortalidad_por_cancer_en_Colombia.pdf

Posso, D. (2012). *Análisis de la operación y mantenimiento de la filtración en gravas de flujo ascendentes a escala real* [Undergraduate thesis, Universidad del Valle]. https://hdl.handle.net/10893/7675

Premoli, G., González, A., Millán-mendoza, B., Percoco, T., and Vielma, A. (2004). Diagnóstico de *Helicobacter pylori* mediante la reacción en cadena de la polimerasa. *Revista Cubana de Medicina Tropical*, *56*(2), 85-90. http://scielo.sld.cu/pdf/mtr/v56n2/mtr01204.pdf

Raman, A., Paramasivam, R., Heijnen, H. A., Visscher, J. (1992). *Filtracion lenta en arena : tratamiento de agua para comunidades : planeacion, diseno, construccion, operacion y mantenimiento*. https://es.ircwash.org/resources/filtracion-lenta-en-arena-tratamiento-de-agua-para-comunidades-planeacion-diseno

Ramos, A. R., and Sánchez, R. S. (2009), Contribución de Latinoamérica al estudio del *Helicobacter pylori*. *Acta Gastroenterologica Latinoamericana, 39*(3), 197-218. https://www.redalyc.org/exportarcita.oa?id=199317345011

Samra, Z. Q., Javaid, U., Ghafoor, S., Batool, A., Dar, N., and Athar, M. A. (2011). PCR assay targeting virulence genes of *Helicobacter pylori* isolated from drinking water and clinical samples in Lahore metropolitan, Pakistan. *Journal of Water and Health*, *9*(1), 208-216. https://doi.org/10.2166/wh.2010.169

Sánchez, L. D., Sánchez, A., Galvis, G., and Latorre, J. (2007). *Filtración en Múltiples Etapas*. IRC. https://es.ircwash.org/resources/filtraci%C3%B3n-en-m%C3%BAltiples-etapas

Sánchez, L. D. (2017). Determinación del factor de adherencia en filtros de grava de flujo ascendente en capas. *Ingeniería Y Competitividad*, *19*(2), 121-130. https://doi.org/10.25100/iyc.v19i2.5299

Santiago, P. (2016). *Transmisión de Helicobacter pylori a través del agua: estudio de la presencia del patógeno e identificación de formas viables mediante técnicas moleculares* [Doctoral thesis, Universidad Politécnica de Valencia]. https://doi.org/10.4995/Thesis/10251/75086

Vega Serrano, H. A. (2013). *Evaluación del sistema de filtración en múltiples etapas FiME en tanques plásticos con pre-sedimentación y retro-lavado en la hacienda Majavita* (*Socorro, Santander*) [Master's thesis, Universidad de Manizales]. https://ridum.umanizales.edu.co/xmlui/bitstream/handle/20.500.12746/151/402_Vega_Serrano_Haimar_Ariel_2013_Documento.pdf

Vesga, F. (2018). *Detección y viabilidad de Helicobacter pylori en aguas crudas y potables en tres plantas de potabilización en la ciudad de Bogotá* [Doctoral thesis, Pontificia Universidad Javeriana, Universidad Politécnica de Barcelona]. https://doi.org/10.11144/Javeriana.10554.41877

Watson, C. L., Owen, R. J., Said, B., Lai, S., Lee, J. V., Surman-Lee, S., and Nichols, G. (2004). Detection of *Helicobacter pylori* by PCR but not culture in water and biofilm samples from drinking water distribution systems in England. *Journal of Applied Microbiology*, *97*(4), 690-698. https://doi.org/10.1111/j.1365-2672.2004.02360.x

Wegelin, M. (1996). *Surface water treatment by roughing filters: A design, construction and operation manual*. SKAT. https://www.ircwash.org/resources/surface-water-treatment-roughing-filters-design-construction-and-operation-manual

World Health Organization (WHO) (2017). *Guidelines for drinking-water quality: fourth edition incorporating the first addendum*. WHO. https://www.who.int/publications/m/item/guidelines-for-drinking-water-quality-4th-ed.-incorporating-the-1st-addendum-(chapters)

Yamaoka, Y., Kodama, T., Gutiérrez, O., Kim, J. G., Kashima, K., and Graham, D. Y. (1999). Relationship between *Helicobacter priori* iceA, cagA, and vacA status and clinical outcome: Studies in four different countries. *Journal of Clinical Microbiology*, *37*(7), 2274-2279. https://doi.org/10.1128/jcm.37.7.2274-2279.1999

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e107462)

Research Article/Civil and Sanitary Engineering http://doi.org/10.15446/ing.investig.107462

Evaluating the Use of Recycled Brick Powder as a Partial Replacement for Portland Cement in Concrete

Evaluación del uso del polvo reciclado de ladrillo como reemplazo parcial al cemento Portland en el hormigón

Joaquin Humberto Aquino Rocha[1], Boris Marcelo Morales Ruiz[2], and Romildo Dias Toledo Filho[3]

ABSTRACT

Portland cement is one of the most used construction materials. However, its production represents between 5 and 7% of the total CO_2 emissions. On the other hand, during construction and demolition activities, different wastes are produced, including recycled brick powder (RBP), whose potential as a supplementary cementitious material (SCM) has been demonstrated in the literature. This research aims to evaluate RBP as a partial replacement for Portland cement in concrete. 5 to 10% of Portland cement was replaced with RBP in two strength designs (20 and 25 MPa) in order to propose concretes that meet the requirements for use in construction. Tests involving slump, compressive strength, tensile strength by diametrical compression, absorption, density, and void content were performed. The results show that a 5% RBP replacement does not affect workability in concrete mixes, as it maintains their mechanical resistance and slightly improves their physical properties. On the other hand, 10% RBP replacements adversely affect workability and reduce tensile strength. These results are attributed to pozzolanic activity and the physical effect caused by RBP, whose performance may be improved by reducing RBP particles and increasing their specific surface area (SSA). Using RBP as a replacement for Portland cement to produce concrete is a viable alternative with a sustainable approach.

Keywords: workability, mechanical properties, physical properties, sustainability

RESUMEN

El cemento Portland es uno de los materiales de construcción más utilizados. Sin embargo, su producción representa entre el 5 y el 7 % de las emisiones totales de CO_2. Por otro lado, durante las actividades de construcción y demolición, se producen diferentes residuos, entre ellos el polvo de ladrillo reciclado (PLR), cuyo potencial como material cementante suplementario (MCS) ha sido demostrado en la literatura. El objetivo de la presente investigación es evaluar el PLR como reemplazo parcial del cemento Portland en el hormigón. Se sustituyó 5 y 10 % de cemento Portland por PLR para dos resistencias de diseño (20 y 25 MPa), a fin de proponer concretos que cumplan con los requerimientos para ser utilizados en la construcción. Se realizaron ensayos de asentamiento, resistencia a la compresión, resistencia a la tracción por compresión diametral, absorción, densidad y contenido de vacíos. Los resultados muestran que la sustitución del 5 % de PLR no afecta la trabajabilidad de las mezclas, pues mantiene las resistencias mecánicas y mejora levemente las propiedades físicas. Por otro lado, las sustituciones del 10 % de PLR afectan negativamente la trabajabilidad y reducen la resistencia a la tracción. Estos resultados se atribuyen a la actividad puzolánica y al efecto físico del PLR, cuyo desempeño puede ser mejorado si se reducen las partículas de PLR y se incrementa el área superficial específica (ASE). El uso de PLR como reemplazo del cemento Portland para la elaboración de hormigón es una alternativa viable con enfoque sostenible.

Palabras clave: trabajabilidad, propiedades mecánicas, propiedades físicas, sostenibilidad

Received: February 25th, 2023
Accepted: July 29th, 2023

Introduction

Portland cement is one of the most used materials in the construction industry, especially for concrete production (Habert *et al.*, 2020). Due to high demand, particularly from developing countries, Portland cement production has been increasing, and an increasing trend is projected for the coming years (UN Environment *et al.*, 2018). However, manufacturing Portland cement requires a high consumption of natural resources and energy, so it constitutes an important source of carbon dioxide (CO_2) emissions, approximately 5-7% of the total anthropogenic CO_2 emissions (Singh and Middendorf, 2020; Sousa and Bogas 2021). This is due to limestone decomposition ($CaCO_3 \rightarrow CaO + CO_2$) at high temperatures (~1 450 °C) in order to produce clinker, Portland cement's main component. This process releases CO_2 and consumes significant amounts of fossil fuel (Gao *et al.*, 2021; Arif *et al.*, 2021). To reduce the environmental impacts of using Portland cement, a partial replacement with

---

[1] Master of Science in Civil Engineering, University of Pernambuco, Brazil. Affiliation: PhD student, Federal University of Rio de Janeiro, Brazil. Email: joaquin.rocha@coc.ufrj.br

[2] Civil engineer, Universidad Privada del Valle, Bolivia. Affiliation: Researcher, Universidad Privada del Valle, Bolivia. Email: mrb2018052@est.univalle.edu

[3] PhD in Civil Engineering, Pontifical Catholic University of Rio de Janeiro, Brazil. Affiliation: Researcher-professor, Federal University of Rio de Janeiro, Brazil. Email: toledo@coc.ufrj.br

supplementary cementitious materials (SCMs) has been proposed (Jiang *et al*., 2020; Rocha *et al*., 2022; Panesar and Zhang, 2020).

Using SCMs in concrete is a widely spread strategy to reduce CO_2 emissions which may improve material durability and constitutes a sustainable approach in the construction industry depending on the characteristics and content of SCMs (Panesar and Zhang, 2020). However, the availability of these materials hinders their implementation (Juenger and Siddique, 2015). In this sense, using local waste such as municipal solid waste (MSW) and agro-industrial waste as SCMs has been proposed (Thomas *et al*., 2021; Tripathy and Acharya, 2022), as well as construction and demolition waste (CDW) (Tang *et al*., 2020; Likes *et al*., 2022), among others (Juenger *et al*., 2019; Jiang *et al*., 2022; Zhou *et al*., 2021).

On the other hand, a rapid population growth has influenced the increase in construction activities, thus generating more CDW, *i.e.*, waste made up of sand, gravel, concrete, mortar, bricks, and glass, among others (Wong *et al*., 2018; Vieira *et al*., 2020). Recent studies have repurposed CDW as recycled aggregates to produce concrete and mortar, demonstrating its technical and environmental feasibility (Da Silva *et al*., 2022; Borges *et al*., 2023). Da Silva *et al*. (2023) pointed out that the use of CDW as a recycled aggregate, in conjunction with fly ash (FA) and hydrated lime (HL), has a synergistic effect that could enhance the mechanical properties and durability of concrete. In this sense, Da Silva and Andrade (2022) argued that the use of both CWD and SCMs such as FA, metakaolin (MK), and rice husk ash (RHA), among others, could provide economic, environmental, and energy benefits in the production of new cement-based materials.

Brick is quite widely used in construction, followed only by concrete (Adamson *et al*., 2015). Several studies have proposed the utilization of brick residue as a recycled aggregate in the construction industry (Mohammed *et al*., 2015; Ge *et al*., 2021; Nguyen *et al*., 2023). However, in order to provide a sustainable solution for increasing cement production and CDW generation, the use of recycled brick powder (RBP) has been studied in recent years (He *et al*., 2021). Several studies have shown that RBP is a pozzolanic material given its high SiO_2 and Al_2O_3 contents, in addition to the amorphous compounds formed during the brick-making process (Reig *et al*., 2013; Ortega *et al*., 2018). On the other hand, Luo *et al*. (2022) not only highlighted the chemical characteristics of RBP, but also its physical effects, such as filling and nucleation.

Arif *et al*. (2021) replaced Portland cement with 5 and 10% RPB in order to produce concrete. These authors reported an increase in both workability and mechanical properties, attributing this behavior to the morphology and pozzolanic activity of RBP particles. Liu *et al*. (2020) also found that substituting Portland cement with 10% RBP may improve mechanical strength in mortars. Naceri and Hamina (2009) indicated that up to 20% RBP may maintain or improve the compressive strength of concrete and mortar, due to the amorphous SiO_2 content that promotes pozzolanic reaction, thus forming hydrated calcium silicate (C-S-H). Likes *et al*. (2022) indicated that the use of 20% RBP resulted in a strength activity index (SAI) of 98% for concrete after 28 days, as well as an improvement in durability when compared to the reference.

Although there is literature on the use of RBP as a partial replacement for Portland cement, each study has considered different material characteristics. In order to propose RBP as an application-focused SCM, this study intends to replace Portland cement with RBP while considering two concrete strength designs (20 and 25 MPa) and local materials from the city of Cochabamba, Bolivia. Thus, this study intends to evaluate workability and both physical and mechanical concrete properties, replacing Portland cement with 5 and 10% RBP. The substitution percentages were selected with the purpose of proposing concretes with RBP that have the same physical-mechanical properties as conventional concrete, thus providing an alternative for the reuse of RBP in concrete for structural purposes.

Materials and methods

Materials

For this research, RBP was obtained from a demolition site in Cochabamba, Bolivia. The RBP went through a sieving process using a fraction-exceeding #100 sieve (150 μm). IP 40 cement was used, corresponding to Portland modified with pozzolan, according to the ASTM C595 standard (American Society for Testing and Materials, 2021a).

The chemical composition for both the RBP and the IP 40 cement was determined via X-ray fluorescence spectrometry (XRF). The results are presented in Table 1. Regarding the RBP, the sum of chemical components $SiO_2+Al_2O_3+Fe_2O_3$ was higher than 70%, meeting the requirements to be classified as a pozzolanic material in accordance with ASTM C618 (American Society for Testing and Materials, 2022).

Table 1. Chemical composition for RBP and IP 40 cement (in %)

Chemical compounds	IP 40 cement	RBP
CaO	70,69	6,90
SiO_2	12,95	44,22
Fe_2O_3	5,04	26,06
Al_2O_3	4,61	8,37
SO_3	3,11	6,02
K_2O	0,45	4,14
TiO_2	0,34	1,74
MnO	0,24	0,13
SrO	0,56	0,05
ZrO_2	-	0,05
LOI	2,01	2,32

Source: Authors

The physical properties of both the RBP and the IP 40 cement were determined. Their particle size distribution was determined via laser diffraction, their density by means of helium gas pycnometry, and their specific surface area (SSA) according to the ASTM C204 (American Society for Testing and Materials, 2019). Figure 1 and Table 2 present the materials' particle size distribution and a summary of their physical properties, respectively.

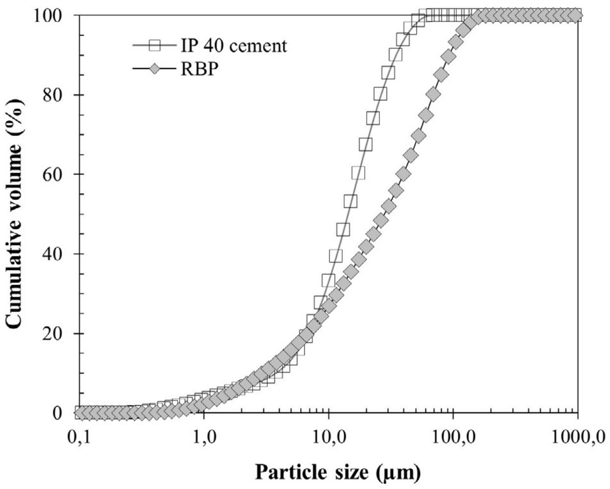


Figure 1. Particle size distribution of the RBP and the IP 40 cement
Source: Authors

Table 2. Physical properties of the RBP and the IP 40 cement

Parameter	RBP	IP 40 Cement
D10 (µm)	2,93	3,75
D50 (µm)	27,99	14,26
D90 (µm)	92,40	34,60
Specific gravity (g/cm^3)	2,69	3,07
Specific surface area (cm^2/g)	3 443,90	3 642,51

Source: Authors

The RBP's mineral composition was determined via X-ray diffraction (XRD). The XRD pattern for the RBP is presented in Figure 2, where quartz is the largest crystalline phase, followed by other compounds: hematite, ilite, albite, and gypsum – the latter due to possible contamination at the demolition site. The proportion corresponding to the RBP's amorphous phase is 60,7%, as determined by refinement according to the Rietveld method.

The granulometric distribution of coarse and fine aggregate was determined by following the ASTM C136 standard (American Society for Testing and Materials, 2015a) (Figure 3). The fineness modulus of coarse and fine aggregate was 7,37 and 2,65, respectively, with a maximum aggregate size of 19 mm. Following the ASTM C127 (American Society for Testing and Materials, 2015b) and C128 (American Society for Testing and Materials, 2015c) methods, the relative density of coarse and fine aggregate was determined, obtaining 2,63 and 2,77 g/cm^3, respectively.

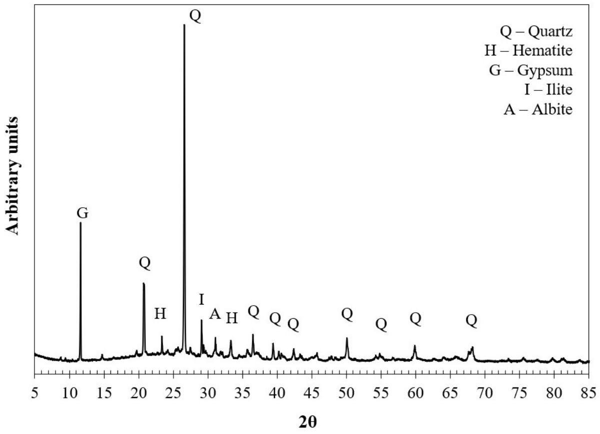


Figure 2. X-ray diffraction pattern of the RBP
Source: Authors

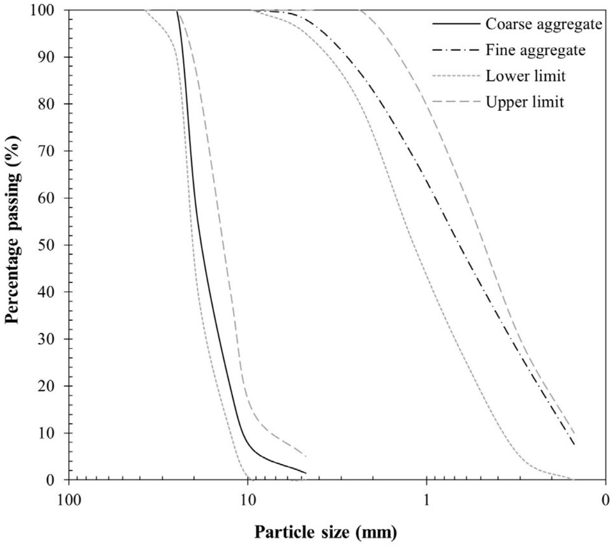


Figure 3. Granulometric distribution of coarse and fine aggregate
Source: Authors

Methods

Definition of mixtures. Two compressive strengths, *i.e.*, 20 (C20) and 25 MPa (C25), were considered since the CBH 87 Bolivian standard establishes them for use in building structures (IBNORCA, 1987). The ACI 211 guidelines (ACI, 1991) were followed as the design mixture method for C20 and C25 concretes. For each design strength, IP 40 cement was replaced with 5 and 10% RBP (by weight). Table 3 summarizes the material amounts for the proposed concretes.

Concrete was prepared in a mechanical mixer, and 10 x 20 cm cylindrical molds were used; the specimens were cured in a humid chamber until the day of their test. The mixing procedure was carried out in accordance with the recommendations of CBH 87 (IBNORCA, 1987). The first half of the water was placed in the mixer, followed by the fine aggregate and binders. The coarse aggregate was then added, and the remaining water was added last. The mixing time was a minimum of 3 minutes.

Table 3. Mix proportions for each concrete strength design

Materials	Mixtures					
	C20			C25		
	Reference (0%)	5% RBP	10% RBP	Reference (0%)	5% RBP	10% RPB
Cement Portland	306,9	291,6	276,2	341,6	324,5	307,4
RBP	0	15,3	30,7	0	17,1	34,2
Water	199,5	199,5	199,5	187,9	187,9	187,9
Sand	903,6	903,6	903,6	1 024,6	1 024,6	1 024,6
Coarse aggregate	1 071,1	1 071,1	1 071,1	1 202,2	1 205.2	1 205.2

Source: Authors

Scanning Electron Microscopy (*SEM*). A microstructural analysis of RBP particle shape was performed by SEM. Hydration products were also detected in concrete mixtures including RBP.

Workability. Workability of the mixtures was evaluated by Abrams's cone slump test, following the UNE-EN 12350-2 standard (Spanish Association for Standardization, 2020). Four measurements were taken per mixture in order to calculate the mean and standard deviation.

Compressive and tensile strength. The compressive and tensile strength of every mixture (by diametrical compression) were determined after 7, 28, and 56 days using ASTM C39 (American Society for Testing and Materials, 2021b) and NBR 7222 (ABNT, 2011), respectively. For both properties, 10 x 20 cm cylindrical molds were used, and four samples for each mixture and age were tested.

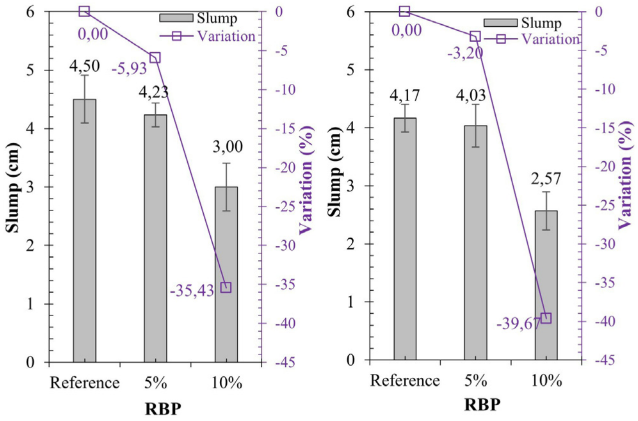
Water absorption, specific mass and void index. Following the ASTM C642 standard (American Society for Testing and Materials, 2021c), water absorption, specific mass, and void index were determined after 28 days. 10 x 20 cm cylindrical molds were used, and four samples were considered for each mixture and physical property.

The data were compared using a one-way analysis of variance (ANOVA) with Tukey's test (p≤0,05).

Results and discussion

Slump test

The slump results for mixtures including RBP are presented in Figure 4. A tendency towards reducing workability for both design resistances was observed. For C20 mixes, the reductions were 5,93 and 35,43% for replacements of 5 and 10% RBP, respectively. Similar results were obtained for the C25 mixes: 3,20 and 39,67% for 5 and 10% RBP.



(a) (b)

Figure 4. Slump for a) C20 and b) C25
Source: Authors

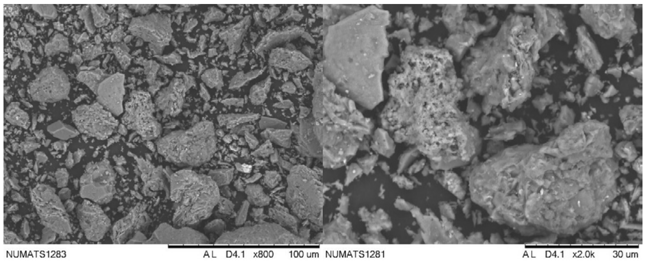
Through the ANOVA, significant differences were found for both C20 (p-value=0,011) and C25 (p-value=0,004). Table 4 presents Tukey's test results. There is no difference between the slump reference and 5% RBP for both C20 and C25, thus confirming that an RBP replacement of up to 5% has no influence on the slump. On the other hand, 10% RBP significantly reduces the slump by more than 35% in both design resistances. Similar results were found by Ge *et al.* (2015). According to them, for low RBP percentages, the workability is similar to the reference, but it is drastically reduced at 20-30% RBP.

Table 4. Tukey's test for the slump

Group 1	Group 2	p-value	
		C20	C25
Reference	5% RBP	0,7427	0,9081
Reference	10% RBP	0,0129	0,0055
5% RBP	10% RBP	0,0302	0,0084

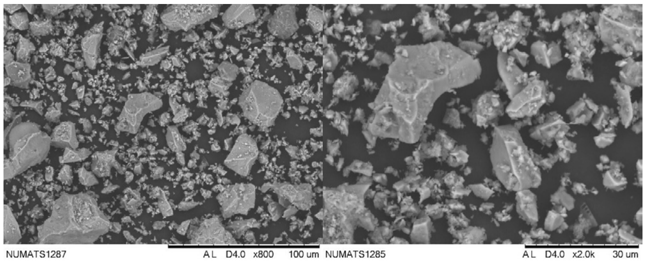
Source: Authors

As the cement replacement with RBP increases, the slump decreases. This behavior is also reported in the literature (Schackow *et al.*, 2015; Luo *et al.*, 2022; Ma *et al.*, 2020a). The workability loss is attributed to irregular RBP particles, thus increasing the water demand and reducing the air content in the mix (Liu *et al.*, 2020; Tang *et al.*, 2020). However, it may be observed that the addition of RBP in low proportions (≤5%) has no influence on the slump. This behavior can be attributed to the fact that RBP (D50) particles are larger than those of cement, in addition to having a lower SSA (Table 2). These characteristics reduce the water demand. However, as shown in Figure 5, RBP particles are porous when compared to IP 40 cement (Figure 6), which could have a negative effect on the workability of the mixes. For 5% RBP, the particle size and the SSA were higher than the RBP porosity, so there were no significant differences. On the other hand, the high slump reduction for 10% RPB may be explained by a) the porosity of the RBP particles, which, for a 10% substitution, had a significant adverse effect on workability; b) the high amount of RBP particles that did not generate an adequate packing, which could entail friction between particles; and c) the shape and rough texture of RBP particles in comparison with IP 40 cement.



(a) (b)

Figure 5. SEM images of RBP at different magnifications: a) 800 and b) 2k
Source: Authors



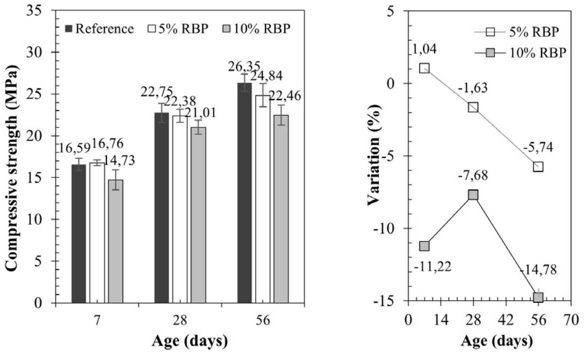
(a) (b)

Figure 6. SEM images of IP 40 cement at different magnifications: a) 800 and b) 2k
Source: Authors

Schackow *et al.* (2015), Ma *et al.* (2020a), and He *et al.* (2021) point out that RBP particles reduce the workability of cement-based materials due to their irregular microstructure and size, thus increasing the water demand. Luo *et al.* (2022) indicate that workability reductions are also related to the SSA of RBP, which, in this case, is similar to that of IP 40 cement (Table 2). Adding more water content and/or using superplasticizers to offset this effect is recommended. Zhao *et al.* (2020) point out that, if the RBP particle size is reduced, the microstructure becomes regular and non-angular, thus generating a lubricating effect and compensating for the water requirements.

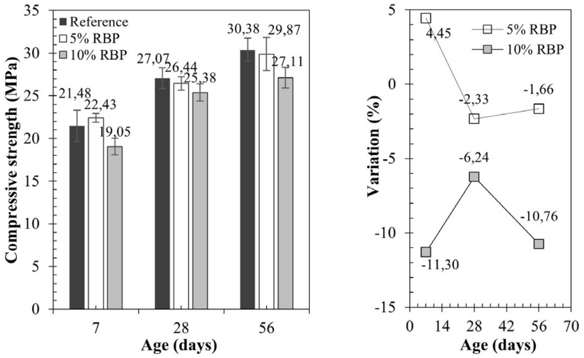
Compressive strength

The compressive strength results are presented in Figures 7 and 8 for C20 and C25, respectively. Regarding C20, a reduction trend is observed when adding RBP. 5% RBP shows the smallest negative variation, which is even higher than the reference for 7 days, *i.e.*, 1,04%. 10% RBP exhibits a greater reduction in compressive strength, with the maximum value being 14,78% for 56 days. However, through the ANOVA, it can be concluded that the compressive strength means are statistically the same for every period, *i.e.*, for 7 (p-value=0,356), 28 (p-value=0,133), and 56 days (p-value=0,054), thus indicating that an addition of 5 and 10% RBP does not have a significant influence on compressive strength. It is important to emphasize that, in all cases, the p-value is greater than 0,05, meaning that the mixtures have statistically equal means. Therefore, the addition of 5 and 10% RBP would not affect this property for C20 concrete.



(a) (b)

Figure 7. a) Compressive strength for C20. b) Compressive strength variation regarding the reference.
Source: Authors



(a) (b)

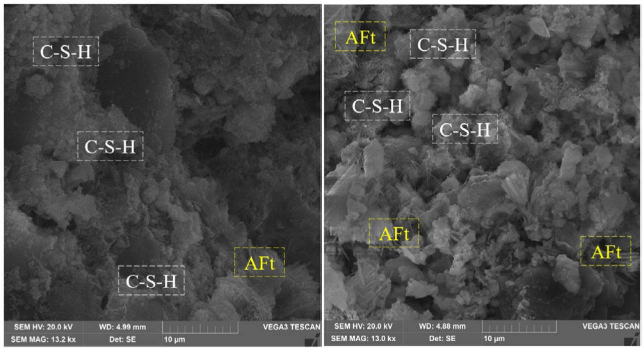
Figure 8. a) Compressive strength for C25. b) Compressive strength variation regarding the reference.
Source: Authors

Likewise, every C25 mix exhibits a negative trend when replacing IP 40 cement with RBP. For 5% RBP, an increase in compressive strength after 7 days (4,45%) is observed. Therefore, later ages exhibit smaller reductions (<2,5%). In the case of 10% RBP, the reductions are greater, with a maximum negative variation of 11,30% for 7 days. Despite a tendency towards reducing compressive strength, the ANOVA confirms that there are no significant differences between the means of mixtures including RBP and the reference, considering stages of 7 (p-value=0,080), 28 (p-value=0,314), and 56 days (p-value=0,150). The p-values for C25 concrete were greater than 0,05, indicating that the means were statistically equal. Therefore, it can also be stated that the addition of 5 and 10% RBP does not affect the compressive strength of C25 concrete at 7, 28, and 56 days, a result similar to that of C20 concrete.

It is observed that every mixture shows an increasing resistance to compression over time due to cement hydration and pozzolanic reaction. The latter is caused by the active particles of SiO_2, Fe_2O_3, and Al_2O_3 (Table 1) present in RBP, which react with calcium hydroxide (CH) to form C-S-H, C-A-H, and C-A-S-H gel (Shao *et al*., 2019). In most cases, it is observed that the compressive strength of the reference is higher on average than that of the mixtures including RBP. This is due to the clinker dilution effect, which reduces the amount of hydrated products (Luo *et al*., 2022).

On the other hand, it can be observed that mixtures including 5% RBP at the seven-day stage exhibit a higher compressive strength (on average) than the reference. Shao *et al*. (2019) point out that RBP has a filler effect, which develops at the first stages of cement-based materials. In this case, it was up to 7 days and only for 5% RBP. Subsequently, the filler effect was reduced as the stages went by.

Figure 9 shows the microstructure of mixtures including RBP at 28 days, where hydrated products are observed, mainly C-S-H and ettringite (AFt). However, CH could not be distinguished, which would indicate its partial consumption by pozzolanic reaction (Ortega *et al*., 2018; Luo *et al*., 2022; Wong *et al*., 2018).



(a) (b)

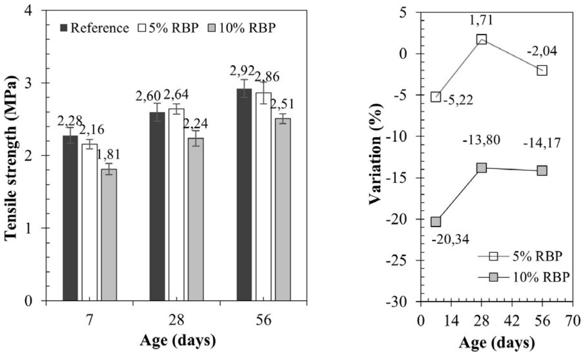
Figure 9. Microstructure of mixtures including a) 5% RBP and b) 10% RBP
Source: Authors

The literature reports that cement-based materials decrease in compressive strength as the RBP proportion increases (Shao *et al*., 2019; Duan *et al*., 2020; He *et al*., 2021). However, the hydration products obtained by pozzolanic reaction were beneficial, as the compressive strength of every mixture statistically remained similar to the reference (up to 10% RBP). This may be due to pore refinement by pozzolanic activity and the filler effect (Ortega *et al*., 2018). In this sense, several authors suggest using lower replacement percentages (~10-20% RBP) to avoid reduction – and even increase compressive strength (Schackow *et al*., 2015; Toledo Filho *et al*., 2007; Ma *et al*., 2020b).

The results are similar to those reported by Arif *et al*. (2021), who indicated a positive variation of 5,43% in the compressive strength of 5% RBP after 28 days. Liu *et al*. (2020) also reported that the compressive strength is comparable to the reference in replacements of up to 10% RBP, highlighting the influence of particle size in maintaining or improving the mechanical properties of mortars. Similarly, Wang *et al*. (2022) indicated that, for advanced ages, 10% RBP only reduces the compressive strength by 5%, which suggests that the pozzolanic reaction of RBP has an influence at later ages. Although literature may report variations in compressive strength with the use of RBP (for replacements of up to 10%), these differences are not significant, which was also verified in our study. However, higher substitution percentages and a larger particle size can significantly affect compressive strength. On the other hand, it is important to mention that the RBP used did not require additional treatment, for example, chemical activation, which highlights the benefits of its use in terms of energy expenditure and economic feasibility.

Tensile strength by diametrical compression

Figures 10 and 11 present the results obtained regarding tensile strength by diametrical compression for the C20 and C25 mixes. For C20, this aspect shows a tendency to decrease, except for 5% RBP after 28 days, where an increase of 1,71% is observed. Through the ANOVA, significant differences are noted between the means of the mixes after 7 (p-value=0,004), 28 (p-value=0,014), and 56 days (p-value=0,028). Table 5 shows that the reference means and those of 5% RBP are the same, thus indicating that no reduction in tensile strength takes place for this RBP percentage.



(a) (b)

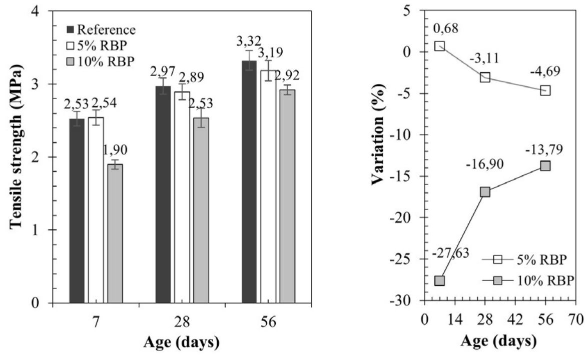
Figure 10. a) Tensile strength for C20. b) Tensile strength variation regarding the reference.
Source: Authors

Table 5. Tukey's test for the tensile strength of C20

Group 1	Group 2	p-value		
		7 days	28 days	56 days
Reference	5% RBP	0,408	0,902	0,877
Reference	10% RBP	0,004	0,029	0,032
5% RBP	10% RBP	0,017	0,018	0,059

Source: Authors

Every C25 mixture shows a trend similar to the results obtained with C20, as well as reductions in tensile strength. Through the ANOVA, significant differences for the ages of 7 (p-value=0,001), 14 (p-value=0,021), and 28 days (p-value=0,037) are found. It can be stated that the tensile strength means of the reference and those of 5% RBP are the same (Table 6). However, note that, after 28 and 56 days, there is no difference in average tensile strength for 5 and 10% RBP.



(a) (b)

Figure 11. a) Tensile strength for C25. b) Tensile strength variation regarding the reference.
Source: Authors

Table 6. Tukey's test for the tensile strength of C25

Group 1	Group 2	p-value		
		7 days	28 days	56 days
Reference	5% RBP	0,984	0,779	0,516
Reference	10% RBP	0,001	0,022	0,033
5% RBP	10% RBP	0,001	0,051	0,141

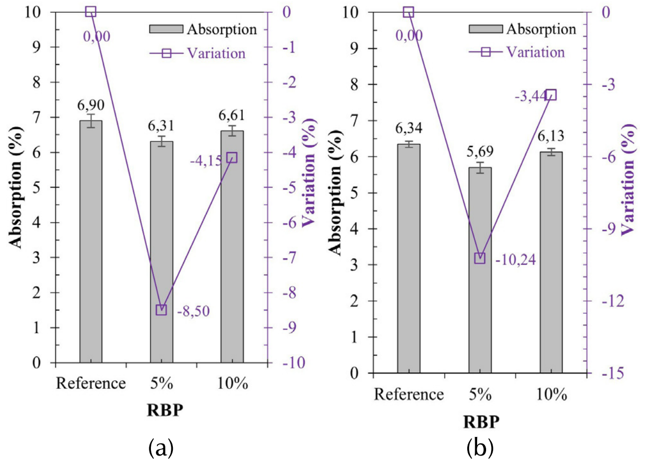
Source: Authors

In general, it has been reported that using RBP in smaller amounts does not significantly influence tensile strength (Ge *et al*., 2015; Ortega *et al*., 2018). The results follow the same trend regarding compressive strength, where IP 40 cement dilution is observed. However, RBP's pozzolanic reaction leads to forming C-S-H gel (He *et al*., 2021), which compensates for the dilution effect in the 5% replacement. A replacement of 10% RBP entails a significant reduction, indicating that the mechanical resistance of the cement matrix was negatively affected. Irki *et al*. (2018) pointed out that tensile strength increases as the RBP fineness and SSA increase, but the RBP used has a similar SSA to that of IP 40 cement as well as a higher D50, so this effect was not observed.

Water absorption, specific mass, and void index

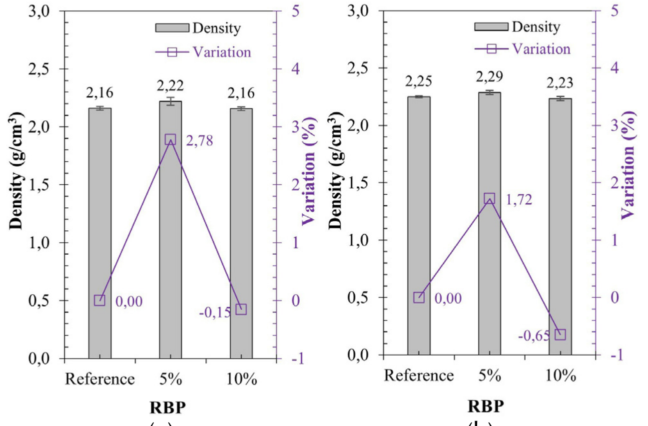
Figures 12, 13 and 14 present the physical properties of C20 and C25. On the one hand, it is observed that a 5% RBP replacement reduces absorption, increases density, and reduces the void content of every mixture. On the other hand, a 10% RBP replacement reduces absorption, density, and voids to a lesser extent when compared to the former, regardless of design strength.

Although a decrease in water absorption is observed when including RBP, there is a predominance for 5% RBP, especially in C25, with a reduction of 10,24%. Through the ANOVA, it is confirmed that there are differences between means of C20 (p-value=0,028) and C25 (p-value=0,004). In the case of C20, only a difference between the reference and 5% RBP is noted (Table 7). For C25, differences between two pairs of mixes are evidenced (Table 8). These results follow the same trend as those of C20, indicating that a 5% RBP replacement has a positive influence on absorption (reduction) and that higher RBP percentages do not show variations in this regard.



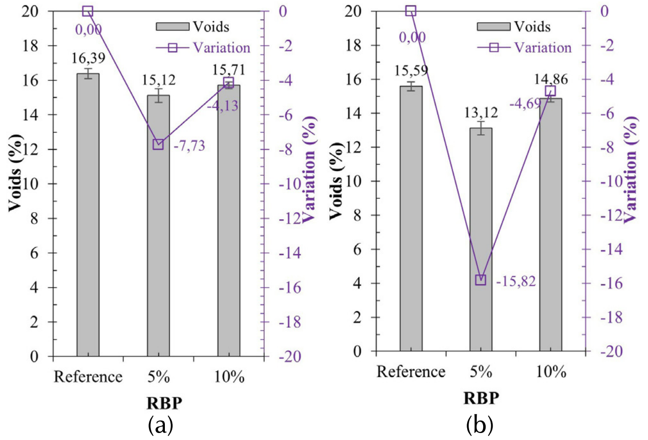
(a) (b)

Figure 12. Water absorption for a) C20 and b) C25
Source: Authors



(a) (b)

Figure 13. Density for a) C20 and b) C25
Source: Authors



(a) (b)

Figure 14. Voids for a) C20 and b) C25
Source: Authors

A 5% RBP replacement increases the density by 2,78 and 1,72% in C20 and C25, respectively, while 10% RBP minimally reduces it to 0,15 and 0,65%. Through the ANOVA, it is confirmed that there are differences between the means of the C25 mixes (p-value=0,037), as opposed to C20 (p-value=0,073). However, for C25, only 5 and 10% RBP mixtures show differences (Table 8). The results indicate that cement replaced by up to 10% RBP does not significantly impact density.

