Universidad Nacional de Trujillo

Journal of Advanced Mining Modeling (JAMM)

Journal website: https://revistas.unitru.edu.pe/index.php/jamm/index
Vol. 2, Nº 1, pp. 9-23, January – June 2026

Journal of Advanced Mining Modeling (JAMM)

Prediction of concrete compressive strength using multiple lineal regression (OLS) and data preprocessing

Adrian Ruiz-Victorio[1*], Rosas Simón Cruz-Romero[2], Orlando Alex Siccha-Ruiz[3]

[1.] Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de, Trujillo, Trujillo, Perú
[2.] Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de, Trujillo, Trujillo, Perú
[3.] Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de, Trujillo, Trujillo, Perú

* Corresponding author: h5053401521@unitru.edu.pe (A. Ruiz-Victorio)

Received: December 19, 2025 Revised: December 29, 2025 Accepted: December 1, 2026 Published online: December 1, 2026

Abstract

The safety and stability of ground support and lining systems are critical aspects of mining engineering projects, due to the high structural demands imposed by operational conditions, particularly in underground mining. The main objective of this study was to evaluate the capability of multiple linear regression models (OLS) to predict concrete compressive strength, analyzing the effect of different data preprocessing techniques on their predictive performance. A quantitative approach was adopted, based on the analysis of an experimental dataset comprising more than 1,000 concrete mix designs, considering variables related to material proportions and curing age. Three modeling configurations were developed and compared: untransformed predictors, standardized predictors, and predictors transformed using the log(x+1) function. Model performance was assessed using Mean Squared Error (MSE) and the coefficient of determination (R^2). The results indicate that models built with original and standardized data exhibited similar behavior, with R^2 values close to 0.351 on the test dataset. In contrast, logarithmic transformation significantly improved model performance, reducing the test MSE to approximately 55 MPa^2 and increasing R^2 to values close to 0.750, demonstrating a substantial enhancement in accuracy and generalization capability. These results support the use of the proposed model as a predictive tool for concrete quality control in mining applications. In conclusion, multiple linear regression combined with appropriate logarithmic preprocessing represents an efficient and reliable alternative for estimating concrete compressive strength, preserving model interpretability and providing practical support for technical decision-making in mining engineering.

Keywords: Compressive strength; Concrete; Linear regression; Data preprocessing; Machine learning.

1. Introduction

Concrete is an essential material in the mining industry, valued for its accessibility, reasonable cost, and adequate structural performance, particularly for the foundations of concentrator plants, leaching platforms, ramps, reinforced slopes, and other mining-related applications. According to technical studies, the compressive strength of concrete is one of the most critical parameters in mining engineering. In this context, concrete, especially shotcrete, plays a vital role in operational safety, as its mechanical performance directly affects the stability of the rock mass and the mitigation of geomechanical risks [1]. In underground mining, variability in concrete quality can lead to significant consequences, ranging from premature support failures to increased operational costs due to material overuse or the need for rework. Therefore, reliably predicting concrete strength based on mix proportions and curing conditions is a highly relevant technical challenge for mining and geotechnical engineers [2]. Traditionally, this parameter has been assessed through experimental testing methods which, although accurate, require time, resources, and planning constraints that are not always compatible with the operational dynamics of a productive mine [3]. Numerous studies have demonstrated that

concrete strength depends on many interrelated variables, such as cement content, water-to-binder ratio, the use of mineral additives, chemical admixtures, and curing age. In mining applications, these variables tend to exhibit greater variability due to logistical constraints, underground environmental conditions, and differences in placement methods. This increases the uncertainty associated with the mechanical performance of concrete [4]. Given these considerations, the use of statistical models and machine learning techniques has emerged as a viable alternative for estimating compressive strength based on historical mix records.
Numerous previous studies have demonstrated that regression models can identify meaningful relationships between the components of concrete and its final strength, provided that the dataset is properly prepared and analyzed [5]. Among these models, linear regression stands out for its conceptual simplicity and ease of implementation in engineering applications, particularly in industrial environments where transparency in results is essential [6]. In the mining context, the interpretability of predictive models takes on particular importance, as it allows engineers to directly understand the influence of each variable on material behavior and to support technical decisions related to mix design and quality control. Unlike complex "black-box" models, linear approaches facilitate technical validation and integration into standard operating procedures typical of mining projects [7]. However, the direct use of linear models without prior treatment often presents limitations associated with statistical issues, which can undermine the model's generalization capability and result in unreliable strength estimates [8]. To address this, data preprocessing techniques such as standardization and logarithmic transformations have been proposed to enhance the performance of linear models [9]. Research focused on concrete mix design has shown that logarithmic transformation helps reduce data dispersion, stabilize variance, and reveal relationships that may not be evident in the original data scale [10]. This approach is particularly relevant in mining, where real operating conditions rarely meet ideal assumptions, and where model robustness under unbalanced datasets is a key factor for practical application [11]. In underground mining, shotcrete plays a fundamental role as an immediate support element. Therefore, its performance is evaluated not only in terms of final strength but also by early strength criteria established in mine operational standards. In practice, minimum values of approximately 2 MPa at 2 hours and final strengths exceeding 28 MPa at 28 days are commonly required to ensure both initial and long-term excavation stability. In this context, the development of predictive tools based on historical mix data enables anticipation of such performance criteria, optimization of quality control, and reduction of exclusive reliance on destructive testing providing a direct operational advantage in underground environments where time and safety are critical factors. This study aims to evaluate the performance and predictive capacity of classical linear regression models for estimating the compressive strength of concrete used in mining engineering applications, by systematically comparing the effects of different data preprocessing methods. Using a widely adopted and validated dataset from the literature, the goal is to identify which transformations enhance model accuracy, generalization, and interpretability, ultimately offering a practical analytical tool for concrete mix design and quality control in contemporary mining operations [12].