Table 7. Tukey's test for the physical features of C20

Group 1	Group 2	p-value		
		Absorption	Density	Voids
Reference	5% RBP	0,024	0,111	0,015
Reference	10% RBP	0,248	0,989	0,153
5% RBP	10% RBP	0,223	0,093	0,217

Source: Authors

Table 8. Tukey's test for the physical features of C25

Group 1	Group 2	p-value		
		Absorption	Density	Voids
Reference	5% RBP	0,004	0,111	0,001
Reference	10% RBP	0,235	0,651	0,103
5% RBP	10% RBP	0,025	0,035	0,002

Source: Authors

The void content also shows the same trend as the absorption results. A greater reduction in mixes including 5% RBP is noted, highlighting 15,82% for C25. A 10% RBP proportion yields lower reduction results when compared to those of 5% RBP. The ANOVA shows void content means with significant differences for C20 (p-value=0,018) and C25 (p-value=0,001). For both design resistances, there is a difference between the reference and 5% RBP, as well as a significant difference between 5 and 10% RBP only for C25.

Schackow *et al.* (2015) reported that the values of apparent porosity and water absorption in mixes including RBP remain within the same range as the reference. In our case, the mixes including 5% RBP show a less porous structure. This is also corroborated by the density results, where, on average, the mixes with 5% RBP exhibit an increase, indicating that RBP particles can fill and reduce the volume of the pores due to physical effects and the pozzolanic reaction of amorphous compounds present in RBP. Likes *et al.* (2022) also point out that RBP yields mixes with higher density, which is due to the granulometry of RBP creating a complete particle distribution when combined with Portland cement.

Note that 5% RBP improved the physical properties of the concretes, such as absorption and their void index, which were reduced. This is attributed to the filling effect and pozzolanic reactions, whereby the new hydration products are located in the capillary pores, refining them, resulting in an improved cementitious matrix (Ma *et al.*, 2020a). On the contrary, the use of 10% RBP did not yield these positive effects; the physical properties were similar to those of the reference, indicating a compensation between the dilution effect of Portland cement and the filling effect and pozzolanic reactions. Liu *et al.* (2020) also pointed out that the porosity of mortars with RBP is equal to those of the reference; even a replacement of up to 30% RBP could show the same tendency, albeit with the use of finer RBP particles (0-0,045 mm). Regarding density, these authors also indicated that there is no significant difference with the use of RBP. Wang *et al.* (2022) noted that, despite the dilution effect of RBP (up to 20% RBP), secondary hydration (pozzolanic reaction) could refine the matrix pores. Ma *et al.* (2020a) reported the same results for a particle size D50 = 42 μm, where the water absorption of replacements of up to 15% RBP was similar to the reference. Therefore, physical properties depend on the percentages of fineness and substitution. In this case, the percentages (up to 10%) and the particle size (D50=27,99 μm) employed could maintain and/or improve the physical properties of the concretes.

Conclusions

Directly extracted from demolition and construction sites, and without the need for pre-treatments, RBP may be used as an SCM, replacing Portland cement to produce concrete. RBP stands out for its pozzolanic activity, it reduces cement consumption, and its mechanical resistance is similar to that of the reference, especially its resistance to compression. Tensile strength is compromised by a 10% RBP replacement, which was observed in both C20 and C25 strength designs.

Mixtures including RBP lose workability, which is due to the size, shape, and rough texture of RBP particles. However, 5% RBP did not entail changes in workability, being statistically similar to the reference. For 10% RBP, using superplasticizers could be considered in order to improve this property.

Replacing Portland cement with 5% RBP benefits the physical properties of concrete. Its absorption and void content are reduced, and the density of any mix is slightly improved, which indicates a refining of the porous structure. On the other hand, using 10% RBP maintains the physical characteristics of the reference.

Both 5 and 10% RBP are suitable for use in structural concrete, regardless of the design resistance (C20 or C25). Although 10% RBP reduces the tensile strength, this property is not considered when calculating reinforced concrete structures. Therefore, the proposed RBP-added concretes are viable for use in the construction industry, as their physical-mechanical properties were equal to the reference. However, in order to increase their mechanical resistance, the RBP particle size could be reduced and, therefore, its SSA could be increased, thus improving the physical effect.

Using RBP as a replacement for Portland cement represents a sustainable approach in the construction industry. On the one hand, reducing the amount of Portland cement results in less CO_2 emissions. On the other hand, any waste would be reused and would prevent disposal in landfills. Nevertheless, future works could determine the CO_2 reductions derived from the use of RBP as a partial substitute for Portland cement, which will make it possible to more accurately determine the sustainability of the proposal.

Acknowledgements

The authors would like to acknowledge CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for their financial support and Universidad Privada del Valle for their institutional support.

CRediT author statement

JHAR conceived the idea and did the background research. BMMR collected the data, developed the workflow, and performed assessments. RDTF provided critical feedback. JHAR led the drafting process and wrote the main part of the manuscript, to which all authors contributed.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

Adamson, M., Razmjoo, A., and Poursaee, A. (2015). Durability of concrete incorporating crushed brick as coarse aggregate. *Construction and Building Materials, 94*, 426-432. https://doi.org/10.1016/j.conbuildmat.2015.07.056

American Society for Testing and Materials (2015a). *ASTM C136-06. Standard test method for sieve analysis of fine and coarse aggregates*. ASTM. https://doi.org/10.1520/C0136-06

American Society for Testing and Materials (2015b). *ASTM C127-15. Standard test method for relative density (specific gravity) and absorption of coarse aggregate*. ASTM. https://doi.org/10.1520/C0127-15

American Society for Testing and Materials (2015c). *ASTM C128-15. Standard test method for relative density (specific gravity) and absorption of fine aggregate*. ASTM. https://doi.org/10.1520/C0128-15

American Society for Testing and Materials (2019). *ASTM C204. Standard test methods for fineness of hydraulic cement by air-permeability apparatus*. ASTM. https://doi.org/10.1520/C0204-18E01

American Society for Testing and Materials (2021a). *ASTM C595/C595M-21. standard specification for blended hydraulic cements*. ASTM. https://doi.org/10.1520/C0595_C0595M-21

American Society for Testing and Materials (2021b). *ASTM C39/C39M-21. Standard test method for compressive strength of cylindrical concrete specimens*. ASTM. https://doi.org/10.1520/C0039_C0039M-21

American Society for Testing and Materials (2021c). *ASTM C642-21. Standard test method for density, absorption, and voids in hardened concrete*. ASTM. https://doi.org/10.1520/C0642-21

American Society for Testing and Materials (2022). *ASTM C618-22. Standard specification for coal fly ash and raw or calcined natural pozzolan for use in concrete*. ASTM: West Conshohocken, PA, USA, 2022. https://doi.org/10.1520/C0618-22

Arif, R., Khitab, A., Kırgız, M. S., Khan, R. B. N., Tayyab, S., Khan, R. A., Tayyab, S., Akhtar Khan, R., Anwar, W., and Arshad, M. T. (2021). Experimental analysis on partial replacement of cement with brick powder in concrete. *Case Studies in Construction Materials, 15*, e00749. https://doi.org/10.1016/j.cscm.2021.e00749

Asociación Española de Normalización (2020). *UNE-EN 12350-2: Ensayos de hormigón fresco. Parte 2: ensayo de asentamiento*. UNE.

Associação Brasileira de Normas Técnicas (2011). *NBR 7222: Concreto e argamassa — Determinação da resistência à tração por compressão diametral de corpos de prova cilíndricos*. ABNT.

Borges, P. M., Schiavon, J. Z., da Silva, S. R., Rigo, E., Junior, A. N., Possan, E., and de Oliveira Andrade, J. J. (2023). Mortars with recycled aggregate of construction and demolition waste: Mechanical properties and carbon uptake. *Construction and Building Materials, 387*, 131600. https://doi.org/10.1016/j.conbuildmat.2023.131600

Da Silva, S. R., and Andrade, J. J. D. O. (2022). A review on the effect of mechanical properties and durability of concrete with construction and demolition waste (CDW) and fly ash in the production of new cement concrete. *Sustainability, 14*(11), 6740. https://doi.org/10.3390/su14116740

da Silva, S. R., Cimadon, F. N., Borges, P. M., Schiavon, J. Z., Possan, E., and de Oliveira Andrade, J. J. (2022). Relationship between the mechanical properties and carbonation of concretes with construction and demolition waste. *Case Studies in Construction Materials, 16*, e00860. https://doi.org/10.1016/j.cscm.2021.e00860

da Silva, S. R., de Brito, J., and de Oliveira Andrade, J. J. (2023). Synergic effect of recycled aggregate, fly ash, and hydrated lime in concrete production. *Journal of Building Engineering, 70*, 106370. https://doi.org/10.1016/j.jobe.2023.106370

Duan, Z., Hou, S., Xiao, J., and Li, B. (2020). Study on the essential properties of recycled powders from construction and demolition waste. *Journal of Cleaner Production, 253*, 119865. https://doi.org/10.1016/j.jclepro.2019.119865

Gao, T., Dai, T., Shen, L., and Jiang, L. (2021). Benefits of using steel slag in cement clinker production for environmental conservation and economic revenue generation. *Journal Of Cleaner Production, 282*, 124538. https://doi.org/10.1016/j.jclepro.2020.124538

Ge, P., Huang, W., Zhang, J., Quan, W., and Guo, Y. (2021). Mix proportion design method of recycled brick aggregate concrete based on aggregate skeleton theory. *Construction and Building Materials, 304*, 124584. https://doi.org/10.1016/j.conbuildmat.2021.124584

Ge, Z., Wang, Y., Sun, R., Wu, X., and Guan, Y. (2015). Influence of ground waste clay brick on properties of fresh and hardened concrete. *Construction and Building Materials, 98*, 128-136. https://doi.org/10.1016/j.conbuildmat.2015.08.100

Habert, G., Miller, S. A., John, V. M., Provis, J. L., Favier, A., Horvath, A., and Scrivener, K. L. (2020). Environmental impacts and decarbonization strategies in the cement and concrete industries. *Nature Reviews Earth & Environment, 1*(11), 559-573. https://doi.org/10.1038/s43017-020-0093-3

He, Z., Shen, A., Wu, H., Wang, W., Wang, L., Yao, C., and Wu, J. (2021). Research progress on recycled clay brick waste as an alternative to cement for sustainable construction materials. *Construction and Building Materials, 274*, 122113. https://doi.org/10.1016/j.conbuildmat.2020.122113

IBNORCA (1987). *CBH 87: Código boliviano del hormigón*. Instituto Boliviano de Normalización y Calidad.

ACI (1991). *Standard practice for selecting proportions for normal, heavyweight and mass concrete* (*ACI 211.1-91*). ACI.

Irki, I., Debieb, F., Ouzadid, S., Dilmi, H. L., Settari, C., and Boukhelkhel, D. (2018). Effect of Blaine fineness of recycling brick powder replacing cementitious materials in self compacting mortar. *Journal of adhesion science and Technology, 32*(9), 963-975. https://doi.org/10.1080/01694243.2017.1393202

Jiang, W., Li, X., Lv, Y., Jiang, D., Liu, Z., and He, C. (2020). Mechanical and hydration properties of low clinker cement containing high volume superfine blast furnace slag and nano silica. *Construction and Building Materials, 238*, 117683. https://doi.org/10.1016/j.conbuildmat.2019.117683

Jiang, X., Xiao, R., Bai, Y., Huang, B., and Ma, Y. (2022). Influence of waste glass powder as a supplementary cementitious material (SCM) on physical and mechanical properties of cement paste under high temperatures. *Journal of Cleaner Production, 340*, 130778. https://doi.org/10.1016/j.jclepro.2022.130778

Juenger, M. C., and Siddique, R. (2015). Recent advances in understanding the role of supplementary cementitious materials in concrete. *Cement and Concrete Research, 78*, 71-80. https://doi.org/10.1016/j.cemconres.2015.03.018

Juenger, M. C., Snellings, R., and Bernal, S. A. (2019). Supplementary cementitious materials: New sources, characterization, and performance insights. *Cement and Concrete Research, 122*, 257-273. https://doi.org/10.1016/j.cemconres.2019.05.008

Likes, L., Markandeya, A., Haider, M. M., Bollinger, D., McCloy, J. S., and Nassiri, S. (2022). Recycled concrete and brick powders as supplements to Portland cement for more sustainable concrete. *Journal of Cleaner Production, 364*, 132651. https://doi.org/10.1016/j.jclepro.2022.132651

Likes, L., Markandeya, A., Haider, M. M., Bollinger, D., McCloy, J. S., and Nassiri, S. (2022). Recycled concrete and brick powders as supplements to Portland cement for more sustainable concrete. *Journal of Cleaner Production, 364*, 132651. https://doi.org/10.1016/j.jclepro.2022.132651

Liu, Q., Li, B., Xiao, J., and Singh, A. (2020). Utilization potential of aerated concrete block powder and clay brick powder from C&D waste. *Construction and Building Materials, 238*, 117721. https://doi.org/10.1016/j.conbuildmat.2019.117721

Luo, X., Gao, J., Li, S., Xu, Z., and Chen, G. (2022). Experimental study on the early-age properties of cement pastes with recycled brick powder. *Construction and Building Materials, 347*, 128584. https://doi.org/10.1016/j.conbuildmat.2022.128584

Ma, Z., Tang, Q., Wu, H., Xu, J., and Liang, C. (2020a). Mechanical properties and water absorption of cement composites with various fineness and contents of waste brick powder from C&D waste. *Cement and Concrete Composites, 114*, 103758. https://doi.org/10.1016/j.cemconcomp.2020.103758

Ma, Z., Liu, M., Duan, Z., Liang, C., and Wu, H. (2020b). Effects of active waste powder obtained from C&D waste on the microproperties and water permeability of concrete. *Journal of Cleaner Production, 257*, 120518. https://doi.org/10.1016/j.jclepro.2020.120518

Mohammed, T. U., Hasnat, A., Awal, M. A., and Bosunia, S. Z. (2015). Recycling of brick aggregate concrete as coarse aggregate. *Journal of Materials in Civil Engineering, 27*(7), B4014005. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001043

Naceri, A., and Hamina, M. C. (2009). Use of waste brick as a partial replacement of cement in mortar. *Waste Management, 29*(8), 2378-2384. https://doi.org/10.1016/j.wasman.2009.03.026

Nguyen, M. H., Trinh, S. H., and Ly, H. B. (2023). Toward improved prediction of recycled brick aggregate concrete compressive strength by designing ensemble machine learning models. *Construction and Building Materials, 369*, 130613. https://doi.org/10.1016/j.conbuildmat.2023.130613

Ortega, J. M., Letelier, V., Solas, C., Moriconi, G., Climent, M. Á., and Sánchez, I. (2018). Long-term effects of waste brick powder addition in the microstructure and service properties of mortars. *Construction and Building Materials, 182*, 691-702. https://doi.org/10.1016/j.conbuildmat.2018.06.161

Panesar, D. K., and Zhang, R. (2020). Performance comparison of cement replacing materials in concrete: Limestone fillers and supplementary cementing materials – A review. *Construction and Building Materials, 251*, 118866. https://doi.org/10.1016/j.conbuildmat.2020.118866

Reig, L., Tashima, M. M., Borrachero, M. V., Monzó, J., Cheeseman, C. R., and Payá, J. (2013). Properties and microstructure of alkali-activated red clay brick waste. *Construction and Building Materials, 43*, 98-106. https://doi.org/10.1016/j.conbuildmat.2013.01.031

Rocha, J. H. A., Toledo Filho, R. D., and Cayo-Chileno, N. G. (2022). Sustainable alternatives to CO2 reduction in the cement industry: A short review. *Materials Today, Proceedings, 52*(2), 436-439. https://doi.org/10.1016/j.matpr.2021.12.565

Schackow, A., Stringari, D., Senff, L., Correia, S. L., and Segadães, A. M. (2015). Influence of fired clay brick waste additions on the durability of mortars. *Cement and Concrete Composites, 62*, 82-89. https://doi.org/10.1016/j.cemconcomp.2015.04.019

Shao, J., Gao, J., Zhao, Y., and Chen, X. (2019). Study on the pozzolanic reaction of clay brick powder in blended cement pastes. *Construction and Building Materials, 213*, 209-215. https://doi.org/10.1016/j.conbuildmat.2019.03.307

Singh, N. B., and Middendorf, B. (2020). Geopolymers as an alternative to Portland cement: An overview. *Construction and Building Materials, 237*, 117455. https://doi.org/10.1016/j.conbuildmat.2019.117455

Sousa, V., and Bogas, J. A. (2021). Comparison of energy consumption and carbon emissions from clinker and recycled cement production. *Journal of Cleaner Production, 306*, 127277. https://doi.org/10.1016/j.jclepro.2021.127277

Tang, Q., Ma, Z., Wu, H., and Wang, W. (2020). The utilization of eco-friendly recycled powder from concrete and brick waste in new concrete: A critical review. *Cement and Concrete Composites, 114*, 103807. https://doi.org/10.1016/j.cemconcomp.2020.103807

Thomas, B. S., Yang, J., Bahurudeen, A., Abdalla, J. A., Hawileh, R. A., Hamada, H. M., Nazar, S., Jittin, V., and Ashish, D. K. (2021). Sugarcane bagasse ash as supplementary cementitious material in concrete–A review. *Materials Today, Sustainability, 15*, 100086. https://doi.org/10.1016/j.mtsust.2021.100086

Toledo Filho, R. D., Gonçalves, J. P., Americano, B. B., and Fairbairn, E. M. R. (2007). Potential for use of crushed waste calcined-clay brick as a supplementary cementitious material in Brazil. *Cement and Concrete Research, 37*(9), 1357-1365. https://doi.org/10.1016/j.cemconres.2007.06.005

Tripathy, A., and Acharya, P. K. (2022). Characterization of bagasse ash and its sustainable use in concrete as a supplementary binder – A review. *Construction and Building Materials, 322*, 126391. https://doi.org/10.1016/j.conbuildmat.2022.126391

UN Environment, Scrivener, K. L., John, V. M., and Gartner, E. M. (2018). Eco-efficient cements: Potential economically viable solutions for a low-CO2 cement-based materials industry. *Cement and Concrete Research, 114*, 2-26. https://doi.org/10.1016/j.cemconres.2018.03.015

Vieira, C. S. (2020). Valorization of fine-grain construction and demolition (C&D) waste in geosynthetic reinforced structures. *Waste and Biomass Valorization, 11*(4), 1615-1626. https://doi.org/10.1007/s12649-018-0480-x

Wang, H., Wang, L., Qian, X., Cao, K., Xu, Y., Fang, Y., and Cui, L. (2022). Hydration, compressive strength and durability of eco-friendly cement mortars containing recycled brick powder and metakaolin. *KSCE Journal of Civil Engineering, 26*(9), 4023-4037. https://doi.org/10.1007/s12205-022-0035-3

Wong, C. L., Mo, K. H., Yap, S. P., Alengaram, U. J., and Ling, T. C. (2018). Potential use of brick waste as alternate concrete-making materials: A review. *Journal of Cleaner Production, 195*, 226-239. https://doi.org/10.1016/j.jclepro.2018.05.193

Zhao, Y., Gao, J., Liu, C., Chen, X., and Xu, Z. (2020). The particle-size effect of waste clay brick powder on its pozzolanic activity and properties of blended cement. *Journal of Cleaner Production, 242*, 118521. https://doi.org/10.1016/j.jclepro.2019.118521

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e106174)

Research Article/Civil and Sanitary Engineering http://doi.org/10.15446/ing.investig.106174

Influence of Expanded Clay Aggregate on the Engineering Properties of Lightweight Concrete

Influencia del agregado de arcilla expandida en las propiedades de ingeniería del hormigón ligero

As'at Pujianto[1], Hakas Prayuda[2], Farrel Asani[3], Muji Basuki Santoso[4], and Fahriza Wirawan[5]

ABSTRACT

In seismically active locations, civil infrastructures, such as buildings, bridges, and dams, are frequently subjected to earthquakes. Using lightweight construction materials is one method for enhancing the seismic resistance of infrastructure. This study examined the engineering properties of lightweight concrete manufactured using expanded clay aggregate, with the purpose of developing sustainable and environmentally friendly building materials. Laboratory tests focused on the effects of the aggregate shape and the supplementary superplasticizer, as well as on the influence of the concrete age. Experimental studies were conducted to measure fresh (slump) and hardened properties (compressive strength, splitting tensile strength, and density). The expanded clay aggregate was produced by burning at a temperature of 800 to 1 200 °C. Cubic, oval, and round aggregate shapes with a maximum size of 20 mm were evaluated. This study also examined the effect of superplasticizers on the engineering properties of lightweight concrete. The composition of the superplasticizer varied from 0 to 2,5%. According to the experimental results, the engineering properties of lightweight concrete made with oval aggregates are advantageous in comparison with those using cubic and round shapes. It is also demonstrated that optimal amounts of superplasticizer are necessary to develop materials with adequate properties. It can be concluded that expanded clay aggregate can be used as an alternative material to produce lightweight concrete.

Keywords: expanded clay aggregate, lightweight concrete, compressive strength, splitting tensile strength, superplasticizer

RESUMEN

En lugares sísmicamente activos, las infraestructuras civiles, como edificios, puentes y represas, están frecuentemente sujetas a terremotos. El uso de materiales de construcción livianos es un método para mejorar la resistencia sísmica de la infraestructura. Este estudio examinó las propiedades de ingeniería del hormigón ligero fabricado con agregado de arcilla expandida, con el objetivo de desarrollar materiales de construcción sostenibles y respetuosos con el medio ambiente. Las pruebas de laboratorio se enfocaron en los efectos de la forma del agregado y el superplastificante suplementario, así como en la influencia de la edad del concreto. Se realizaron estudios experimentales para medir las propiedades en estado fresco (asentamiento) y endurecido (resistencia a la compresión, resistencia a la tracción por división y densidad). El agregado de arcilla expandida se produjo mediante incineración a una temperatura de 800 a 1 200 °C. Se evaluaron agregados de forma cúbica, ovalada y redonda, con un tamaño máximo de 20 mm. Este estudio también examinó el efecto de los superplastificantes en las propiedades de ingeniería del hormigón ligero. La composición del superplastificante varió de 0 a 2,5 %. De acuerdo con los resultados experimentales, las propiedades de ingeniería del hormigón ligero hecho con formas ovaladas son ventajosas en comparación con los que utilizan formas cúbicas y redondas. También se demuestra que se necesitan cantidades óptimas de superplastificante para desarrollar materiales con propiedades adecuadas. Se puede concluir que el agregado de arcilla expandida se puede utilizar como material alternativo para producir hormigón liviano.

Palabras clave: agregado de arcilla expandida, hormigón ligero, resistencia a la compresión, resistencia a la tracción por división, superplastificante

Received: December 2nd, 2022
Accepted: August 17th, 2023

Introduction

Concrete is one of the most used construction materials in the world. This is due to the fact that the components for producing concrete are accessible in a wide range of locations and relatively inexpensive compared to other materials. Concrete has several benefits as a construction material, including high resistance to compressive stress, fire resistance, low cost, and long-term serviceability. However, concrete also has several drawbacks, which have recently become the main concern of various countries, especially in terms of sustainability and its impact on the environment (Monika *et al.*, 2022; Saleh *et al.*, 2022).

---

[1] Corresponding author, Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Tamantiroto, Kasihan, Bantul, 55182, Indonesia. Email: pujiantoasat@umy.ac.id

[2] Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Tamantiroto, Kasihan, Bantul, 55182, Indonesia. Email: hakasprayuda@umy.ac.id

[3] Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Tamantiroto, Kasihan, Bantul, 55182, Indonesia. Email: farrel.asani.ft18@mail.umy.ac.id

[4] Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Tamantiroto, Kasihan, Bantul, 55182, Indonesia. Email: muji.basuki.ft18@mail.umy.ac.id

[5] Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Tamantiroto, Kasihan, Bantul, 55182, Indonesia. Email: fahriza.wirawan.ft18@mail.umy.ac.id

Concrete construction always generates waste and gas emissions. It starts with collecting constituent materials, extending through the construction stage, service life, the monitoring stage, and the demolition stage. In addition, producing concrete requires using natural resources, such as sand and rock, which may cause damage to the ecosystem, most notably the functional deterioration of river systems. According to the findings of previous studies, the yearly manufacture of concrete exceeds 10 billion metric tons, with aggregate consumption responsible for almost 70% (7,5 billion metric tons) and cement consumption accounting for roughly 1,2 billion metric tons (Bogas *et al*., 2015; Meyer, 2009; Bogas *et al*., 2014). Therefore, extensive innovation and the study of more sustainable and environmentally friendly alternative materials must be conducted in order to contribute to the development of the global construction industry.

In addition, the mass density of normal concrete is between 2 200 and 2 400 kg/m^3. This mass of concrete has also been widely studied by researchers and is still an open issue. Especially in earthquake-prone areas, the collapse of concrete structures with a high density might result in a fatal disaster. Several types of construction require a lighter structure to reduce density. The development of lightweight concrete is closely linked to the option of using lightweight materials as a substitute for aggregate, especially considering that coarse aggregate has the most significant proportion in concrete composition. Various innovations to produce lighter concrete have been explored over the years, such as using coarse aggregate replacement material with waste from the agriculture industry (Shafigh *et al*., 2011; Shafigh *et al*., 2014; Aslam *et al*., 2016; Mannan *et al*., 2006), using aggregate from plastic waste (Alqahtani *et al*., 2017; Akcaozoglu *et al*., 2010; Castillo *et al*., 2020), using pumice aggregate (Pujianto and Prayuda, 2021; Rashad, 2019; Karthika *et al*., 2021), using recycled aggregate (Hassan *et al*., 2015; Posi *et al*., 2013; Saha *et al*., 2021), and using expanded clay aggregate (Ahmad *et al*., 2019; Ahmad and Chen, 2019; Rumsys *et al*., 2017).

Lightweight cement-based materials have recently been the focus of research on lightweight materials for concrete production. This project aims to develop cement substitutes that can provide lightweight concrete and reduce the consumption of Portland cement, whose production process emits excessive amounts of carbon dioxide (Liu *et al*., 2021; Pelisser *et al*., 2012). There are several studies on lightweight cement-based materials, some of which use cenosphere and fly ash (Arunachalam *et al*., 2023; Kavinkumar *et al*., 2023; Souza *et al*., 2019; Chen and Huang, 2019; Shi *et al*., 2022; Hanif *et al*., 2017; Brooks *et al*., 2021; Aungatichart *et al*., 2022), pumice powder (Hossain, 2004; Kurt *et al*., 2016), nano-silica materials (Sikora *et al*., 2020; Du *et al*., 2015; Adhikary *et al*., 2021; Elrahman *et al*., 2019; Adhikary *et al*., 2020; Gayathiri and Praveenkumar, 2022; Zhang *et al*., 2018). Using cement-based materials can produce lightweight concrete that is environmentally friendly, of high quality, and similar to normal concrete. However, when compared to conventional concrete or lightweight concrete using coarse aggregate replacement, the manufacturing process for lightweight concrete employing cement-based material replacement requires skills, a high level of accuracy, and higher production costs.

Each alternative material for lightweight concrete has different advantages and disadvantages. Producing lightweight concrete from plastic waste requires a long process to produce aggregates that can be used as lightweight raw material. Additionally, the utilization of agricultural waste is still debatable due to its durability and resistance to fire. This research aims to utilize expanded clay aggregate as a lightweight artificial aggregate to produce more sustainable and environmentally friendly lightweight concrete. There is research on expanded clay aggregate for light concrete, focusing on physical and mechanical properties (Campione *et al*., 2001; Kulkarni and Muthadhi, 2020; Dabbaghi *et al*., 2021) and durability (Hubertova and Hela, 2013; Nahhab and Ketab, 2020; Nawel *et al*., 2017). Based on the results of said research, it can be concluded that this expanded clay aggregate material is suitable for use in the production of lightweight concrete. In addition, using this material can improve thermal performance, which may reduce energy consumption in civil infrastructures (Vijayalakshmi and Ramanagopal, 2018). However, the compressive strength obtained is always lower than that of normal concrete. Some research has been conducted on utilizing the combination of expanded clay aggregate and other light materials, such as expanded glass aggregate (Rumsys *et al*., 2018; Adhikary *et al*., 2020; Augonis *et al*., 2022), metakaolin, silica fume, and fly ash (Tawfik *et al*., 2021; Vivek *et al*., 2022; Mohammed *et al*., 2022). It has also been determined that expanded clay aggregate could be combined with other materials, but several studies have found that the clay source and processing method significantly affect the performance of lightweight concrete.

Although expanded clay aggregate has been widely accepted by the research community as one of the materials for producing lightweight concrete, its effectiveness is significantly affected by various factors, with the manufacturing method being the most important. In addition, it is suspected that the shape and size of expanded clay aggregate influence the engineering properties of lightweight concrete, and admixtures are frequently used to improve the properties of concrete by enhancing workability and accelerating performance. The effect of admixtures on the properties of lightweight concrete made with expanded clay aggregate has not been thoroughly investigated. Due to this lack of information, it is necessary to conduct additional research on the effect of the aggregate shape and the effect of admixtures on the engineering properties of lightweight concrete made from expanded clay aggregate.

This research consists of three series based on several mix proportions. The first series focuses on the effect of the three different clay aggregate shapes: cubic, oval, and round. It should be noted that the manufacturing method is identical for all shapes, with grain distributions ranging from 5 to 20 mm. In addition, the source of the clay employed in this research is Yogyakarta, Indonesia. It should be emphasized that the location from which the clay is collected, as well as its type, significantly affects the aggregate production process and the properties of concrete. Thus, the novelty of this study is the use of three shape variations of expanded clay aggregate from the Yogyakarta area.

The second series evaluates the effects of superplasticizers on the properties of lightweight concrete. The studied variations were 0, 1, 1,5, 2, and 2,5%. Compressive and split tensile strength testing constituted the majority of the tests conducted in this research. In addition, a slump test was carried out to determine workability and mass density. This, in order to manage the index level of the lightweight concrete produced.

Series III focuses on the index corresponding to the development of split tensile and compressive strength in relation to the age of concrete. Here, a compressive strength conversion factor is generated for a specific age. The conversion factor of concrete properties to age has been commonly utilized for normal concrete, but it has not been studied for lightweight concrete made from expanded clay aggregate. Therefore, this study aims to establish a correlation between the age and the properties of lightweight concrete, such as compressive and split tensile strength, as there is no previous research in this regard. Given the relationship between the age of concrete and its properties, it is expected that it will no longer be necessary to conduct compressive strength tests with various ages in the future. Thus, it is possible to reduce the number of specimens in order to evaluate, based on age, the compressive and split tensile strength of lightweight concrete including expanded clay aggregate.

Experimental program

Materials

The materials used in this study were cement, water, fine aggregate, coarse aggregate, and superplasticizer. The cement used in this study was in line with the ASTM C150 standard, *i.e.*, type 1 cement (ordinary Portland cement) (ASTM International, 2015). The fine aggregate used in this study was river sand from Kali Progo, Yogyakarta, Indonesia. Its physical and mechanical properties were examined before use as a constituent material of concrete. The evaluated mechanical parameters were water content, absorption, specific gravity, and mud content. The results are summarized in Table 1. In addition, the fine aggregate's grain particle size distribution was measured (Figure 1). These measurements were taken based on ASTM C136 (ASTM International, 2019a).

Table 1. Properties of the fine aggregate

Properties	Value
Specific gravity	2,45
Water absorption (%)	2,57
Water content (%)	2,22
Mud content (%)	2,00
Mass density (g/cm^3)	1,52
Fineness modulus	2,38

Source: Authors

Lightweight expanded clay aggregate (LECA) was the coarse aggregate employed in this study. The source of this clay is Sleman, Yogyakarta, Indonesia. As mentioned before, our research comprised three series. The first was focused on coarse aggregate shapes, *i.e.*, cubic, oval, and round. Meanwhile, series II and III only employed round coarse aggregate. Figure 2 shows the shape of each aggregate. The production process of expanded clay aggregate began by mixing clay with sufficient water. The same amount of water was used for each aggregate shape. The dough was then molded into the desired size, with the aggregate size distribution ranging between 5 and 20 mm. The aggregate was then burned at temperatures between 800 and 1 200 °C. Subsequently, the properties of the coarse aggregate were evaluated, *i.e.*, specific gravity, water content, water absorption, density, mud content, and roughness level. The results regarding these properties are displayed in Table 2 for all shapes.

In addition, the effect of admixtures was explored in this study. The superplasticizer used was either Viscocrete 3115N from SIKA or type D superplasticizer for water reduction and retarding, in accordance with ASTM C494 (ASTM International, 2019b).

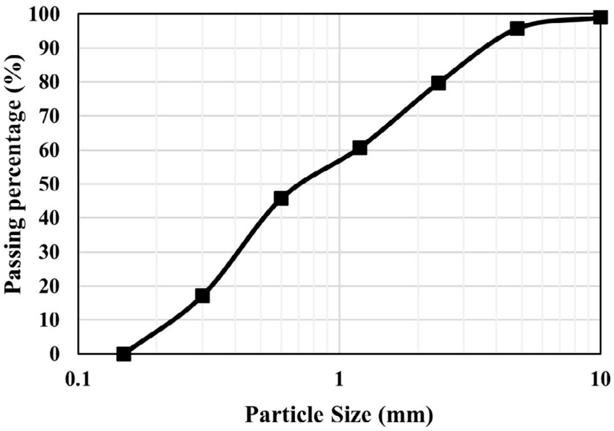


Figure 1. Size distribution of the fine aggregate
Source: Authors

Figure 2. Shape of the coarse aggregates
Source: Authors

Table 2. Properties of the coarse aggregate

Properties	Value
Specific gravity	1,98
Water absorption (%)	18,91
Water content (%)	1,67
Mud content (%)	2,00
Mass density (g/cm^3)	0,98
Roughness (%)	94,00

Source: Authors

Mix proportions

This research aimed to evaluate the effect of aggregate shape, admixtures, and age on concrete. Table 3 shows the mix proportions for series I. Table 4 presents the mix proportions for series II, which corresponds to the amount of superplasticizer used. The superplasticizer percentage was derived from the total binder used to cast the specimen.

Table 3. Mix proportions for series I

ID	SI.A	SI.B	SI.C
Aggregate shape	Cubic	Oval	Round
Cement (kg/m^3)	342,41	342,41	342,41
Fine aggregate (kg/m^3)	679,02	679,02	679,02
Coarse aggregate (kg/m^3)	918,67	918,67	918,67
Water (l)	184,90	184,90	184,90

Source: Authors

Table 4. Mix proportions for series II

ID	SII.A	SII.B	SII.C	SII.D	SII.E
Superplasticizer (%)	0	1	1,5	2	2,5
Cement (kg/m^3)	342,41	342,41	342,41	342,41	342,41
Fine aggregate (kg/m^3)	679,02	679,02	679,02	679,02	679,02
Coarse aggregate (kg/m^3)	918,67	918,67	918,67	918,67	918,67
Water (l)	184,90	147,92	147,92	147,92	147,92

Source: Authors

Table 5 shows the mix proportions used in examining the development of concrete's mechanical properties with age. In series II and III, expanded round clay aggregate was employed. It should be noted that all test objects were cured with water before testing.

Table 5. Mix proportion for Series III

ID	SIII.A
Superplasticizer (%)	0
Cement (kg/m^3)	342,41
Fine aggregate (kg/m^3)	679,02
Coarse aggregate (kg/m^3)	918,67
Water (l)	184,90
Aggregate shape	Round

Source: Authors

Test method

This study included four tests: a fresh properties (workability) test, a density test, a compressive strength test, and a splitting tensile strength test. The slump test was performed in series I and II, but not in series III, since it utilized the same proportion of materials as the ID SII.A specimens. The slump test followed ASTM C143 (ASTM International, 2020), the standard test procedure for slump testing in hydraulic cement concrete. Moreover, a density test was conducted on 28-day-old concrete that had already solidified and aimed to determine whether the concrete mass met the requirements for lightweight concrete. This test followed ASTM C642 (ASTM International, 2021) and was only performed on series I and II specimens.

In this study, compressive and splitting tensile strength tests were conducted for all series. They were carried out with 28-day-old concrete in series I and II, whereas, in series III, age variations of 3, 7, 14, 21, and 28 days were examined. The dimensions of the cylindrical specimen for compressive strength testing were 10 x 20 cm, while that for split tensile strength had dimensions of 15 x 30 cm. The average of the three specimens is obtained from each variation in compressive and splitting tensile strength. There are 18 specimens in series I, 30 in series II, and 30 in series III.

Thus, a total of 78 specimens were used in this study, out of which 39 were used to measure compressive strength and the other half to measure splitting tensile strength. This study followed the ASTM C39 standard for compressive strength testing (ASTM International, 2021), with a constant loading rate of 0,15 MPa/sec, as well as ASTM C496 for splitting tensile strength testing (ASTM International, 2017b), with a constant loading rate of 0,7 MPa/sec. To preserve the hydration process, all cast specimens underwent water curing for 28 days. Additionally, all specimens were placed in a controlled room with a fixed temperature and relative humidity during the curing process.

Results and discussion

Influence of the coarse aggregate shape

Various studies regarding the effect of expanded clay aggregate size have been carried out by several researchers (Ozguven and Gunduz, 2012; Rashad, 2018). However, there is limited information regarding the effect of LECA shapes on the engineering properties of concrete. As previously mentioned, this study examines this effect on the properties of lightweight concrete. These experiments were conducted since expanded clay aggregate is typically produced with a uniform shape but with different sizes.

Four types of tests were carried out in this regard: a slump test, density test, a compressive strength test, and a splitting tensile strength test. Figure 3 shows the fresh concrete slump test results for series I, indicating that the aggregate shape does not affect the slump value. This is due to the fact that the three mix proportions used are identical, except for the aggregate shape. The specimens in series I (with a round coarse aggregate, SI.C) yield lower slump values compared to other shapes (cubic and oval), but the difference is insignificant. Therefore, it can be concluded that the shape of the expanded clay aggregate has no significant effect on the slump value of fresh concrete.

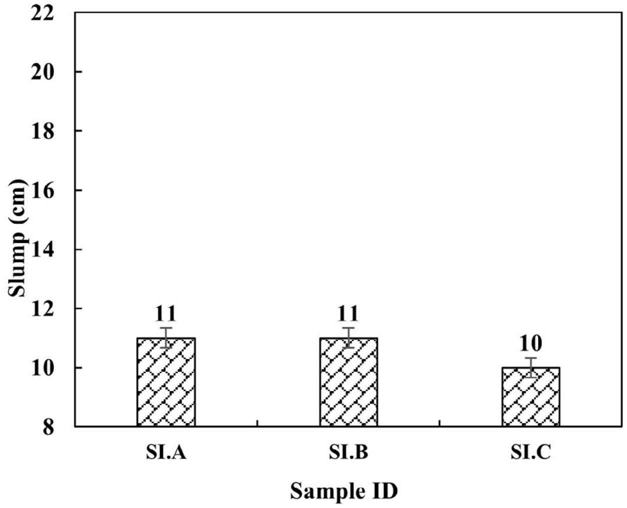


Figure 3. Slump results for series I
Source: Authors

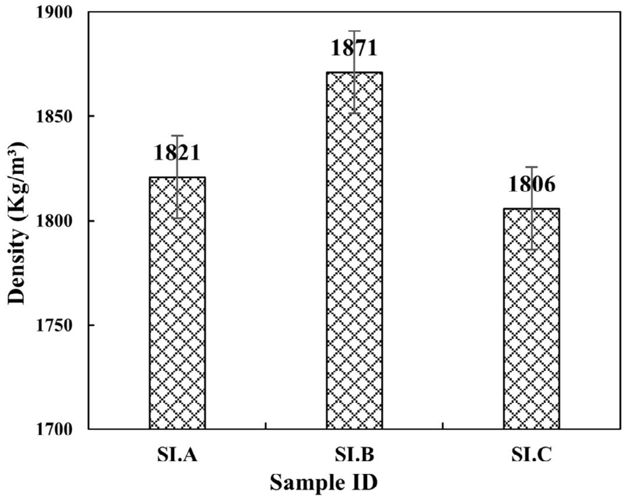


Figure 4. Density results for series I
Source: Authors

Density is a key aspect in determining the classification of concrete. Lightweight concrete corresponds to density values ranging from 320 to 1 920 kg/m^3. This, according to ACI 213 and ASTM C330 (ACI Committee, 2014; ASTM International, 2017a). The results of the laboratory tests indicate that all specimens in series I correspond to lightweight concrete (Figure 4). The SI.B specimens (utilizing oval expanded clay aggregates) reported the highest density compared to other specimens. Nonetheless, the density of SI.B does not exceed 1 900 kg/m^3. In contrast, SI.C specimens (with round expanded clay aggregate) showed the lowest density: only 1 806 kg/m^3. The difference in hardened density between these samples may be due to the influence of the expanded clay aggregate's size and shape. However, these differences are not significant, and all specimens can be classified as lightweight concrete.

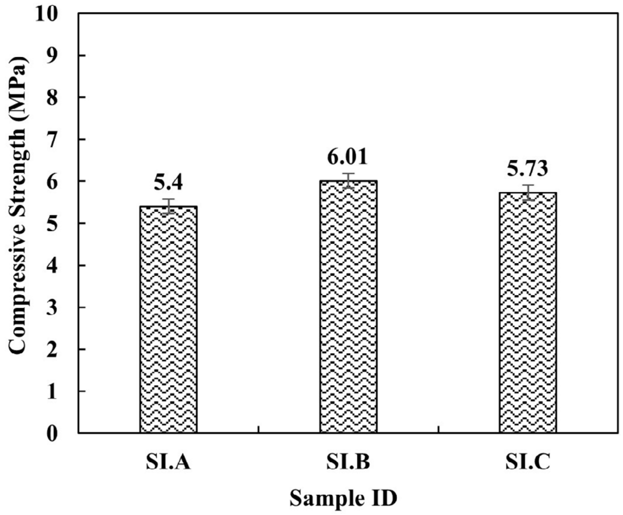


Figure 5. Compressive strength for series I
Source: Authors

Figure 5 shows the compressive strength of concrete after 28 days, with a comparison regarding the coarse aggregate shape. These values range between 5 and 6 MPa. Series I used the same mix proportions for all specimens, but a different aggregate shape. No additional superplasticizer was used in this series. This demonstrates that the aggregate shape has no significant effect on the compressive strength of 28-day-old concrete. Th oval expanded clay aggregate provided the specimens with the highest compressive strength, while the round coarse aggregate reported the specimens with the lowest values. This is because round coarse aggregates have a weaker compressive stress resistance compared to those with cubic and oval shapes, which are more stable in withstanding compressive forces.

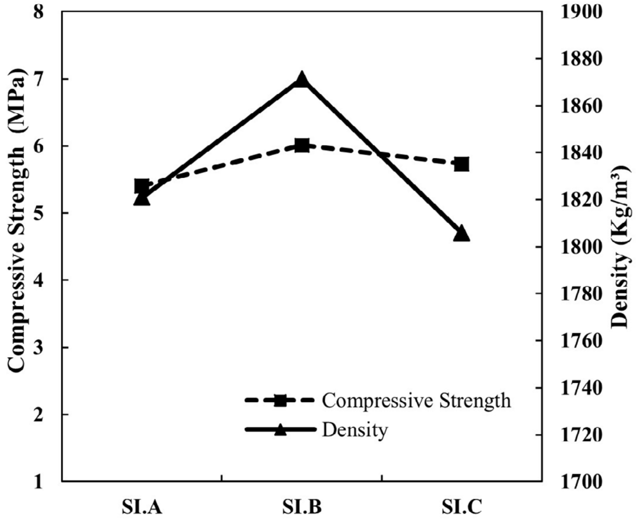


Figure 6. Comparison between compressive strength and density for series I
Source: Authors

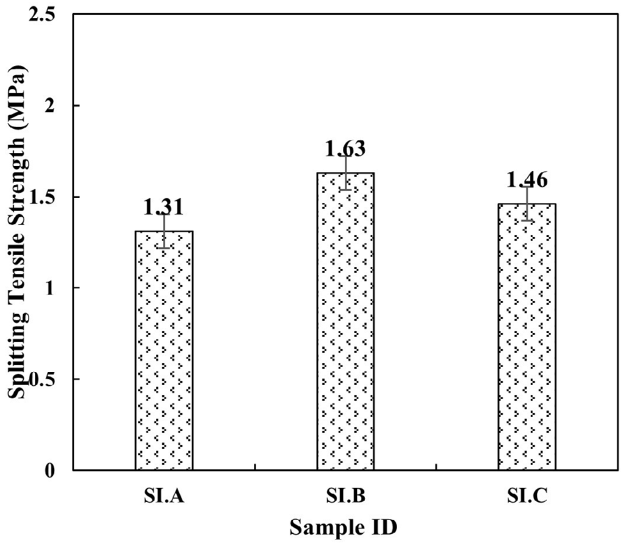


Figure 7. Splitting tensile strength for series I
Source: Authors

This study also evaluated the relationship between compressive strength and concrete density. Figure 6 shows a comparison between compressive strength and density with different coarse aggregate shapes. At the age of 28 days, a high density provides a higher compressive strength, specifically in specimens with oval coarse aggregates. This is due to the influence of concrete density. Consequently, the greater the density, the smaller the pore volume, and *vice versa*. Thus, higher density concrete has higher compressive strength, even with the same mix proportions. Figure 7 provides information about another test regarding hardened qualities, which involved the splitting tensile strength after 28 days. The results show a linear trend comparable to that of the compressive strength, with the highest values obtained by concrete with oval particles. This indicates that the oval aggregate is more resistant to load deformation than the other shapes evaluated.

Influence of chemical admixture

This research also studied the effect of admixture addition on lightweight concrete's fresh and hardened properties. The corresponding experiments used a homogeneous spherical aggregate with dimensions ranging from 5 to 20 mm. After 28 days, concrete was subjected to testing with regard to its hardened properties. Figure 8 shows results of the slump test for series II concrete, which was conducted to examine the influence of chemical admixture. These results indicate a significantly increased slump value due to the additional superplasticizer. It should be noted that, in this variation, the water content was reduced by 20% when compared to normal concrete without a superplasticizer. However, this test demonstrates that superplasticizers can improve the performance of fresh concrete by increasing workability. Previous studies have shown similar results, even with different types of superplasticizers. According to them, increasing the amount of superplasticizer increases the fluidity of fresh concrete (Nahhab and Ketab, 2020; Nepomuceno *et al.*, 2018; Tang *et al.*, 2020).

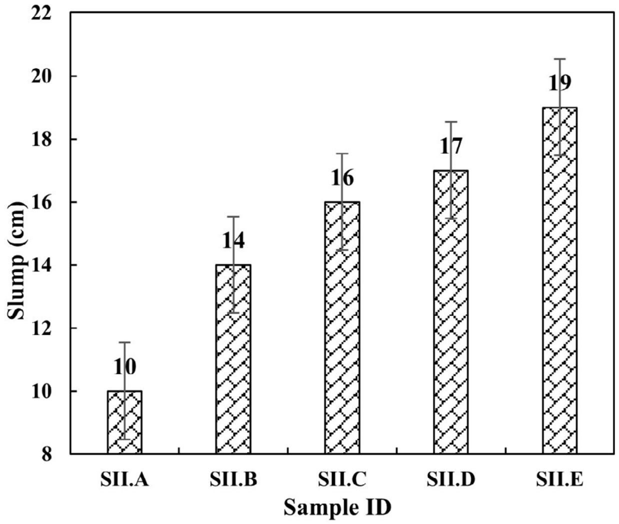


Figure 8. Slump results for series II
Source: Authors

The results of the density examination carried out for each specimen in series II are shown in Figure 9. Here, the density of hardened concrete with chemical admixture ranges from 1 750 to 1 850 kg/m^3. This shows no significant differences in the density of each specimen when compared to normal concrete without admixture (SII.A). However, with respect to concrete made with oval and cubic coarse aggregate, the density of concrete made with this superplasticizer is significantly lower. In addition, all specimens containing this superplasticizer were classified under the lightweight concrete category. The relationship between the compressive strength of concrete and the amount of admixture after 28 days can be seen in Figure 10.

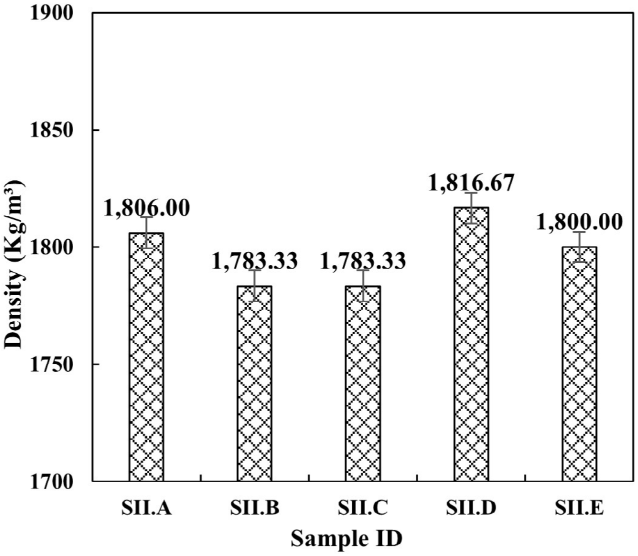


Figure 9. Density results for series II
Source: Authors

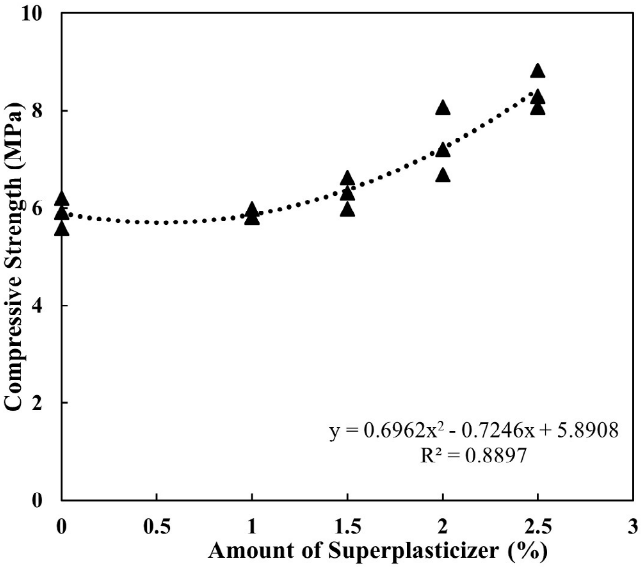


Figure 10. Compressive strength after 28 days for series II
Source: Authors

Curve fitting with polynomial regression is an efficient and simple method to study the relationship between the engineering properties of concrete, its age, and the amount of superplasticizer used. This approach has also been used in previous research to evaluate the effects of superplasticizers on the compressive and split tensile strength of concrete (Gu *et al.*, 2021; Faraj *et al.*, 2022a, 2022b, 2022c), as well as the effects of age on these properties (Varga *et al.*, 2012; Elaty, 2014). Our analysis determined a polynomial relationship with a correlation coefficient of 0,8897 (88,97%). This value indicates a significant relationship between the compressive strength of concrete and the amount of superplasticizer used. The test findings indicate that compressive strength increased along with the amount of superplasticizer. One of the factors contributing to this trend is the decreased proportion of water in the mixture (due to the use of a superplasticizer). Note that reducing the water content of concrete decreases the water-to-binder ratio, and cement requires a minimal amount of water to react. Due to evaporation during the hydration process, water is retained in the concrete when a sufficient volume is used. This allows it to expand the pores upon evaporation, resulting in a decreased compressive strength. However, water is required for the processing of fresh concrete. With the aid of this additive ingredient, the amount of water can be reduced significantly, and fresh concrete can become more plastic.

Figure 11 illustrates the results obtained regarding the splitting tensile strength of 28-day-old concrete with various concentrations of superplasticizer. The correlation coefficient between the splitting tensile strength and the amount of admixture is relatively high, reaching 95,10%. The trend identified in our study indicates that tensile strength increases along with the amount of superplasticizer. This pattern closely resembles the compressive strength testing results. It is important to recall that the addition of superplasticizers can help to reduce the water content: even though the density of concrete decreases, its compressive and splitting tensile strength increase, as less water is trapped in it and the volume of the pores is reduced. Increasing the amount of superplasticizer also increases the workability of the concrete, resulting in a more uniform aggregate distribution. This is extremely helpful for the production of lightweight, high-performance concrete. It should be noted that the resulting compressive strength is still lower than that of conventional concrete made with coarse aggregate from natural resources such as rock. However, this lightweight aggregate can also be utilized in various non-structural concrete constructions.

Influence of the age of concrete

According to our results, compressive strength increases with age (Figure 12). This resembles the patterns shown by concrete with typical aggregate.

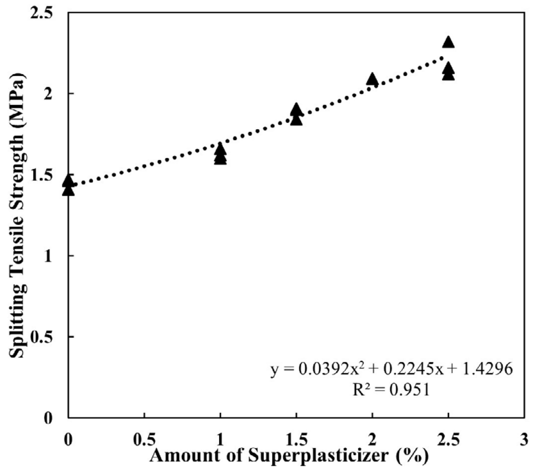


Figure 11. Splitting tensile strength after 28 days for series II
Source: Authors

Table 6. Strength increase rate

Concrete age (Days)	3 to 7	7 to 14	14 to 21	21 to 28
Compressive strength (%)	59,24	23,14	15,99	10,51
Splitting tensile strength (%)	71,42	21,12	17,44	5,49

Source: Authors

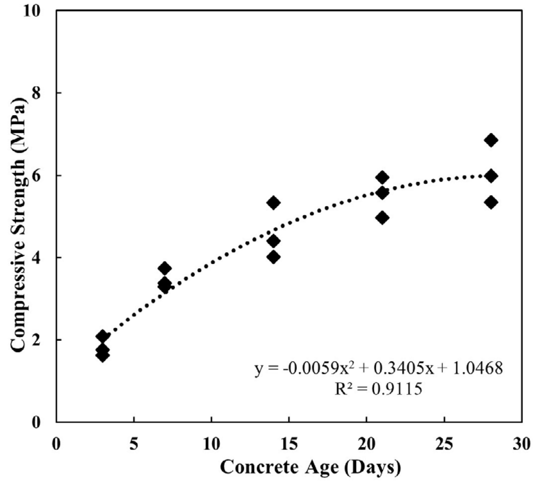


Figure 12. Relationship between age and compressive strength (series III)
Source: Authors

The relationship between age and splitting tensile strength is presented in Figure 13, and the evolution of compressive and splitting tensile strength is shown in Table 6. These findings indicate that the splitting tensile strength of concrete increases with age. Furthermore, the strength increase rate decreases with age, which is caused by the nearly complete hydration of the concrete, whose strength approaches ideal conditions. The conversion factor for the compressive and splitting tensile strength at a given age is presented in Table 7. With this factor, there will be no need to test for age differences, thus reducing the number of required samples during experimentation. This conversion factor applies to three shapes of expanded clay aggregate: cubic, oval, and round. There is a need for additional study on a more effective formula for producing improved hardened properties, such as compressive and splitting tensile strength. In addition, future studies aiming to use this aggregate for structural components must also conduct another types of experiments, such as flexural strength testing.

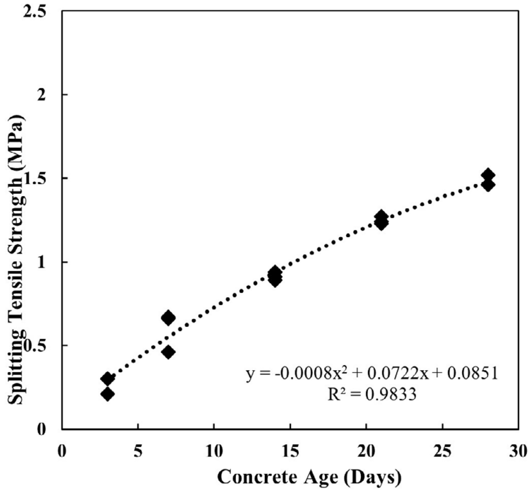


Figure 13. Relationship between age and splitting tensile strength (series III)
Source: Authors

Table 7. Conversion factor regarding compressive and splitting tensile strength

Concrete age (Days)	3	7	14	21	28
Compressive strength	0,30	0,57	0,76	0,91	1,00
Splitting tensile strength	0,18	0,40	0,62	0,84	1,00

Source: Authors

Several specimens are shown in Figure 14 after testing. It is evident that the expanded clay aggregate component of the specimen can fail, alongside the surface mortar area. As a result, a compressive strength that is often lower than that of normal concrete is reported.

Figure 14. Specimens after testing
Source: Authors

Conclusions

Based on the results and discussion, the following conclusions can be drawn.

1. The shape of expanded clay aggregate affects the qualities of hardened concrete. However, the obtained difference is insignificant. An oval aggregate improves the compressive and splitting tensile strength.
2. The incorporation of superplasticizers can lower the amount of water required for the casting of concrete. This reduces the amount of water consumed while improving the workability of fresh properties. In addition, the decrease in water content increases the compressive and flexural strength of lightweight concrete and decreases its density.
3. This study presents a conversion factor for predicting the compressive and splitting tensile strength of lightweight concrete made with expanded clay aggregate.
4. Using expanded clay aggregate as the coarse aggregate in concrete can result in the production of lightweight concrete.

Acknowledgments

The authors would like to express appreciation to the Department of Civil Engineering, Faculty of Engineering, Universitas Muhmmadiyah Yogyakarta, for providing materials and a laboratory for this experiment. This study was supported by skema hibah kolaboratif luar negeri oleh Lembaga Riset dan Inovasi (LRI) Universitas Muhammadiyah Yogyakarta Tahun Anggaran 2022 dengan nomor hibah 01/RIS-LRUU2022.

CRediT author statement

As'at Pujianto: conceptualization, data curation, funding acquisition, methodology, supervision, writing (review and editing). *Hakas Prayuda*: data curation, project administration, writing (review and editing). *Farrel Asani*: data curation, investigation, writing (original draft). *Muji Basuki Santoso*: data curation, investigation, writing (original draft). *Fahriza Wirawan*: data curation, investigation, writing (original draft).

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Not applicable.

References

ACI Committee (2014). *ACI PRC-213-14: Guide for structural lightweight aggregate concrete*. American Concrete Institute.

Adhikary, S. K., Rudzionis, Z., Tuckute, S., and Ashish, D. K. (2021). Effects of carbon nanotubes on expanded glass and silica aerogel based lightweight concrete. *Scientific Report*, *11*, 1-11. https://doi.org/10.1038/s41598-021-81665-y

Adhikary, S. K., Rudzionis, Z., and Vaiciukyniene, D. (2020). Development of flowable ultra-lightweight concrete using expanded glass aggregate, silica aerogel, and prefabricated plastic bubbles. *Journal of Building Engineering*, *31*, 1-10. https://doi.org/10.1016/j.jobe.2020.101399

Ahmad, M. R., and Chen, B. (2019). Experimental research on the performance of lightweight concrete containing foam and expanded clay aggregate. *Composites Part B: Engineering*, *171*, 44-60. https://doi.org/10.1016/j.compositesb.2019.04.025

Ahmad, M. R., Chen, B., and Shah, S. F. A. (2019). Investigate the influence of expanded clay aggregate and silica fume on the properties of lightweight concrete. *Construction and Building Materials*, *220*, 253-266. https://doi.org/10.1016/j.conbuildmat.2019.05.171

Akcaozoglu, S., Atis, C. D., and Akcaozoglu, K. (2010). In investigation on the use of shredded waste PET bottles as aggregate in lightweight concrete. *Waste Management*, *30*(2), 285-290. https://doi.org/10.1016/j.wasman.2009.09.033

Alqahtani, F. K., Ghataora, G., Khan, M. I., and Dirar, S. (2017). Novel lightweight concrete containing manufactured plastic aggregate. *Construction and Building Materials*, 148, 386-397. https://doi.org/10.1016/j.conbuildmat.2017.05.011

Arunchalam, KP., Avudaiappan, A., Flores, E. I. S., and Parra, P. F. (2023). Experimental study on the mechanical properties and microstructures of cenosphere concrete. *Materials*, *16*(9), 1-19. https://doi.org/10.3390/ma16093518

Aslam, M., Shafigh, P., and Jumaat, M. Z. (2016). Oil palm by products as lightweight aggregate in concrete mixture: a review. *Journal of Cleaner Production*, *126*, 56-73. https://doi.org/10.1016/j.jclepro.2016.03.100

ASTM International (2015). *ASTM C150/C150M-16: Standard specification for Portland cement*. ASTM International. https://doi.org/10.1520/C0150

ASTM International (2017a). *ASTM C330/C330M-17a: Standard specification for lightweight aggregates for structural concrete*. ASTM International. https://doi.org/10.1520/C0330_C0330M-17A

ASTM International (2017b). *ASTM C496/C496M-17: Standard test method for splitting tensile strength of cylindrical concrete specimens*. ASTM International. https://doi.org/10.1520/C0496_C0496M-17

ASTM International (2019a). *ASTM C136/C136M-19: Standard test method for sieve analysis of fine and coarse aggregates*. ASTM International. https://doi.org/10.1520/C0136_C0136M-19

ASTM International (2019b). *ASTM C494/C494M-19e1: Standard specification for chemical admixtures for concrete*. ASTM International. https://doi.org/10.1520/C0494_C0494M-19E01

ASTM International (2020). *ASTM C143/C143M-20: Standard test method for slump of hydraulic cement concrete*. ASTM International. https://doi.org/10.1520/C0143_C0143M-20

ASTM International (2021). *ASTM C39/C39M-21: Standard test method for compressive strength of cylindrical concrete specimens*. ASTM International. https://doi.org/10.1520/C0039_C0039M-21

ASTM International (2021). *ASTM C642-21: Standard test method for density absorption and voids in hardened concrete*. ASTM International. https://doi.org/10.1520/C0642-21

Augonis, A., Ivanauskas, E., Bocullo, V., Kantautas, A., and Vaiciukyniene. (2022). The influence of expanded glass and expanded clay on lightweight aggregate shotcrete properties. *Materials*, *15*(5), 1-13. https://doi.org/10.3390/ma15051674

Aungatichart, O., Nawaukkaratharnant, N., and Wasanapiarnpong, T. (2022). The potential use of cold-bonded lightweight aggregate derived from various types of biomass fly ash for preparation of lightweight concrete. *Materials Letters*, *327*, 1-12. https://doi.org/10.1016/j.matlet.2022.133019

Bogas, J. A., Brito, J. D., and Cabaco, J. (2014). Long term behaviour of concrete produced with recycled lightweight expanded clay aggregate concrete. *Construction and Building Materials*, *65*, 470-479. http://dx.doi.org/10.1016/j.conbuildmat.2014.05.003

Bogas J. A., Brito, J. D., and Figueiredo, J. M. (2015). Mechanical characterization of concrete produced with recycled lightweight expanded clay aggregate concrete. *Journal of Cleaner Production*, *89*, 187-195. http://dx.doi.org/10.1016/j.jclepro.2014.11.015

Brooks, A. L., Fang, Y., Shen, Z., Wang, J., and Shou, H. (2021). Enabling high-strength cement-based materials for thermal energy storage via fly-ash cenosphere encapsulated phase change materials. *Cement and Concrete Composites*, *120*, 1-15. https://doi.org/10.1016/j.cemconcomp.2021.104033

Campione, G., Miraglia, N., and Papia, M. (2001). Mechanical properties of steel fibre reinforced lightweight concrete with pumice stone or expanded clay aggregates. *Materials and Structures*, *34*, 210-210. https://doi.org/10.1007/BF02480589

Castillo, E. D. R., Almesfer, N., Saggi, O., and Ingham, J. M. (2020). Lightweight concrete with artificial aggregate manufactured from plastic waste. *Construction and Building Materials*, *265*, 1-10. https://doi.org/10.1016/j.conbuildmat.2020.120199

Chen, W., and Huang, Z. (2019). Experimental study of the mechanical properties and microstructures of lightweight toughness cement-based composites. *Materials*, *12*(23), 1-19. https://doi.org/10.3390/ma12233891

Dabbaghi, F., Dehestani, M., and Yousefpour, H. (2021). Residual mechanical properties of concrete containing lightweight expanded clay aggregate (LECA) after exposure to elevated temperatures. *Structural Concrete*, *23*(4), 2162-2184. https://doi.org/10.1002/suco.202000821

Du, H., Du, S., and Liu, X. (2015). Effect of nano-silica on the mechanical and transport properties of lightweight concrete. *Construction and Building Materials*, *82*, 114-122. https://doi.org/10.1016/j.conbuildmat.2015.02.026

Elaty, M. A. A. A. (2014). Compressive strength prediction of Portland cement concrete with age using a new model. *Housing and Building National Research Center Journal*, *10*(2), 145-155. https://doi.org/10.1016/j.hbrcj.2013.09.005

Elrahman, M. A., Chung, S. Y., Sikora, P., Rucinska, T., and Stephan, D. (2019). Influence of nanosilica on mechanical properties, sorptivity, and microstructure of lightweight concrete. *Materials*, *12*(19), 1-16. https://doi.org/10.3390/ma12193078

Faraj, R. H., Mohammed, A. A., Omer, K. H., and Ahmed, H. U. (2022a). Soft computing techniques to predict the compressive strength of green self-compacting concrete incorporating recycled plastic aggregates and industrial waste ashes. *Clean Technologies and Environmental Policy*, *24*, 2253-2281. https://doi.org/10.1007/s10098-022-02318-w

Faraj, R. H., Mohammed, A. A., and Omer, K. M. (2022b). Modeling the compressive strength of eco-friendly self-compacting concrete incorporating ground granulated blast furnace slag using soft computing techniques. *Environmental Science and Pollution Research*, *29*, 71338-71357. https://doi.org/10.1007/s11356-022-20889-5

Faraj, R. H., Mohammed, A. A., Mohammed, A., Omer, K. M., and Ahmed, H. U. (2022c). Systematic multiscale models to predict the compressive strength of self-compacting concretes modified with nanosilica at different curing ages. *Engineering with Computers*, *38*, 2365-2388. https://doi.org/10.1007/s00366-021-01385-9

Gayathiri, K., and Praveenkumar, S. (2022). Influence of nano silica on fresh and hardened properties of cement-based materials – A review. *Silicon*, *14*, 8327-8357. https://doi.org/10.1007/s12633-021-01598-z

Gu, C., Yao, J., Yang, Y., Huang, J., Ma, L., Ni, T., and Liu, J. (2021). The relationship of compressive strength and chemically bound water content of high-volume fly ash-cement mortar. *Materials*, *14*(21), 1-16. https://doi.org/10.3390/ma14216273

Hanif, A., Lu, Z., and Li, Z. (2017). Utilization of fly ash cenosphere as lightweight filler in cement-based composites – A review. *Construction and Building Materials*, *144*, 373-384. https://doi.org/10.1016/j.conbuildmat.2017.03.188

Hassan, A. A. A., Ismail, M. K., and Mayo, J. (2015). Mechanical properties of self-consolidating concrete containing lightweight recycled aggregate in different mixture compositions. *Journal of Building Engineering*, *4*, 113-126. https://doi.org/10.1016/j.jobe.2015.09.005

Hossain, K. M. A. (2004). Properties of volcanic pumice based cement and lightweight concrete. *Cement and Concrete Research*, *34*(2), 283-291. https://doi.org/10.1016/j.cemconres.2003.08.004

Hubertova, M., and Hela, R. (2013). Durability of lightweight expanded clay aggregate concrete. *Procedia Engineering*, *65*, 2-6, https://doi.org/10.1016/j.proeng.2013.09.002

Karthika, R. B., Vidyapriya, V., Sri, K. V. N., Beaula, K. M. G., Harini, R., and Sriram, M. (2021). Experimental study on lightweight concrete using pumice aggregate, *Materials Today: Proceedings* *43*(2), 1060-1613. https://doi.org/10.1016/j.matpr.2020.09.762

Kavinkumar, V., Priya, A. K., and Praneeth, R. (2023). Strength of light weight concrete containing fly ash cenosphere. *Materials Today: Proceedings*, *17*, 1-6. https://doi.org/10.1016/j.matpr.2023.04.094

Kulkarni, P., and Muthadhi, A. (2020). Improving thermal and mechanical property of lightweight concrete using N-Butyl stearate/expanded clay aggregate with alccofine1203. *International Journal of Engineering Transactions A: Basics*, *33*(10), 1842-1851. https://doi.org/10.5829/ije.2020.33.10a.03

Kurt, M., Gul, M. S., Gul, R., Aydin, A. C., and Kotan, T. (2016). The effect of pumice powder on the self-compactability of pumice aggregate lightweight concrete. *Construction and Building Materials*, *103*, 36-46. https://doi.org/10.1016/j.conbuildmat.2015.11.043

Liu, B., Shi, J., Jiang. J., Wu, X., and He, Z. (2021). New perspectives on utilization of CO_2 sequestration technologies in cement-based materials. *Construction and Building Materials*, *272*, 1-17. https://doi.org/10.1016/j.conbuildmat.2020.121660

Mannan, m. A., Alexander, J., Ganapathy, C., and Teo, D. C. L. (2006). Quality improvement of oil palm shell (OPS) as coarse aggregate in lightweight concrete. *Building and Environment*, *41*(9), 1239-1242. https://doi.org/10.1016/j.buildenv.2005.05.018

Meyer, C. (2009). The greening of the concrete industry. *Cement and Concrete Composites*, *31*, 601-605. https://doi.org/10.1016/j.cemconcomp.2008.12.010

Mohammed, T., Aguayo, F., Nodehi, M., and Ozbakkaloglu, T. (2022). Engineering properties of structural lightweight concrete containing expanded shale and clay with high volume class F and class C fly ash. *Structural Concrete*, *24*(3), 4029-4046. https://doi.org/10.1002/suco.202200562

Monika, F., Prayuda, H., Cahyati, M. D., Augustin, e. N., Rahman H. A., and Prasintasari, A. D. (2022). Engineering properties of concrete made with coal bottom ash as sustainable construction materials. *Civil Engineering Journal*, *8*(1), 181-194. http://dx.doi.org/10.28991/CEJ-2022-08-01-014

Nahhab, A. H., and Ketab, A. K. (2020). Influence of content and maximum size of light expanded clay aggregate on the fresh, strength and durability properties of self-compacting lightweight concrete reinforced with micro steel fibres. *Construction and Building Materials*, *233*, 1-14. https://doi.org/10.1016/j.conbuildmat.2019.117922

Nawel, S., Mounir, L., and Hedi, H. (2017). Characterisation of lightweight concrete of Tunisian expanded clay: Mechanical and durability study. *European Journal of Environmental and Civil Engineering*, *21*(6), 670-695. https://doi.org/10.1080/19648189.2016.1150893

Nepomuceno, M. C. S., Oliveira, L. A. P., and Pereira, S. F. (2018). Mix design of structural lightweight self compacting concrete incorporating coarse lightweight expanded clay aggregates. *Construction and Building Materials*, *116*, 373-385. https://doi.org/10.1016/j.conbuildmat.2018.01.161

Ozguven, A., and Gunduz, L. (2012). Examination of effective parameters for the production of expanded clay aggregate. *Cement and concrete Composites*, *34*, 781-787. http://dx.doi.org/10.1016/j.cemconcomp.2012.02.007

Pelisser, F., Barcodes, A., Santos, D., Peterson, M., and Bernardin, A. M. (2012). Lightweight concrete production with low Portland cement consumption. *Journal of Cleaner Production*, *23*(1), 68-74. https://doi.org/10.1016/j.jclepro.2011.10.010

Posi, P., Teerachanwit, C., Tanutong, C., Limkamoltip, S., Lertnimoolchai, S., Sata, V., and Chindaprasirt, P. (2013). Lightweight geopolymer concrete containing aggregate from recycle lightweight block. *Materials and Design*, *52*, 580-586. https://doi.org/10.1016/j.matdes.2013.06.001

Pujianto, A., and Prayuda, H. (2012). Fresh and hardened properties of lightweight concrete made with pumice as coarse aggregate. *International Journal of Geomate*, *21*(87), 110-117, https://doi.org/10.21660/2021.87.j2365

Rashad, A. M. (2018). Lightweight expanded clay aggregate as a building material – An overview. *Construction and Building Materials*, *170*, 757-775. https://doi.org/10.1016/j.conbuildmat.2018.03.009

Rashad, A. M. (2019). A short manual on natural pumice as a lightweight aggregate. *Journal of Building Engineering*, *25*, 1-10. https://doi.org/10.1016/j.jobe.2019.100802

Rumsys, D., Bacinskas D., Spudulis, E., and Meskenas, A. (2017). Comparison of material properties of lightweight concrete with recycled polyethylene and expanded clay aggregates. *Procedia Engineering*, *172*, 937-944. https://doi.org/10.1016/j.proeng.2017.02.105

Rumsys, D., Spudulis, E., Bacinkas, D., and Kaklauskas, G. (2018). Compressive strength and durability properties of structural lightweight concrete with fine expanded glass and/or clay aggregates. *Materials*, *11*(12), 1-14. https://doi.org/10.3390/ma11122434

Saha, A. K., Majhi, S., Sarker, P. K., Mukherjee, A., Siddika, A., Aslani, F., and Zhuge. Y. (2021). Non-destructive prediction of strength of concrete made by lightweight recycled aggregates and nickel slag. *Journal of Building Engineering*, *33*, 1-10, https://doi.org/10.1016/j.jobe.2020.101614

Saleh, F., Gunawan, M. A., Yolanda, T. I., Monika F., Prayuda, H., Cahyati, M. D., and Pratama, M. M. A. (2022). Properties of mortar made with bottom ash and silica fume as sustainable construction materials. *World Journal of Engineering*, *20*(5), 835-845. http://dx.doi.org/10.1108/WJE-08-2021-0481

Shafigh, P., Jumaat, M. Z., and Mahmud, H. (2011). Oil palm shell as a lightweight aggregate for production high strength lightweight concrete. *Construction and Building Materials*, *25*(4), 1848-1853. https://doi.org/10.1016/j.conbuildmat.2010.11.075

Shafigh, P., Mahmud, H. B., Jumaat, M. Z. B., Ahmmad, R., and Bahri, S. (2014). Structural lightweight aggregate concrete using two types of waste from the palm oil industry as aggregate. *Journal of Cleaner Production*, *81*, 187-196. https://doi.org/10.1016/j.jclepro.2014.05.051

Shi, J., Liu, Y., Wang, E., Wang, L., Li, C., Xu, H., Zheng, X., and Yuan, Q. (2022). Physico-mechanical, thermal properties and durability of foamed geopolymer concrete containing cenospheres. *Construction and Building Materials*, *325*, 1-13. https://doi.org/10.1016/j.conbuildmat.2022.126841

Sikora, P., Rucinska, T., Stephan, D., Chung, S. Y., and Elrahman, M. A. (2020). Evaluating the effects of nanosilica on the material properties of lightweight and ultra-lightweight concrete using image-based approaches. *Construction and Building Materials*, *264*, 1-15. https://doi.org/10.1016/j.conbuildmat.2020.120241

Souza, F. B. D., Montedo, O. R. K., Grassi, R. L., and Antunes, E. G. P. (2019). Lightweight high-strength concrete with the use of waste cenosphere as fine aggregate. *Revista Materia*, *24*(4), 1-12. https://doi.org/10.1590/S1517-707620190004.0834

Tang, X., Zhao, C., Yang, Y., Dong, F., and Lu, X. (2020). Amphoteric polycarboxylate superplasticizers with enhanced clay tolerance: preparation, performance and mechanism. *Construction and Building Materials*. *252*, 1-9. https://doi.org/10.1016/j.conbuildmat.2020.119052

Tawfik, T. A., Alsaffar, D. M., Tayeh, B. A., Metwally, K. A., and Elkattan I. M. (2021). Role of expanded clay aggregate, metakaolin, and silica fume on the of modified lightweight concrete properties. *Geosystem Engineering*, *24*(3), 146-156. https://doi.org/10.1080/12269328.2021.1887002

Varga, I. D. L., Bentz, D. P., Weiss, W. J., Castro, J., Barrett, T. J., Spragg, R. P., Bella, C. D., Obla, K. H., Kim, H., Schindler, A. Keith, K. P., and Sato, T. (2012) *Increased use of fly ash in hydraulic cement concrete (HCC) for pavement layers and transportation structures*. Purdue University. https://doi.org/10.5703/1288284316554

Vijayalakshi, R., and Ramanagopal, S. (2018). Structural concrete using expanded clay aggregate: S review. *Indian Journal od Science and Technology*, *11*(16), 1-12. https://doi.org/10.17485/ijst/2018/v11i16/121888

Vivek, S. S., Karthikeyan, B., Kanna, G. R., Selvaraj, S. K., Jose, S., Palanisamy, P., and Adane, T. M. (2022). Study on fresh and mechanical properties of polyblend self-compacting concrete with metakaolin, lightweight expanded clay aggregate, and SAP as alternative resources. *Advances in Civil Engineering*, *2022*, 2350447. https://doi.org/10.1155/2022/2350447

Zhang, P., Xie, N., Cheng X., Feng L., Hou, P., and Wu, Y. (2018). Low dosage nano-silica modification on lightweight aggregate concrete. *Nanomaterials and Nanotechnology*, *8*, 1-8. https://doi.org/10.1177/1847980418761283

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e99526)

Research Article/Civil and Sanitary Engineering http://doi.org/10.15446/ing.investig.99526

A New Methodology Based on Artificial Intelligence for Estimating the Compressive Strength of Concrete from Surface Images

Una nueva metodología basada en inteligencia artificial para estimar la resistencia a la compresión del hormigón a partir de imágenes superficiales

Gamze Doğan[1], Ahmet Özkış[2], and Musa Hakan Arslan[3]

ABSTRACT

This study used digital image processing and an artificial neural network (ANN) to determine the compressive strength of concrete in reinforced concrete buildings without coring. First, 32 concrete samples were produced in the laboratory, with different water-to-cement ratios, aggregate types, amounts of binder, compression values applied to fresh concrete, and amounts of additive. Next, the locations of 192 cores were visualized, and the compressive strengths of their corresponding core samples were matched with the surface images of the concrete, which were then digitized by image processing. The digitized images were the input layer, and the training and testing procedures were performed using the ANN as an output layer. After testing, the model was validated in existing reinforced concrete buildings. For the verification process, 20 cores taken from randomly selected concrete buildings were used. Although the results obtained from the samples produced in the laboratory were satisfactory, the success rate of the samples taken from the field was limited. Finally, the findings of this study are compared against the literature on this subject, especially from the last two decades.

Keywords: reinforced concrete, building, digital image processing, intelligent system, compressive strength, experimentation

RESUMEN

En este estudio se utilizó procesamiento de imágenes digitales y una red neuronal artificial (ANN) para determinar la resistencia a la compresión del hormigón en edificios de hormigón armado sin tomar núcleos. Primero, se generaron 32 muestras de concreto en el laboratorio con diferentes proporciones de agua a cemento, tipos de agregado, cantidades de aglutinante, valores de compresión aplicada al concreto fresco y cantidades de aditivo. A continuación, se visualizaron las ubicaciones de 192 núcleos, y las resistencias a la compresión de sus correspondientes muestras se compararon con las imágenes de la superficie del hormigón, que se digitalizaron mediante procesamiento de imágenes. Si las imágenes digitalizadas fueron la capa de entrada, y los procedimientos de entrenamiento y prueba se realizaron utilizando la ANN como capa de salida. Después de las pruebas, el modelo se validó en edificios reales de hormigón armado. Para el proceso de verificación, se utilizaron 20 núcleos tomados de edificios de hormigón seleccionados al azar. Si bien los resultados obtenidos de las muestras producidas en el laboratorio fueron satisfactorios, el porcentaje de éxito de las muestras tomadas en campo fue limitado. Por último, se comparan los hallazgos del estudio con la literatura sobre este tema, especialmente de las últimas dos décadas.