2. Materials and methods

This study is based on a real experimental dataset comprising more than 1,000 concrete mixes, designed to analyze and predict the compressive strength of the material in technical contexts related to mining engineering, such as underground or surface support, tunnel lining, and auxiliary mining structures. Each entry in the dataset represents a concrete mix characterized by its composition and curing age, along with the measured value of compressive strength in megapascals (MPa), a fundamental mechanical property for assessing structural stability and safety in mining excavations [13]. The included variables reflect components commonly used in concrete mixes for mining operations, where placement conditions, curing, and quality control often exhibit greater variability compared to conventional environments [14].

2.1 Data and variables

2.1.1 Variables in the experimental dataset

Table 1 provides a description of the predictor variables and the target variable considered in this study.

Table 1. Variables in the experimental dataset

Variable	Description	Unit

Cement	Cement content in the mix	Kg/m³
Slag	Ground granulated blast furnace slag	Kg/m³
FlyAsh	Fly ash	Kg/m³
Water	Water content	Kg/m³
Superplasticizer	Superplasticizer admixture	Kg/m³
CoarseAgg	Coarse aggregate	Kg/m³
FineAgg	Fine aggregate	Kg/m³
Age	Curing age	Días
Strength	Concrete compressive strength	MPa

2.2. Data preparation and partitioning

Prior to model development, the dataset underwent a numerical verification process to ensure consistency of magnitudes and variable coherence. Subsequently, the dataset was split into two subsets: training and testing, following a fixed partitioning scheme that enables evaluation of the model's performance on observations not used during fitting [15].
This procedure replicates a realistic scenario, in which predictive models must maintain their generalization capacity when exposed to new mix conditions and operational variations inherent to the mining environment [16].

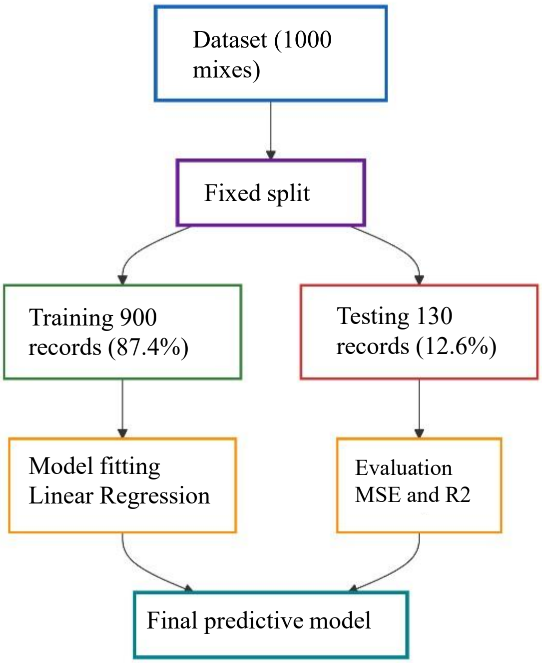

Figure 1. Schematic of the data partitioning process

2.3 Data preprocessing

To assess the impact of preprocessing techniques on the predictive performance of the model, three distinct input data configurations were evaluated: (i) untransformed predictors, (ii) standardized predictors, and (iii) predictors transformed using the log(x+1) function. Standardization was performed using the following expression:

$$Z_i = \frac{X_i - \mu}{\sigma} \quad (1)$$

Where:

Z_i – standardized value,
X_i – original value of the variable,
μ – mean of the variable,
σ – standard deviation of the variable.

2.3.1 Logarithmic transformation applied to predictors

This transformation is used to reduce data skewness and stabilize variance, thereby improving the performance of linear models on real-world experimental datasets [18].

$$X'_i = \log(X_i 1) \tag{2}$$

Where:
X'_i – logarithmic transformation,
log – logarithm.

2.4 Multiple linear regression modeling (OLS)

The estimation of concrete compressive strength was carried out using a multivariate linear regression model, fitted using an Ordinary Least Squares (OLS) algorithm [19]. The general form of the model is expressed as:

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \varepsilon \tag{3}$$

Where:
y – concrete compressive strength (MPa),
β_o – intercept term,
β_i – regression coefficients,
X_i – predictor variables,
ε – error term.
This approach was chosen for its high interpretability an essential feature in mining engineering since it allows for the direct identification of how each concrete component influences mechanical performance, thus supporting decision-making in mix design and quality control [20].

2.5 Evaluation metrics

These metrics allow for the quantification of model accuracy and generalization capability critical aspects for mining applications, where prediction errors can compromise both structural safety and operational efficiency [21]. The performance of the models was evaluated using Mean Squared Error (MSE) and the coefficient of determination (R^2), both of which are discussed below.