Palabras clave: edificios de hormigón armado, procesamiento de imágenes digitales, sistema inteligente, resistencia a la compresión, experimentación

Received: April 1st, 2022
Accepted: April 10th, 2023

Nomenclature

Abbreviation	Meaning
ANN	Artificial neural networks
DIP	Digital Image processing
CS	Compressive strength
RC	Reinforced concrete
f_{cmin}	Minimum compressive strength of concrete
f_{cmean}	Average compressive strength of concrete
s_s	Standard deviation

---

[1] Faculty of Engineering and Natural Sciences, Department of Civil Engineering, Konya Technical University, Konya 42130, Turkey, gdogan@ktun.edu.tr

[2] Necmettin Erbakan University, Faculty of Engineering, Department of Computer Forensics Engineering, Konya 42090, Turkey, aozkis@erbakan.edu.tr

[3] Faculty of Engineering and Natural Sciences, Department of Civil Engineering, Konya Technical University, Konya 42130, Turkey, mharslan@ktun.edu.tr

Introduction

Determining the compressive strength (CS) of concrete in existing buildings is time-consuming, costly, and complicated. The most precise method for doing so is the destructive method of coring, whose disadvantages include damage to the area from which the core is taken, the loss of section capacity, after-coring repair and modification costs, testing costs, and a delay of at least one day between taking the core and obtaining the CS value.

Building codes – such as ACI-318-19 (2018), TS-500-2000 (2000), and TBEC-2018 (2018) – provide some rules for taking concrete core samples. They must be taken from an existing building if the supervising engineer requires tests to determine the CS of the concrete, and the coring must be done from locations that do not reduce the strength of the structure. Moreover, building codes allow the use of non-destructive testing methods (*e.g.*, surface hardness and sound velocity) for correlation with the concrete coring results, but these alternative methods are less effective than taking core samples. For example, the success rate of the ultrasonic sound velocity method, which is frequently cited in the literature, is known to be only 65-75% (Sbartai *et al.*, 2012; Breysse, 2012; Ferreira and Jalali, 2010).

There have been numerous studies involving different methods for determining the CS of concrete, and the success of each method has been assessed in comparison with that reported for the concrete pressure test in the literature (Bogas *et al.*, 2013; Sbartai *et al.*, 2012; Breysse, 2012; Ferreira and Jalali, 2010; Bingöl and Çavdar, 2018; Thapa *et al.*, 2019). It has been established that each method has its own success rate, which is unfortunately inversely related to its ease of application. Therefore, it is highly necessary to devise a method that is both easily applicable and cost-effective while providing results both quickly and as close as possible to those of coring.

In recent years, images of the surfaces of newly produced (Doğan *et al.*, 2015, 2017; Al-Kamaki *et al.*, 2017; Beskopylny *et al.*, 2022; Shiuly *et al.*, 2022) or hardened cylinder and cube samples have been taken to calculate the CS of concrete via image processing technology. As an alternative to mechanical methods, image processing technology (Jang *et al.*, 2019; Naderpour *et al.*, 2018; López *et al.*, 2009; Chang *et al.*, 2009; Gencturk *et al.*, 2014) allows measuring without direct contact with the surface of the material or requiring a complicated and expensive experimental setup.

Unlike previous studies, this study assesses the feasibility of determining the CS of reinforced concrete (RC) in existing buildings not by taking cores but via concrete-surface images analyzed using digital image processing (DIP) and an artificial neural network (ANN). The main objectives of this research are as follows: (i) to develop a new test method that engineers can use with an acceptable accuracy without damaging the structure; (ii) to make it easier to quickly reach results in the most time-consuming part of fieldwork regarding processes such as risk analysis in existing RC buildings and earthquake performance calculations; (iii) to provide significant cost and time savings by ensuring that concrete CS (one of the most critical factors in the behavior of a structure) can be determined both rapidly and economically; (iv) to offer an alternative and innovative method that uses expert systems instead of an uneconomical and time-consuming process that damages the structure (*e.g.*, coring).

Compressive strength (f_c) is one the main parameters of concrete since it directly affects other mechanical parameters. Especially in the last major earthquakes (such as the 2023 Kahramanmaraş-Turkey earthquakes, in which approximately more than 50 000 people died and hundreds of thousands of RC buildings completely collapsed), it has been clearly seen that a low concrete strength may cause structure collapse. The method proposed in this study shows the extent to which concrete strength can be determined in existing RC buildings with an artificial intelligence-based algorithm and without the need for destructive methods. The novelty of this study is the ability to determine the concrete strength detectability ratio from the surface in buildings without taking core samples via a technology used in many engineering problems. This work fills the gap on this subject while showing huge differences between the experimental studies carried out in a laboratory environment and the tests carried out in real buildings.

To this effect, 32 different samples were produced while varying five parameters (water/cement ratio, aggregate type, the amount of binder, the compression to be applied to fresh concrete, and the amount of additive) that affected the CS and surface appearance of the concrete. The locations of 192 cores to be taken from these concrete samples were determined, and these regions were visualized before coring. The CS values of the core samples taken from these regions were matched with the surface images of the concrete, which were digitized via image processing. The DIP images formed the input layer of an ANN, whose output layer was the core CS of the imaged regions. The 192 data were used to train and test the ANN, and the tested model was validated in existing RC buildings. For the verification process, 20 different concrete images taken from randomly selected concrete buildings were used.

A brief literature review

In this section, current studies on concrete CS estimation in the literature are briefly summarized (Table 1). Note that intelligent systems and algorithms have been frequently used in research on concrete samples in recent years.

Digital image processing and artificial neural network

ANNs are a type of machine learning algorithm inspired by the structure and functioning of the human brain. They are designed to recognize patterns and learn from data inputs in

Table 1. Literature survey on the use of intelligent systems in concrete

Researcher / Methods	Brief description	Main findings
Doğan *et al*. (2017) ANN and image processing	A new non-destructive experimental method that uses image processing techniques and ANN	The method showed a very high prediction success (between 97,18 and 99,87%).
Behnood and Golafshani (2018) ANN, MOGWO	A predictive model for the compressive strength of silica fume concrete	A total of 31 non-dominated ANN models on the Pareto front with different architectures and accuracies was achieved. 22 of them had a Pearson correlation coefficient higher than 0,95.
Golafshani *et al*. (2020) ANN, ANFIS, and GWO	A reliable model for the prediction of concrete compressive strength.	Using metaheuristic optimization algorithms in the training phase of ANN and ANFIS can reduce the weakness of classical optimization algorithms and lead to more reliable results.
Kandiri *et al*. (2020) ANN, multi-objective salp swarm algorithm, M5P model tree	A database for the compressive strength of concrete was elaborated via GGBFS, modeling the compressive strength of concrete with ground granulated blast furnace slag (GGBFS) using machine learning.	19 ANN models with various structures and precisions were obtained, which provides alternatives according to the required simplicity and complexity.
Feng *et al*. (2020) Adaptive boosting approach	The AdaBoost algorithm was adopted to predict the compressive strength of concrete.	The AdaBoost model can accurately and efficiently predict the compressive strength of concrete with given input variables.
Omran *et al*. (2016) Multilayer perceptron, support vector machines, Gaussian processes regression, M5P, REPTree, M5-Rules, and decision stump, additive regression, and bagging	A comparison of the performance of nine data mining models in predicting the compressive strength of a new type of concrete containing three alternative materials (fly ash, haydite lightweight aggregate, and Portland limestone cement).	The analytical results show that, with appropriate parameter settings, all of these models except decision stump can achieve an acceptable prediction performance.
Yaseen *et al*. (2018) Extreme Learning Machine (ELM)	The regression-based soft computing model called *Extreme Learning Machine* was used for predicting the compressive strength of lightweight foamed concrete.	Four models achieved superior predictive capability for the effects of various influential input variables on the compressive strength of concrete.
Dao *et al*. (2019) ANFIS, ANN	Two artificial intelligence approaches, ANFIS and ANN, were used to predict the compressive strength of GPC aggregates.	Both ANN and ANFIS models have strong potential for predicting the compressive strength of GPC.
Bui *et al*. (2018) modified firefly algorithm (MFA), ANN	An expert system to predict the compressive and tensile strength of HPC.	Rapid convergence to an optimal value is a key advantage of the MFA-ANN hybrid expert system, facilitated by the systematic updating, memorization, and optimization of firefly details along with weight and bias parameters.
Başyiğit *et al*. (2012)	IP methods.	The compressive strength of concrete was successfully estimated through the use of IP, demonstrating an impressive accuracy rate of 94,8%.
Waris *et al.* (2022)	Combined use of image pro-cess-ing, ANN and ANFIS computational tech-niques.	The accuracy achieved through the combination of IP and ANN stands at an impressive 99,7%, an accuracy of 97,8% was obtained with IP and ANFIS.

Source: Authors

a way that is similar to how the humans process information. An ANN consists of a large number of interconnected processing nodes, which are organized in layers. Each node receives an input from other nodes or from external data, processes it via a mathematical function, and then passes the output to other nodes in the network. The connections between the nodes are weighted, which means that some inputs are more important than others. During the training process, the network is presented with a large number of input-output pairs, and the weights of the connections between the nodes are adjusted in order to minimize the difference between the actual output and the desired output. This is typically done using an optimization algorithm such as a gradient descent.

ANNs are used in a wide variety of applications, including image and speech recognition, natural language processing, and predictive modeling. They have proven to be particularly effective in applications where the input data is high-dimensional and complex, and where traditional rule-based approaches are difficult to apply.

Digital image processing means processing images via computers or other digital devices. This process uses a range of mathematical, algorithmic, and hardware techniques to capture, process, analyze, refine, and ultimately make images more useful. Digital image processing is used in a variety of industries, such as medicine, security, computer vision, robotics, manufacturing, remote sensing, and entertainment. This technology is used to take image data, make them more meaningful, and facilitate analysis. Examples of this include X-ray images, air traffic control systems, image-based sensors in automated vehicles, *etc*.

Image processing involves emulating, in a computer environment, the visual mechanism of the human eye and the image-interpretation mechanism of the brain. With this technique, certain information that cannot be accessed directly from an image of either an entire object or a region of it can be accessed indirectly. An image can be defined by the function f (x, y) with values in the xy coordinate plane, each of which represents the color value of the image at those coordinates. The color values of monochrome images are defined either as black and white (0 and 1, respectively) or as grey levels (0-255) (González *et al*., 2004). Continuous density values in continuous space are expressed as discrete values using a digitization method, thereby allowing images to be displayed, improved, segmented, and transformed by computers (Easton Jr., 2010). This process is known as DIP, which, in this study was performed using the MATLAB software (MATrix LABoratory).

Materials and methods

This section explains the production and acquisition stages of the core sample data and the application of DIP and the ANN, which constitutes the analytical part of this study.

Training and production of the test set

First, a dataset was created for the training and testing procedures. Therefore, different concrete samples were produced with binary changes in five different parameters in a laboratory setting. The parameters and their ranges are presented in Table 2. Each parameter affected both the surface appearance and the mechanical properties of the concrete.

Table 2. Brief description of the parameters

Parameter Number	Parameter	Parameter properties
1	Water/ cement ratio	0.93-0.60
2	Aggregate	Normal aggregate – fine sand additive aggregate
3	Compaction	Yes/No
4	Standard curing	Applied – implemented
5	Admixture materials (plasticizer)	Yes/No

The 32 concrete samples were produced with the aim of taking a total of 192 core samples (six from each concrete sample). Figure 1 shows a schematic of the dimensions of the concrete samples, which were prepared according to the calculated mixture ratios, poured into rectangular molds and then grouped in sets (Figure 2). Table 3 shows the mixing ratio for each concrete sample. The cement used was selected according to TS-EN-197-1 (2012). Because the desired target concrete strengths were in the range of 15-35 MPa, the cement dosage used in the study varied. Inel *et al.* (2008) showed that the CS of the concrete in the RC buildings that account for a significant portion of the existing building stock in Turkey is quite variable. Therefore, half of the concrete samples prepared in the present study were subjected to TS EN 206 (2014), while the other half were not subjected to compression. Similarly, half of the samples were subjected to standard curing conditions according to TS EN 206 (2014), while the other half were not. The additive Sikament-MR 50SR was used in proportions of up to 1,0% by weight of cement

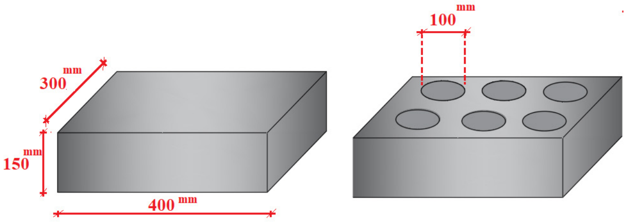


Figure 1. Schematic of the produced concrete masses
Source: Authors

Table 3. General properties of parameters.

Parameter Number	Parameters		Variation range	
			Code:1	Code:0
1	Cement (CEM I 42.5)		220 kg/m^3	360 kg/m^3
	Water		205 kg/m^3	209 kg/m^3
	Water/cement ratio		0.60	0,93
2	Aggregate granulite (Dmax=22,4 mm)	15-22,5	230	200
		0-5 Sand	510	675
		8-16	600	600
		0-5 Sand	510	500
3	Compaction		Complying with TS-EN 206	Non-complying TS-EN 206
4	Standard curing		Complying with TS-EN 206	Non-complying TS-EN 206
5	Admixture materials (plasticizer)		Sikament-MR 50SR (Complying with TS-EN 934-2)	No

Figure 2. Produced concrete masses
Source: Authors

28 days after the concrete samples were made, concrete faces were taken so that the cores could be taken. Having cleaned the surfaces of the concrete samples, the regions from which the cores were to be taken were marked and visualized under the same conditions. The surface images were taken at a constant distance under the same light intensity and with a fixed angle and resolution ratio. After imaging the locations of the cores, the latter were taken from the concrete samples. The imaging and coring process is shown schematically in Figure 3, and the coring process and core samples in the laboratory are shown in Figure 4. The core length-to-diameter ratio was set at 1,5. The process of making and breaking the cores was compliant with ASTM C42 (2016), TS EN 13791 (2010), and TS EN 12504-1 (2010). In the experiments on the core samples, a uniaxial vertical loading test was carried out by means of an automatic controlled hydraulic press with a vertical loading capacity of 3 000 kN. Table 4 presents the properties of the samples and their minimum and maximum strengths. The test results show that the core CS values were in the range of 15,51-41,84 MPa (fc).

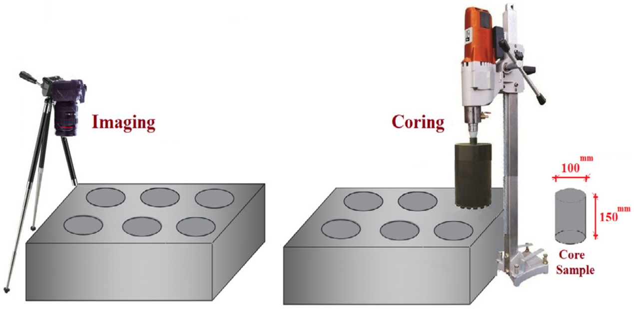


Figure 3. Imaging and coring of a concrete mass
Source: Authors

Figure 4. Data preparation: a) coring process; b) core samples whose surface images were taken
Source: Authors

Obtaining a validation set

To verify the results from the concrete samples to be used in the intelligent system for training and testing purposes, a validation dataset was created independently from the dataset. To create the validation set, six different existing RC buildings from different years in Turkey were randomly selected, and core sampling was carried out on them. The fact that they were all low-rise buildings is related to the distribution of the building stock in the selected region. Figure 5 shows the buildings from which the cores were taken, the year in which each building was constructed, and the number of cores taken from each building.

(a) (b)

(c) (d)
(e) (f)

Figure 5. Construction years of the buildings and the number of core samples taken from each one: a) building-1 (2001, three concrete cores); b) building-2 (2009, five concrete cores); c) building-3 (2011, three concrete cores); d) building-4 (2017, three concrete cores); e) building-5 (1989, three concrete cores); f) building-6, industrial structure (2015, three concrete cores)
Source: Authors

Table 4. Sample types, parameter codes, and strength ranges

Sample name*	Sample number*	W/C	Aggregate	Compaction	Standard curing	Admixture materials	fc (MPa)	
		Codes					min	max
A1	6	1	1	1	1	1	26.17	34.33
A2	6	0	1	1	1	1	23.56	33.99
A3	6	0	0	1	1	1	15.51	33.80
A4	6	0	0	0	1	1	20.85	29.48
B1	6	0	0	0	0	1	17.64	29.75
B2	6	0	0	0	0	0	21.65	32.00
B3	6	1	0	1	1	1	18.45	29.03
B4	6	1	1	0	1	1	23.56	30.21
C1	6	1	1	1	0	1	20.70	27.04
C2	6	1	1	1	1	0	17.19	21.25
C3	6	0	0	1	0	1	18.16	27.10
C4	6	0	1	0	1	0	18.26	23.67
D1	6	1	0	0	1	1	33.59	36.82
D2	6	1	0	0	0	1	31.41	41.84
D3	6	1	0	0	0	0	25.08	36.20
D4	6	1	1	0	0	1	26.20	33.95
E1	6	1	1	0	0	0	30.01	36.23
E2	6	1	1	1	0	0	25.14	36.17
E3	6	0	1	0	0	0	25.15	35.39
E4	6	0	1	1	0	0	28.88	32.93
G1	6	0	1	1	1	0	32.77	35.29
G2	6	0	0	1	0	0	27.89	35.46
G3	6	0	0	1	1	0	18.31	32.31
G4	6	0	1	1	0	1	26.80	36.86
H1	6	0	0	0	1	0	29.78	36.20
H2	6	1	1	0	1	0	28.39	34.82
H3	6	0	1	0	0	1	20.60	38.95
H4	6	0	1	0	1	1	27.25	35.88
K1	6	1	0	1	0	0	22.42	37.32
K2	6	1	0	1	0	1	34.11	41.84
K3	6	1	0	1	1	0	27.33	36.26
K4	6	1	0	0	1	0	30.57	37.44

*: A total of 8 different groups (A-K), 4 different types (1-2-3 and 4) and 192 different samples in each group

Figure 6. Coring from buildings and imaging of cores
Source: Authors

The photographs of the core surfaces were taken at a fixed distance, and the CS values of the cores was then determined by subjecting them to a uniaxial vertical loading test. Because the buildings that were used were constructed in different years and with different classes of concrete, the strengths of the concrete samples were in the range of 13,41-35,63 MPa.

The process of obtaining the cores from the RC structures and the images of the cored samples is shown step by step in Figure 5. First, the areas in the RC elements from which the cores were to be taken were determined, for which an X-ray device was used, in order to identify any reinforcements inside the elements (Figure 6a). Having determined a non-reinforced area in an RC element, the area to be cored was prepared for imaging (Figure 6b). To obtain an image of the concrete surface in the coring area, the layers of paint and plaster on the column were dug through to reach the plain concrete surface (Figure 6c). The camera was secured on a tripod to obtain an image from a fixed distance (40 cm away) (Figure 6d). A coring device was used to cut a core sample from the RC column (Figure 6d and 6f), and the core sample was also photographed. Here, the aim was to determine whether the photographs of the core samples before cutting yield different results from those after cutting (Figure 6g). The core samples were then brought to the laboratory, and their CS values were measured in a uniaxial pressure test using a hydraulic press (Figure 6h).

Analytical work for the core samples

In this study, the concrete dataset used for training and the test data of the algorithm comprised the 192 concrete samples obtained by taking six cores from 32 different concrete types. Because the CS of the concrete was a random variable, the frequency factor was taken to be 1,28 according to the probability distribution of the concrete CS (assuming an exceedance probability of 10%). The standard deviation of the distribution was calculated, and elements with either very low or very high strength were removed from the 192 concrete data (Figure 7). This process left 160 concrete samples in the data set.

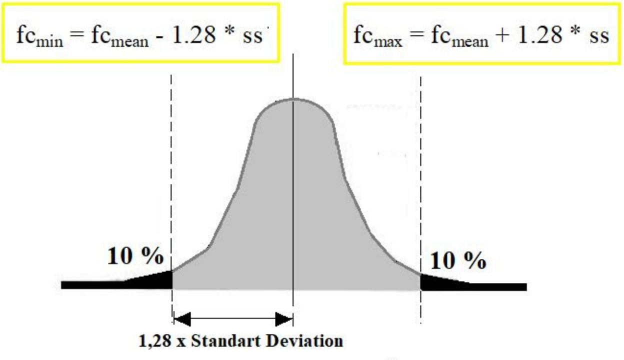


Figure 7. Elimination of the data set
Source: Authors

A 779 × 1 069 pixel image was taken from each of the 160 concrete samples and then converted to grey scale. Then, the average, standard deviation and median values of all the columns in each of the concrete images were calculated, thus yielding data comprising 3 × 1 069 = 3 207 pixels for each concrete image. Thus, a 160 × 3 207 pixel dataset was created for the 160 concrete samples.

The fact that the number of inputs for the ANN was as large as 3 207 had a negative effect on the training process in terms of time and calculation. To counter that problem, a feature-extraction algorithm based on principal components analysis (PCA) was used to reduce the number of properties for each concrete sample to 50, resulting in a data set in the form of a 160 × 50 matrix. The 160 data were divided into 128 training and 32 testing data to perform five-fold cross-validation (a separate 32-group for each cross-validation process was divided as testing data).

In this study, the Levenberg-Marquardt ANN model was used, with 50 neurons in the input layer, 52 in the interlayer, and one in the output layer (Figure 8). Table 5 presents the values of the other network parameters. In the literature, commonly used activation functions are the linear, step, sigmoid, Gaussian, and hyperbolic tangent. In this study, the selected activation function was the (logarithmic) sigmoid one. The ANN training process was terminated either after 50 iterations or if the training error became less than 10-6. The latter tended to happen after 10 to 12 epochs (Figure 9).

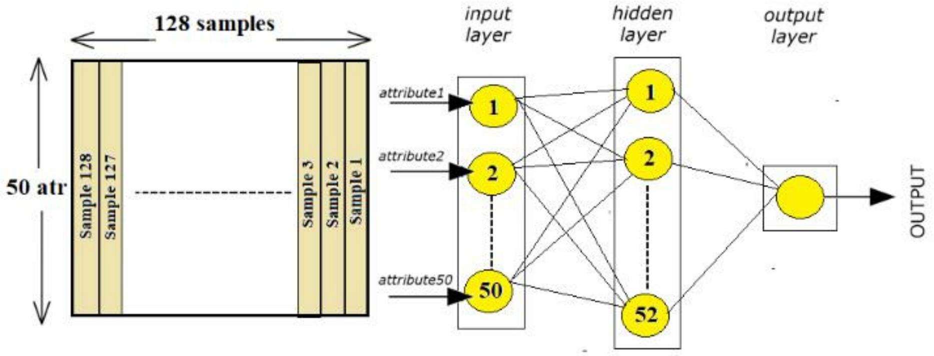


Figure 8. ANN structure used for training
Source: Authors

Table 5. ANN model and used parameters

Properties	Variation range
Activation function	Logarithmic sigmoid
Epoch number	50
Stopping criterion	1e-6
Maximum validation error	6
Minimum error tolerance	1e-10
Marquardt setting parameter	0,005
Marquardt setting parameters reduction factor	0,1
Marquardt setting parameters growth factor	10
Marquardt setting parameters max value	1e10
How many epochs will be updated in the image	50

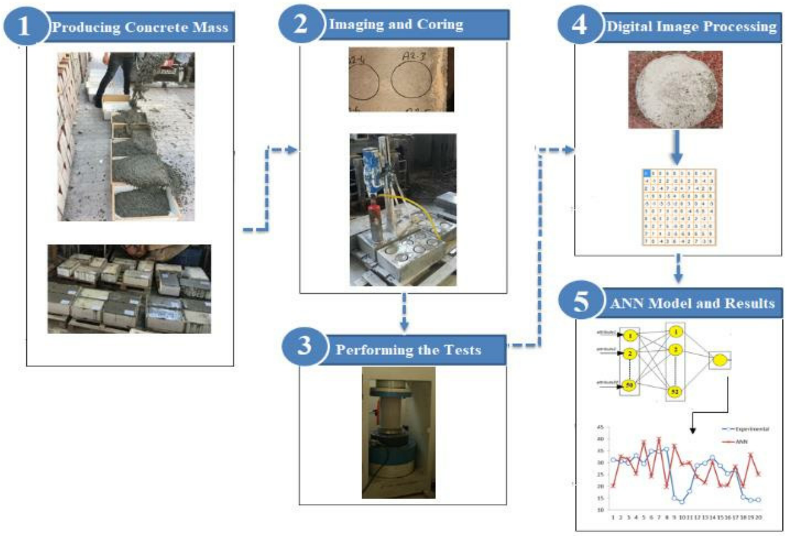


Figure 9. Process steps
Source: Authors

Results

The 160 data were divided into 128 training and 32 testing data to perform different five-fold cross-validations. For each cross-validation process, a separate group of 32 was determined as the testing data. After completing the training, the estimated values produced by the network for the 32 testing data were compared against the expected values. If the margin of error was less than 5% (between the estimate and the expected value), then the value was considered to be correct. The margin of error was then increased by 5%, and the correct predicted value ratio was determined. In light of the fact that destructive test methods include error levels of 30-40%, Table 6 shows (i) an accuracy of 73,6% with a 30% margin of error and (ii) an accuracy of 86,4% with a 40% margin of error. Table 6 shows the accuracy and error ratios, and Figure 10 shows a graph of the mean accuracy of the five-fold cross-validation results.

After training and testing the ANN model, it was applied to the validation set. The laboratory results for the cores obtained from the buildings were compared to those obtained from the ANN. Table 7 lists the experimental CS values obtained from the core samples and the estimated CS values obtained via DIP and the ANN. Figure 11 compares the results graphically.

Table 6. Error ratio (%) – Accuracy rate (%)

Error ratio (%)	Accuracy Rate (%)					Average accuracy rate (%)
	Cross fold 1	Cross fold 2	Cross fold 3	Cross fold 4	Cross fold 5	
30%	81	53	**84**	72	78	73.6
35%	84	69	**88**	84	78	80.6
40%	**94**	72	**94**	84	88	86.4

Source: Authors

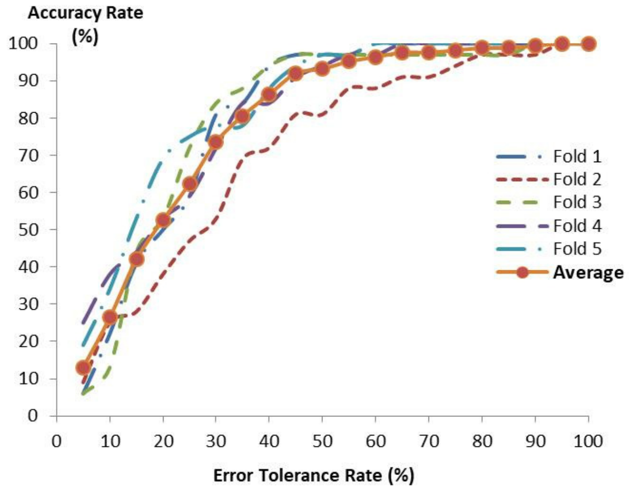


Figure 10. Accuracy of 5-fold cross validation results for the test set
Source: Authors

Table 7. Compressive strength values (MPa) of the core samples (experimental, EXP – analytical, ANN)

N-1		N-2		N-3		N-4	
							
EXP	ANN	EXP	ANN	EXP	ANN	EXP	ANN
31,18	20,13	30,61	32,53	29,78	31,47	33,06	25,35
N-5		N-6		N-7		N-8	
							
EXP	ANN	EXP	ANN	EXP	ANN	EXP	ANN
29,53	38,60	34,91	24,17	34,81	39,93	35,63	19,80
N-9		N-10		N-11		N-12	
							
EXP	ANN	EXP	ANN	EXP	ANN	EXP	ANN
15,04	37,02	13,41	29,28	17,86	29,89	28,81	23.96
N-13		N-14		N-15		N-16	
							
EXP,	ANN	EXP	ANN	EXP	ANN	EXP	ANN
29,76	21,47	32,14	30,19	28,66	20,23	25,37	20,45
N-17		N-18		N-19		N-20	
							
EXP	ANN	EXP	ANN	EXP	ANN	EXP	ANN
26,94	28,26	15,49	19,83	14,08	33,21	14,24	25,05

Source: Authors

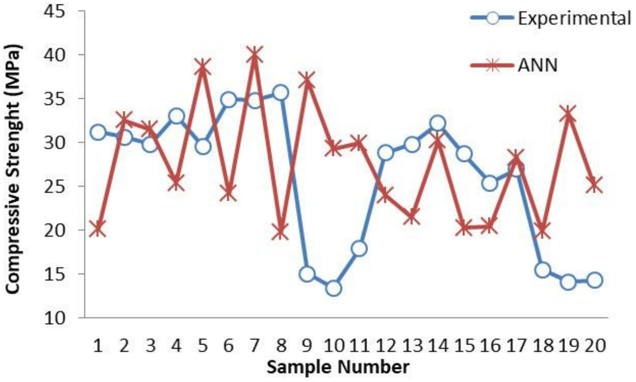


Figure 11. Compressive strength comparison chart for real buildings (experimental-ANN)
Source: Authors

Table 7 shows that, when the estimated and measured CS values (MPa) of the core samples are compared, some are close, and the results of different experiments were reached in most samples.

Discussion

The results obtained in this study are somewhat poor when compared to similar studies in the literature (such as Başyiğit *et al*., 2012, and Waris *et al*., 2022, who found success rates close to 95% in their studies). However, the main reason for this is that the images were not taken from the surfaces of the concrete samples, which were tested by cutting (slicing) in a laboratory environment. However, there is not much chance of such a production being practically carried out in field studies. Although the results obtained for the laboratory-produced concrete samples are quite satisfactory, the success for the fieldwork samples in the validation set was limited. The proposed method had a success percentage of less than 50% for these samples. In this case, the training, testing, and validation set yield quite different results. The fact is that the laboratory-produced samples do not fully reflect the concrete in the real RC buildings; the images of the cores taken from the existing buildings are different from those of the concrete samples obtained in the laboratory. The results were affected by changes in the image quality of the samples and the difference in the amount of light. The reason for this was fluctuations in the images due to the roughness of the core samples. In this study, because the prediction success for the existing buildings was low, the authors conclude that using only surface images is insufficient for estimating the CS of concrete. It is thought that using images of internal concrete surfaces revealed by cutting the core samples will increase the success of predicting the CS. Similarly, high success was achieved from new concrete cylinder and cube samples (Jang *et al*., 2019) taken using the appropriate method, while the success in core samples was not satisfactory.

According to the results obtained in previous studies involving artificial intelligence and examples of various engineering problems, this study's results are poor. The main reasons for this are given below.

1. Although a comprehensive dataset was prepared for the laboratory test and a wide parameter group was used, the surface images obtained in the laboratory did not represent the results obtained from a real structure.

2. In the literature, the training and testing datasets are similar in many engineering problems to which artificial intelligence is applied. Moreover, in many applications, the testing and training data are produced in a laboratory environment, in which case the results tend to be more accurate. However, in a branch of science such as civil engineering, where laboratory and field conditions vary greatly, it would not be right to perform evaluations only from laboratory results. The results show how field and laboratory studies differ in structural engineering. Sufficient engineering service (proper supervision to reduce human error) is also vital, especially during concrete production.

3. There are many parameters affecting concrete CS. In this study, parameters such as the water/cement ratio, aggregate type, compaction, curing, and admixture materials (plasticizer) were used to create the dataset. 28 days after the production of concrete in the laboratory, the surface images were taken, and compression tests were conducted. On the other hand, the construction years of the randomly selected buildings in the field study vary between 1989 and 2017. Here, it can be said that the image data changed a lot, as the carbonation formed on the concrete surface changes depending on the time. This is thought to be the most crucial reason that affects the results.

4. Another cause is surface defects due to the type of RC formwork, design errors, and the improper use of formwork oils. In practice, mold surfaces tend to be made of wood, plywood, metal, or plastic, and impermeable mold surfaces that do not absorb water cause gaps and possibly surface roughness on the concrete.

5. In the core application of this study, the plaster was scraped before the concrete surface was photographed. To remove dust, it would be useful to clean the photographed area using either compressed air or pressurized water.

6. The direction in which concrete is cast also affects its placement and therefore its CS. The samples produced in the laboratory were cast horizontally, whereas the real building samples that were tested were taken from columns that had been cast vertically.

7. While extracting the concrete core samples, the possible reinforcement bar parts in the core samples can significantly change the compression test results.

8. Increasing the dataset used in the study will positively affect the results. However, the fact that the surface image quality is standard (i.e., photographed distance, light level, resolution, shading) is fundamental. These parameters affect the result.

9. In the study, in the image analysis step, a region of 779 x 1 069 pixels was taken from the concrete core image. The whole area of the core affecting compressive strength was not used.

10. Regarding the core samples taken from the RC in the buildings, parts of the reinforcement bars that could have been in the core samples would have significantly changed the results of the compression test.

In the literature, for the determination of concrete CS, the success rates of the methods proposed in comparison with the core test results do not exceed this level. When evaluated from this point of view, it can be said that only the analyses made on laboratory samples are satisfactory. The authors think that a combined non-destructive testing method will be an important tool for determining CS in the future, not only with images, but also with other parameters such as sound permeability and surface hardness.

Conclusions

In this study, an experimental and analytical study was conducted using DIP and ANN together to determine the compressive strength of concrete in existing reinforced concrete buildings without cores. According to the study, the findings presented in the following items were reached.

1. The laboratory samples obtained an accuracy of 73,6% with a 30% margin of error and an accuracy of 86,4% with a 40% margin.

2. In this study, the test procedures had a high success rate, which was lower for random samples collected from real buildings.

3. The success rate of other non-destructive testing methods such as surface hardness and sound absorption is known to be only 65-75% The prediction success of the proposed method in this study is slightly better than that of traditional methods.

Acknowledgements

This study was carried out within the framework of research project no. 16401038, supported by the Selcuk University, Scientific Research Projects Coordination.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

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

References

ACI 318-19 (2018). *Building code requirements for structural concrete and commentary*. American Concrete Institute. https://www.usb.ac.ir/FileStaff/5526_2020-1-25-11-12-7.pdf

Al-Kamaki, Y. S. S., Al-Mahaidi, R., and Bennetts, I. (2017). Strain efficiency of carbon fibre reinforced polymer-confined RC columns. *Journal Of Duhok University*, *20*(1), 484-497. https://Doi.Org/10.26682/Sjuod.2017.20.1.43

ASTM C42/C42M (2016). *Standard test method for obtaining and testing drilled cores and sawed beams of concrete*. ASTM. https://www.astm.org/c0042_c0042m-13.html

Başyiğit C., Çomak B., Kiliçarslan S., and Üncü I. S. (2012). Assessment of concrete compressive strength by image processing technique. *Construction and Building Materials, 37*, 526-532. https://doi.org/10.1016/j.conbuildmat.2012.07.055

Behnood, A., and Golafshani E. M., 2018, Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves, *Journal of Cleaner Production*, *202*, 54-64. https://doi.org/10.1016/j.jclepro.2018.08.065

Beskopylny, A. N., Stel'makh, S. A., Shcherban', E. M., Mailyan, L. R., Meskhi, B., Razveeva, I., Chernil'nik, A., and Beskopylny, N. (2022). Concrete strength prediction using machine learning methods CatBoost, k-Nearest Neighbors, Support Vector Regression. *Applied Science, 12*, 10864. https://doi.org/10.3390/app122110864

Bingöl, S., and Çavdar, A. (2016). *A new nomogram proposal to determine concrete compressive strength by combined nondestructive testing methods. Research in Nondestructive Evaluation*, *29*(1), 1-17. https://www.tandfonline.com/doi/abs/10.1080/09349847.2016.1195466

Bogas J. A., Gomes, M. G., and Gomes, A. (2013). *Compressive strength evaluation of structural lightweight concrete by non-destructive ultrasonic pulse velocity method*. *Ultrasonics*, *53*, 962-97. https://pubmed.ncbi.nlm.nih.gov/23351273/

Breysse, D. (2012). Nondestructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods. *Construction and Building Materials, 33*, 139-163, https://doi.org/10.1016/j.conbuildmat.2011.12.103

Bui, D. K., Nguyen, T., Chou, J. S., Xuan, H. N., and Ngo T. C. (2018). A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. *Construction and Building Materials*, *180*, 320-333. https://doi.org/10.1016/j.conbuildmat.2018.05.201

Chang, C. W., Chen, P. H., and Lien, H. S. (2009). Evaluation of residual stress in pre-stressed concrete material by digital image processing photoelastic coating and hole drilling method. *Measurement*, *42*(4), 552-558. https://doi.org/10.1016/j.measurement.2008.10.004

Dao, D. V., Ly, H. B., Trinh, S. H., Le, T. T., and Pham, B. T. (2019). Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials*, *12*(6), 983. https://doi.org/10.3390/ma12060983

Doğan, G., Arslan, M. H., and Ceylan, M. (2015). Statistical feature extraction based on an ANN approach for estimating the compressive strength of concrete. *Neural Network World*, *25*(3), 301-318. http://www.nnw.cz/doi/2015/NNW.2015.25.016.pdf

Doğan, G., Arslan, M. H., and Ceylan, M. (2017).Concrete compressive strength detection using image processing based new test method. *Measurement, 109*, 137-148. https://doi.org/10.1016/j.measurement.2017.05.051

Easton Jr., R. L. (2010). *Fundamentals of digital image processing.* https://www.cis.rit.edu/class/simg361/Notes_11222010.pdf

Feng, D. C., Liu, Z. T., Wang, X. D., Chen, Y., Chang, J. Q., Wei D. F., and Jiang, Z. M. (2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach, *Construction and Building Materials*, *230*, 117000. https://doi.org/10.1016/j.conbuildmat.2019.117000

Ferreira, R. M., and Jalali, S. (2010). NDT measurements for the prediction of 28-day compressive strength. *NDT&E International, 43*, 55-61. https://doi.org/10.1016/j.ndteint.2009.09.003

Gencturk, B., Hossain, K., Kapadia, A., Labib, E., and Mo, Y. L. (2014). Use of digital image correlation technique in full-scale testing of prestressed concrete structures. *Measurement, 47*, 505-515. http://dx.doi.org/10.1016/j.measurement.2013.09.018.

González, R. C., Eddins, S. L., and Woods, R. E. (2004). *Digital image processing using MATLAB* (vol. 624). Pearson-Prentice-Hall.

Golafshani, E. M, Behnood, A., and Arashpour, M. (2020). Predicting the compressive strength of normal and high-performance concretes using ANN and ANFIS hybridized with grey wolf optimizer. *Construction and Building Materials*, *232*, 117266. https://doi.org/10.1016/j.conbuildmat.2019.117266

Inel, M., Bilgin, H., and Ozmen, H. B. (2008). Seismic capacity evaluation of school buildings in Turkey. *Structures & Buildings, 161*(3),147-159. https://doi.org/10.12989/sem.2008.30.5.535.

Jang, Y., Ahn, Y., and Kim, H. Y. (2019). Estimating compressive strength of concrete using deep convolutional neural networks with digital microscope images. *Journal of Computing in Civil Engineering, 33*(3), 0000837. https://ascelibrary.org/doi/10.1061/%28ASCE%29CP.1943-5487.0000837

Kandiri, A., Golafshani, E. M., and Behnood, A. (2020). Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm. *Construction and Building Materials*, *248*, 118676. https://doi.org/10.1016/j.conbuildmat.2020.118676

López, M., Kahn, L. F., and Kurtis, K. E. (2009). Characterization of elastic and time-dependent deformations in high performance light weight concrete by image analysis. *Cement and Concrete Research, 39*, 610-619. http://dx.doi.org/10.1016/j.cemconres.2009.03.015

Naderpour, H., Rafiean, A. H., and Fakharian, P. (2018). Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *Journal of Building Engineering, 16*, 213-219. https://doi.org/10.1016/j.jobe.2018.01.007

Omran, B. O., Chen, Q., and Jin, R. (2016). Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete. *Journal of Computing in Civil Engineering*, *30*(6), 0000596. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000596

Sbartai, A. M., Breysse, D., Larget, M., and Balayssac, J. P. (2012). Combining NDT techniques for improved evaluation of concrete properties. *Cement and Concrete Composites, 34*, 725-733. https://doi.org/10.1016/j.cemconcomp.2012.03.005

Shiuly, A., Dutta, D., and Mondal, A. (2022). Assessing compressive strengths of mortar and concrete from digital images by machine learning techniques. *Frontiers of Structural and Civil Engineering*, *16*, 347-358. https://doi.org/10.1007/s11709-022-0819-z

Thapa, S., Halder, L., and Dutta, S. C. (2019). *Evaluation of concrete made with stone and brick aggregate using non-destructive testing. Municipal Engineer*, *174*(1), 43-50. https://doi.org/10.1680/jmuen.18.00030

TBEC-2018 (2018). *Turkey building earthquake code*. Ministry of the Interior, Disaster and Emergency Management Authority.

TS EN 197-1 (2012). Cement- Part 1: *Compositions and conformity criteria for common cements.* TSE.

TS EN 206- (2002). Concrete – Part 1: *Specification, performance, production and conformity*. TSE.

TS EN 12504-1 (2010). *Testing concrete in structures – Part 1: Cored specimens – taking, examining and testing in compression*. TSE.

TS EN 13791 (2010). *Assessment of in-situ compressive strength in structures and precast concrete components.* TSE.

TS-500 (2000). *Turkish standards, design and construction rules of concrete.* TSE.

Waris, M. I., Plevris, V., Mir, J., Chairman, N., and Ahmad, A. (2022). An alternative approach for measuring the mechanical properties of hybrid concrete through image processing and machine learning. *Construction and Building Materials*, *328*, 12689. https://doi.org/10.1016/j.conbuildmat.2022.126899

Yaseen, Z. M., Deo, R. C., Hilal, A., Abd, A. M., Cornejo Bueno, L., Salcedo-Sanz, S., and Nehdi, M. L. (2018) Predicting Compressive Strength of Lightweight Foamed Concrete Using Extreme Learning Machine Model. *Advances in Engineering Software*, *115*, 112-125. https://doi.org/10.1016/j.advengsoft.2017.09.004

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e107160)

Research Article/Civil Engineering http://doi.org/10.15446/ing.investig.107160

Lessons Learned from Construction Site Layout Planning Practices

Lecciones aprendidas de las prácticas de planificación del diseño del sitio de construcción

Seng Hansen[1]

ABSTRACT

Proper layout planning is necessary to prepare a site for construction activities, prevent potential delays, and ensure project safety and security. Despite the considerable amount of studies regarding construction site layout planning, most research has relied on fabricated data using simulations, while, in fact, site layout is unique for each construction project, and its planning requires previous experience. This paper presents an in-depth study on site layout planning practices, as observed in four actual project cases. Using an empirical case study approach, these practices are collected through site observations, document examinations, and semi-structured interviews. Several discussions related to site layout planning practices are described, including the considerations, types, tasks and benefits related to proper site layout planning. The contribution of this paper is the ability to capture some lessons learned regarding construction site layout planning from actual project cases, which can be utilized by other parties for the benefit of future projects. The lessons learned allow developing appropriate strategies required by construction project teams in order to increase their response capability and hence be able to pro-actively make decisions regarding proper site layout planning.

Keywords: construction site, empirical case study, Indonesia, project planning

RESUMEN

La planificación adecuada de diseño es necesaria para preparar un sitio para actividades de construcción, prevenir posibles demoras y garantizar la seguridad de un proyecto. A pesar de la cantidad considerable de estudios sobre la planeación del diseño de sitios de construcción, la mayoría de las investigaciones se han basado en datos fabricados mediante simulaciones, si bien, de hecho, el diseño de sitio es único para cada proyecto de construcción y su planeación requiere experiencia previa. Este artículo presenta un estudio a profundidad sobre prácticas de planeación para el diseño de sitio en cuatro casos de proyectos reales, Empleando un enfoque de estudio de caso empírico, se recolectan estas prácticas por medio de observaciones en sitio, examen documental y entrevistas semiestructuradas. Se describen varias discusiones relacionadas con las prácticas de planeación para el diseño de sitio, incluyendo las consideraciones, los tipos, las tareas y los beneficios relacionados con una planeación adecuada para diseño de sitio. La contribución de este artículo radica en la posibilidad de capturar algunas lecciones aprendidas con respecto a la planeación del diseño de sitios de construcción de casos reales de proyectos, las cuales pueden ser aprovechadas por otras partes para el beneficio de proyectos futuros. Las lecciones aprendidas permiten desarrollar estrategias apropiadas según lo requieran los equipos de proyectos de construcción para aumentar su capacidad de respuesta y, por tanto, ser capaces de tomar decisiones de manera proactiva respecto a la planeación adecuada del diseño de sitio.

Palabras clave: sitio de construcción, estudio de caso empírico, Indonesia, planificación de proyectos

Received: February 6th, 2023
Accepted: July 25th, 2023

Introduction

Construction projects are labor-intensive, so construction sites are prone to becoming a source of delays and conflicts if not properly managed. Researchers have recognized site layout as one of the main factors affecting construction project performance (Liang *et al.*, 2018; Su *et al.*, 2012; Zolfagarian and Irizarry, 2014; Tsegay *et al.*, 2023a). Therefore, proper management is necessary to prepare a site for construction activities, as well as to prevent and overcome potential delays and ensure compliance with project occupational safety and health (OSH). As site layout affects productivity, cost, quality, duration, and work safety (Alavi and Rizk, 2021; Pradana, 2021; Sanad *et al.*, 2008; Xu and Li, 2012; Zolfagarian and Irizarry, 2014), it must be properly planned. However, this is the most ignored aspect by project managers and site engineers (Kamil and Adnan, 2020; Sadeghpour *et al.*, 2006; Sanad *et al.*, 2008). In most cases, it is carried out on a first-come-first-served basis through human judgment (Andayesh and Sadeghpour, 2013; El-Rayes and Said, 2009; Kim *et al.*, 2021; Tam and Tong, 2003). This could result in chaotic site management, which ultimately leads to poor project performance and accidents (Mawdesley *et al.*, 2002; Sadeghpour *et al.*, 2006).

---

[1] PhD from RMIT University, Australia. Affiliation: Full time lecturer, Universitas Agung Podomoro, Indonesia. E-mail: seng.hansen@gmail.com

Construction site layout is unique for each project (Tam *et al*., 2002; Vashishtha, 2022; Zolfagarian and Irizarry, 2014). This concept refers to the graphical representation of the construction site, which shows the area and arrangement of the necessary facilities and resources for construction work. A site layout presents a lot of information that can be considered by site engineers and contractors in managing construction activities. Meanwhile, site layout planning is defined as the process of arranging materials, facilities, and equipment within a construction site (Pheng and Hui, 1999). A site layout plan may consist of the planned building, existing buildings, temporary facilities, utilities, and access roads (Marx and Konig, 2011).

Although there is a lot of published literature that examines site layout planning problems, most works consider data fabricated using simulations and mathematical models (Alavi and Rizk, 2017; Binhomaid, 2019; El-Rayes and Said, 2009; Fadhlan *et al*., 2020; Hammad *et al*., 2016; Kim *et al*., 2021; Pradana, 2021; Sanad *et al*., 2008; Su *et al*., 2012; Tsegay *et al*., 2023a). A model (simulation or optimization) for construction site layout planning must comply with the rules used to create it. It is very difficult to create an entirely realistic model due to a lack of research and past data. Consequently, these models may rely on many assumptions that can affect the quality of site layout planning decisions (Marx and Konig, 2011). Therefore, previous project experience can be a good lesson to obtain input information.

Lessons learned is a technique or approach aimed at gaining knowledge and understanding through experience (Weber *et al*. 2000). It can be used on all types of construction to improve project quality, manage changes, and control expectations (Resende *et al*., 2020). To obtain these lessons, an evaluation of project phase or performance is carried out by the organization. The results regarding the lessons learned must be recorded within a knowledge management process which will then be transferred to other parties (Resende *et al*., 2020). Considering their importance in construction projects, this paper aims to investigate the practices of site layout planning from actual project cases. Using an empirical case study approach, experience from these projects is recorded, analyzed, and shared with the objective of improving the performance of construction projects in the future.

Literature review

Construction site management

Construction site management refers to planning, organizing, and monitoring construction facilities and resources in an integrated manner for a smooth workflow. It involves many tasks such as site investigation, site layout planning, materials delivery, site housekeeping, and site communication. As construction workers spend most of their time in construction sites (Zolfagarian and Irizarry, 2014), proper site management is critical for the successful completion of construction projects. Thus, if these workers can move quickly, easily, and safely within the site, productivity will increase, and work will be completed more quickly (Sadeghpour *et al*., 2006; Zolfagarian and Irizarry, 2014).

On the other hand, construction space is a limited resource in construction projects (Sadeghpour and Andayesh, 2015; Tsegay *et al*., 2023b; Vashishtha, 2022). It is equal to other construction resources such as money, time, materials, labor, and equipment (Hegazy and Elbetagi, 1999; Sadeghpour *et al*., 2006). However, as an integral part of construction site management, it is a dynamic environment which ideally changes its layout in order for each project phase to be effective (Liang *et al*., 2018). Therefore, site layout planning is a crucial task that has a major impact on construction productivity, costs, and safety (El-Rayes and Said, 2009; Kamil and Adnan, 2020; Sjøbakk and Skjelstad, 2015).

Site layout planning is an essential task of construction site management; one of the main causes of project failure is poor site layout planning practices (Mawdesley *et al*., 2002; Vashishtha, 2022). On the contrary, careful planning of the site layout, including the proper placement of temporary facilities, can help project managers to achieve significant gains in the form of increased productivity, decreased travel times, and better working conditions (Fadhlan *et al*., 2020; Tsegay *et al*., 2023a; Vashishtha, 2022). According to Skejlbred *et al*. (2015), proper site layout planning, simultaneously considering logistics and resource allocation, must be implemented for an effective project performance.

Construction site layout

Site layout planning includes the process of identifying the position and size of temporary and permanent facilities in construction sites (Alavi and Rizk, 2017; Sanad *et al*., 2008; Zouein *et al*., 2002). Temporary facilities are those needed in carrying out work but are not a physical part of the project. This may include but is not limited to fabrication yards, warehouses, material storages, parking lots, and temporary offices (El-Rayes and Said, 2009; Sanad *et al*., 2008). In contrast to temporary facilities, permanent facilities will become an integral part of the finished building.

Sanad *et al*. (2008) highlighted the importance of site layout planning to ensure project safety and the effectiveness and efficiency of project operations. Site layout planning depends on many factors and constraints. Traditionally, site layout plans were mainly developed based on personal experience and common sense (Cheng and O'Connor, 1996; Sanad *et al*., 2008; Xu *et al*., 2020). However, complex terrain conditions cause difficulties in site layout planning (Xu *et al*., 2020). Moreover, site layout plans often do not reflect the actual construction process (Xu *et al*., 2020) because they are not updated. This causes them to be ineffective in controlling work on site.

Ignoring site layout planning in the initial phase can result in an improper site layout that requires corrections during the execution of a project. These on-site corrections imply a greater cost than preventive measures in the early stages of planning. Therefore, a considerable amount of research has attempted to improve construction site layout planning performance by introducing systematic site layout planning approaches. Cheng and O'Connor (1996) proposed an automated site layout system using a Geographical Information System (GIS) and a database management system.

Optimization techniques that consider the dynamic interdependencies of construction sites and facilities in various project phases must therefore be developed (Pteroutsatou *et al*., 2021). Zouein and Tommelein (1999) presented a dynamic site layout planning model using a hybrid incremental solution method. Osman *et al*. (2003) developed an automated computer system based on an evolutionary optimization genetic algorithm (GA) engine. Khalafallah and El-Rayes (2006) presented an optimization GA model for airport construction site layout planning in order to minimize the hazards of wildlife attractants and costs. Sanad *et al*. (2008) provided an optimization model for construction site layout planning that considers safety and environmental issues.

El-Rayes and Said (2009) studied a dynamic site layout planning process using approximate dynamic programming. They further developed a model to optimize construction materials storage layout for cost-effectiveness purposes (Said and El-Rayes, 2011). Through proper site layout planning (including logistics planning), poor inventories of construction materials, people movement, material transportation, and waiting times can be effectively addressed (Alavi and Rizk, 2021; Sezer and Fredriksson, 2021). Similarly, Zhang and Yu (2021) developed a mathematical model using the particle swarm optimization algorithm in order to solve the dynamic constraints of site layout planning for prefabricated components by enhancing site coordination.

Zolfagharian and Irizarry (2014) emphasized the importance of a rule-based site layout checking system. Hammad *et al*. (2016) developed a computational framework for estimating the travel frequencies required in site layout planning. Alavi and Rizk (2017) developed an integrated framework for construction site layout planning using GA simulation. Ning *et al*. (2018) found concerns related to site waste and risks resulting from the interactions between facilities, and they developed an assessment model to evaluate the safety risks in construction site layout planning. Several research works have used simulation-based techniques to solve the adjustment of materials delivery and construction schedules in addressing the dynamic site layout planning problem (Alanjari *et al*., 2014; Ma *et al*., 2005).

In recent years, some contemporary technologies have been utilized for construction site layout planning (Alavi and Rizk, 2021). For instance, Liang *et al*. (2018) developed a real-time construction site layout monitoring system using building information modeling (BIM). Similarly, Xu *et al*. (2020) proposed a platform that allows project managers to interactively place temporary facilities in a virtual environment by integrating BIM and virtual reality (VR) for highway construction. Meanwhile, Jiang *et al*. (2020) used drones to reconstruct the construction site in real time for the purposes of hoist and facility layout planning.

Lessons learned in construction management research

Lessons learned is a technique that can be used to improve the performance of the construction industry. This is done by studying actual experiences from previous project implementations (Hansen *et al*., 2018), which may be good or bad (Weber *et al*., 2000). Harrison (2003) sees it as an innovative approach to capture and share, either to repeat the application of good practices or to avoid the recurrence of bad practices. It is defined as "the shared knowledge that allows project managers and the organization to grow and projects to mature" (Gordon and Curlee, 2011, p. 177). According to Caldas *et al*. (2009), lessons learned is closely related to knowledge management, wherein the lessons learned are stored, studied, and transferred to other parties who need them (Resende *et al*., 2020). Von Zedtwitz (2003) called it a *post-project review* in which project lessons are studied and used for the benefit of future projects. In addition, Rowe and Sikes (2006) presented the lessons learned workflow, consisting of five steps, namely (1) identifying the comments and recommendations derived from a project, (2) documenting and sharing the findings, (3) analyzing and organizing the lessons learned to apply the results, (4) storing the lessons learned in a repository, and (5) retrieving them for use in subsequent projects.

This technique has been used in many construction management studies to increase the efficiency of construction projects. It can be used on all construction types to improve project quality, manage changes, and control expectations (Resende *et al*., 2020). For instance, Hansen *et al*. (2020) utilized the lessons learned from the GBK Aquatic Stadium project to develop a framework for minimizing construction change orders. Lessons learned from a high-speed rail project in Spain were used to promote owner-contractor collaboration in large-scale infrastructure projects (Hetemi *et al*., 2020). Shiferaw and Klakegg (2013) employed case studies to assess five aspects of project evaluation (efficiency, effectiveness, relevance, sustainability, and effects) of housing development projects in Ethiopia. The lessons learned from the Shanghai Expo Construction were used by Hu *et al*. (2012) to improve the design and management of multicriteria incentives in megaprojects. A benchmarking initiative focused on construction project performance measures was developed by Costa *et al*. (2006) using this technique.

Similarly, the lessons learned from bad practices have also been utilized to increase the knowledge and performance of construction projects. For example, the lessons learned from a cancelled urban transport project in Jakarta were provided by Hansen *et al.* (2018), highlighting the importance of the front-end planning (FEP) phase in infrastructure projects. Meanwhile, three examples of incidents that can impact construction projects and lead to extensive delays and cost overruns were discussed in detail by Baroncini and Wong (2022).

Methodology

This study applies a qualitative method with a case study approach. This approach can be effectively applied to research focusing on 'how' and 'why' questions about a contemporary set of events where the researcher has little or no control over the research object (Yin, 2003). In this research context, it is useful for exploring insights from actual construction projects regarding previously implemented site layout practices. Four project cases were successfully collected, whose brief profiles are presented in Table 1. They are actual construction projects with different backgrounds and geographies, aiming to provide researchers with more diverse data.

Table 1. Project profiles

Profile	Project A	Project B	Project C	Project D
Type	College hall building	Fast-food restaurant building	Car showroom building	Mixed-use apartment
Location (city)	Padang	Tangerang	Bandung	Serpong
Contract value (IDR)	3,678 billion	6,168 billion	12 billion	313,5 billion
Duration	4 months	4 months	6 months	5 years, 7 months
Site area (m^2)	1 320	2 742	37 868	146 900

Source: Author

The data collection techniques consisted of site observations, document examinations, and semi-structured interviews. Site observations were carried out by visiting construction sites directly in order to gain an understanding of actual field conditions. The data collected come in the form of photos and videos. Document examinations were carried out, requesting a site layout image of each project. Meanwhile, semi-structured interviews were conducted with four project managers (one for each project). This is a small sample size, considering the narrow research scope (focused on construction site layout practices), but it includes homogeneous interviewees (project managers) with a high level of expertise (Hansen, 2021). The list of interview questions is presented in Table 2. The analysis of the interview results involved the process of familiarizing oneself with and reviewing the data. Familiarization was carried out by reading the interview transcripts repeatedly, and data analysis used coding analysis techniques.

Table 2. Interview questions matrix

No	Questions	References
1	How do you plan the site layout?	Alavi and Rizk (2017), Binhomaid (2019), Dio and Frenky (2023), Liang *et al.* (2018), Ma *et al.* (2005), Mawdesley *et al.* (2002), Su *et al.* (2012), Tam *et al.* (2002), Xu *et al.* (2020), Zolfagarian and Irizarry (2014), Zouein *et al.* (2002)
2	What type of site layout planning do you apply?	Alavi and Rizk (2017), El-Rayes and Said (2009), Osman *et al.* (2003), Xu and Li (2012), Zolfagarian and Irizarry (2014), Zouein and Tommelein (1999)
3	What are some parameters that you consider in planning the site layout?	Dio and Frenky (2023), Hammad *et al.* (2016), Kumar and Bansal (2015), Mawdesley *et al.* (2002), Sadeghpour and Andayesh (2015), Sanad *et al.* (2008), Su *et al.* (2012), Tam *et al.* (2002), Zolfagarian and Irizarry (2014)
4	What are some best practices in planning and managing site layouts?	Binhomaid (2019), El-Rayes and Said (2009), Kumar and Bansal (2015), Mawdesley *et al.* (2002), Sadeghpour and Andayesh (2015), Sanad *et al.* (2008), Su *et al.* (2012), Tam *et al.* (2002), Zolfagarian and Irizarry (2014)
5	How can they influence the construction work performance?	Binhomaid (2019), Dio and Frenky (2023), El-Rayes and Said (2009), Mawdesley *et al.* (2002), Sadeghpour and Andayesh (2015), Sanad *et al.* (2008), Su *et al.* (2012), Tam *et al.* (2002), Zolfagarian and Irizarry (2014)
6	What are some bad practices or problems that you encounter in planning and managing site layouts?	Binhomaid (2019), El-Rayes and Said (2009), Kumar and Bansal (2015), Sadeghpour and Andayesh (2015), Su *et al.* (2012)
7	What are the consequences of these bad practices or problems?	Binhomaid (2019), El-Rayes and Said (2009), Kumar and Bansal (2015), Sadeghpour and Andayesh (2015), Su *et al.* (2012)
8	What can be done to overcome these problems?	El-Rayes and Said (2009), Kumar and Bansal (2015), Mawdesley *et al.* (2002), Sadeghpour and Andayesh (2015), Su *et al.* (2012)

Source: Author

Results and discussion

Case study 1 – College hall building

Project A was a small-scale project aimed at expanding an existing college building with the addition of a hall made of reinforced concrete and steel structures. The site area was very limited, and hence there was little space for storage. Due to the small scale and short duration of the project, the contractor prepared only one site layout plan (Figure 1). The results of the interviews describe the five steps of site layout planning: (1) planning access, mobilizing heavy equipment and materials; (2) determining the location of beds, storage, and workshop areas; (3) managing work hours and traffic because the project took place within a campus environment; (4) determining the position of the mobile crane; and (5) planning the location of the OSH signs on site.

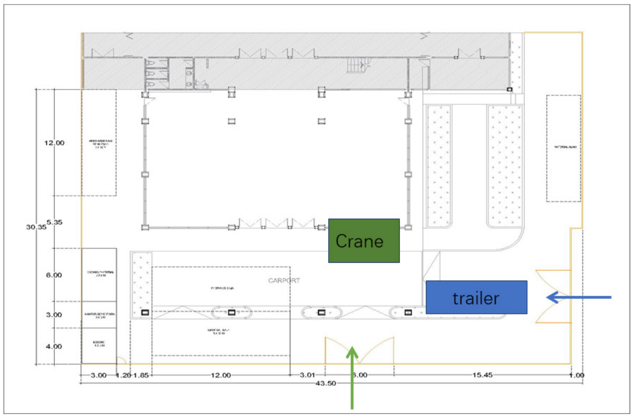


Figure 1. Site layout plan for Project A
Source: Author, based on project documentation

The contractor acknowledged that there were several challenges in developing the site layout plan due to space limitations. Structural steel is a major material that must be considered with regard to the delivery process and its placement on site. Even though the OSH signs were planned, site observation reported no signs on the project site or around the campus area, potentially endangering workers and passers-by in the surrounding area. Considering that the project is in a public area, traffic and OSH signs must be provided.

Case study 2 – Fast-food restaurant building

Project B was a fast-food restaurant project with a short duration and a simple design. As it would function as a fast-food restaurant, the large parking area allocated for the project aided the site layout planning. Therefore, the contractor applied static planning with only one layout (Figure 2). Workshop and storage areas could be placed in the parking area. Based on the interview, apart from paying attention to temporary storage facilities and workshops, the contractor had to prepare housing for the workers. However, due to the short duration of the work, they decided not to build temporary housing, but to rent houses around the project area.

Site observation showed a bad practice: some materials were eventually stored in the finished building when the storage material was removed because paving block work would be carried out in the parking area. This has the potential to cause damage to the completed architectural work. Therefore, it is recommended that at least two site layout plans be made, so that the contractor can consider and arrange any storage areas that are still needed. Inappropriate site layout planning can hinder project performance and quality, which was also recognized by the contractor.

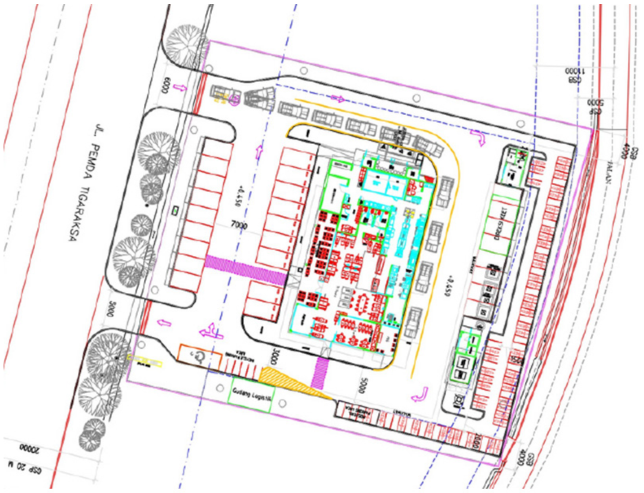


Figure 2. Site layout plan for Project B
Source: Author, based on project documentation

Case study 3 – Car showroom building

Project C was a 3-storey car showroom building project covering an area of 37 868 m^2. It had no site layout problems because there was enough space for material storage and workshops (Figure 3). Other temporary facilities available were a workers' mess, a site office, a meeting room, a prayer room, and parking lots. According on the results of the interview, the work implementation was carried out in stages, which were divided into five zones due to the large area of the project.

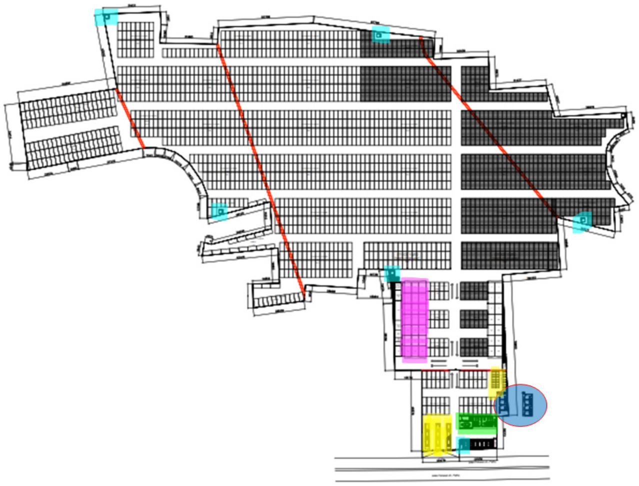


Figure 3. Site layout plan for Project C
Source: Author, based on project documentation

Other issues that had to be considered by the contractor were the type and conditions of the soil (as it was a paddy field) and the relocation of the warehouse area and the workers' dormitories. The type and conditions of the soil had to be improved because the project would be traversed by heavy equipment. The workers' dormitories and the warehouse were also moved from their initial location in the southwest area to the northeast area from mid-project until its completion. This aimed to make mobility more efficient due to the large project area and the division of work zones. Site observation showed no canteen for workers, so they workers had to look for food vendors, which were quite far away. A temporary canteen or catering distribution area could have been built for workers in the project area, so they would not have to waste time looking for food during break times.

Case study 4 – Mixed-use building

Project D was an apartment with a superblock concept and luxury offices. It is located close to various shopping centers, hospitals, schools, and universities. Due to the complexity of the work, site layout planning was carried out in stages consisting of six site layout plans, namely phase 1 for excavation work (Figure 4a), phase 2 for substructure work (Figure 4b), phase 3 for tower construction (Figure 4c), phase 4 for tower structure, architecture, and MEP works (Figure 4d), phase 5 for post-structure topping off (Figure 4e), and phase 6 for MEP and fitting out works (Figure 4f).

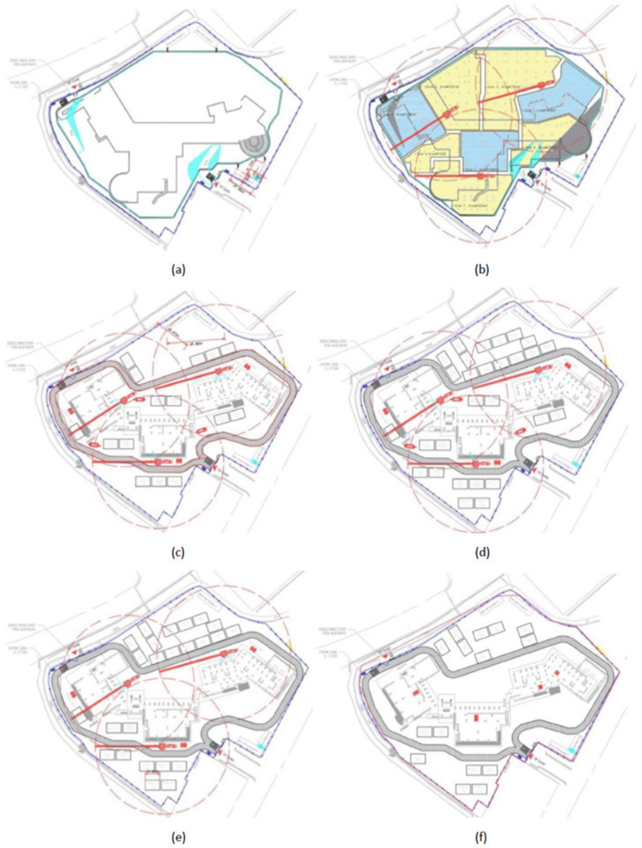

(a) (b) (c) (d) (e) (f)

Figure 4. Site layout plan for Project D
Source: Author, based on project documentation

As seen in these pictures, there was a gradual change in the site layout planning following the developments and changes in work over time. The first site layout plan displays site access, work area boundaries, and building plan spots. The second layout plan displays the number and location of the tower cranes used as well as the distribution of work zones for substructure works. The third layout plan maps traffic routes and the locations of storages and workshop areas. In the fourth layout plan, there are more storages for architecture and mechanical, electrical, and plumbing (MEP) works. The fifth layout plan is like the fourth one, albeit with reduced heavy equipment except for tower cranes. Meanwhile, in the sixth layout plan, the tower cranes and several workshops and storages have been dismantled.

The interview results showed that the limited space for the project was one of the main reasons for implementing the dynamic site layout planning method. Even though this planning method is efficient, the contractor needs to devote time and resources to repositioning warehouses, workshops, and other temporary facilities. Site observation reported a good housekeeping practice by the contractor. In general, the site layout planning facilitated the coordination and execution of work due to faster mobility and an organized environment.

Planning construction site layout: good and bad practices

Construction site layout planning is a complex task involving many interrelated factors and possible solutions to choose from (Abdel-Fattah, 2013; Kumar and Bansal, 2015). It is a step-by-step planning process that must consider the interactions in the flow of materials, equipment, and workers (Binhomaid, 2019) across a construction site, which is a limited resource (Sadeghpour and Andayesh, 2015). This is especially evident in Projects A and D, where limited space is the main issue. Similarly, Tsegay *et al*. (2023b) argues that, as space is limited, especially for construction projects in the city center, it must be optimally utilized. The narrower the site area, the more considerations that must be a concern for the site engineer and contractor. Therefore, it is crucial for the site engineer and contractor to collect all the necessary information about the project, the site, and the environment (including topography, soil conditions, and existing buildings and access). For example, in Project A the mobilization of heavy equipment is an important consideration since it will pass through existing buildings. In Project C, the type and conditions of the soil are a major consideration because the site used to be a paddy field, so soil strengthening and stabilization had to be performed. This becomes even more complex in Project D, which is in a busy city center with heavy traffic. Thus, that apart from internal site management, contractors also had to pay attention to public safety during site layout planning.

Ma *et al*. (2005) argue that construction site layout planning should be a dynamic planning activity that varies from one phase to the next. Static site layout planning identifies static locations for all facilities on site in a single site layout plan (El-Rayes and Said, 2009). Dynamic site layout planning requires updating the positions of all or part of the temporary facilities over the entire project duration (El-Rayes and Said, 2009). Unlike the former, it allows for the reuse of space, the relocation of temporary facilities, and changing space needs

(El-Rayes and Said, 2009). The benefit of implementing dynamic site layout planning is that it significantly lowers operating costs for the contractors (Petroutsatou *et al.*, 2021). On complex or large-scale projects, multiple site layout plans are required, so dynamic site layout planning is recommended. Out of the four project cases, Project D is the most complex, large-scale project, so the contractor created six site layout plans. Meanwhile, in Projects A and B, there is only one site layout plan, and, in Project C, there are only two site layout plans. It is recommended that at least three site layout plans be made, following the three main stages of a project: substructure, upper structure, and finishing works. On small scale projects, a minimum of two site layout plans is recommended, one for substructure works and another one for upper structure works.

There are several important tasks in site layout planning, namely identifying the facilities to be arranged, determining acceptable travel distances, and safety considerations (Alavi and Rizk, 2021). Facility identification is carried out to determine the facilities available at the project site and to determine their distribution locations. Identification is also performed to determine the type of facility being reviewed. The type of facility to be moved is taken into consideration in site layout optimization. In small and simple projects (*i.e.*, Project A), there may not be many facilities that need to be arranged, but, in large and complex ones (*e.g.*, Project D) there are many facilities that must be properly managed for the smooth progress of work on site.

In general, facilities in construction projects can be grouped into two types: fixed and temporary (Fadhlan *et al.*, 2020). Fixed or permanent facilities are those that are part of the project and its all-associated elements, such as fixtures, drainage, retaining walls, *etc.* Meanwhile, temporary facilities are those needed to support safe and efficient construction operations, including guard posts, warehouses, parking lots, wood fabrication, worker toilets, canteens, scaffolding stockyards, sand and aggregate stockyards, and site offices. This study also found a correlation between site space and facility placement, where the wider a construction site, the higher the tendency towards static layout planning, as seen in Project C. Meanwhile, the more congested a construction site, the higher the tendency towards dynamic layout planning as seen in Project D.

Traveling distance means the distance reached during the movement of materials, workers, and equipment from one facility to another (Abdel-Fattah, 2013; Adhika and Nurcahyo, 2017; Fadhlan *et al.*, 2020). It expresses the preference for a facility or an object to be located within a certain distance of another facility or object on site (Sadeghpour *et al.*, 2006). Zhou (2006) argued that rectilinear or Euclidean distance can be used to measure a distance on site. On projects with a large site area, the placement of facilities and objects becomes increasingly dispersed, which makes reducing travel distance and trip frequency the main goal (Fadhlan *et al.*, 2020). For instance, Project C implemented a work-zoning strategy to achieve this goal.

Safety and security considerations are another important criterion in site layout planning (Abdel-Fattah, 2013; Adhika and Nurcahyo, 2017; Kumar and Bansal, 2015; Sanad *et al.*, 2008; Tsegay *et al.*, 2023a). Irregularities in the arrangement of site layouts can affect site safety and security for workers and the public. The risk level that can arise is not the same from one facility to another. Therefore, the contractor must make efforts to identify the hazard zones and calculate the safety index (Adhika and Nurcahyo, 2017; Kumar and Bansal, 2015). The risk level is divided into several zones, obtaining one value for each one (Adhika and Nurcahyo, 2017). The safety index is calculated through the relationship between the value of the safety level between facilities and the frequency of movements between facilities (Adhika and Nurcahyo, 2017). At the project site, workers can pass through several hazard zones at the same time when moving from one facility to another, so the safety level between facilities is calculated using the distance proportion. Even though safety considerations are an important issue in the four project cases, the contractors did not calculate the safety index to optimize the site layout.

Finally, proper site layout planning provides benefits for contractors to manage work effectively and efficiently. A well-planned site layout will increase productivity and safety as well as minimize the travel time spent on materials handling (Kumar and Bansal, 2015; Sadeghpour *et al.*, 2006). This was agreed by all interviewees. Project C interviewee added that, with a proper site layout plan, project access and the location and number of temporary facilities become clear, making site coordination more efficient. On the other hand, these plans must be updated regularly according to arising needs. This is essential for site engineers and contractors to make better decisions faster, *e.g.*, decisions related to work changes (as in Projects B and D) and the need to divide working zones (as in Project C).

Research implications

Construction site layout planning is one of the most crucial tasks that affect project success (Fadhlan *et al.*, 2020; Su, 2013). It can influence project cost, duration, productivity, safety, quality, and conflicts (Dio and Frenky, 2022; Kumar and Bansal, 2015; Sadeghpour *et al.*, 2006; Tsegay *et al.*, 2023a). However, construction site layout planning practices, especially in Indonesia, are still rarely studied. Thus, this research highlights the importance of construction site layout planning as a crucial force in construction project implementation. Research on construction site layout planning practices is carried out using a case study approach to obtain lessons learned empirically. By studying previous experiences, project teams can create a suitable environment for project success (Hansen *et al.*, 2018).

The demand to make better organization and project decisions has become a challenge in the construction business, where lessons learned techniques can be applied. Increasing the use of project experiences to improve organizational performance is a practice widely adopted in previous studies. Anbari *et al*. (2008) emphasized the regular collection of lessons learned and their meaningful utilization in future projects as critical aspects of organizational competitiveness and project success.

On the other hand, Williams (2007) and Rezania and Lingham (2009) noted that any learning from a project will be largely dissipated, considering that projects are temporary organizations by nature. The lessons learned process requires specific attention so that it can be properly performed. Therefore, this paper contributes by accumulating lessons learned from construction site layout planning related to actual project cases, which can be utilized by other parties for the benefit of future projects.

In addition, the complexity of construction sites requires site engineers and contractors to be able to properly perform site layout planning (Kamil and Adnan, 2020). Construction site layout planning is a decision-making problem related to the allocation of facilities which has a significant impact in reducing the construction cost (Hawarneh *et al*., 2021). Making such decisions is a challenging task due to the interdependence of variables and affecting factors (Alavi and Rizk, 2021). In actual practice, site layout planning in construction projects still relies heavily on the intuition and experience of site engineers, who make decisions considering various project parameters and characteristics (Kim *et al*., 2021). As a decision-making problem, it may need both expert judgment and engineering analysis to aid in decision-making and guide the planning process. This study offers technical advice and examples of site layout planning practices in various construction projects, which can be used as lessons learned.

Construction site layout planning practice belongs to the disciplines of construction management and civil engineering. Thus, research on best practices for site layout planning is very important for site engineers and contractors to be able to manage their sites appropriately. This study offers empirical knowledge based on lessons learned from four actual project cases, aiming to improve construction site layout planning practices. As a proper site layout planning influences project performance, this research contributes to the study of engineering design by assessing the site layout planning procedures and processes of real-world cases in order to improve the efficiency and effectiveness of the decision-making process.

While this research examines construction site layout practices in the Indonesian construction sector, the findings may be of significance to other countries as well. As construction site layout planning is a complex task that requires human interpretation of site characteristics (Kim *et al*., 2021), the findings of this study might allow others to understand the risks of improper site layout planning, as well as to assist in making better decisions. The lessons learned from these project cases enabled the development of appropriate strategies that project teams require in order to increase their response capabilities and hence be able to pro-actively make engineering decisions regarding proper site layout planning. Thus, the findings of this study can be used as a reference for more efficient and effective project site performance, especially with advances in construction technology and research on site layout planning, which are increasingly sophisticated and diverse (Pradana, 2021), making it easier for site engineers and contractors to carry out their tasks.

Conclusions

Site layout planning is essential for all construction projects because it has a significant impact on productivity, safety, and the smooth execution of construction activities. This study focuses on site layout planning practices as observed in four actual construction project cases in Indonesia. These four projects have different site characteristics that affect site layout planning by contractors. This study found that many factors cause site layout planning to be complex, including the size and condition of the site area, the type and scale of the project, the number and types of facilities and resources to be arranged, and the site layout planning type and technique applied.

The main contribution of this study lies in establishing some lessons learned about site layout planning practices. The findings are relevant to all parties involved, and the relationship between the research object, the research context, and the parties involved is described in a coherent and in-depth manner. Thus, this study can become a reference or best practice that can be used to enable better decision-making and optimal outcomes related to construction site layout planning.

On the other hand, this research has some limitations. As a context-specific case study, it requires a small sample size to be effective. Therefore, the findings, corresponding to Indonesian construction projects, may be difficult to generalize to other research settings. In addition, lessons learned collected from actual project practices must continue to be consistently reviewed to become effective knowledge. Therefore, we recommend capturing more real-world cases regarding construction site layout planning practices, especially from projects with different geographical backgrounds and characteristics, in order to obtain a knowledge database with depth and breadth of information that is available for site engineers and contractors who will manage future projects.

Acknowledgements

The author would like to thank Arif Prayoga, Riri Vermatasari, Naila Ramadhani, Ghina Yovelia, Intan Ermariza, Siti Helena Erizti, Fitri Amaliny, Dika Syabela Yuhastri, Al Fajra, Natalia Christina, Austin Steven Febliko, Ronald Dwi Kurniawan, William, Steven Setiawan, Ivana Marcella Sumar, and Derrick Libertio for their help in the data collection process.

CRediT author statement

The author conducted the entirety of this research, from conceptualization to writing.

Conflicts of interest

The author reports no potential conflict of interest.