2.5.1 Mean squared error (MSE)

MSE measures the overall prediction error of the model. It quantifies how far the model's predicted values deviate from the actual compressive strength values of the concrete. This is especially important in mining, as underestimations may compromise support safety, while overestimations may lead to unnecessary material costs.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - y'_i)^2 \tag{4}$$

Where:

MSE – mean squared error,
n – total number of evaluated observations,
y_i – actual measured compressive strength,
y'_i – predicted compressive strength,
$(y_i - y'_i)^2$ – individual squared error.

2.5.2 Coefficient of determination (R^2)

R^2 is a key metric that quantifies the explanatory power of the model i.e., the proportion of variability in compressive strength that can be explained by the set of predictor variables. Studies have shown that a value close to 1 indicates the model successfully explains most of the mechanical behavior of the concrete, which is fundamental when evaluating the structural reliability of support systems in mining operations subject to geomechanical stresses.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y'_i)^2}{\sum_{i=1}^{n}(y_i - y''_i)^2} \quad (5)$$

Where:
R^2 – coefficient of determination,
y_i – actual measured compressive strength,
y'_i – predicted compressive strength,
y''_i – mean of the actual compressive strength,
$(y_i - y'_i)^2$ – sum of squared prediction errors,
$(y_i - y''_i)^2$ – total variance in the dataset.

3. Results

3.1. Predictive performance of the models

The performance of the three evaluated models untransformed predictors, standardized predictors, and predictors transformed using log(x+1) is presented in Table 2, which reports the Mean Squared Error (MSE) and R^2 values for both training and test sets.

Table 2. Model performance in terms of MSE and R^2

Model configuration	Train MSE	Test MSE	Train R^2	Test R^2
Untransformed predictors	105.20	142.67	0.624	0.351
Standardized predictors	105.20	142.67	0.624	0.351
Log(x+1) predictors	56.39	55.06	0.799	0.750

The results indicate that the models using untransformed and standardized predictors perform almost identically. In both cases, the test set prediction error is high, and the R^2 values suggest a limited ability to explain the variability in compressive strength.

In contrast, the model using log(x+1) transformed predictors demonstrates a substantial reduction in error, with MSE values nearly half those obtained from the other approaches. Additionally, the increase in R^2 on the test set highlights a significant improvement in the model's generalization capability.

3.2. Graphical comparison of mean squared error

Figure 2 presents a visual comparison of the MSE values for the three models, in both the training and test sets. It is clearly observed that the models with untransformed and standardized predictors exhibit high error values and a pronounced gap between training and test performance, suggesting limited robustness.
In contrast, the model using the logarithmic transformation shows similar MSE values across both datasets, indicating more stable and consistent behavior when applied to unseen data.

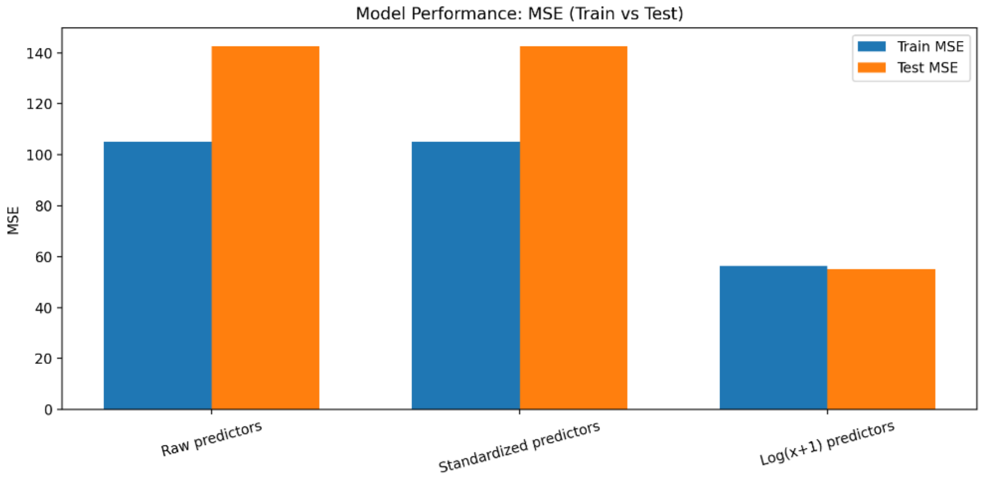

Figure 2. Comparison of mean squared error (MSE) for the three evaluated models

3.3. Analysis of the coefficient of determination

Figure 3 presents a comparison of the R^2 coefficient for the three modeling approaches. The models using untransformed and standardized predictors explain only a limited portion of the variability in compressive strength within the test set, with values around 0.351. In contrast, the model employing log(x+1) transformed predictors explains approximately 75% of the variability observed in the test data, indicating a significantly better fit and a more accurate representation of the relationships between input variables and compressive strength.