References

Abdel-Fattah, A. (2013). *Dynamic site layout planning model* [Doctoral thesis, University of Calgary]. https://prism.ucalgary.ca/bitstream/handle/11023/545/ucalgary_2013_abdel-fattah_aly.pdf?sequence=2&isAllowed=y

Adhika, H. D., and Nurcahyo, C. B. (2017). Optimasi site layout menggunakan multi-objectives function pada proyek pembangunan transmart rungkut surabaya. *Jurnal Teknik ITS, 6*(1), C22-C26. https://ejurnal.its.ac.id/index.php/teknik/article/viewFile/21535/3510

Alanjari, P., Razavialavi, S., and AbouRizk, S. (2014). A simulation-based approach for material yard laydown planning. *Automation in Construction*, 40, 1-8. https://doi.org/10.1016/j.autcon.2013.12.010

Alavi, S. R. R., and Rizk, S. A. (2017). Site layout and construction plan optimization using an integrated genetic algorithm simulation framework. *Journal of Computing in Civil Engineering, 31*(4), 04017011. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000653

Alavi, S. R. R., and Rizk, S. A. (2021). Construction site layout planning using a simulation-based decision support tool. *Logistics*, 5(65), 1-25. https://doi.org/10.3390/logistics5040065

Anbari, F., Carayannis, E., and Voetsch, R. (2008). Post-project reviews as a key project management competence. *Technovation, 28*(10), 633-643. https://doi.org/10.1016/j.technovation.2007.12.001

Andayesh, M., and Sadeghpour, F. (2013). Dynamic site layout planning through minimization of total potential energy. *Automation in Construction*, *31*, 92-102. https://doi.org/10.1016/j.autcon.2012.11.039

Baroncini, M., and Wong, C. (2022). *Lessons learnt in construction projects*. Swiss Re Corporate Solutions Ltd. https://corporatesolutions.swissre.com/dam/jcr:f330d163-b624-4b29-9dbd-f7632be30dba/lessons-learnt-in-construction-projects.pdf

Binhomaid, O. S. (2019). *Construction*. [Doctoral thesis, University of Waterloo] https://core.ac.uk/download/pdf/185636963.pdf

Caldas, C. H., Gibson, G. E., Weerasooriya, R., and Yohe, A. M. (2009). Identification of effective management practices and technologies for lessons learned programs in the construction industry. *Journal of Construction Engineering and Management, 135*(6), 531-539. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000011

Cheng, M.Y., and O'Connor, J.T. (1996). ArcSite: Enhanced GIS for construction site layout. *Journal of Construction Engineering and Management, 122*(4), 329-336. https://doi.org/10.1061/(ASCE)0733-9364(1996)122:4(329)

Costa, D. B., Formoso, C. T., Kagioglou, M., Alarcon, L. F., and Caldas, C. H. (2006). Benchmarking initiatives in the construction industry: lessons learned and improvement opportunities. *Journal of Management in Engineering, 22*(4), 158-167. https://doi.org/10.1061/(ASCE)0742-597X(2006)22:4(158)

Dio, A. H., and Frenky, E. P. (2022). *Evaluasi penataan pelaksanaan pekerjaan pada sebuah site layout proyek konstruksi dari aspek waktu dan biaya dengan building information modeling 4d* (*studi kasus proyek X dan proyek Y*) [Bachelor thesis, Universitas Katolik Soegijapranata]. http://repository.unika.ac.id/29392/1/21.B1.0060-Evi%20Parnius%20Frenky-COVER_a.pdf

El-Rayes, K., and Said, H. (2009). Dynamic site layout planning using approximate dynamic programming. *Journal of Computing in Civil Engineering, 23*(2), 119-127. https://doi.org/10.1061/(ASCE)0887-3801(2009)23:2(119)

Fadhlan, D., Wiguna, I. P. A., and Rohman, M. A. (2020). Optimasi penataan site layout pada proyek grand dharmahusada surabaya dengan metode logika fuzzy AHP. *Jurnal Teknik ITS*, 9(2), D250-D256. https://doi.org/10.12962/j23373539.v9i2.58373

Gordon, R. L, and Curlee, W. (2011). *The virtual project management office: Best practices, proven methods*. Berrett-Koehler Publishers.

Hammad, A. W. A., Akbamezhad, A., Rey, D., and Waller, S. T. (2016). A computational method for estimating travel frequencies in site layout planning. *Journal of Construction Engineering and Management, 142*(5), 04015102. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001086

Hansen, S. (2021). Characterizing interview-based studies in construction management research: Analysis of empirical literature evidences. *International Conference on Innovations in Social Sciences Education and Engineering* (*ICOISSEE*), *1*(1), 1-12. https://proceedings.conference.unpas.ac.id/index.php/icoissee/article/view/724

Hansen, S., Rostiyanti, S.F., and Rif'at, A. (2020). Causes, effects, and mitigations framework of contract change orders: Lessons learned from GBK aquatic stadium project. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 12*(1), 05019008. https://doi.org/10.1061/(ASCE)LA.1943-4170.0000341

Hansen, S., Too, E., and Le, T. (2018). Lessons learned from a cancelled urban transport project in a developing country: The importance of the front-end planning phase. *International Journal of Technology, 9*(5), 898-909. https://doi.org/10.14716/ijtech.v9i5.1559

Harrison, W. (2003). *A software engineering lessons learned repository* [Conference presentation]. 27th Annual NASA Goddard/IEEE Software Engineering Workshop, Los Alamitos, California.

Hawarneh, A. A., Bendak, S., and Ghanim, F. (2021). Construction site layout planning problem: Past, present and future. *Expert Systems with Application, 168,* 114247. https://doi.org/10.1016/j.eswa.2020.114247

Hegazy, T. M., and Elbeltagi, E. (1999). EvoSite: Evolution-based model for site layout planning. *Journal of Computing in*

Civil Engineering, 13(3), 198-206. https://doi.org/10.1061/(ASCE)0887-3801(1999)13:3(198)

Hetemi, E., Gemunden, H. G., and Mere, J. O. (2020). Embeddedness and actors' behaviors in large-scale project life cycle: Lessons learned from a high-speed rail project in Spain. *Journal of Management in Engineering, 36*(6), 05020014. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000849

Hu, Y., Chan, A. P. C., Le, Y., Jiang, W. P., Xie, L. L., and Hon, C. H. K. (2012). Improving megasite management performance through incentives: Lessons learned from the Shanghai Expo construction. *Journal of Management in Engineering, 28*(3), 330-337. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000102

Jiang, W., Zhou, Y., Ding, L., Zhou, C., and Ning, X. (2020). UAV-based 3D reconstruction for hoist site mapping and layout planning in petrochemical construction. *Automation in Construction*, *113*, 103137. https://doi.org/10.1016/j.autcon.2020.103137

Kamil, A. K. A., and Adnan, A. Z. A. (2020). Planning of construction site layout: A review. *Journal of Science and Engineering Application*, *4*(1), 43-51. https://journals.iunajaf.edu.iq/index.php/JSEA/article/view/1

Khalafallah, A., and El-Rayes, K. (2006). Optimizing airport construction site layouts to minimize wildlife hazards. *Journal of Management in Engineering, 22*(4), 176-185. https://doi.org/10.1061/(ASCE)0742-597X(2006)22:4(176)

Kim, M., Ryu, H. G., and Kim, T. W. (2021). A typology model of temporary facility constraints for automated construction site layout planning. *Applied Sciences*, *11*(3), 1027. https://doi.org/10.3390/app11031027

Kumar, S., and Bansal, V. K. (2015). Framework for safe site layout planning in hilly regions. *European Journal of Advances in Engineering and Technology, 2*(4), 14-19. https://ejaet.com/PDF/2-4/EJAET-2-4-14-19.pdf

Liang, C. J., Kamat, V. R., and Menassa, C. M. (2018). *Real-time construction site layout and equipment monitoring* [Conference presentation]. Construction Research Congress, Reston, VA, USA.

Ma, Z., Shen, Q., and Zhang, J. (2005). Application of 4D for dynamic site layout and management of construction projects. *Automation in Construction, 14*(3), 369-381. https://doi.org/10.1016/j.autcon.2004.08.011

Marx, A., and Konig, M. (2011). *Preparation of constraints for construction simulation* [Conference presentation]. International Workshop on Computing in Civil Engineering, Reston, VA, USA.

Mawdesley, M. J., Al-Jibouri, S. H., and Yang, H. (2002). Genetic algorithms for construction site layout in project planning. *Journal of Construction Engineering and Management, 128*(5), 418-426. https://doi.org/10.1061/(ASCE)0733-9364(2002)128:5(418)

Ning, X., Qi, J., and Wu, C. (2018). A quantitative safety risk assessment model for construction site layout planning. *Safety Science*, *104*, 246-259. https://doi.org/10.1016/j.ssci.2018.01.016

Osman, H., Georgy, M., and Ibrahim, M. (2003). *An automated system for dynamic construction site layout planning* [Conference presentation]. 10th International Colloquium in Structural and Geotechnical Engineering (ICSGE 2003), Cairo, Egypt.

Petroutsatou, K., Apostolidis, N., Zarkada, A., and Ntokou, A. (2021). Dynamic planning of construction site for linear projects. *Infrastructures*, 6(21), 1-22. https://doi.org/10.3390/infrastructures6020021

Pheng, L. S., and Hui, M. S. (1999). The application of JIT philosophy to construction: A case study in site layout. *Construction Management & Economics, 17*(5), 657-668. https://doi.org/10.1080/014461999371268

Pradana, H. Y. (2021). *Site layout optimization using multi-objectives function on dormitory building project* (*studi kasus: Proyek Pembangunan Gedung Asrama MTSN 1 Kebumen*) [Bachelor thesis, Universitas Islam Indonesia]. https://dspace.uii.ac.id/handle/123456789/35883

Resende, C. B., Volk, M. J., and Shane, J. S. (2020). *Post-project evaluation and lessons learned* [Conference presentation]. Construction Research Congress 2020, Reston, VA, USA. https://ascelibrary.org/doi/10.1061/9780784482889.046

Rezania, D., and Lingham, T. (2009). Towards a method to disseminate knowledge from the post project review. *Knowledge Management Research & Practice, 7*(2), 172-177. https://link.springer.com/article/10.1057/kmrp.2009.9

Rowe, S. F, and Sikes, S. (2006). *Lessons learned: Taking it to the next level* [Conference presentation]. PMI Global Congress 2006, Seattle, Washington. https://www.pmi.org/learning/library/lessons-learned-next-level-communicating-7991

Sadeghpour, F., and Andayesh, M. (2015). The constructs of site layout modeling: An overview. *Canadian Journal of Civil Engineering, 42*(3), 199-212. https://doi.org/10.1139/cjce-2014-030

Sadeghpour, F., Moselhi, O., and Alkass, S. T. (2006). Computer-aided site layout planning. *Journal of Construction Engineering and Management, 132*(2), 143-151. https://doi.org/10.1061/(ASCE)0733-9364(2006)132:2(143)

Said, H., and El-Rayes, K. A. (2011). Optimizing material procurement and storage on construction sites. *Journal of Construction Engineering and Management*, *137*(6), 421-431. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000307

Sanad, H. M., Ammar, M. A., and Ibrahim, M. E. (2008). Optimal construction site layout considering safety and environmental aspects. *Journal of Construction Engineering and Management, 134*(7), 536-544. https://doi.org/10.1061/(ASCE)0733-9364(2008)134:7(536)

Sezer, A. A., and Fredriksson, A. (2021). Environmental impact of construction transport and the effects of building certification schemes. *Resources, Conservation and Recycling*, *172*, 105688. https://doi.org/10.1016/j.resconrec.2021.105688

Shiferaw, A. T., and Klakegg, O. J. (2013). Project evaluation: Accomplishments, shortfalls, and lessons learned in housing development projects in Ethiopia. *Journal of Management in Engineering, 29*(3), 289-301. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000138

Sjøbakk, B., and Skjelstad, L. (2015). Proposing a standard template for construction site layout: A case study of a Norwegian contractor. In S. Umeda, M. Nakano, H. Mizuyama, N. Hibino, D. Kiritsis, and G. von Cieminski (Eds.), *Advances in Production Management Systems: Innovative Production*

Management Towards Sustainable Growth (APMS 2015) (pp, 316-323). Springer. https://doi.org/10.1007/978-3-319-22756-6_39

Skejlbred, S., Fossheim, M. E., and Drevland, F. (2015, July 13-22). *Comparing site organization and logistics in the construction industry and the oil industry: A case study* [Conference presentation]. 23rd Annual Conference of the International Group for Lean Construction, Perth, Australia. https://iglc.net/Papers/Details/1178

Su, X. (2013). *Spatial temporal information model for construction planning* [Doctoral thesis, Purdue University]. https://docs.lib.purdue.edu/dissertations/AAI3605150/

Su, X., Andoh, A.r., Cai, H., Pan, J., Kandil, A., and Said, H. M. (2012). GIS-based dynamic construction site material layout evaluation for building renovation projects. *Automation in Construction, 27*, 40-49. https://doi.org/10.1016/j.autcon.2012.04.007

Tam, C. M., and Tong, T. K. L. (2003). GA-ANN model for optimizing the locations of tower crane and supply points for high-rise public housing construction. *Construction Management and Economics, 21*(3), 257-266. https://doi.org/10.1080/0144619032000049665

Tam, C. M., Tong, T. K. L., Leung, A. W. T., and Chiu, G. W. C. (2002). Site layout planning using nonstructural fuzzy decision support system. *Journal of Construction Engineering and Management, 128*(3), 220-231. https://doi.org/10.1061/(ASCE)0733-9364(2002)128:3(220)

Tsegay, F. G., Mwanaumo, E., and Mwiya, B. (2023a). Mathematical optimization methods for inner-city construction site layout planning: A systematic review. *Asian Journal of Civil Engineering, 24*, 3781-3795. https://doi.org/10.1007/s42107-023-00713-2

Tsegay, F. G., Mwanaumo, E., and Mwiya, B. (2023b). Construction site layout planning practices in inner-city building projects: Space requirement variables, classification and relationship. *Urban, Planning and Transport Research, 11*(*1*), 2190793. https://doi.org/10.1080/21650020.2023.2190793

Vashishtha, A. U. (2022). A Framework to Optimize Construction Site Layout Planning. *International Research Journal of Engineering and Technology*, 9(3), 1911-1919. https://www.irjet.net/archives/V9/i3/IRJET-V9I3347.pdf

von Zedtwitz, M. (2003). Post-project reviews in R&D. *Research-Technology Management, 46*(5), 43-49.

Weber, R., Aha, D. W., Munoz-Avila, H., and Breslow, L. A. (2000). An intelligent lessons learned process. In Z. W. Raś and S. Ohsuga (Eds.) *Foundations of Intelligent Systems* (pp. 358–367). Springer. https://link.springer.com/chapter/10.1007/3-540-39963-1_38

Williams, T. (2007). *Post-project reviews to gain effective lessons learned*. PMI.

Xu, J., and Li, Z. (2012). Multi-objective dynamic construction site layout planning in fuzzy random environment. *Automation in Construction, 27*, 155-169. https://doi.org/10.1016/j.autcon.2012.05.017

Xu, S., Fu, D., Xie, Y., Hou, L., and Bu, S. (2020). *Integrating BIM and VR for highway construction site layout planning* [Conference presentation]. 20th COTA International Conference of Transportation Professional, Reston, VA, USA. https://doi.org/10.1061/9780784482933.093

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications.

Zhang, H., and Yu, L. (2021). Site layout planning for prefabricated components subject to dynamic and interactive constraints. *Automation in Construction*, 126, 103693. https://doi.org/10.1016/j.autcon.2021.103693

Zhou, F. (2006). *An integrated framework for tunnel shaft construction and site layout optimization* [Master's thesis, University of Alberta]. https://era.library.ualberta.ca/items/ea82c2a6-b006-4618-b911-3a3c5587a541

Zolfagarian, S., and Irizarry, J. (2014). *Current trends in construction site layout planning* [Conference presentation]. Construction Research Congress 2014, Reston, VA, USA. https://doi.org/10.1061/9780784413517.176

Zouein, P., Harmanani, H., and Hajar, A. (2002). genetic algorithm for solving site layout problem with unequal-size and constrained facilities. *Journal of Computing in Civil Engineering, 16*(2), 143-151. https://doi.org/10.1061/(ASCE)0887-3801(2002)16:2(143)

Zouein, P., and Tommelein, I. (1999). Dynamic layout planning using a hybrid incremental solution method. *Journal of Construction Engineering and Management, 125*(6), 400-408. https://doi.org/10.1061/(ASCE)0733-9364(1999)125:6(400)

Research Article/CivilEngineering
http://doi.org/10.15446/ing.investig.105622

Studying the Repeatability of Measurements Obtained via Network Real-Time Kinematic Positioning at Different Times of the Day

Estudio de la repetibilidad de mediciones obtenidas mediante posicionamiento cinemático de red en tiempo real para diferentes momentos del día

Kutalmış Gümüş[1]

ABSTRACT

The network real-time kinematic (NRTK) positioning technique is currently used in numerous applications. The aim of this study was to better understand the process of obtaining accurate positions by statistically evaluating the significance of differences between repeated measurements for a single point at different times of the day (morning, noon, and evening) using the Virtual Reference Station (VRS), Flächen Korrektur Parameter (FKP), and Master Auxiliary Concept (MAC) correction methods. An analysis of variance (ANOVA) was used to this effect. Further analysis was carried out to determine the accuracy and precision of the coordinate differences obtained via a static GNSS (global navigation satellite system) and by averaging the repeated measurements. It was determined that the accuracy and precision of the vertical component of the coordinates were lower than that of the horizontal component. The FKP correction method yielded the best results. It was observed that the accuracy and precision of the measurements taken at noon were the lowest. The ANOVA showed that the differences between repeated measurements were statistically significant and that there were outlier measurements. The results of this study are important for NRTK users to be able to statistically evaluate different measurement configurations and obtain positions with the desired accuracy and precision.

Keywords: ANOVA, FKP, MAC, NRTK, static GNSS, VRS

RESUMEN

La técnica de posicionamiento cinemático de red en tiempo real (NRTK) se utiliza actualmente en numerosas aplicaciones. El objetivo de este estudio fue comprender mejor el proceso de obtención de posiciones precisas evaluando estadísticamente la importancia de las diferencias entre mediciones repetidas para un solo punto en diferentes momentos del día (mañana, mediodía y tarde), utilizando los métodos de corrección de Estación Virtual de Referencia (VRS), Parámetro de Corrección de Área (FKP) y Concepto Maestro Auxiliar (MAC). Para ello, se utilizó un análisis de varianza (ANOVA). Se realizaron análisis adicionales para determinar la exactitud y la precisión de las diferencias de coordenadas obtenidas mediante un GNSS (sistema global de navegación por satélite) estático y promediando las mediciones repetidas. Se determinó que la exactitud y la precisión del componente vertical de las coordenadas eran menores que las del componente horizontal. El método de corrección FKP proporcionó los mejores resultados. Se observó que la exactitud y la precisión de las mediciones realizadas al mediodía eran las más bajas. El ANOVA mostró que las diferencias entre mediciones repetidas eran estadísticamente significativas y que había mediciones atípicas. Los resultados de este estudio son importantes para que los usuarios de NRTK puedan evaluar estadísticamente diferentes configuraciones de medición y obtener posiciones con la exactitud y precisión deseadas.

Palabras clave: ANOVA, FKP, GNSS estático, MAC, NRTK, VRS

Received: November 03th, 2022
Accepted: July 27th, 2023

Introduction

GNSS use satellite systems to determine absolute positions via GNSS receivers. The position of a receiver can be identified by observing at least four satellites. Positioning technologies such as classical RTK, network RTK, and PPP (precise point positioning) have been developed with the advances made in GNSS technology, and they offer more accurate, efficient, and comprehensive solutions in the field of precise positioning to meet the needs of different applications. Classical RTK has achieved high accuracy by utilizing GNSS signals in real time. Network RTK is a development that uses multiple base stations and provides wide coverage. PPP is a method that emerged from a deeper use of GNSS analysis. In real-time kinematic GNSS applications, important progress has been made in modeling error sources affecting GNSS

---

[1] PhD in Geomatics Engineering, Yıldız Technical University, Turkey, Affiliation: Associate Professor, Niğde Ömer Halisdemir University, Engineering Faculty, Geomatics Engineering Department, Turkey. , Email: kgumus@ohu.edu.tr

systems. The coordinates of new positions are momentarily calculated using classical RTK applications. Due to the increase in distance-dependent errors, the distance between the reference receiver and the mobile receiver is set to be between 10 to 15 km (Gao *et al*., 1997; Han, 1997; Colombo *et al*., 1999). To eliminate limitations based on distance and increase the positioning accuracy, GNSS reference stations have been put into operation, generating continuous static and real-time observations. GNSS reference stations, as well as data processing stations and NRTK designs, have made significant contributions to GNSS users in terms of cost and speed. In the NRTK approach, corrections based on information obtained from multiple reference stations are calculated by using correction techniques like VRS, FKP, and MAC, which are then sent to the mobile receiver (Raquet, 1998; Chen, 2000; Dai *et al*., 2001; Cannon *et al*., 2001; Wanninger, 2002). The positions of GNSS receivers are determined with centimeter accuracy via the relative positioning technique, using correction techniques and observations from reference stations (Wanninger, 2002; Landau, 2002).

PPP relies on a single receiver and performs positioning based on continuous GNSS data and analyses, including precise orbit and clock information. In Ge *et al*. (2008), the importance of integer uncertainty resolution for improving the eastern component was emphasized within the framework of the PPP method. This was confirmed via analysis using IGS data, and it was concluded that PPP can be improved in terms of accuracy and repeatability.

Li *et al*. (2015) presented a model and service system for real-time precise orbit determination, time estimation, and positioning using observations of the four satellite systems (GPS, GLONASS, BeiDou, and Galileo). Statistical analyses showed the good accuracy of these satellite systems. It was emphasized that multi-GNSS PPP can provide continuous position estimates up to a 40° elevation angle limit with a high availability rate.

In Dardanell *et al*. (2021), a comparison between GNSS position determination methods (PPP, NRTK, and static mode) was performed. NRTK and static solutions were obtained by analyzing a local GNSS CORS network. The PPP solution was calculated using a different software. A statistical approach was used to check whether the coordinate differences belonged to a normal distribution. The results showed that the static mode and the NRTK pair provide the best fit.

Gandolfi *et al*. (2017) evaluated the performance of the PPP method in static positioning. Accuracies ranging from 5 mm to 10 cm were obtained through analyses using 24, 12, 6, 3, and 1 h data obtained from European GNSS stations. It was determined that accuracy decreases as the data duration decreases.

NRTK applications have recently become important in real-time positioning and have been developed in parallel with modern technology. Rizos *et al*. (2002) evaluated the theoretical and practical problems of NRTK applications in terms of geodesic points. Landau *et al*. (2002) explained the general functioning of the VRS method, and Hu *et al*. (2003) studied its performance. In a different study conducted by Hu *et al*. (2003), several tests were carried out at different locations in Singapore, with the purpose of determining the start time and positioning accuracy of VRS RTK positioning. Wielgosz *et al*. (2003) revealed that, regarding accuracy, the average distance between reference stations should be between 30 and 40 km. Herbert *et al*. (2003) discussed the advantages and disadvantages of the VRS method when compared to the FKP method. In a work conducted by Grejner-Brzezinska *et al*. (2005), the accuracy of NRTK correction techniques was evaluated using different strategies, obtaining a 2 cm accuracy in the horizontal plane and a 4 cm accuracy in the vertical plane. Moreover, Kun and Yonh (2006) studied the applications and developments of the continuously operating reference station (CORS).

In a study by Brown (2006), different aspects of the MAC, VRS, and FKP methods were compared. Eren *et al*. (2009) thoroughly explained the TUSAGA-Aktif (CORS-TR) project. Pirti *et al*. (2009) presented a cadastral application for the use of RTK-GPS in boundary measurements. Janssen (2009) compared the VRS and MAC techniques. In another study carried out by Janssen *et al*. (2011), a methodology was presented for the measurement and control of RTK-GPS-based infrastructure observations in cadastral studies.

Garrido *et al*. (2014) studied the accuracy and precision of the results obtained via the VRS and MAC correction methods within an NRTK approach for two active GNSS networks in Spain. The positioning and evaluation of wide-area GNSS networks were carried out by Tarig in 2012. Martin (2012) tested the performance of NRTK in Ireland, and Bisnath *et al*. (2013) determined the accuracy of horizontal NRTK applications to be in the order of cm. Cina *et al*. (2015) discussed the applications, benefits, limitations, and problems of NRTK and tested its performance. Dabove *et al*. (2014) emphasized that there were no limitations for network products, and that, when the distance between stations increases, the results obtained from postprocessing are limited. Kudrys *et al*. (2011) conducted a comparative analysis of the positions of the points determined using current network/surface corrections in the VRS, MAC, and FKP formats in the NAWGEO service of the ASG-EUPOS system.

In Turkey, several studies on NRTK GNSS applications have been carried out in recent years. In Öğütcü and Kalaycı (2018), for example, measurements were carried out by taking the baseline lengths of approximately 5-20-40-50 km according to the nearest CORS station. This study aimed to establish the empirical accuracy and precision models of the FKP, MAC, and VRS NRTK techniques as a function of baseline distance and occupation time. In Atinc and Hosbas (2019), measurements from disabled and unobstructed environments (where sky vision is limited) were compared in terms of accuracy. The performance and use of the CORS were examined based on the results obtained near or below the wooded area. In Gökdaş and Özlüdemir (2020), two CORS networks in Turkey were employed to study the effect of the FKP and VRS correction methods on accuracy and precision at different base lengths in NRTK positioning. An experimental mathematical model was developed to express this effect.

In general, some changes are observed in NRTK applications, in the form of an increase or decrease in the 3D point position values obtained at different times during the day. These changes are due to the different directions and the intensity of various factors on the observation values obtained. Atmospheric conditions, environmental factors, the number of visible satellites and their different configurations, satellite geometry, the NRTK correction method used, signal interruptions, reflective surfaces, design features and the quality of GNSS receivers and antennas, different measurement configurations, the user's location, and the distance from fixed stations in the network, among others, are the different factors that affect NRTK GNSS accuracy. The effects of some of these factors on positioning accuracy are greatly reduced or eliminated by modeling. Studies on solutions to problems encountered during applications of NRTK are continuously being published. The most suitable measurement configuration should be selected to obtain accurate and sensitive results. The use of statistical methods can be very useful for interpreting results and selecting measurement times.

In this study, repeated three-dimensional position data were obtained from a NRTK GNSS application at different times of the day, at the same point, and using the same satellite system and measurement configurations, the same GNSS receiver, and antenna. Atmospheric conditions, environmental factors, the number of visible satellites, and satellite geometry changed while measurements were being made. The effects of changes in these factors on positioning accuracy during the day were statistically tested. Since our work studied the point position accuracy of the repeated measurements obtained according to different correction methods at different time intervals of the day, the factors that most affected this relationship and that actually changed over time were taken into account as a whole. These differences were determined statistically via an analysis of variance (ANOVA) and multiple comparisons.

This study aims to contribute to a better understanding of accurate positioning by statistically evaluating the differences between repeated measurements made at different measurement times and using different correction methods in NRTK. Significant differences were further investigated using the correction methods at different times. The results of this study allow people using NRTK to statistically evaluate different measurement configurations and obtain position information with the desired accuracy and precision. The materials and methods section provides general information about the NRTK technique, different correction techniques (VRS, FKP, and MAC), the factors affecting accuracy, the working area, the TUSAGA Aktif Network, measurement configurations and methods, and statistical analysis methods.

Materials and methods

In light of the above, the NRTK technique was developed, through which systematic and atmospheric errors can be modeled via repeated observations based on many GNSS reference stations. While position information is obtained by using GNSS reference stations in a CORS structure, different data transmission methods are used to model corrections in the observation space.

The data transmission methods that are widely used around the world, as well as in this application (VRS, FKP, and MAC), are completely different from each other. In the first technique, an imaginary VRS point at a 1 m distance to the mobile GNSS receiver is identified. Observations at the VRS points are then obtained via surrounding GNSS reference stations. Correction parameters for mobile receivers are calculated by interpolating numerous GNSS reference station data inside the study area. Assuming that the CORS closest to the mobile GNSS receiver is the main reference station (A), the mathematical formulation of the VRS model, including inter-satellite binary differences (for m, n satellites) and geometric displacements for A and the VRS station, is defined below (Vollath *et al*., 2000a; Vollath *et al*., 2000b; Wanninger, 2002; Landau *et al*., 2002; Wei *et al*., 2006):

$$\nabla\Delta\varphi^{m,n}_{A,VRS} = \left(\Delta\varphi^{m,n}_{VRS} - \Delta\varphi^{m,n}_{A}\right) \tag{1}$$

$$\nabla\Delta\varphi^{m,n}_{A,VRS} = \frac{1}{\lambda}\left(\nabla\Delta\rho^{m,n}_{A,VRS} - \nabla\Delta I^{m,n}_{A,VRS} + \nabla\Delta T^{m,n}_{A,VRS}\right)\nabla\Delta N^{m,n}_{A,VRS} \tag{2}$$

$$\nabla\Delta\rho^{m,n}_{A,VRS} = \left(\rho^{n}_{VRS} - \rho^{n}_{A}\right) - \left(\rho^{m}_{VRS} - \rho^{m}_{A}\right) \tag{3}$$

∇ : single differences
$\nabla\Delta$: double differences
$\Delta\varphi^{m,n}_{A}$: single differences
$\Delta\varphi^{m,n}_{VRS}$: single differences
$\nabla\Delta\varphi^{m,n}_{A,VRS}$: double differences
$\nabla\Delta\rho^{m,n}_{A,VRS}$: geometric distance between satellites and stations
$\nabla\Delta I^{m,n}_{A,VRS}$: ionospheric delay
$\nabla\Delta T^{m,n}_{A,VRS}$: tropospheric delay
$\nabla\Delta N^{m,n}_{A,VRS}$: carrier phase initial ambiguity

FKP is an area-correction parameter method developed by the German SAPOS group. To calculate correction parameters using this technique, a simple plane formed by connecting three points is used as a surface reference. Correction parameters between three points, whose positions are accurately known, are defined as changes in North-South and East-West directions. These surface parameters are expressed as area correction parameters. A separate FKP plane is constructed for each reference surface. In this approach, atmospheric and carrier phase corrections for each permanent station within the entire GNSS network are calculated (Wübbenna *et al.*, 1996; Wübbena *et al.*, 2001; Park *et al.*, 2010). The FKP model can be described as (Jin, 2012; Wübbena and Bagge, 2006):

$$\delta r_0 = 6.37H(N_0((\varphi-\varphi_R)+E_0\left(\lambda-\lambda_R\right)cos\varphi_R \quad (4)$$

$$\delta r_1 = 6.37H(N_1((\varphi-\varphi_R)+E_1\left(\lambda-\lambda_R\right)cos\varphi_R \quad (5)$$

$$H = 1+16\left(0.53-E/\pi\right)^3 \quad (6)$$

N_0 : FKP parameter N-S direction, ionospheric delay [ppm]

E_0: FKP parameter E-O direction, ionospheric delay [ppm]

N_1 : FKP parameter N-S direction, tropospehric delay

E_1: FKP parameter E-O direction, tropospheric delay

φ_R, λ_R : geographic coordinates of the reference station

φ, λ : the rover approximate coordinates

E: satellite elevation [°]

$\delta r0$: distance dependent error of geometric signal combination

$\delta r1$: distance dependent error of ionospheric signal

The corrections on L_1 and L_2 are:

$$\delta r_{1=}\delta r_0 + \left(\frac{120}{154}\right)\delta r_1 \quad (7)$$

$$\delta r_{2=}\delta r_0 + \left(\frac{154}{120}\right)\delta r_1 \quad (8)$$

$$R_{correct} = R-\delta r_i\left(i=1\div 2\right) \quad (9)$$

The concept of *MAC* was proposed by Euler (2001) and is regarded by the RTCM SC104 working groups as a standard for network correction. In this NRTK application, the GNSS receiver sends its position to the calculation center. The closest station to the receiver is selected as the *main station*, while the others are called *auxiliary stations*. Observations at the main and auxiliary stations are combined to obtain the correction differences. In this method, the points forming the network are separated into clusters, and corrections for each region are then calculated from the main and auxiliary reference points (Euler *et al.*, 2001; Brown *et al.*, 2006; Leica Geosystems, 2005). NRTK-based positioning applications are designed according to one of the different correction techniques mentioned above. The MAC model for the satellite (i) between the main reference station (A) and the auxiliary reference station (B) with respect to single differences is defined as follows (Jin, 2012; Euler, 2006; Lin, 2006):

$$\lambda\Delta\varphi^i_{AB} = \Delta\rho^i_{AB} + c\Delta\delta t_{AB} + \lambda\Delta N^i_{AB} + \Delta T^i_{AB} - \Delta I^i_{AB} + \Delta Y^i_{AB} \quad (10)$$

In the Equations above,

$\Delta\varphi_i$: single differences of the phase measurements between A and B

ρ^i_{AB} : geometric distance difference between two stations

δt_{AB} : satellite clock errors

N^i_{AB} : initial uncertainty of the frequency-dependent phase

T^i_{AB} : tropospheric errors

I^i_{AB} : ionospheric errors

Y^i_{AB} : satellite orbit errors

Moreover, there are many factors that affect positioning accuracy in NRTK applications, one of them being satellite geometry, which expresses the distribution of satellites relative to each other and to the receiver on the earth. Accuracy is better when GNSS satellites are dispersed in a wide area in the sky. In addition, interruptions during GNSS observation may be due to objects that prevent signals from reaching the receiver, a low signal-to-noise (S/N) ratio due to poor ionospheric conditions, the signal reflection effect (multipath), and faults in the receiver software. GNSS receiver antennas can simultaneously receive satellite signals from all directions depending on the terrain structure and the signal elevation angle. The signal reflection (multipath) effect, which involves the arrival of signals broadcasted from satellites to the receiving antenna while following one or more paths, is also a factor that affects positioning accuracy.

One of the most important factors affecting NRTK positioning accuracy corresponds to ionospheric and tropospheric effects. Signals from satellites reach the receiver by passing through different layers of the atmosphere (*e.g.*, ionosphere and troposphere). Thus, they are exposed to different effects. The ionosphere is the upper region of the atmosphere, where air molecules are highly concentrated and gain electrical conductivity. Since the position of the ionosphere relative to the sun changes over time, ionization levels also change. Therefore, the magnitude of the ionospheric effect varies depending on the frequency of the satellite signal, time, location, season, and solar radiation movements. In general, the effect of the ionosphere on electromagnetic waves is greater during the daytime than at night (Klobuchar and Doherty 1990; Kahveci *et al.*, 2021). In the troposphere, which is close to the ground, signals are delayed especially due to water vapor. Tropospheric refraction has two components. Out of these, the dry air component causes 90% of the total fracture. The air pressure at the observation point is used to model the dry component of the troposphere. The other component, wet air, has little effect but is more difficult to model (Davis *et al.*, 1985).

Ambiguity is another factor affecting the accuracy of NRTK GNSS positioning since it is uncertain how many full waves are present in the carrier wave phase between the satellite and the receiver during the first epoch (first instant). Many methods (such as P-code assisted, antenna displacement, modern QIF (Quality Indicator for Fix), and OTF (On-The-Fly) have been developed to solve this problem (Brown, 2006).

One of the most important factors affecting positioning accuracy corresponds to the hardware features of the GNSS receiver and antenna. In addition, environmental effects must be taken into account in selecting the location of GNSS receivers. These receivers should not be installed too close to the ground, and it is important that there are no reflective surfaces causing signal reflection at the measuring point. In addition, given the increased distance between the fixed stations forming the network and the application points in NRTK, the atmosphere, satellites, and the surroundings of the stations affect positioning accuracy; as the distance between points increases, these effects become independent from each other, but they can be neutralized via modeling.

In this study, the effects of the various factors involved in NRTK GNSS positioning accuracy during the day were statistically studied with repeated measurements. The CORS network, which is called *TUSAGA Aktif* and is used for determining locations in Turkey (especially in Istanbul), was employed in this study (Figure 1). The differences between the position values obtained at different times of the day and according to different correction methods were evaluated via an ANOVA and multiple comparisons.

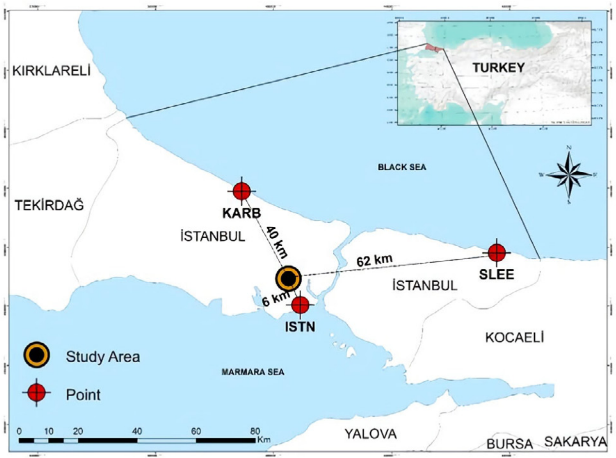


Figure 1. TUSAGA Aktif NRTK control station positions and the study area in Istanbul
Source: Author

The TUSAGA Aktif Network consists of 158 reference stations over the entire country, three of which are located in Istanbul. The distances between stations range from 80 to 100 km, and static and real-time coordinate positions can be determined in an accurate, quick, and economical way. The control center sends corrections obtained via the VRS, FKP, and MAC techniques to mobile receivers. TUSAGA Aktif Network coordinates were obtained from the 2005.00 reference epoch of the ITRF 96 Datum (Eren *et al.*, 2009). In this study, the coordinates obtained via the NRTK application correspond to the Gauss-Krüger cartographic representation (ITRF 96/GRS 80), which is based on the Transversal Mercator (TM) projection and divided into 3° slices with values of 27°, 30°, 33°, 36°, 39°, 42°, and 45° mid meridian longitude (λ0) for Turkey. The scale factor is 1. The horizontal axis of the coordinate system is the equator, and the vertical axis is the slice median meridian. The orthogonal coordinates of the origin were taken as (500 000, 0) to avoid negative values on the left of the slice. The geographic coordinates of the origin were (λ0, 0) (Basaran, 2019). In this coordinate system, Y and X represent the orthogonal coordinates (right value, upper value), and h represents the ellipsoidal height.

The selected study area was the UZEL location of the Davutpaşa Campus, Yıldız Technical University (Figure 2). The KARB, SLEE, and ISTN points, which carry out 24/7 observations within the TUSAGA Aktif Network, were used as fixed locations. The distances between these fixed stations and the application point are given in Figure 1. Measurements were carried out with an observation period of at least 5 h. The elevation angle was 10° and the recording interval was 10 s. Satellite signals were recorded using the GPS/GLONASS satellite configurations. The final IGS ephemeris was used in post-processing GNSS solutions. Based on these fixed points, adjustments were made according to the least squares method. The coordinates and rms values of the adjusted UZEL point are given in Table 1. The location of UZEL (X, Y, h), with millimeter precision, was determined from the 2005.00 reference epoch of the ITRF 96 Datum, using the GNSS solution software for static GNSS applications. The resulting coordinates were used as actual values to determine the accuracy of the measurements obtained with the NRTK application at different times of the day.

The UZEL locations corresponding to the morning, noon, and evening were determined via the VRS, FKP, and MAC correction techniques used in NRTK. These measurements were taken since the factors affecting NRTK positioning accuracy are different throughout the day. Our measurements were taken at a single point on a sunny summer day when the weather conditions were stable and there was no wind or rain. It is important to mention that the ionization of the ionosphere changes over time according to its position relative to the sun, so the effect of the ionosphere on electromagnetic waves is greater during the day than at night.

In this study, a Topcon HyperPro GNSS receiver was used for the NRTK and static GNSS applications, and the Topcon Link v.8.2.3 software was used for evaluating the NRTK applications. The accuracy of this equipment, as declared by the manufacturer, is 10 mm + 1,0 ppm for horizontal positioning and 15 mm + 1,0 ppm for vertical positioning in environments without high PDOP and ionospheric activity. Furthermore, Topcon Hyper Pro has a tracking feature (GPS+GLONASS) which provides 40% more satellite coverage, eliminating downtime due to poor satellite coverage or obstructions. Multipath correction was applied to the repeated measurements. The Topcon Hyper Pro RTK Receiver has a feature that reduces the effects caused by signal reflection, which was used in our study. Thus, Table 1 shows that the multipath effect was considered, as multipath correction was included in the measurements.

For each correction technique, the UZEL location was measured in consecutive time periods (interval of 1 second) averaged over 45 minutes (Table 1). Accuracy and precision were evaluated via the average of repeated measurements and coordinate differences obtained from static GNSS.

Boxplots of the coordinate differences obtained from the averages were elaborated in order to determine whether these differences are sensitive with respect to each other. We also evaluated whether the X, Y, and h coordinates obtained from the NRTK application are significantly different from each other in statistical terms. If a significant difference was found, an ANOVA was performed with the purpose of determining its time period and correction method.

If a difference between groups was found as a result of the ANOVA, *post hoc* tests were conducted to identify the cause of this difference. If variance homogeneity was achieved, the Tukey test was used. However, in the opposite case, Tamhane's T2 was applied. The SPSS software was used to perform analyses of variance. This software offers a p-value (Sig) of 95% to confirm the significance of the F-test. If the p-value is lower than 0,05, the H_0 hypothesis is rejected. Moreover, SPSS constructs sub-groups based on dependent variables (Kalaycı, 2010). These sub-groups are determined according to the correction methods and the different times of the day analyzed. The following section shows the results obtained in this regard, which are statistically interpreted with the help of tables and graphs.

Results and discussion

This study considers the effects of many factors involved in the positioning accuracy of NRTK applications. In general, as the geometry of the satellite changes during the morning, noon, and evening hours of the day, positioning accuracy is affected. In addition, the number of satellites that are visible at certain times of the day changes according to their

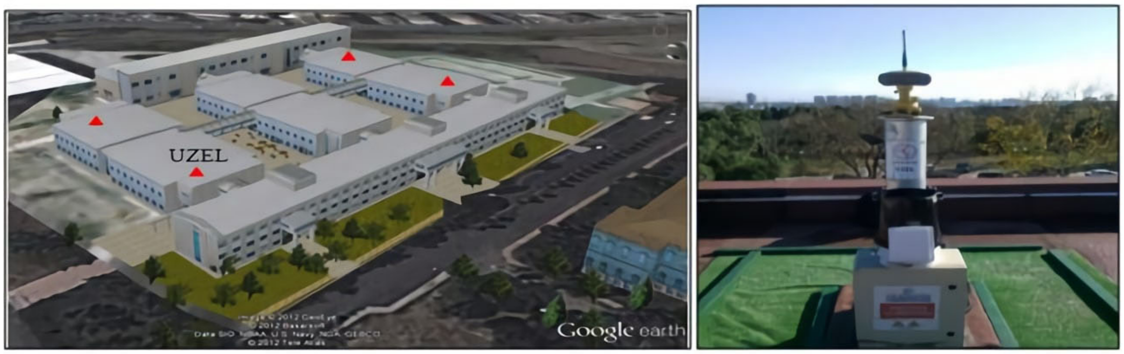


Figure 2. UZEL location
Source: Author

Table 1. Measurement data for NRTK and static GNSS

	Correction method	Start and finish time – epoch number		
		Morning	Noon	Evening
		07:00 -10:00	12:00-15:00	19:00-21:00
NRTK	**MAC**	2 564	2 560	2 527
	FKP	2 510	2 506	2 632
	VRS	2 519	2 022	3 012
		1 s (multipath effect)		
Static GNSS	**Post-processing**	**Three points (control station) – 24 h**		
		KARB	SLEE	ISTN
		UZEL point adjusted coordinates and rms		
		X(m)	**Y(m)**	**h(m)**
	Coordinates (m)	4543749,092	406541,0325	126,9082
	rms (m)	0,002	0,003	0,007

Source: Author

different configurations (GPS, GLONASS). The number of visible satellites should be as large as possible, as many measurements are required to increase positioning accuracy. To be able to see more satellites with GNSS receivers, specific elevation angles are necessary. More satellites can be seen at lower elevation angles, but they negatively affect positioning accuracy due to atmospheric refraction. In this study, the elevation angle as set at 10° for the NRTK applications.

The multipath effect varies according to environmental conditions. The hardware features of receiver antennas are crucial in increasing the accuracy of the height component. In order to eliminate ionospheric delay, a dual-frequency Topcon Hyper Pro GNSS receiver was used, considering the fact that waves of different frequencies will encounter different resistances. In addition, due to seasonal effects and instantaneous changes in the atmosphere, unexpected results can be obtained because of the troposphere.

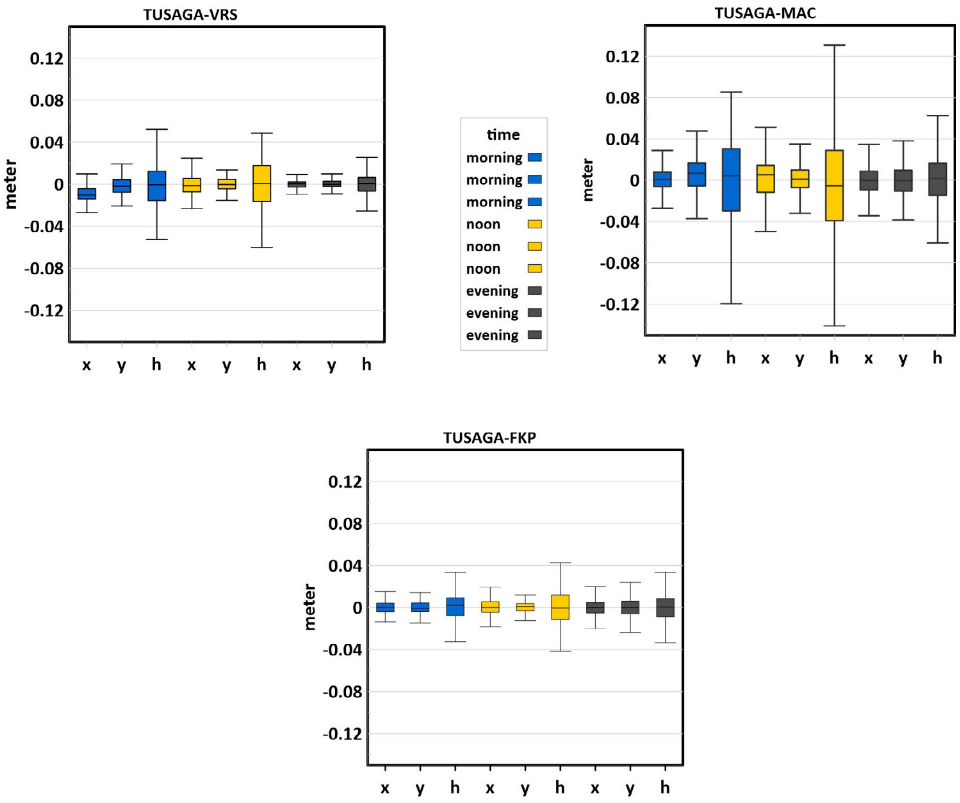
The data and the corrections were received from the same fixed stations in the NRTK network during measurement. We attempted to ensure the same conditions for different times of the day in order to minimize the effects associated with variability. In this study, repeated observations of a single point were made upon the basis of more than one GNSS reference station. All this, by using the NRTK technique, through which systematic and atmospheric errors can be modeled and different correction methods.

This section shows the results obtained using the materials and methods explained in the previous section. The coordinate differences obtained from the average of repeated measurements were examined to determine if they were sensitive to each other. Boxplots, which contribute to an interpretation of the results, were elaborated to clearly visualize the differences (Figure 3).

Figure 3 shows that the differences in the h coordinate are greater than those in the X and Y coordinates for the three correction methods. The values obtained via the VRS correction method are close to the average values. In addition, outlier measurements were observed. All boxplots in Figure 3 show a normal distribution. A comparison of repeated measurements was carried out via an ANOVA.

The accuracy and precision values of measurements obtained via NRTK positioning at the UZEL location are shown with different measurement configurations. Precision values were determined by averaging the measures obtained for each correction method at different times of the day. The coordinates of the UZEL point found via the static GNSS method were regarded as an absolute value. Accuracy values were then determined by subtracting the coordinates found by the static GNSS method from those of each correction method. This is shown in Figure 4.

It can be observed that the accuracy and precision values along the h coordinate axes are lower than those in the


TUSAGA-VRS

TUSAGA-MAC

TUSAGA-FKP

Figure 3. Boxplots of the differences obtained from repeated measurements
Source: Author

other two axes. In general, the results obtained with the FKP method are better. Differences were observed between the results for different times of the day. The accuracy and precision values of the measurements obtained at noon are lower than those obtained in the morning and evening. The results obtained through the MAC correction method seem to be lower than those of the other two. In the TUSAGA Aktif system, there is no GLONASS satellite data for the MAC method. In this method, the initial lockout time can be extended for network corrections. When the first correction dataset in the network arrives at the GNSS receiver, complex algorithms for fault modeling are used depending on the software employed. The location resolution may take time depending on processor speeds. Since all calculations are carried out on the rover, the appropriate software must be installed. This is why the accuracy of the MAC method may be lower than that of others.

Gumus (2016) attempted to determine whether there is a statistically significant difference between the coordinates obtained with different correction methods at different elevation angles and in epochs in NRTK applications. When comparing statistical data on different variables, the accuracy regarding the h coordinate axis was shown to be lower. The FKP and VRS methods showed different characteristics, while the MAC method showed similar features. By comparing the results of our study against those of Gumus (2016), close values can be observed, albeit with differences in the order of millimeters for different times of the day. In this study, the MAC method differs from the other two along the x and y coordinate axes. This shows that using different correction methods at different times of the day affects the results to a small extent.

We evaluated whether there was a significant difference between the X, Y, and h coordinate values obtained. If there was a difference, its corresponding time period and correction method were identified. The ANOVA showed a normal distribution for each test group while statistically confirming the hypothesis that the variance values of the groups were homogeneous. Table 2 shows the mean and standard deviation values of the X, Y, and h coordinates obtained for different time intervals and correction methods.

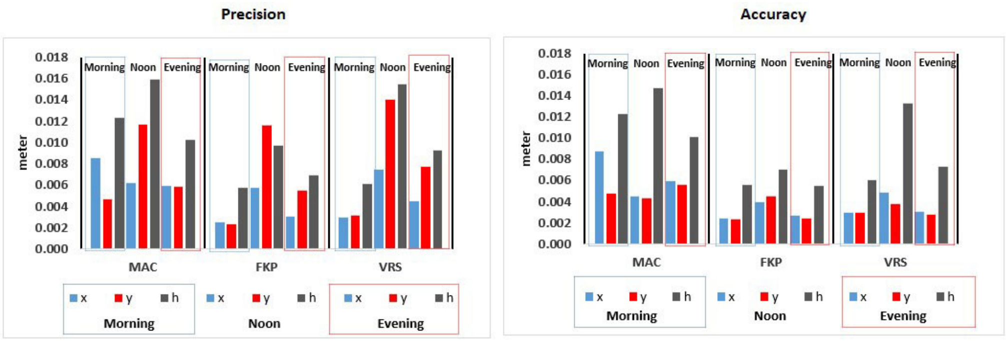


Figure 4. Accuracy and precision values
Source: Author

Table 2. Mean and standard deviation values of the X, Y, and h coordinates (m)

Time Interval	Correction Method	X(m) Mean	X(m) Std. Dev.	Y(m) Mean	Y(m) Std. Dev.	h(m) Mean	h(m) Std. Dev.
Morning	**FKP**	0,000	0,005	0,000	0,005	0,000	0,013
	MAC	0,000	0,013	0,002	0,026	-0,002	0,044
	VRS	-0,009	0,008	-0,001	0,008	-0,001	0,021
Noon	**FKP**	0,001	0,009	0,000	0,006	0,000	0,015
	MAC	0,000	0,024	0,001	0,016	0,001	0,061
	VRS	0,000	0,010	0,000	0,006	0,000	0,023
Evening	**FKP**	0,000	0,008	0,000	0,009	0,000	0,016
	MAC	0,000	0,013	-0,001	0,014	0,001	0,022
	VRS	0,000	0,004	0,000	0,004	0,000	0,011

Source: Author

The average of the differences for the VRS method in the x coordinate axis obtained during the morning hours is -0,009 m (Table 2), which indicates that the results from the NRTK network vary throughout the day in different coordinate axes. Upon inspecting the boxplots, it was observed that the repeated measurements made using VRS are closely grouped and exhibit little scatter. Many factors may have affected the VRS measurements for the morning hours, such as atmospheric conditions, environmental factors, the number of visible satellites, satellite geometry, signal interruptions, multipath effect, *etc*.

The ANOVA showed statistically significant differences in the results. *Post hoc* tests were performed to identify the cause of these differences. Time intervals were determined via pairwise comparisons between the morning, noon, and evening times of the day, as shown in Table 3. To this effect, the mean differences between the averages of the measures obtained were used. The p-value corresponds to the value of F at a 95% significance level. According to this value, there is a significant difference (below 0,05) between the groups.

Table 3 shows no significant difference between the measurements obtained along the X and h coordinate axes when comparing the noon and evening intervals. For the measurements obtained along the Y coordinate axis, there is no significant difference between the morning and noon hours.

In addition, the averages of the subgroups obtained via the correction methods for different time intervals were compared using means for groups in homogeneous subsets. We tested for statistical significance. Table 4 shows these subsets, including the X, Y, and h coordinate axes and the time intervals. Note that this was done using the SPSS software.

For the TUSAGA Aktif Network, along the coordinate axes, the morning interval was defined as a single subset, while the noon and evening periods were incorporated into another one. The latter indicates that these two intervals have similar characteristics, unlike the morning interval. An ANOVA showed that the results were statistically significant. *Post hoc* tests were performed to identify the cause of differences. Via pairwise comparisons between the average values obtained, the VRS, FKP, and MAC methods were shown to be responsible for the differences (Table 5).

Table 4. Subsets of coordinates for different times of the day

	Tukey HSD					
	X coordinate (m)		Y coordinate (m)		h coordinate (m)	
Time	Subset		Subset		Subset	
	1	2	1	2	1	2
Morning	-0,002		0,000		-0,001	
Evening		0,000		0,000		0,000
Noon		0,000		0,001		0,000
Sig.	1,000	0,819	1,000	0,776	1,000	0,974

Means for groups are displayed in homogeneous subsets.

Subset for alpha = 0,05, Sig. = p-value. The sig. value is not in meters.

Source: Author

Table 3. Pairwise comparisons for different times of the day along the X, Y, and h coordinate axes (m)

DV = Time		X coordinate				Y coordinate				h coordinate			
Time (I)	Time (J)	MD (I-J)	Sig.	CI Low.	CI Up.	MD (I-J)	Sig.	CI Low.	CI Up.	MD (I-J)	Sig.	CI Low.	CI Up.
Morning	**Noon**	-0,003	0,000	-0,003	-0,002	0,000	0,916	-0,001	0,000	-0,002	0,042	-0,003	0,000
	Evening	-0,002	0,000	-0,003	-0,002	0,001	0,040	0,000	0,001	-0,001	0,002	-0,002	0,000
Noon	**Morning**	0,003	0,000	0,002	0,003	0,000	0,916	0,000	0,001	0,002	0,042	0,000	0,003
	Evening	0,000	0,936	0,000	0,001	0,001	0,000	0,000	0,001	0,000	0,996	-0,001	0,001
Evening	**Morning**	0,002	0,000	0,002	0,003	-0,001	0,040	-0,001	0,000	0,001	0,002	0,000	0,002
	Noon	0,000	0,936	-0,001	0,000	-0,001	0,000	-0,001	0,000	0,000	0,996	-0,001	0,001

▨ There is a significant difference DV =Dependent Variable

MD = Mean difference, Sig. = P-value, CI = confidence interval (95%), Low. = lower bound, Up. = upper bound

Source: Author

Table 5 shows that there is no significant difference between the FKP and VRS methods regarding the measurements obtained for the X coordinate axes. There are statistically significant differences between the measurements along the Y coordinate axis, unlike those along the h coordinate axis. In general, the MAC correction method shows a statistical difference when compared to the FKP and VRS methods. For Table 6, subsets containing X, Y, and h coordinates from correction methods were created using SPSS.

In a study conducted by Gumus and Selbesoğlu (2019), the distances (simulated space changes) for the X, Y, and h coordinate axes were measured using NRTK methods and the statistical GNSS technique. The performance of this system was compared to the space changes measured using an interferometer with a precision of 0,1 mm. The performance of the FKP correction method was better, especially in the horizontal axis. In light of the above, the results obtained in our study with the FKP and VRS methods were good for all coordinate axes when compared to the MAC technique. The results regarding measurements taken at different times of the day appear to be influenced by many factors.

The study by Garrido *et al.* (2018) examined different scenarios (both adequate and inadequate) for a local active network under a variety of topographic, environmental, and meteorological conditions in four different test areas of southern Spain. Multiple observation sessions were conducted over three days in each test area with the aim of assessing the repeatability of RTK solutions. These sessions took place in the morning, noon, and afternoon hours, implying different satellite configurations and atmospheric conditions. Numerical results showed that the observation sessions, satellite configurations, and changes in atmospheric conditions are factors that affect positioning precision and accuracy. Furthermore, it was found that part of the ionospheric effect is now stronger during times of high solar activity. These results support ours. This consistency is important for contributing to the accumulation of knowledge.

Table 5. Pairwise comparisons of correction methods along the X, Y, and h coordinate axes

DV = Correction method		**X coordinate**				**Y coordinate**				**h coordinate**			
Corr. Met. (I)	**Corr. Met. (J)**	**M.D. (I-J)**	**Sig.**	**C.I. Low.**	**C.I. Up.**	**M.D. (I-J)**	**Sig.**	**C.I. Low.**	**C.I. Up.**	**M.D. (I-J)**	**Sig.**	**C.I. Low.**	**C.I. Up.**
FKP	**VRS**	0,000	0,761	0,000	0,001	-0,001	0,012	-0,001	0,000	0,000	0,837	-0,001	0,002
	MAC	0,003	0,000	0,002	0,003	0,001	0,000	0,000	0,001	0,000	0,339	0,000	0,001
VRS	**FKP**	0,000	0,761	-0,001	0,000	0,001	0,012	0,000	0,001	0,000	0,837	-0,002	0,001
	MAC	0,003	0,000	0,002	0,003	0,001	0,000	0,001	0,002	0,000	1,000	-0,001	0,001
MAC	**FKP**	-0,003	0,000	-0,003	-0,002	-0,001	0,000	-0,001	0,000	0,000	0,339	-0,001	0,000
	VRS	-0,003	0,000	-0,003	-0,002	-0,001	0,000	-0,002	-0,001	0,000	1,000	-0,001	0,001

▨ There is a significant difference DV = Dependent Variable

MD = mean difference, Sig. = p-value, CI = confidence interval (95%), Low. = lower bound, Up. = upper bound

Source: Author

Table 6. Subsets of coordinates for each correction method (meter)

	Tukey HSD					
	X coordinate (m)		**Y coordinate (m)**			**h coordinate(m)**
Correction Method	**Subset 1**	**Subset 2**	**Subset 1**	**Subset 2**	**Subset 3**	**Subset 1**
VRS	-0,002		-0,003			-0,003
MAC		0,000		0,002		-0,003
FKP		0,000			0,001	0,001
Sig.	1,000	0,578	1,000	1,000	1,000	0,670

Means for groups in homogeneous subsets are displayed.

Subset for alpha = 0,05, Sig. = p-value. The sig. value is not in meters.

Source: Author

Conclusion

This study statistically evaluated the significance of the differences between repeated measurements for a single point obtained at different times of the day (morning, noon, and evening) using the VRS, FKP, and MAC correction methods within the TUSAGA Aktif Network in Turkey. In NRTK applications, some changes were observed in the form of increases or decreases in 3D point positioning values. Many different factors affect the accuracy of NRTK GNSS positioning.

These statistical significances were tested using an ANOVA. By analyzing the results obtained, it was seen that the differences for the averages of the h coordinate axis were higher than those for the other two axes. The values for this axis were less accurate and precise. Although the values obtained with the VRS correction method were close to the average values, the results obtained via the FKP method were generally better. Outlier measurements were observed between repeated measurements. Differences were observed between the morning, noon, and evening intervals. In general, the accuracy and precision values of the morning measurements were lower than those of the other two periods analyzed.

As per the ANOVA, the differences between time intervals were statistically significant. While the noon and evening periods showed similar characteristics, the morning hours showed different characteristics. The results also showed that the differences in terms of the correction methods used were statistically significant. Along the X coordinate axis, the VRS method exhibited a different behavior from MAC and FKP, which performed similarly. All correction methods showed different characteristics along the Y coordinate axis, as well as similarities regarding the h axis.

The results of this study are important for NRTK users to be able to statistically evaluate different measurement configurations and provide the necessary positional reproduction with the desired accuracy and precision.

Acknowledgements

I would like to thank my students, as well as the Land Registry and Cadastre General Directorate for providing the TUSAGA data.

CRediT author statement

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

References

Basaran, M. (2019). CBS 1 - ek - projeksiyon koordinat sistemleri ve coğrafi bilgi sistemi ortamında kullanımı, Yıldız Teknik Üniversitesi Akademik Veri Yönetimi. https://avesis.yildiz.edu.tr/mbasaran/dokumanlar

Bisnath, S., Saeidi, A., Wang, J. G., and Seepersad, G. (2013). Evaluation of Network RTK performance and elements of certification – A Southern Ontario case study. *Geomatica, 67*(4), 243-251. https://doi.org/10.5623/cig2013-050

Brown, N. E. (2006). *Sequential phased estimation of ionospheric path delays for improved ambiguity resolution over long GPS baselines* [PhD thesis, University of Melbourne].

Brown, N., Geisler, I., and Troyer, L. (2006). RTK rover performance using the Master Auxiliary Concept. *Journal of Global Positioning Systems, 5*(1-2), 135-144.

Brown, N., Keenan, R., Rıchter, B., and Troyer, L. (2005). *Advances in ambiguity resolution for RTK applications using the new RTCM V3.0 Master-Auxiliary messages* [Conference presentation]. ION GNSS 2005, Long Beach, CA, USA.

Cannon, M. E., Lachapelle, G., Fortes, L. P., Alves, P., and Townsend, B. (2001). *The use of multiple reference station VRS for precise kinematic positioning* [Conference presentation]. Japan Institute of Navigation, GPS Symposium 2001, Tokyo, Japan.

Chen, X. M., Han, S. W., Rizos, C., and Goh, P. C. (2000). *Improving real-time positioning efficiency using the Singapore integrated multiple reference station network (SIMRSN)* [Conference presentation]. 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000), Salt Lake City, UT, USA.

Cina, A., Dabove, P., Manzino, A. M., and Piras, M. (2015). Network Real Time Kinematic (NRTK) Positioning - Description, Architectures and Performances. In S. Jin (Eds.), *Satellite Positioning – Methods, Models and Applications* (pp. 23-45). IntechOpen. https://doi.org/10.5772/59083

Colombo, O. L., Hernández-Pajares, M., Juan, J. M., Sanz, J., and Talaya, J. (1999). *Resolving carrier phase ambiguities on the fly, at more than 100 km from nearest reference site, with the help of ionospheric tomography* [Conference presentation]. 12th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1999), Nashville, TN, USA.

Dabove, P., Manzino, A. M., and Taglioretti, C. (2014). GNSS network products for post-processing positioning: Limitations and peculiarities. *Applied Geomatics, 6*, 27-36. https://doi.org/10.1007/s12518-014-0122-3

Dai, L., Han, S. W., Wang, J. L., and Rizos, C. (2001). *A study on GPS/GLONASS multiple reference station technique for precise realtime carrier phase-based positioning* [Conference presentation]. 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2001), Salt Lake City, UT, USA.

Dardanelli, G., Maltese, A., Pipitone, C., Pisciotta, A., and Lo Brutto, M. (2021). NRTK, PPP or static, that is the question. Testing different positioning solutions for GNSS survey. *Remote Sensing, 13*(7), 1406. https://doi.org/10.3390/rs13071406

Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E., and Elgered, G. (1985). Geodesy by radio interferometry: Effects of atmospheric modelling errors on estimates of baseline length, *Radio Science, 20*, 1593-1607. https://doi.org/10.1029/RS020i006p01593

Eren, K., Uzel, T., Gulal, E., Yildirim, O., and Cingoz, A. (2009). Results from a comprehensive Global Navigation Satellite

System test in the CORS-TR network: Case study. *Journal of Surveying Engineering, 135*(1), 10-18. https://doi.org/10.1061/(ASCE)0733-9453(2009)135:1(10)

Euler, H. J., Keenan, C. R., Zebhauser, B. E., and Wübbena, G. (2001). Study of a simplified approach in utilizing information from permanent reference station arrays [Conference presentation] ION GPS-01, Salt Lake City, UT, USA.

Euler, H. J. (2006). *Real-time RTK messages for permanent reference station applications standardized by RTCM* [Conference presentation]. IAG Symposium, Munich, Germany.

Gao, Y., Li, Z., and McLellan, J. F. (1997). *Carrier phase based regional area differential GPS for decimeter-level positioning and navigation* [Conference presentation]. ION GPS-97: 10th International Technical Meeting of the Satellite Division of the Institute of Navigation, Kansas City, MO, USA.

Gandolfi, S., Tavasci, L., and Poluzzi, L. (2017). Study on GPS-PPP precision for short observation sessions. *GPS solutions*, *21*, 887-896. https://doi.org/10.1007/s10291-016-0575-4

Garrido, M. S., Giménez, E., Lacy, M. C., and Gil, A. J. (2011). Testing precise positioning using RTK and NRTK corrections provided by MAC and VRS approaches in SE Spain. *Journal of Spatial Science, 56*(2), 169-184. https://doi.org/10.1080/14498596.2011.623341

Garrido, M. S., de Lacy, M. C., and Rojas, A. M. (2018). Impact of tropospheric modelling on GNSS vertical precision: An empirical analysis based on a local active network. *International Journal of Digital Earth*, *11*(9), 880-896. https://doi.org/10.1080/17538947 .2017.1367040

Ge, M., Gendt, G., Rothacher, M. A., Shi, C., and Liu, J. (2008). Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. *Journal of Geodesy, 82*(7), 389. https://doi.org/10.1007/s00190-007-0187-4

Gumus, K. (2016). A research on the effect of different measuring configurations in Network RTK applications. *Measurement*, *78*, 334-343. https://doi.org/10.1016/j.measurement.2015.10.022

Gumus, K., and Selbesoglu, M. O. (2019). Evaluation of NRTK GNSS positioning methods for displacement detection by a newly designed displacement monitoring system. *Measurement*, *142*, 131-137. https://doi.org/10.1016/j.measurement.2019.04.041

Gökdaş, Ö., and Özlüdemir, M. T. (2020). A variance model in NRTK-based geodetic positioning as a function of baseline length. *Geosciences, 10*(7), 262. https://doi.org/10.3390/geosciences10070262

Grejner-Brzezinska, D. A., Kashani, I., and Wielgosz, P. (2005). On accuracy and reliability of instantaneous network RTK as a function of network geometry, station separation and data processing strategy. *GPS Solutions, 93*, 179-193. https://doi.org/10.1007/s10291-005-0130-1

Han, S. W. (1997). *Carrier phase-based long-range GPS kinematic positioning* [PhD thesis, University of New South Wales].

Hu, G. R., Khoo, H. S., Goh, P. C., and Law, C. L. (2003). Development and assessment of GPS virtual reference stations for RTK positioning. *Journal of Geodesy, 77*(5), 202-302. https://doi.org/10.1007/s00190-003-0327-4

Janssen, V. (2009). *A comparison of the VRS and MAC principles for network RTK* [Conference presentation]. IGNSS Symposium, Surfers Paradise, Queensland, Australia.

Janssen, V., Grinter, T., and Roberts, C. (2011). Can RTK GPS be used to improve cadastral infrastructure? *Engineerıng Journal, 15*(1), 43. https://doi.org/10.4186/ej.2011.15.1.43

Jin, S. (Ed.) (2012). *Global navigation satellite systems: Signal, theory and applications*. Books on Demand.

Kahveci, M., Alioglu, D., and Çetin, G. (2021). Tek frekansli GNSS alıcılarında kullanılan iyonosferik etki düzeltme modellerinin karşılaştırılması. *Konya Journal of Engineering Sciences, 9*(2), 428-441. https://doi.org/10.36306/konjes.849391

Kalaycı, Ş. (2010). *SPSS uygulamali çok değişkenli istatistik teknikleri*. Asil yayin dağitim Ltd. Şti.

Klobuchar, J. A., and Doherty, P. H. (1990). *The statistics of ionospheric time delay for GPS ranging on L1* [Conference presentation]. ION GPS-90, 3rd International Technical Meeting of the Satellite Division of the Institute of Navigation, Colorado Springs, CO, USA.

Kudrys, J., and Krzyek R. (2011). Analysis of coordinates time series obtained using the NAWGEO service of the ASG-EUPOS system. *Geomatics and Environmental Engineering, 5*(4), 39-46.

Kun, S., and Yong, W. (2006). The development and application of CORS technology and its value in the solution for environment protect in the GMS. *GMSARN International Journal, 1*, 69-74. https://gmsarnjournal.com/home/wp-content/uploads/2015/08/vol1no2-3.pdf

Landau, H., Vollath, U., and Chen, X. (2002). Virtual reference station systems. *Journal of Global Positioning Systems, 1*(2), 137-143.

Landau, H., Vollath, U., and Chen, X. (2003). *Virtual reference stations versus broadcast solutions in network RTK – Advantages and limitations* [Conference presentation]. GNSS 2003 – The European Navigation Conference, Graz, Austria.

Leica Geosystems (2005). *Networked Reference Stations: Take it to the MAX*. Leica Geosystems.

Li, X., Ge, M., Dai, X., Ren, X., Fritsche, M., Wickert, J., and Schuh, H. (2015). Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. *Journal of Geodesy*, *89*(6), 607-635. https://doi.org/10.1007/s00190-015-0802-8

Lin, M. (2006). *RTCM 3.0 Implementation in network RTK and performance analysis* [Master's thesis, University of Calgary]. http://hdl.handle.net/1880/101496

Martin, A., and McGovern, E. (2012). *An evaluation of the performance of network RTK GNSS services in Ireland* [Conference paper]. https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1001&context=dsiscon

Öğütcü, S., and Kalaycı, İ. (2018). Accuracy and precision of network-based RTK techniques as a function of baseline distance and occupation *time. Arabian Journal of Geosciences, 11*, 354. https://doi.org/10.1007/s12517-018-3712-2

Park, B., and Kee C. (2010). The Compact Network RTK Method: An effective solution to reduce GNSS temporal and spatial decorrelation error. *Journal of Navigation, 63*(2), 343-362. https://doi.org/10.1017/S0373463309990440

Pirti, A., Arslan, N., Deveci, B., Aydin, O., Erkaya, H., and Hosbas, R. G. (2009). Real-time kinematic GPS for cadastral surveying. *Survey Review, 41*(314), 339-351. https://doi.org/10.1179/003962609X451582

Pirti, A., and Hosbas, R. G. (2019). Role of CORS RTK (Network RTK) mode measurements in determination of the forest boundary: A case study of ISKI-CORS. *Forestry Ideas, 25*(2), 394-403. https://forestry-ideas.info/issues/issues_Index.php?pageNum_rsIssue=2&totalRows_rsIssue=19&journalFilter=64

Raquet, J. (1998). *Development of a method for kinematic gps carrier-phase ambiguity resolution using multiple reference receivers* [PhD thesis, University of Calgary].

Rizos, C. (2002). Network RTK research and implementation – A geodetic perspective. *Journal of Global Positioning Systems, 1*(2), 144-150.

Tarig, A. (2012). Positioning with wide-area GNSS Networks: Concept and application positioning. *Scientific Research, 3*, 1-6. https://doi.org/10.4236/pos.2012.31001

Vollath, U., Buecherl, A., and Landau, H. (2000a). *Long range RTK positioning using virtual reference stations* [Conference presentation]. 2000 National Technical Meeting of The Institute of Navigation, Anaheim, CA, USA.

Vollath, U., Buecherl, A., Landau, H., Pagels, C., and Wagner, B. (2000b). *Multi base RTK positioning using virtual reference stations* [Conference presentation]. 2000 National Technical Meeting of The Institute of Navigation, Anaheim, CA, USA.

Wanninger, L. (2002). *Virtual reference stations for centimeter level kinematic positioning* [Conference presentation]. 15th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2002), Portland, OR, USA.

Wanninger, L. (2003). Virtual reference stations (VRS). *GPS Solutions, 7*, 143-144. https://doi.org/10.1007/s10291-003-0060-8

Wei, E., Chai, H., and An, Z. (2006). VRS: virtual observations generation algorithm. *Journal of Global Positioning Systems, 5*(1-2), 76-81.

Wielgosz, P., Grejner-Brzezinska, D. A., and Kashani, I. (2003). *Network approach to precise GPS navigation* [Conference presentation]. 59th Annual Meeting of The Institute of Navigation and CIGTF 22nd Guidance Test Symposium, Albuquerque, NM, USA.

Wübbena, G., Bagge, A., Seeber, G., Volker, B., and Hankemeier, P. (1996). *Reducing distance dependent errors for real time precise DGPS applications by establishing reference station networks* [Conference presentation]. 1996 National Technical Meeting of The Institute of Navigation, Santa Monica, CA.

Wübbena , G., Bagge, A. and Schmitz, M. (2001). *Network based techniques for RTK applications* [Conference presentation]. GPS JIN 2001, Tokyo, Japan.

Wübbena, G., and Bagge, A. (2002). *RTCM Message Type 59-FKP for transmission of FKP*. http://www.geopp.com/pdf/geopp-rtcm-fkp59.pdf

Ingeniería e Investigación vol. 44 No. 1, January - 2024 (e105143)

Research Article/Industrial Engineering

http://doi.org/10.15446/ing.investig.105143

Optimization in Territorial Partitioning to Improve the Performance of a Common Building Maintenance Service Contract: A Case Study of a Public Agency in Paraná State, Brazil

Optimización en la partición territorial para mejorar el desempeño de un contrato común de servicios de mantenimiento de edificios: un estudio de caso de una agencia pública en el Estado de Paraná, Brasil

Alexandre A. Steiner[1], David G. de B. Franco[2], Elpídio O. B. Nara[3], and Maria T. A. Steiner[4]

ABSTRACT

All public administration contracts must be evaluated in order to improve their performance while respecting the limits established by laws and regulations. The purpose of this article is to apply an approach to improve the performance of a common building maintenance service contract for the Paraná State Court of Justice (PRCJ), which has a total built area of 544 283,79 m² in 224 buildings distributed over 161 counties, by optimizing the territorial partitioning of Paraná State. To partition the state into 14 regions, a binary integer linear programming (BILP) mathematical model is applied to the facilities location problem (FLP) in three scenarios. The results show that Scenario 3 (in which the location of the 14 maintenance offices and the configuration of their areas of activity were optimized) is the best in terms of minimizing the distances traveled by maintenance teams. In this scenario, the total distance traveled would be 9 775 km per day (instead of the current 11 150 km), achieving savings of around 12,3% when compared to the current solution. With this solution, in addition to the distance, the direct and indirect costs associated with the displacement of work teams and the time spent on their corresponding trips would be minimized. Furthermore, the users of maintenance services could be served more quickly, resulting in a higher number of services and greater satisfaction for the target audience of the contract.

Keywords: public administration contracts, maintenance management, binary integer linear programming, facility location problem

RESUMEN

Todos los contratos de la administración pública deben ser evaluados a fin de mejorar su ejecución, respetando los límites establecidos por las leyes y los reglamentos. El objetivo de este artículo es aplicar un enfoque para mejorar el desempeño de un contrato de servicio común de mantenimiento de edificios en la Corte de Justicia del Estado de Paraná (PRCJ), que tiene un área total construida de 544 283,79 m² en 224 edificios distribuidos en 161 municipios, optimizando la partición territorial del Estado de Paraná. Para dividir el estado en 14 regiones, se aplicó un modelo matemático de programación lineal entera binaria (BILP) al problema de ubicación de instalaciones (FLP) en tres escenarios. Los resultados muestran que el Escenario 3 (en el que se optimizó la ubicación de las 14 oficinas de mantenimiento y la configuración de sus áreas de actividad) es el mejor escenario en cuanto a la minimización de las distancias recorridas por los equipos de mantenimiento. En este escenario, la distancia total recorrida sería de 9 775 km diarios (en lugar de los 11 150 km actuales), consiguiendo un ahorro en torno al 12,3 % respecto a la solución actual. Con esta solución, además de la distancia, se minimizarían los costos directos e indirectos relacionados con el desplazamiento de los equipos de trabajo y el tiempo empleado en sus respectivos viajes. Además, los usuarios de los servicios de mantenimiento podrían ser atendidos más rápidamente, lo que redundaría en un mayor número de servicios y una mayor satisfacción del público objetivo del contrato.

Palabras clave: contratos de administración pública, gestión de mantenimiento, programación lineal entera binaria, problema de ubicación de instalaciones

Received: October 6th, 2022
Accepted: July 24th, 2023

---

[1] Civil engineer, Federal University of Paraná, Brazil. MSc Industrial Engineering, Pontifical Catholic University of Paraná, Brazil. Affiliation: Civil engineer, Paraná State Court of Justice, Brazil. E-mail: alexandre.arns@pucpr.edu.br

[2] Industrial engineer, Pontifical Catholic University of Paraná, Brazil. PhD Industrial Engineering, Pontifical Catholic University of Paraná, Brazil. Affiliation: Associate professor, Federal University of North Tocantins, Brazil. E-mail: david.franco@ufnt.edu.br

[3] Mechanical industrial engineer, Federal University of Santa Maria, Brazil. PhD Industrial Engineering, Federal University of Santa Catarina, Brazil. Affiliation: Full professor, Pontifical Catholic University of Paraná, Brazil. E-mail: elpidio.nara@pucpr.br

[4] Civil engineer, Federal University of Paraná, Brazil. PhD Industrial Engineering, Federal University of Santa Catarina, Brazil. Affiliation: Full professor, Pontifical Catholic University of Paraná, Brazil. E-mail: tere@ufpr.br

Introduction

Contracts signed by public authorities at the municipal, state, and federal levels must comply with legal obligations that are governed by the National Bidding Law, federal and state decrees, and by the rulings handed down by the Court of Accounts.

These contracts must be based on the use of traditional management methods, applying the principles of governance to provide adequate instructions for managers regarding the reasons for and the purpose of the contract, as well as matters pertaining to taxation. Consideration should also be given to compliance with the legislation concerning the way in which the bidding process is conducted.

Furthermore, all public administration contracts must be evaluated with the aim of improving their performance while respecting the limits established by laws and norms.

The aim of this article is to implement an approach aimed at improving the performance of a common building maintenance service contract. This is done by optimizing the territorial partitioning of Paraná State.

In addition to this introduction, a literature review is presented in the next section, with a description of common building maintenance services and works related to the subject under study. Then, the case study is addressed, including, among other aspects, a description of the scenario currently adopted by the Paraná State Court of Justice (PRCJ) regarding the partitioning of the state into regions. Subsequently, the methodology is described, and the results achieved through it are outlined. The final section presents the conclusions of the study.

Literature review

This section contains a brief description of building maintenance experiences and articles related to the studied topic, addressing territorial partitioning applied to a wide range of situations.

Building maintenance services

Every building must have effective maintenance processes and methods that should be systematically planned throughout its lifespan. For public buildings, this need is even more important because, in addition to being public property, they are the nerve center where civil servants perform their duties.

Buildings have a profound impact on people's quality of life, providing the basic structure required for conducting their productive activities, in addition to playing an important social role (Brandão and Santos 2020).

According to Barbosa and Pusch (2011), depending on the form of use, conservation, and especially the frequency of adequate maintenance, a building may exhibit early signs of wear and performance capacity loss.

According to Fontoura *et al.* (2019), maintenance should always be regarded as one of the precepts for ensuring the safety, health, and well-being of a building's users. Non-compliance with legislation on the compulsory maintenance of equipment and systems can lead to putting people's safety at risk and holding managers accountable.

In a study of public buildings conducted by the researchers, it was found that 66% of the probable causes of accidents involving equipment are related to inadequate maintenance, early loss of performance, and rapid deterioration. All of these aspects can be directly addressed with good and correct maintenance. Meanwhile, the cause and origin of the remaining 34% of accidents are related to so-called *constructive defects*, or endogenous anomalies.

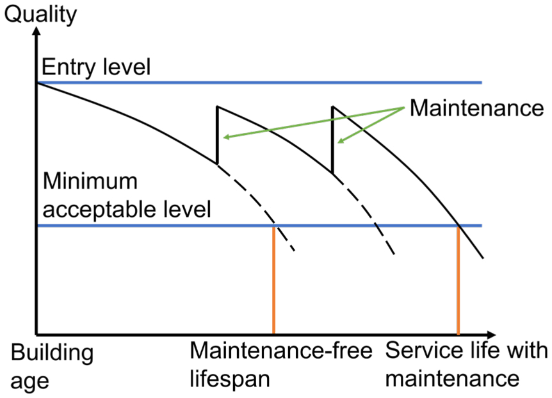


Figure 1. Building performance, lifespan, and maintenance
Source: Poli (2017)

From a legal viewpoint, it is important to highlight the NBR 14037 technical norms (ABNT, 2011) regarding the drafting of manuals on the use and maintenance of buildings, in addition to norm NBR 5674 (ABNT, 2012), which addresses the requirements for building maintenance management systems.