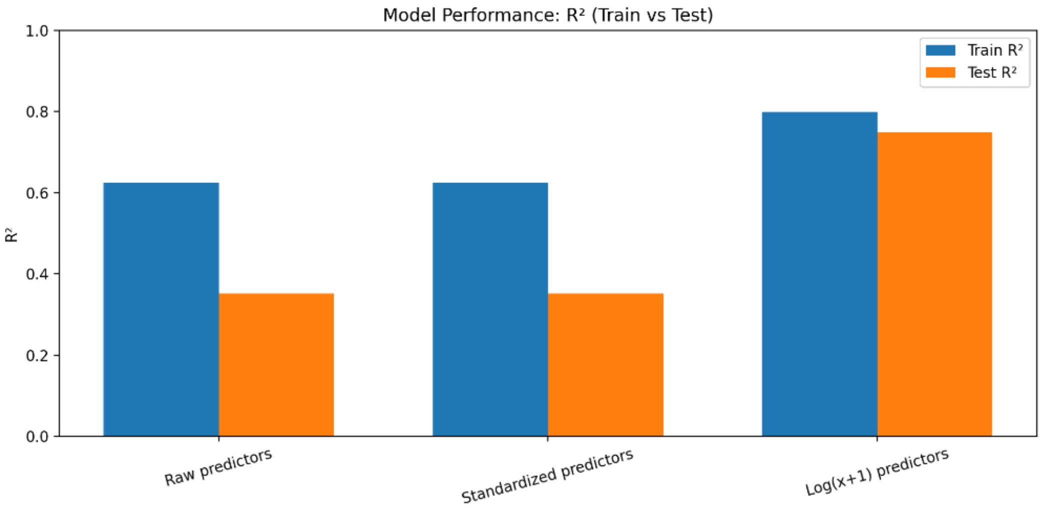

Figure 3. Comparison of the coefficient of determination (R^2) for training and test sets

3.4. Statistical significance

The statistical significance of the predictor variables was evaluated based on the models fitted using Ordinary Least Squares (OLS). For each preprocessing configuration, p-value tables were generated corresponding to the models with untransformed, standardized, and log(x+1) transformed predictors.

Figure 4 displays the $-\log_{10}$(p-value) plot for the model with standardized predictors, where a clear distinction in the statistical relevance of the variables can be observed. Similarly, graphical results for the models using untransformed and logarithmically transformed predictors are presented in Figures 5 and 6, respectively.

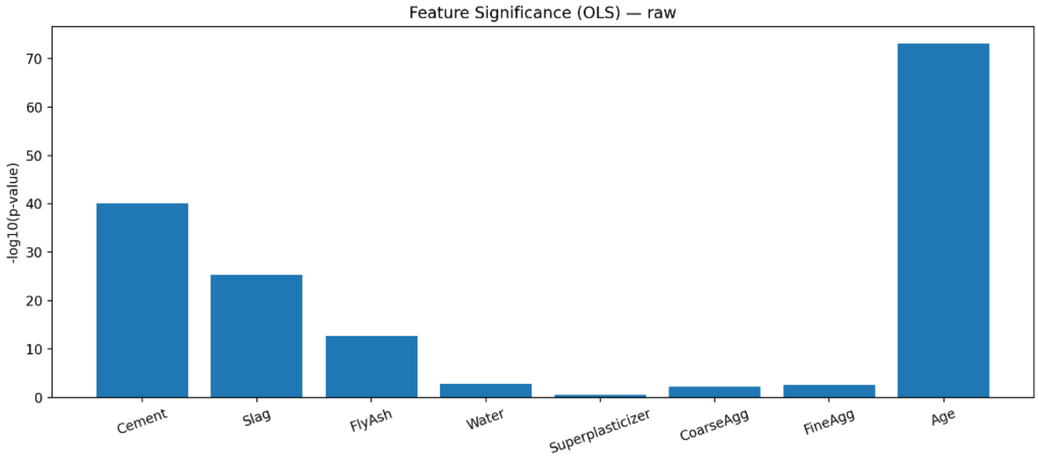

Figure 4. Statistical significance of predictor variables in the model using untransformed predictors, expressed as $-\log_{10}$(p-value)

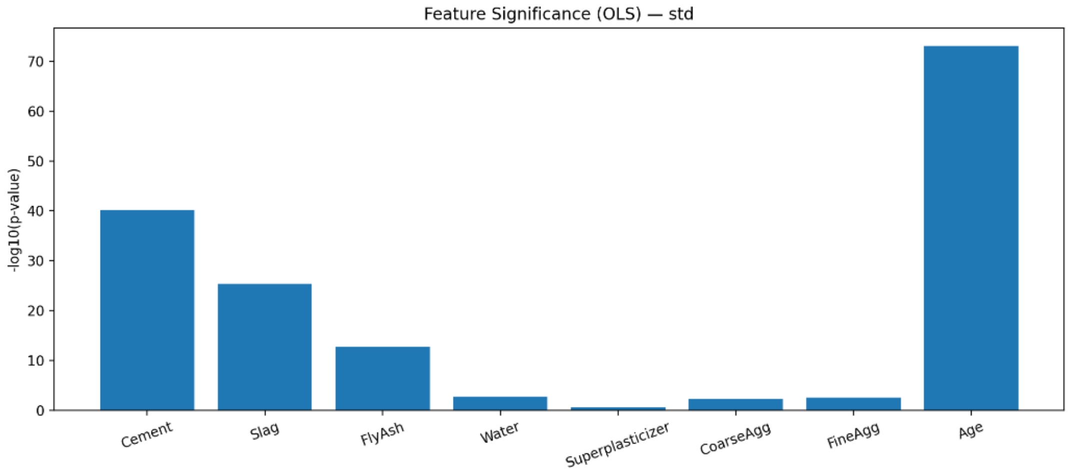

Figure 5. Statistical significance of predictor variables in the model using standardized predictors, expressed as $-\log_{10}$(p-value)