The definition of maintenance, identifying its different forms, is that of Dos Santos (2010), with the basic purpose of keeping equipment running most of the time at a low cost. The author also understands that maintenance, depending on how it is done, can be divided into corrective, preventive, and predictive aspects.

Meanwhile, technical norm NBR 5674 (ABNT, 1999) divides maintenance management into three groups: routine maintenance (characterized by a constant flow of standardized and cyclical services); corrective maintenance (characterized by services that require immediate action or intervention to enable the continuous use of the systems, elements, or components of buildings, or to avoid serious risks to people and/or the property of users or owners); and preventive maintenance (characterized by services that are scheduled in advance, prioritizing user requests, estimates of the expected durability of systems, the elements or components of buildings in use, severity and urgency, and periodic reports to gauge their state of disrepair).

In addition to the maintenance services themselves, there is also a need to optimally define the regions of Paraná State that each building maintenance team must serve. When defining this aspect, it must be considered that teams should travel the minimum distance from their place of origin to the locations that require maintenance, in order to guarantee a rapid service for users, which is the aim of this article.

Related works

Within the broad scope of the territorial partitioning problem, several mathematical models have been presented to group smaller territorial units into larger units with application to different domains. Although mixed integer linear programming (MILP) is the most suitable approach to deal with these problems, it has proven to be inadequate for large real-world problems. Therefore, different metaheuristic procedures have been developed for the territorial partitioning problem. Some of these approaches, which have been reported in the specialized literature over the last decade, are presented below in chronological order.

D'Amico *et al*. (2002) applied the simulated annealing metaheuristic to the problem of police management in the city of Buffalo (USA), with the goal of elaborating a map with optimized regions to minimize the disparity between the extreme workloads of police officers, specifically in the process of determining the number of patrol vehicles in use at various times of the day. Muyldermans *et al*. (2002) studied the salt spreading operation in the province of Antwerp (Belgium) through multi-purpose heuristic search procedures, where the road network was partitioned to minimize the distance traveled, as well as the number of trucks needed for the operation. Applying a genetic algorithm with multiple objectives to partition the electricity grid of the Republic of Ghana into economically viable districts, Bergey *et al*. (2003) identified some fundamental characteristics to model and solve a district's electric power problem.

Bozkaya *et al*. (2003) proposed an approach using the tabu search metaheuristic with an adaptive memory heuristic procedure for a districting problem in the city of Edmonton (Canada). Galvão *et al*. (2006) presented a Voronoi diagram for a logistics districting problem applied to parcel delivery in the city of São Paulo (Brazil), which resulted in a more balanced time/capacity use when compared to other approaches. A delivery planning system was studied by Haugland *et al*. (2007) in two steps using the tabu search and multi-start metaheuristics with the regions defined in the first stage, while the vehicle routing problem was solved for each district in the second stage. Combining a local search procedure with a multipurpose genetic algorithm, Tavares-Pereira *et al*. (2007) analyzed the public transport system in the region of Paris to suggest a reform of its pricing system.

A multipurpose genetic algorithm was studied by Datta *et al*. (2007) for the problem of land use management in Baixo Alentejo (Portugal), aiming to achieve the natural balance of the environment and financial profit. Ricca and Simeone (2008) compared several heuristic and metaheuristic procedures for the electoral districting problem in a set of regions of Italy with multiple goals, seeking to achieve population equality, compactness, and administrative compliance. Ricca *et al*. (2008) compared two methods for locating Italian polling stations using weighted Voronoi regions with two aims: population equality and compactness.

Also using Voronoi diagrams, but associated with continuous approximation models, Novaes *et al*. (2009) solved location-districting problems related to transport and logistics applied to an urban distribution service covering part of the city of São Paulo (Brazil). Salazar-Aguilar *et al*. (2011) solved the distribution of beverages in the city of Monterrey (Mexico) via a bi-objective programming model, aiming to achieve an equilibrium between dispersion and balance in territories regarding the number of customers and the sales volume. Datta *et al*. (2012), applied a multi-objective genetic algorithm to a problem with census tracts in order to aggregate census units to obtain better compactness and population/area uniformity in the Metropolitan Census area of London (Ontario, Canada).

Shirabe (2012) used a heuristic procedure to solve a MILP model with several illustrative school bus instances. Benzarti and Dallery (2013) addressed home care as a districting problem, modeling it as a MILP to achieve, among other things, compactness, and workload balance. Ríos-Mercado and López-Pérez (2013) used a branch-and-bound approach to solve the problem of a bottled beverage distribution company located in the city of Monterrey (Mexico). A GRASP procedure was used by Assis *et al*. (2014) to solve a territorial partitioning problem applied to the reading of energy meters in the city of São Paulo (Brazil) while striving for compactness and homogeneity.

Aiming to optimize truck routes for urban waste collection, Vecchi *et al*. (2016) presented a sequential approach involving three phases: the clustering of arcs based on a model adapted from the p-median problem, formulated as a binary integer linear programming (BILP) problem; the development of a model for the solution of the capacitated arc routing problem, formulated as a MILP problem; and, finally, the application of an adapted Hierholzer algorithm to sequence the arcs obtained in the previous phase. Ferreira *et al*. (2017) presented a vehicle routing problem solved in two steps: the definition of demand point clusters as a facility location problem (PLF) and the definition of routes within each cluster, *i.e.*, an asymmetric traveling salesman problem (ATSP). For the ATSP, the simulated annealing and tabu search metaheuristics were used, as well as a hybrid algorithm.

Corn and soy production, which has grown continuously in Brazil, was also researched in the context of territorial partitioning with the intention of analyzing how to increase the number of grain silos. To this effect, Steiner Neto *et al*. (2017) proposed aggregating the municipalities of Paraná State into regions for the effective transport of grains. Their work organized the storage regions, aggregating the municipalities as a multi-objective graph (territories) partitioning problem, aiming to maximize the homogeneity of the storage deficit and minimize the cost of inter-region transport from the production sources to the silos. The authors made use, among other procedures, of genetic algorithms. Endler *et al*. (2018) evaluated the spatial distribution of public daycare centers in the city of Curitiba (Paraná, Brazil), based on three types of analysis: the analysis of the current location of daycare centers, the analysis of the possible expansions for existing units, and the analysis of opening new units. The authors solved the problem using mathematical modeling.

Franco *et al*. (2021) used the concept of the FLP to find the best configuration for clusters of municipalities (*consortia*) for the municipal solid waste (MSW) landfill network in Paraná State, Brazil. To this effect, they used the BILP mathematical model. In 2022, Franco *et al.* used a MILP mathematical model to analyze the current and future scenarios (2021 and 2033) for the location of MSW landfills in the State of Rio Grande do Sul (Brazil).

This work, in comparison with those analyzed above, deals with territorial partitioning aimed at improving the performance of a common building maintenance service contract in Paraná State. To this effect, following data collection and processing, a BILP mathematical model was used for several scenarios, the results of which were analyzed.

Case study

The case study involved a common building maintenance contract for the PRCJ. The PRCJ is an organ of the Judiciary that provides jurisdictional services in the counties located within Paraná State.

Ranked as one of the five largest courts of justice in the country, alongside São Paulo, Minas Gerais, Rio de Janeiro, and Rio Grande do Sul, the PRCJ had, in 2020, a total of 929 judges and 18 592 civil servants (CNJ, 2021). Table 1 shows the classification of these Brazilian justice courts with their corresponding numbers of cases, judges, and civil servants.

According to data presented by the PRCJ regarding the Term of Reference published for the Bidding Notice for Electronic Auction No. 87/2022, as shown by the public in its own survey conducted in 2020, the physical structure consisted of a total of 224 occupied buildings, including units owned, leased, and assigned, distributed over 161 counties, consisting of a total built area of 544 283,79 m². This dimension of the area to be served requires careful optimization procedures to be followed by common building maintenance teams.

Table 1. Classification of the major Justice Courts, highlighting the PRCJ

COURT	New cases	Cases pending	Judges	Civil Servants
São Paulo CJ	4 456 839	19 432 935	2 620	65 179
Minas Gerais CJ	1 428 480	3 940 277	1 085	27 334
Rio de Janeiro CJ	1 463 530	7 897 304	877	24 629
Paraná CJ	**1 287 624**	**3 754 090**	**929**	**18 592**
Rio Grande do Sul CJ	1 095 931	3 035 797	759	16 603

Source: CNJ (2021)

According to data presented by the PRCJ regarding the Term of Reference published for the Bidding Notice for Electronic Auction No. 87/2022, as shown by the public in its own survey conducted in 2020, the physical structure consisted of a total of 224 occupied buildings, including units owned, leased, and assigned, distributed over 161 counties, consisting of a total built area of 544 283,79 m². This dimension of the area to be served requires careful optimization procedures to be followed by common building maintenance teams.

Description of the common building maintenance contract

The PRCJ experienced problems due to the retirement of civil servants from its own staff who worked in the maintenance services of its buildings in Paraná State.

With no interest in holding public tenders for the hiring of operational maintenance workforce, the administration opted to hire a company that would supply outsourced labor to establish service stations for preventive and corrective maintenance operations.

As the solution required outsourced labor, the project was led by the technical engineering and maintenance sector, together with the outsourced services sector, as they had detailed knowledge of policies and management related to outsourcing services.

The following criteria were used as the basic premises with regard to labor laws in order to optimize the human resources that were hired: 1) employees could work a maximum of nine hours and 48 minutes per day, with a 1 h break; 2) activities would be carried out exclusively during daytime; 3) any necessary action on the weekends would be compensated for on working days; and 4) it was not possible to pay expenses or fund overnight stays for employees.

Current scenario (solution)

Due to legal requirements, it was determined that the buildings occupied by the PRCJ should be divided into regions, with the outsourced work teams headquartered in one of the counties belonging to the region in which they operated. The division of Paraná State (which comprises 399 municipalities) into 11 regions (a number obtained through preliminary experiments), is shown in Figure 2. These regions are duly named as follows, beginning on the right: the state capital, Curitiba; Santo Antônio da Platina; Ponta Grossa; Telêmaco Borba; Londrina; Guarapuava; Francisco Beltrão; Maringá; Campo Mourão; Umuarama; and Cascavel.

Figure 2. Map of Paraná State partitioned into 11 regions for building maintenance
Source: Authors

Based on the areas and the number of buildings to be serviced in each region, the number of service stations for the outsourced teams (headquarters/maintenance offices) in each region was determined. Each work team was made up of four people: one supervisor and three building maintenance officers with multiple specialties. These data are presented in Table 2.

Due to the substantial number of buildings and areas served, it was decided that six work teams would be allocated to the Curitiba regional office. There are also outsourced professionals distributed in pairs to permanently service buildings that have a high number of requests for service. It was also determined that it would be necessary to have three administrative supervisors from the outsourced company for direct and daily dialogue with the management and inspection team.

Each work team has a company vehicle for traveling between cities, as well as tools, equipment, and smartphones. These smartphones are used for timekeeping, displacement, and service provision through the company's own application for team management regarding outsourced maintenance services, mainly for evaluating and monitoring displacement and the duration of services.

Services are provided primarily through visits for scheduled preventive maintenance. Emergency services are also provided based on local demand and on assessments by the contract management team regarding their urgency. If the demand is not identified as an emergency, the problem will be recorded in the building's list, to be addressed during the next technical visit by the team for preventive and corrective maintenance services.

Table 2. Team offices characteristics

REGION	Number of buildings	Total area (m²)	Team headquarters	Outsourced team	
				Profes-sionals	Supervi-sors
Curitiba	66	244 248,66	Curitiba	40	9
Ponta Grossa	17	23 960,14	Ponta Grossa	3	1
			União da Vitória	3	1
Telêmaco Borba	11	10 503,06	Telêmaco Borba	3	1
Santo Antônio da Platina	16	21 072,09	Ribeirão Claro	3	1
			Cornélio Procópio	3	1
Londrina	19	61 432,68	Londrina	9	3
Umuarama	17	28 853,97	Cruzeiro do Oeste	6	2
Maringá	23	29 755,78	Maringá	9	3
Campo Mourão	13	19 073,27	Campo Mourão	3	1
Cascavel	15	50 460,39	Cascavel	6	2
			Foz do Iguaçu	3	1
Francisco Beltrão	15	28 715,63	Pato Branco	6	2
Guarapuava	12	26 208,12	Guarapuava	6	2
Total	**224**	**544 283,79**		**103**	**30**

Source: Authors

The scheduled service cycle of each maintenance team is 45 to 60 consecutive days at its regional office, which means that a service routine is conducted in each building six to eight times a year. As these are services of a common nature, with no need for specialized knowledge to evaluate them, but rather only the identification of their performance and result, the documentation regarding the performance of the activities is done directly by the requesting customers, in other words, the claimants of the services, represented by the Assistants of the Court Directors or the supervisors of administrative units.

To determine and select the offices of the work teams in each of the 11 regions, an investigation was conducted with the Court Boards to consider the possibility of using a small room (headquarters/office) to accommodate the structure of the team, simply furnished with lockers and shelves for storing tools and equipment. The initial objective was to allocate the teams to the district after which the region is named. However, due to insufficient space in some buildings, this goal was not achieved.

It should be highlighted that a county is defined as the territory over which the first instance judge has jurisdiction, and it may cover one or more municipalities depending on the number of residents and voters, the caseload, and the territorial extension of the municipalities, among other aspects (CNJ, 2016).

Some counties were chosen to function as the headquarters/offices of maintenance teams according to the availability of office space and distributed as shown in Table 2. Other counties were chosen as host municipalities, so that the teams' maintenance facilities could be properly located, in such a way that the travel distance to the buildings to be served did not exceed 161 km. It is important to note that, due to the large built area, the cities of Londrina, Maringá, Cascavel, and Guarapuava have a maintenance team in permanent operation, *i.e.*, they do not have to travel to other cities.

Figure 3 shows the territorial division of the state into municipalities, indicating, in the colored areas, those that house the county offices, *i.e.*, the locations that have at least one building occupied by the PRCJ which must be serviced within the framework of the building maintenance contract. Thus, the problem in question consists of 161 clients to be served (the counties, represented in color in Figure 3). The other 238 municipalities in the state do not have judiciary structures for maintenance.

Early work on the scope of the contract indicated that the region of Santo Antônio da Platina should have two offices instead of one, and the chosen ones were Ribeirão Claro and Cornélio Procópio. The Cascavel region had an additional team allocated to Foz do Iguaçu to better serve a large local building. Cruzeiro do Oeste was chosen as the headquarters for the Umuarama region, and Pato Branco was chosen as the location of the office in Francisco Beltrão. Finally, União da Vitória was chosen to host an additional team in the Ponta Grossa region due to its considerable distance to the main city, which would require long travel times. All these 14 cities for the maintenance teams are represented in the fourth column of Table 2 and the map in Figure 3, consequently generating 14 areas of action. It should be mentioned that the configuration of the map in Figure 3 is the same as in Figure 2, only with the inclusion of the areas of operation of the regions' additional host municipalities. Figure 3 highlights the data necessary for this work: the 161 counties and the 14 offices.

The contracting process was managed by the PRCJ, following a bidding procedure in the electronic trading modality, in February 2021, and services began in April of the same year.

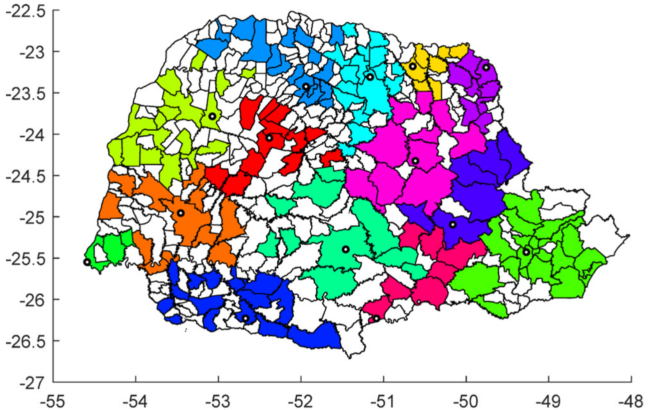


Figure 3. Map of Paraná State, partitioned according to the location of the municipalities for the 14 maintenance offices/headquarters (small circles) and their 14 areas of activity for building maintenance services
Source: Authors

For this scenario, as contracted by the PRCJ, the data regarding the work teams' displacement are presented in Table 3.

Table 3. Displacement of work teams considering the current scenario

	One-way distance (km)
Maximum distance traveled	161,0
Average distance traveled	69,3
Total distance traveled	11 150,0

Source: Authors

Table 3 shows that the sum of all the distances traveled by the maintenance teams from their offices to the serviced buildings is 11 150 km. Out of the distances traveled, the longest was 161 km, and the average was 69,3 km.

Analyzing the clusters in the current solution, we have five offices in Curitiba, Cruzeiro do Oeste, Maringá, Londrina, and Pato Branco, serving around 53% of the counties. The complete distribution of counties regarding the offices in the current solution is presented in Table 4.

Table 4. Characteristics of work teams considering the current scenario

OFFICE	Counties served		Built area (m^2)	
Curitiba	21	13,0%	244 249	44,9%
Cruzeiro do Oeste	17	10,6%	28 854	5,3%
Maringá	17	10,6%	29 756	5,5%
Londrina	16	9,9%	61 433	11,3%
Pato Branco	15	9,3%	28 716	5,3%
Campo Mourão	12	7,5%	19 073	3,5%
Cascavel	12	7,5%	31 995	5,9%
Telêmaco Borba	10	6,2%	10 503	1,9%
Guarapuava	10	6,2%	26 208	4,8%
Ribeirão Claro	8	5,0%	9 251	1,7%
União da Vitória	7	4,3%	9 665	1,8%
Cornélio Procópio	7	4,3%	11 821	2,2%
Ponta Grossa	6	3,7%	14 295	2,6%
Foz do Iguaçu	3	1,9%	18 466	3,4%

Source: Authors

Table 4 also shows that the five offices that serve the fewest counties are Ribeirão Claro, União da Vitória, Cornélio Procópio, Ponta Grossa, and Foz do Iguaçu. Combined, they serve approximately 19,3% of the counties. Regarding the maintenance area, the highest numbers are those in Curitiba, Londrina, and Cascavel, which represent 62,0% of the total area.

Methodology

The methodology used in the study is shown in Figure 4 and comprises four stages: data selection, pre-processing, mathematical modeling, and results evaluation.

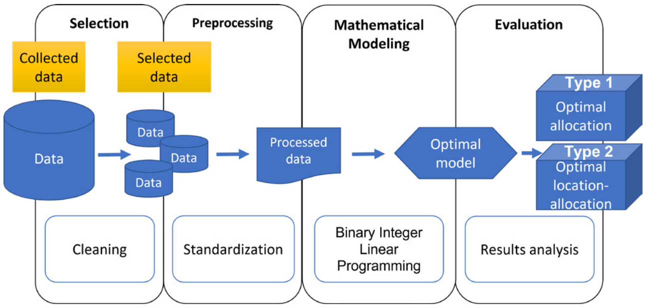


Figure 4. Research methodology
Source: Authors

The collected data include the location of the municipalities, *i.e.*, their geographical centers; the distances between the municipalities; the total areas of building maintenance in the Paraná legal system; and the current configuration of the maintenance teams. With these data, data cleaning and standardization were carried out for the later implementation of the BILP mathematical model for facility location, which was divided into two types: the allocation problem, where only the configuration of maintenance *consortia* is optimized; and the location-allocation problem, where both the location of the offices of the *consortia* and their configuration are optimized.

The data regarding the distances between municipalities were obtained through the Google Distance Matrix API platform, which provides the road distance between a user-defined origin and a destination. A total of 25 760 distances were collected (corresponding to the combination of 161 municipalities, two by two).

Facility location problem

The facility location problem (FLP) is applied in the planning and optimization of the location of production facilities – including public agencies – and considers the distances between origin and destination. Therefore, in addition to determining the number of facilities to be opened, it is also responsible for determining their locations, which makes it an NP-hard problem (Laporte, 2019, Franco *et al.*, 2022, Franco and Steiner, 2023).

The FLP can be solved using BILP, in accordance with the model presented in Equations (1) to (5), where J and I are the sets of consumers (counties) and facilities (headquarters), respectively. The decision variables are x_{ij}, which represents the number of counties being assigned to the headquarters, and y_i, denoting which headquarters are open.

$$\min Z = \sum_{i \in \mathrm{I}} \sum_{j \in \mathrm{J}} c_{ij} x_{ij} \tag{1}$$

$$\sum_{i \in \mathrm{I}} x_{ij} = 1 \qquad \forall j \in \mathrm{J} \tag{2}$$

$$x_{ij} \leq y_i \qquad \forall i \in \mathrm{I}, j \in \mathrm{J} \tag{3}$$

$$x_{ij} \geq 0 \qquad \forall i \in \mathrm{I}, j \in \mathrm{J} \tag{4}$$

$$y_i \in \{0,1\} \qquad \forall i \in \mathrm{I} \tag{5}$$

In this general model, Equation (1) minimizes the cost c_{ij} of transport between the consumers j and the facilities i. Equation (2) guarantees that each consumer (county) j will be served by only one facility (headquarters). Constraint (3) ensures that each client (county) is designated to a facility only if it is open. Finally, Equations (4) and (5) define the decision variables' domain: x_{ij} must be equal to zero or positive, and y_i must be binary (the county can 'receive' or 'not receive' the facility).

For the optimization, LINGO 19 software was used on an Intel Core i7-2600 computer with 16 GB RAM and Windows 10. The execution time varied between 0,14 and 0,87 s.

Results

Regarding the optimal definition of the regions in Paraná State to be served by the building maintenance teams, three scenarios were taken into account. It should be mentioned again that the regional offices had to be defined while minimizing the distances to be covered by the teams from their places of origin (headquarters/offices) to the maintenance locations. Although the centroids of the municipalities were used as a reference for calculating the distances, this does not affect the performance of the solution, since the maintenance teams move mainly between the districts, and it is this relationship between municipalities that affects the choice of the facilities.

Scenario 1

The first scenario optimizes the maintenance offices' areas of operation while keeping their current location unchanged. The result is shown in Figure 5 below.

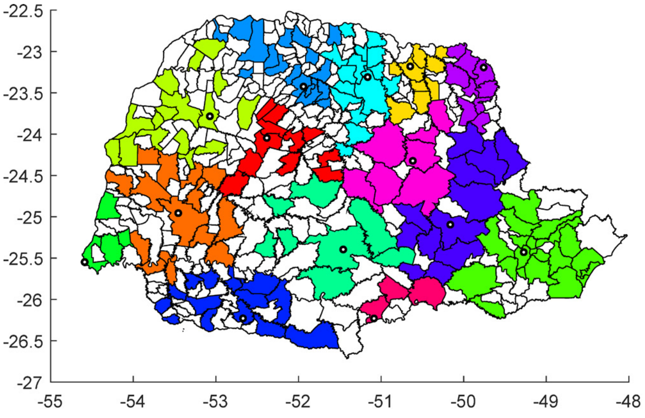


Figure 5. Optimization of the establishment of regional offices considering the current location of the 14 host municipalities of the maintenance offices (Scenario 1)
Source: Authors

The main difference regarding the existing solution (Figure 3) is a better concentration of municipality clusters in certain regions, as well as reductions of 6,20% in the average distance traveled and 6,13% in total distance traveled. The results of optimized Scenario 1 are shown in Table 5.

Table 5. Displacement of work teams under Scenario 1

Parameter	One-way distance (km)	Variation
Maximum distance traveled	155,0	-3,7%
Average distance traveled	65,0	-6,2%
Total distance traveled	10 467,0	-6,1%

Source: Authors

An analysis of the clusters in Scenario 1 shows five offices serving around 52% of the counties: Curitiba, Maringá, Cruzeiro do Oeste, Cascavel, and Londrina, unlike the current solution (Table 4), which had Pato Branco among the five largest instead of Cascavel. The entire distribution of counties among the offices in Scenario 1 is presented in Table 6.

Table 6. Characteristics of work teams under Scenario 1

OFFICE	Counties served		Built area (m²)	
Curitiba	21	13,0%	244 249	44,9%
Maringá	17	10,6%	30 682	5,6%
Cruzeiro do Oeste	16	9,9%	26 463	4,9%
Cascavel	15	9,3%	36 413	6,7%
Londrina	15	9,3%	60 897	11,2%
Pato Branco	14	8,7%	27 736	5,1%
Ponta Grossa	12	7,5%	18 956	3,5%
Campo Mourão	11	6,8%	16 478	3,0%
Cornélio Procópio	9	5,6%	11 422	2,1%
Ribeirão Claro	9	5,6%	11 419	2,1%
Guarapuava	8	5,0%	24 379	4,5%
Telêmaco Borba	7	4,3%	8 401	1,5%
Foz do Iguaçu	4	2,5%	20 065	3,7%
União da Vitória	3	1,9%	6 726	1,2%

Source: Authors

Table 6 also shows that the five offices serving the fewest counties are Ribeirão Claro, Guarapuava, Telêmaco Borba, Foz do Iguaçu, and União da Vitória, which, combined, serve approximately 19,3% of all the counties. Regarding the area in which the building maintenance will be conducted, the highest numbers continue to be those of Curitiba, Londrina, and Cascavel, representing 62,8% of the total area served.

Scenario 1 could be easily implemented in practice, as the municipalities with offices would remain unchanged and there would be few adjustments to be made to the areas of operation. This could mean a reduction of about 6% in the total distance traveled by the teams.

Scenario 2

In the second scenario, the current configuration of the 11 regions (areas of operation), presented in the first column of Table 2, was fixed, aiming to optimize the location of 11 offices rather than 14. In other words, there should only be 11 areas of operation in accordance with the original division of regions to gauge the distances traveled with a lower number of offices. The result is shown in Figure 6.

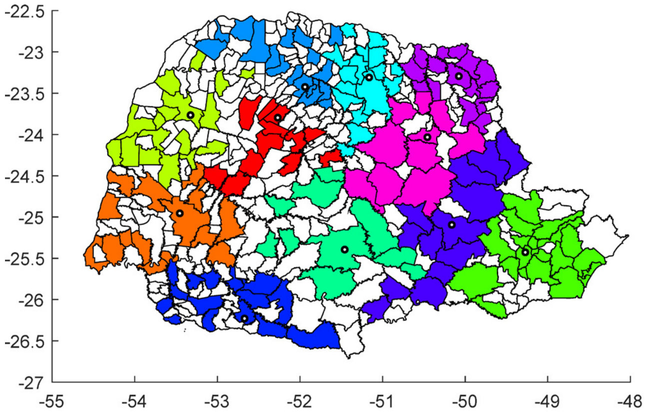


Figure 6. Optimization of the location of the host municipalities for 11 maintenance offices, maintaining the current configuration of regions (Scenario 2)
Source: Authors

By comparing Scenario 2 with the current scenario, an increase can be observed in the total distance traveled, *i.e.*, around 0,56%, due to the decrease in the number of maintenance offices. Analyzing the maintenance cost of the new scenario with 11 offices (rather than the current 14) would make it be possible to identify if this reduction is worthwhile. The findings in terms of the distances traveled are shown in Table 7.

Table 7. Displacement of work teams under Scenario 2

Parameter	One-way distance (km)	Variation
Maximum distance traveled	208,0	+29,2%
Average distance traveled	69,6	+0,43%
Total distance traveled	11 213,0	+0,56%

Source: Authors

In this scenario, there are two routes that do not comply with the constraint on the maximum distance to be covered (161 km). These are the routes between Curiúva and Cândido de Abreu, with 197 km, and between Ponta Grossa and União da Vitória, with 208 km.

An analysis of the clusters in Scenario 2 shows five offices serving around 53,4% of the counties: Curitiba, Maringá, Umuarama, Londrina, and Cascavel. Table 8 presents the complete distribution of counties among the offices in Scenario 2.

Table 8. Characteristics of work teams under Scenario 2

OFFICE	Counties served		Built area (m^2)	
Curitiba	21	13,0%	244 249	44,9%
Maringá	17	10,6%	29 756	5,5%
Umuarama	**17**	**10,6%**	**28 854**	**5,3%**
Londrina	16	9,9%	61 433	11,3%
Cascavel	15	9,3%	50 460	9,3%
Pato Branco	15	9,3%	28 716	5,3%
Santo Antônio da Platina	**15**	**9,3%**	**21 072**	**3,9%**
Ponta Grossa	13	8,1%	23 960	4,4%
Engenheiro Beltrão	**12**	**7,5%**	**19 073**	**3,5%**
Curiúva	**10**	**6,2%**	**10 503**	**1,9%**
Guarapuava	10	6,2%	26 208	4,8%

Source: Authors

The five offices serving the fewest counties in this scenario are União da Vitória, Foz do Iguaçu, Pontal do Paraná, Santo Antônio do Sudoeste, and Guarapuava. Combined, they serve approximately 18,6% of the total number of counties and are responsible for 13,8% of the maintenance area.

Scenario 3 would be the most interesting for implementation in terms of reducing transportation costs. However, as practically nothing was maintained regarding the current scenario (the definition of both the offices and the areas of operation), the entire procedure concerning the location of the offices, distances traveled, and number of teams presented in the Terms of Reference of the bids for the contract would have to be reassessed. Therefore, this case would also require an in-depth study by the PRCJ administration to gauge whether the reformulation work would have legal and technical support, especially considering that it would ensure a reduction of over 12% in the total distance traveled per day by the teams.

Conclusions and future research

The aim of this study was to evaluate the possible application of an approach to improving the performance of a common building maintenance service contract for the PRCJ, which has 224 buildings with a total built area of 544 283,79 m^2, distributed over 161 counties in the State of Paraná.

To this effect, the BILP mathematical model was used, together with the FLP, to optimize the existing territorial partitioning of Paraná State into 11 regions with 14 host municipalities (maintenance offices or headquarters). Three scenarios were considered. In Scenario 1, the areas of operation of the 14 host-municipalities for the maintenance offices were optimized. In Scenario 2, the locations of 11 maintenance offices (instead of 14) were optimized, maintaining the original areas of the regions in order to gauge the distances traveled for this case. Meanwhile, Scenario 3 optimized both the location of the host municipalities for the 14 maintenance offices and their areas of operation. A summary of the solutions found is shown in Table 11.

Table 11. Comparison of costs (distances) regarding the three scenarios

SCENARIO	Maximum distance (km)	Average distance (km)	Total distance (km)
Current solution 11 regions 14 offices	161	69,3	11 150
Scenario 1 11 optimized regions 14 offices maintained	155	65,0	10 467
Scenario 2 11 regions maintained 11 optimized offices	208	69,6	11 213
Scenario 3 14 optimized regions 14 optimized offices	164	60,7	9 775

Source: Authors

Although all the solutions obtained have advantages and disadvantages, it can clearly be perceived that the best scenario in terms of total distance and average distance traveled is Scenario 3, which optimizes the location of the 14 maintenance offices, as well as the configuration of their areas of operation. In this scenario, the teams will have to cover a total distance of 9 775 km per day (instead of the current 11 150 km), from their headquarters to the maintenance locations, achieving a reduction of around 12,3% in the total distance traveled when compared to the solution with the current contract.

With the solution in Scenario 3, the PRCJ would minimize, in addition to the distance, the operating and maintenance costs of displacing the teams, in addition to the time spent on the corresponding trips. Furthermore, the time spent providing maintenance services in each building would increase, which also implies a higher level of satisfaction for the people served through the contract.

In addition to the fact that Scenario 3 is the best, it was shown that starting an optimization procedure 'from scratch' will always yields the best results. However, starting from scratch is costly, and, whenever possible, one should make use of existing facilities/criteria.

Future research could explore the fixed upkeep cost of maintenance offices, with a view to optimizing their number as a function of the trade-off between fixed and variable costs in scenarios for optimizing facility location. Another future research could consider the location of the buildings (instead of the municipalities' geographical center) in the mathematical model which will require a (slightly) larger dataset.

Acknowledgements

The study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

Alexandre A. Steiner: Conceptualization, Data curation, Writing - original draft, Writing - review and editing, Visualization. David G. de B. Franco: Methodology, Software, Formal analysis, Writing - original draft, Writing - review and editing, Visualization. Elpídio O. B. Nara: Writing - review and editing, Validation, Supervision. Maria T. A. Steiner: Project administration, Funding acquisition, Writing - original draft, Writing - review and editing, Validation, Supervision.

References

Associação brasileira de normas técnicas (ABNT) (1999). *NBR 5674. Manutenção de edificações-procedimentos*. ABNT.

Associação brasileira de normas técnicas (ABNT) (2011). *NBR 14037. Diretrizes para elaboração de manuais de uso, operação e manutenção das edificações – Requisitos para elaboração e apresentação dos conteúdos*. ABNT.

Associação brasileira de normas técnicas (ABNT) (2012). *NBR 5674*. Segunda edição. Manutenção de edificações – requisitos para o sistema de gestão de manutenção. ABNT.

Assis, L. S., França, P. M., and Usberti, F. L. (2014). A redistricting problem applied to meter reading in power distribution networks. *Computers & Operations Research*, *41*, 65-75. https://doi.org/10.1016/j.cor.2013.08.002

Barbosa, P. B., and Pusch, J. (2011). *Da intenção de projeto ao uso do edifício*: *a busca da excelência profissional*. Programa de Excelência em Projetos CREA-PR.

Benzarti, E., and Dallery, E. S. (2013). Operations management applied to home care services: Analysis of the districting problem. *Decision Support Systems*, *55*(2), 587-598. https://doi.org/10.1016/j.dss.2012.10.015

Bergey, P. K., Ragsdale, C. T., and Hoskote, M. (2003). A decision support system for the electrical power districting problem. *Decision Support Systems*, *36*(1), 1-17. https://doi.org/10.1016/S1344-6223(02)00033-0

Bozkaya, B., Erkut, E., and Laporte, G. (2003). A tabu search heuristic and adaptative memory procedure for political districting. *European Journal of Operational Research*, 144(1), 12-26. https://doi.org/10.1016/S0377-2217(01)00380-0

Brandão, N. L. S., and Santos, D. G. (2020). Manutenção predial em edificações públicas: um mapeamento sistemático da literatura. *Encontro Nacional de Tecnologia do Ambiente Construído*, *18*(1),1251. https://doi.org/10.46421/entac.v18i.1251

Conselho Nacional de Justiça (CNJ) (2016). *Saiba a diferença entre comarca, vara, entrância e instância*. https://www.cnj.jus.br/cnj-servico-saiba-a-diferenca-entre-comarca-vara-entrancia-e-instancia/

Conselho Nacional de Justiça (CNJ) (2021). *Justiça em números 2021*. https://www.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/

D'Amico, S. J., Wang, S.-J., Batta, R., and Rump, C. M. (2002). A simulated annealing approach to police district design. *Computers & Operations Research*, *29*(6), 667-684. https://doi.org/10.1016/S0305-0548(01)00056-9

Datta, D., Deb, K., Fonseca, C. M., Lobo, F. G., Condado, P. A., and Seixas, J. (2007). Multi-objective evolutionary algorithm for land-use management problem. *International Journal of Computational Intelligence Research*, *3*(4), 371-384.

Datta, D., Malczewski, J., and Figueira, J. R. (2012). Spatial aggregation and compactness of census areas with a multiobjective genetic algorithm: a case study in Canada. *Environment and Planning B: Planning and Design*, *39*, 376-392. https://doi.org/10.1068/b38078

Dos Santos, V. A. (2010). *Prontuário para a Manutenção Mecânica*. Ícone Editora.

Endler, K. D., Scarpin, C. T., and Steiner, M. T. A. (2018). Evaluation of the of public network location day care center: Models and case study. *IEEE Latin America Transactions*, *16*, 2013-2019. https://doi.org/10.1109/TLA.2018.8447370

Ferreira, J. C., Steiner, M. T. A., and Guersola, M. S. (2017). A vehicle routing problem solved through some metaheuristics procedures: A case study. *IEEE Latin America Transactions*, *15*, 943-949. https://doi.org/10.1109/TLA.2017.7910210

Fontoura, L. H. N., Santos, C. H. S., and Oliveira, C. C. (2019). Manutenção de prédios públicos: uma questão de gestão. *Revista eletrônica e Administração*, *18*(2), 322-346. http://periodicos.unifacef.com.br/rea/article/view/1648

Franco, D. G. de B., Steiner, M. T. A., and Assef, F. M. (2021). Optimization in waste landfilling partitioning in Paraná State, Brazil. *Journal of Cleaner Production*, *283*, 125353. https://doi.org/10.1016/j.jclepro.2020.125353

Franco, D. G. B., Steiner, M. T. A., Fernandes, R., and Nascimento, V. F. (2022). Modeling municipal solid waste disposal consortia on a regional scale for present and future scenarios. *Socio-economic Planning Sciences*, *82*(B), 101333. https://doi.org/10.1016/j.seps.2022.101333

Franco, D. G. B., and Steiner, M. T. A. (2023). Optimization of municipal solid waste transportation in the State of Paraná: Rethinking the location of landfills based on mathematical modeling. *Revista Engenharia Sanitaria e Ambiental*, *27*(5), 987-993. https://doi.org/10.1590/S1413-415220210282

Galvão, L. C., Novaes, A. G. N., Cursi, J. E., and Souza, J. C. (2006). A multiplicatively-weighted Voronoi diagram approach to logistics districting. *Computers & Operations Research*, *33*(1), 93-114. https://doi.org/10.1016/j.cor.2004.07.001

Haugland, D., Ho, S. C., and Laporte, G. (2007). Designing delivery districts for the vehicle routing problem with stochastic demands. *European Journal of Operational Research*, *18*(3), 997-1010. https://doi.org/10.1016/j.ejor.2005.11.070

Laporte, G. (2019). *Location science*. Springer International Publishing. https://doi.org/10.1007/978-3-030-32177-2

Muyldermans, L., Cattrysse, D., Oudheusden, D. V., and Lotan, T. (2002). Districting for salt spreading operations. *European Journal of Operational Research*, *139*(3), 521-532. https://doi.org/10.1016/S0377-2217(01)00184-9

Novaes, A. G. N., Cursi, J. E. S., Silva, A. C. L., and Souza, J.C. (2009). Solving continuous location-districting problems with Voronoi diagrams. *Computers & Operations Research*, *36*(1), 40-59. https://doi.org/10.1016/j.cor.2007.07.004

Poli, C. M. B. (2017). *Manual de uso, operação e manutenção das edificações residenciais*: *avaliação do conteúdo a fim de aumentar a utilidade para a construção civil e para o usuário*. UFRGS. http://hdl.handle.net/10183/170982

Ricca, F., and Simeone, B. (2008). Local search algorithms for political districting. *European Journal of Operational Research*, *189*(3), 1409-1426. https://doi.org/10.1016/j.ejor.2006.08.065

Ricca, F., Scorazzi, A., and Simeone, B. (2008). Weighted Voronoi region algorithms for political districting. *Mathematical and Computer Modelling*, *48*(9-10), 1468-1477. https://doi.org/10.1016/j.mcm.2008.05.041

Ríos-Mercado, R. Z., and López-Pérez, F.J. (2013). Commercial territory design planning with realignment and disjoint assignment requirements. *Omega*, *41*(3), 525-535. https://doi.org/10.1016/j.omega.2012.08.002

Salazar-Aguilar, M. A., Ríos-Mercado, R. Z., and González-Velarde, J. L. (2011). A bi-objective programming model for designing compact and balanced territories in commercial districting. *Transportation Research – Part C*, *19*(5), 885-895. https://doi.org/10.1016/j.trc.2010.09.011

Shirabe, T. (2012). Prescriptive modeling with map algebra for multi-zone allocation with size constraints. *Computers, Environment and Urban Systems*, *36*(5), 456-469. https://doi.org/10.1016/j.compenvurbsys.2011.12.003

Steiner Neto, P. J., Datta, D., Steiner, M. T. A., Canciglieri Júnior, O., Figueira, J. R., Detro, S. P., and Scarpin, C. T. (2017). A multi-objective genetic algorithm-based approach for location of grain silos in Paraná State of Brazil. *Computers & Industrial Engineering*, *111*, 381-390. https://doi.org/10.1016/j.cie.2017.07.019

Tavares-Pereira, F., Figueira, J. R., Mousseau, V., and Roy, B. (2007). Multiple criteria districting problems. *Annals of Operations Research*, *154*, 69-92. https://doi.org/10.1007/s10479-007-0181-5

Vecchi, T. P. B., Fukunaga, D. S., Constantino, A., and Steiner, M. T. A. (2016). A sequential approach for the optimization of truck routes for solid waste collection. *Process Safety and Environmental Protection*, *102*, 238-250. https://doi.org/10.1016/j.psep.2016.03.014