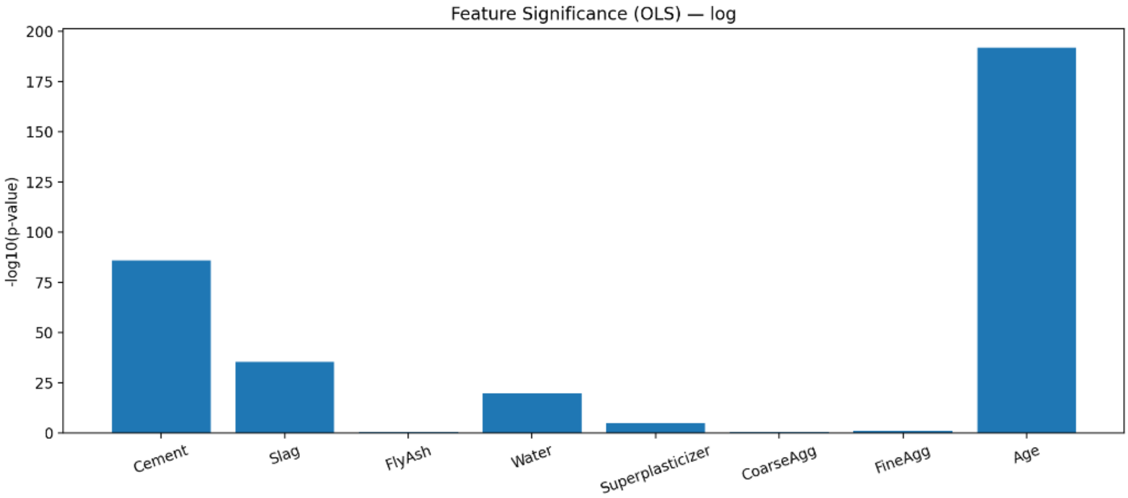

Figure 6. Statistical significance of predictor variables in the model using log(x+1) transformation, expressed as $-\log_{10}$(p-value)

The p-values obtained for each model are detailed in Tables 3, 4, and 5.

Table 3. p-values for the model with untransformed predictors

Predictor	p-value
Const	7.082e-02
Cement	6.971e-41
Slag	4.145e-26
FlyAsh	1.895e-13
Water	1.918e-03
Superplasticizer	2.461e-01
CoarseAgg	5.443e-03
FineAgg	2.738e-03
Age	9.516e-74

Table 4. p-values for the model with standardized predictors

Predictor	p-value
Const	0.000e+00
Cement	6.971e-41
Slag	4.145e-26
FlyAsh	1.895e-13
Water	1.918e-03
Superplasticizer	2.461e-01
CoarseAgg	5.443e-03
FineAgg	2.738e-03
Age	9.516e-74

Table 5. p-values for the model with log(x+1) transformed predictors

Predictor	p-value
Const	2.930e-01
Cement	1.289e-86
Slag	4.634e-36
FlyAsh	2.632e-01
Water	1.837e-20
Superplasticizer	1.021e-05
CoarseAgg	4.085e-01
FineAgg	9.428e-02
Age	1.710e-192

The analysis of these tables shows that in the model using the log(x+1) transformation, cement content and curing age have the lowest p-values, indicating they are highly significant for predicting compressive strength. Water content is also statistically significant, though it has a consistently negative effect on strength. After the logarithmic transformation, superplasticizer becomes more statistically relevant, while fly ash and the aggregates display higher p-values, suggesting their impact is limited in the linear model evaluated.

3.5. Model validation and predictive application

Three regression models untransformed OLS, standardized OLS, and logarithmically transformed OLS were validated using an independent dataset. The untransformed and standardized models showed limited performance ($R^2 = 0.351$) and non-random residual patterns, indicating heteroscedasticity. In contrast, the log(x+1) transformed model demonstrated superior predictive capability ($R^2 = 0.750$) and well-behaved residuals, validating it as the most robust tool, as shown in Figure 7.

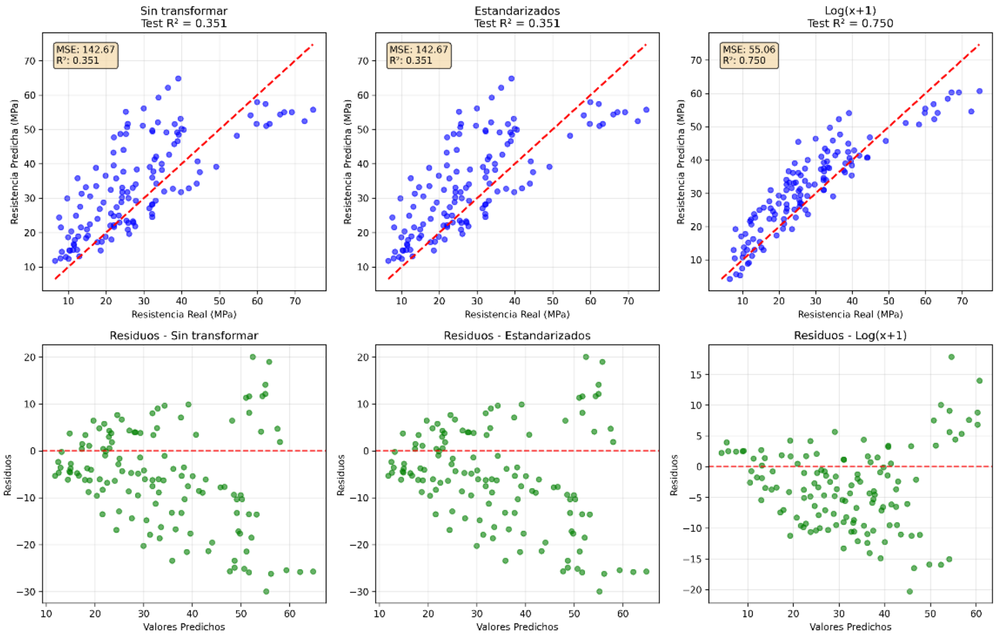

Figure 7. Validation of OLS models

To assess its practical utility, the selected model was applied to three simulated dosage scenarios: a standard mix, a high-cement-content mix, and an optimized mix with admixtures. Figure 8 compares the predicted strength development curves with critical mine standards: early-age strength (2 MPa at 2 hours) and final design strength (>28 MPa at 28 days). The tool enables both visual and quantitative prediction of compliance with these criteria before field application.

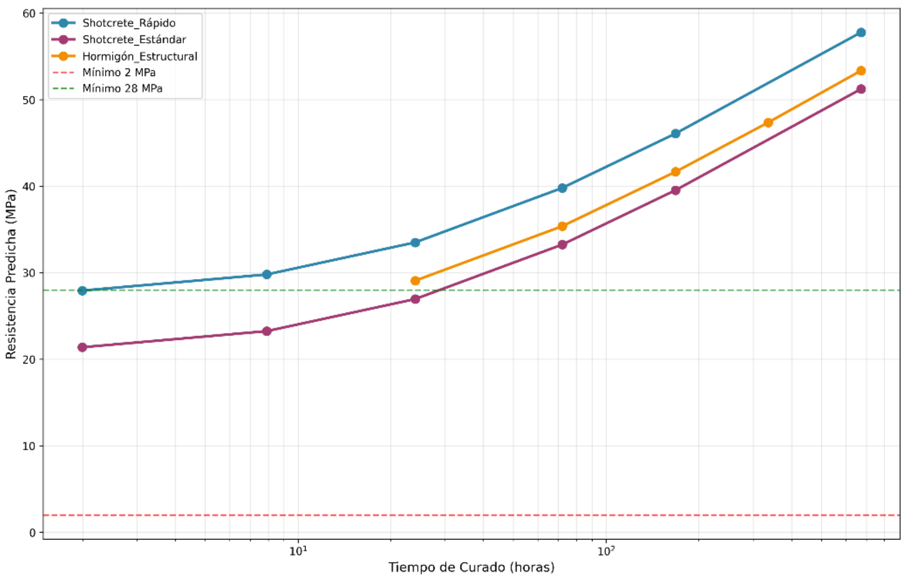

Figure 8. Strength development curves with Log(x+1) transformation

These results confirm that the logarithmic model is not only statistically sound but also a practical predictive tool for optimizing mix designs, ensuring compliance with technical specifications, and reducing uncertainty in underground support systems.

4. Discussion

The results obtained demonstrate that data preprocessing plays a crucial role in the effectiveness of linear regression models used to predict the compressive strength of concrete. Specifically, the model using log(x+1) transformed predictors achieved an R^2 of 0.80 on the training set and 0.75 on the test set, clearly outperforming

the models using untransformed and standardized predictors, which only reached $R^2 \approx 0.35$ on the test set. This marked improvement aligns with the findings of Yeh [22], who observed that the inherent nonlinearity of concrete's mechanical properties limits the performance of linear models when working with data in its original scale.
From the perspective of prediction error, the logarithmically transformed model significantly reduced the Mean Squared Error (MSE) on the test set from values exceeding 140 MPa^2 (raw and standardized models) down to approximately 55 MPa^2, representing a reduction of nearly 60%. This decrease is particularly relevant in engineering, where incorrect strength estimations can lead to overly conservative decisions or worse, pose structural safety risks [23].
The fact that the raw and standardized models showed nearly identical metrics (Train MSE ≈ 105 MPa^2 and Test MSE ≈ 143 MPa^2) indicates that standardization alone does not correct the issues of skewness and heteroscedasticity present in the experimental dataset. This behavior has also been observed by Chou et al. [24], who noted that logarithmic transformations are more effective than classical normalization techniques for mechanical properties of concrete when using linear regression models.
Within the log(x+1) transformation scheme, the cement content and curing age had the lowest p-values, indicating a strong and statistically significant influence on compressive strength. This finding is consistent with prior studies highlighting these two factors as the most influential in the strength development of concrete, both in general applications and in shotcrete used for tunnel reinforcement in mining environments [25]. From a practical perspective, this confirms the importance of carefully managing both cement dosage and curing times.
The water content also had a significant but negative effect, a result that aligns with the traditional understanding of the water-to-binder ratio. In underground mining, where ambient humidity and placement methods can alter the actual water content used, this finding becomes particularly relevant. Neville [26] emphasized that even small increases in water can significantly reduce strength something well captured in the model's estimates.
Notably, superplasticizer reached statistical significance only after the logarithmic transformation, suggesting a nonlinear relationship between this admixture and compressive strength. This pattern has been observed in studies on shotcrete, where superplasticizers improve workability and compaction without reducing strength when used in proper doses [27]. In contrast, fly ash and aggregates showed high p-values, indicating a limited contribution in the linear model, possibly due to interaction effects not captured by a purely linear approach.
The similarity in MSE values between training and test sets for the log(x+1) model (approximately 56 MPa^2 and 55 MPa^2, respectively) reflects strong generalization, which is critical for real-world mining applications. As Asteris et al. [28] point out, a model that remains stable on unseen data is more reliable than one that overfits during training but fails in new scenarios, especially in highly variable material conditions.
From a practical standpoint, a linear regression model offers a cost-effective and interpretable solution for concrete quality control in mining operations. Its simplicity makes it easy to integrate into field control systems and allows for rapid estimation of whether a given mix will meet required strength levels. However, the results also suggest that to capture more complex interactions particularly between aggregates and mineral admixtures nonlinear models should be explored in future research. The results gain further relevance when compared to real-world strength criteria used in underground mining: the log(x+1) model's ability to predict compressive strength with low error enables early assessment of whether a mix is likely to reach critical thresholds such as 2 MPa at 2 hours and >28 MPa at 28 days. This alignment with practical standards highlights the model's potential as a predictive tool for decision support in the field, rather than as a direct replacement for standardized testing.
Finally, while the results obtained are robust within the scope of the dataset used, applying this model in different geological settings will require adjustments and recalibration. As noted by Ahmad et al. [30], variations in material source and mixing conditions can affect the accuracy of prediction models, thus continuous refinement is essential to maintain reliability. Overall, the findings support the idea that combining real experimental data, appropriate preprocessing, and interpretable statistical modeling provides an effective strategy to optimize concrete management in the mining sector.

5. Conclusions

In conclusion, the findings underscore the importance of data preprocessing for linear regression models used to predict the compressive strength of concrete. Specifically, the log(x+1) transformation significantly enhanced the explanatory power of the model, increasing the coefficient of determination on the test set from approximately 0.351 to nearly 0.750. Using log(x+1) transformed predictors, the model reduced the Mean Squared Error (MSE) on the test set by approximately 60% compared to unmodified standardized models, decreasing from over 140 MPa^2 to about 55 MPa^2. This improvement is particularly crucial, as structural stability and safety depend on accurate measurement of concrete strength. The statistical significance analysis revealed that the most influential variables in determining compressive strength are cement content and curing time. Conversely, water content had a consistently negative impact. These results support established mix design practices and emphasize the need for strict control over these factors in concrete used for lining and support in mining operations.

The statistical significance of the superplasticizer increased after the logarithmic transformation, suggesting a nonlinear influence on compressive strength. This finding reinforces the importance of employing appropriate preprocessing techniques.
When evaluating chemical admixtures in concrete intended for mining where balancing workability and mechanical performance is essential the similarity between training and test errors in the log(x+1) model indicates good generalization, suggesting that the model can be reliably applied to fresh concrete mixes within comparable operational ranges. This characteristic is essential for its potential adoption as a quality control support tool in underground mining operations.
Although the performance of the log(x+1) transformed linear regression model was clearly superior, achieving R^2 values near 0.750 and significantly lower prediction errors its accuracy remains strongly dependent on the quality and representativeness of the available experimental data. The model can serve as a complementary predictive tool for staged verification of compliance with concrete strength criteria, such as early strengths of ~2 MPa within the first hours and final strengths above 28 MPa at 28 days. Its practical application can help optimize quality control programs, reduce rework, and enhance proactive geomechanical safety management. It is crucial to recognize that the presence of material variability, outliers from laboratory testing, and potential changes in mixing and curing conditions may affect model stability. Therefore, it is advisable to expand the dataset through new experimental campaigns, include additional variables related to environmental and operational conditions specific to underground mining, and consider hybrid approaches that integrate statistical models with physico-mechanical principles governing concrete behavior. Despite these limitations, the results presented here offer a valuable contribution and validate the potential of data analysis and machine learning techniques as support tools to enhance the efficiency, reliability, and predictability of concrete use in mining and other industrial applications.
Future work may also explore hybrid models that combine statistical approaches with physical-mechanical principles of concrete behavior, aiming to increase model robustness and extend its applicability to diverse operational contexts.

6. Conflict of interest

The authors declare that they have no conflicts of interest.

7. References

[1]. Yeh, I.C. (1998). Modeling of strength of high-performance concrete using artificial neural networks. Cement and Concrete Research, 28(12), 1797–1808.
DOI: https://doi.org/10.1016/S0008-8846(98)00165-3

[2]. Yeh, I.C. (2007). Modeling concrete strength with augment-neuron networks. Cement and Concrete Research, 37(7), 1063–1070.
DOI: https://doi.org/10.1016/j.cemconres.2007.03.013

[3]. Asteris, P.G.; Mokos, V.G. (2020). Concrete compressive strength prediction using machine learning techniques. Neural Computing and Applications, 32, 15807–15828.
DOI: https://doi.org/10.1007/s00521-019-04663-2

[4]. Chou, J.S.; Pham, A.D. (2013). Enhanced artificial intelligence for predicting high-performance concrete compressive strength. Construction and Building Materials, 49, 554–563.
DOI: https://doi.org/10.1016/j.conbuildmat.2013.08.078

[5]. Neville, A.M. (2011). Properties of Concrete (5th ed.). Pearson Education Limited.
[6]. Ahmad, A.; Farooq, F.; Ostrowski, K.A. (2021). Prediction of concrete compressive strength using machine learning algorithms. Materials, 14(15), 4141.
DOI: https://doi.org/10.3390/ma14154141

[7]. Asteris, P.G.; Argyropoulos, I.; Cavaleri, L. (2021). Predicting concrete strength using hybrid machine learning models. Applied Sciences, 11(7), 3036.
DOI: https://doi.org/10.3390/app11073036

[8]. Behnood, A.; Golafshani, E.M. (2018). Predicting the compressive strength of silica fume concrete using hybrid models. Construction and Building Materials, 170, 726–736.
DOI: https://doi.org/10.1016/j.conbuildmat.2018.03.086

[9]. Ashrafian, A.; Nikoo, M.; Kadivar, A. (2020). Prediction of concrete strength using statistical learning methods. Structures, 28, 288–299.
DOI: https://doi.org/10.1016/j.istruc.2020.08.021

[10]. Mansouri, I.; Hu, J.W. (2021). Application of machine learning techniques for predicting concrete strength. Journal of Building Engineering, 43, 102494.
DOI: https://doi.org/10.1016/j.jobe.2021.102494

[11]. UCI Machine Learning Repository. (2019). Concrete Compressive Strength Dataset. University of California, Irvine.
DOI: https://doi.org/10.7717/peerj-cs.154/table-6

[12]. Cortes, C.; Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297.
DOI: https://doi.org/10.1007/BF00994018

[13]. Montgomery, D.C.; Peck, E.A.; Vining, G.G. (2012). Introduction to Linear Regression Analysis (5th ed.). Wiley.
DOI: https://doi.org/10.1111/biom.12129

[14]. Kutner, M.H.; Nachtsheim, C.J.; Neter, J. (2005). Applied Linear Statistical Models. McGraw-Hill.
DOI: https://doi.org/10.2307/1271154

[15]. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. (2021). An Introduction to Statistical Learning (2nd ed.). Springer.
DOI: https://doi.org/10.32614/cran.package.islr2

[16]. Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer.
DOI: https://doi.org/10.1007/978-0-387-45528-0_7

[17]. Kuhn, M.; Johnson, K. (2013). Applied Predictive Modeling. Springer.
DOI: https://doi.org/10.1007/978-1-4614-6849-3
[18]. Hyndman, R.J.; Koehler, A.B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688.
DOI: https://doi.org/10.1016/j.ijforecast.2006.03.001

[19]. Flores, R.; Valderrama, L. (2021). Evaluation of metallurgical process indicators using percentage-based error metrics. Minerals, 11(8), 864.

DOI: https://doi.org/10.3390/min11080864

[20]. Bendaña, D.; Combes, F. (2016). Performance assessment in flotation circuits using error metrics. Minerals Engineering, 95, 57–66.
DOI: https://doi.org/10.1016/j.mineng.2016.05.018

[21]. ASTM C39/C39M. (2021). Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens. ASTM International.

[22]. Yeh, I.C. (2006). Analysis of strength of concrete using data mining techniques. Cement and Concrete Research, 36(5), 879–884.
DOI: https://doi.org/10.1016/j.cemconres.2005.11.005

[23]. Hoang, N.D. (2020). Predicting concrete strength using regression and machine learning. Engineering with Computers, 36, 225–235.
DOI: https://doi.org/10.1007/s00366-018-0660-5

[24]. Gandomi, A.H.; Alavi, A.H. (2012). A new multi-gene genetic programming approach to nonlinear system modeling. Engineering Applications of Artificial Intelligence, 25(3), 577–587.
DOI: https://doi.org/10.1016/j.engappai.2011.09.010

[25]. Topçu, I.B.; Sarıdemir, M. (2008). Prediction of compressive strength of concrete containing fly ash using artificial neural networks. Computers & Materials Science, 41(3), 305–311.
DOI: https://doi.org/10.1016/j.commatsci.2007.04.009

[26]. Erdem, H.; Gül, R. (2012). Prediction of the compressive strength of cement mortar. Construction and Building Materials, 30, 804–812.
DOI: https://doi.org/10.1016/j.conbuildmat.2011.12.011

[27]. Safari, M.; Faradonbeh, R.S. (2018). Reliability-based prediction of concrete strength. Structural Concrete, 19(5), 1425–1436.
DOI: https://doi.org/10.1002/suco.201700234

[28]. Khosravi, K.; Daggupati, P.; Pham, B.T. (2019). Machine learning approaches for predicting concrete strength. Applied Soft Computing, 75, 563–576.
DOI: https://doi.org/10.1016/j.asoc.2018.11.032

[29]. Lin, Y.; Chen, M.; Li, Y. (2020). Data-driven prediction of concrete properties for engineering applications. Automation in Construction, 119, 103325.
DOI: https://doi.org/10.1016/j.autcon.2020.103325

[30]. ACI Committee 211. (2002). Standard Practice for Selecting Proportions for Normal, Heavyweight, and Mass Concrete. American Concrete Institute.
DOI: https://doi.org/10.14359/10987