Applied Chemical Engineering (2025) Volume 8 Issue 2 ISSN: 2578-2010 (O)
doi: 10.59429/ace.v8i2.5639

ORIGINAL RESEARCH ARTICLE

AI-based prediction of the optimal incorporation rate of dredged sediments in concrete: Mechanical performance analysis

Khadija Benhaddou[1], Ayoub Souileh[1,*], Achraf Mabrouk[5], Latifa Ouadif[1], Sabihi Abdelhak[2], Khadija Baba[3], Mustapha Rharouss[1], Azzeddine Imali[4]

[1] *Laboratory of Applied Geophysics, Geotechnics, Engineering Geology, and Environmental (L3GIE), Mohammadia Engineering School, Mohammed V University in Rabat, Avenue Ibn Sina, BP 765, Agdal, Rabat, 10000, Morocco*

[2] *Experimental Center for Major Works, Public Laboratory for Testing and Studies (LPEE), Casablanca, 20250, Morocco*

[3] *Civil and Environmental Engineering Laboratory (LGCE), Mohammadia Engineering School, Mohammed V University in Rabat, Avenue Ibn Sina, BP 765, Agdal, Rabat, 10000, Morocco*

[4] *Regional Directorate of Equipment, Transport, and Logistics of Rabat-Salé-Kénitra (DRETL RSK), Ministry of Equipment and Water, Avenue Al Araar, Hay Riad, Rabat, 10100, Morocco*

[5] *LAFH, Faculty of Sciences and Techniques, Hassan 1st University, BP 577, Settat, 26000, Morocco*

***Corresponding author:** Ayoub Souileh, Ayoub.souileh@research.emi.ac.ma

ARTICLE INFO

Received: 22 March 2025
Accepted: 18 April 2025
Available online: 06 June 2025

ABSTRACT

The management of marine dredged sediments is a critical environmental and economic issue, particularly in port cities where dredging is a necessary activity to maintain navigability. These sediments are typically viewed as waste products and often require costly and environmentally challenging disposal methods. However, repurposing dredged sediments as a component in concrete production presents a promising solution for both waste management and the creation of sustainable construction materials. Despite this potential, determining the optimal percentage of sediment incorporation and accurately predicting the mechanical properties, such as compressive strength, remain significant challenges. This study proposes an artificial intelligence (AI)-based approach to predict the optimal incorporation percentage of marine dredged sediments from Moroccan ports into concrete and to forecast the resulting compressive strength. A dataset consisting of 104 samples, including dune sand and port sediments from JEBHA, was used. The data includes key properties such as granulometry, cleanliness, fineness modulus, and the compressive strength of the concrete mixtures. These experimental data were employed to train and validate several machine learning models, including linear regression, Random Forest, Gradient Boosting, and XGBoost, chosen for their ability to model complex, non-linear relationships between sediment characteristics and concrete performance. The performance of these models was evaluated using two key metrics: the coefficient of determination (R^2) and the root mean square error (RMSE). Among the models tested, the Random Forest Regressor delivered the best results, with an R^2 value greater than 0.98 and an RMSE of less than 0.20 MPa, indicating highly accurate predictions of both the optimal sediment incorporation rate and the compressive strength of the concrete. This model's exceptional performance underscores its potential as a reliable tool for optimizing the use of dredged sediments in concrete production. The findings of this study demonstrate the considerable potential of AI in optimizing the incorporation of marine dredged sediments into concrete. By accurately predicting the mechanical properties of the resulting material, this approach enables the development of more sustainable construction practices while reducing the environmental burden associated with sediment disposal. Moreover, this work illustrates the broader applicability of AI in addressing environmental challenges, offering a pathway to valorize waste materials in the construction industry. The study not only advances our understanding of sediment utilization in concrete but also contributes to the growing field of sustainable material science, offering promising avenues for future research and development.

Nevertheless, further research is needed to validate the model's scalability to other sediment types and assess practical limitations in industrial applications.

Keywords: random forests; experimental data; optimal incorporation percentage; model predictions; artificial intelligence; compressive strength

1. Introduction

The massive exploitation of natural resources, particularly sand, poses a growing threat to coastal ecosystems on a global scale. Sand, although perceived as an abundant resource, is in fact a limited raw material whose intensive extraction leads to severe environmental consequences, such as beach erosion, destruction of marine habitats, and disruption of coastal ecosystems[1]. In Morocco, this issue is exacerbated by port activities, which generate significant amounts of dredged sediments. These sediments, resulting from the cleaning of waterways and port basins, present major challenges in terms of management and valorization. Traditionally considered as waste, they nonetheless represent an opportunity to reduce pressure on natural resources by being reused in industrial applications, particularly in the construction sector[2].

Incorporating dredged sediments as a partial substitute for sand in concrete production offers a promising way to balance economic growth with environmental sustainability. However, the variability in the physicochemical properties of these sediments, which depends on their geographic origin and composition, has a direct impact on the concrete's mechanical performance, especially its compressive strength (CS). This variability makes it challenging to create the ideal concrete mix, highlighting the need for accurate predictive tools that can estimate the mechanical properties of concrete based on the amount of sediments incorporated[3,4].

Recent studies have explored the integration of artificial intelligence in sustainable construction, particularly in optimizing concrete formulations using industrial by-products. However, the use of dredged marine sediments in this context remains underexplored, thus highlighting the novelty of this study.

In this context, the study introduces an innovative machine learning-based approach to predict the optimal incorporation rate of dredged sediments in concrete and to estimate the compressive strength of the resulting material. The objective is to develop formulations that enhance mechanical performance while minimizing environmental impact, particularly by reducing the need for natural sand extraction. To achieve this, a variety of machine learning models were investigated, including linear regression, Random Forest, Gradient Boosting, and XGBoost. These models were trained and tested using a comprehensive dataset containing results from dune sand characterization and sediments collected from the port of JEBHA, alongside the corresponding compressive strength of the concrete mixtures. The models were then validated using additional data from sediments obtained from the ports of Nador, Ras Kebdana, and Mohammedia, three sites that capture Morocco's geological and environmental diversity[5,6].

The evaluation of model performance was conducted using robust statistical indicators, such as the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). These metrics allow for the quantification of prediction accuracy and the comparison of the effectiveness of different models. The results obtained show that machine learning approaches, particularly ensemble methods like XGBoost and Random Forests, offer high accuracy in predicting concrete compressive strength based on sediment composition. These models also enable the identification of key parameters influencing mechanical properties, paving the way for more refined optimization of formulations[7].

In conclusion, this study underscores the potential of machine learning in addressing environmental challenges linked to the exploitation of natural resources. By offering sustainable solutions for the valorization of dredged sediments, it supports the transition to a circular economy within the construction industry while helping to protect sensitive coastal ecosystems. The results also lay the groundwork for future research, which could broaden the application of this approach to other forms of industrial waste or different regions, thereby enhancing the environmental and economic benefits of this innovative method[8].

2. Modeling methodology

The data modeling process is presented in a series of key stages (**Figure 1**). It begins with the collection and preprocessing of data, followed by the application of machine learning algorithms for model development. Finally, the results are rigorously validated to determine the most accurate and effective model, ensuring optimal performance and reliability.

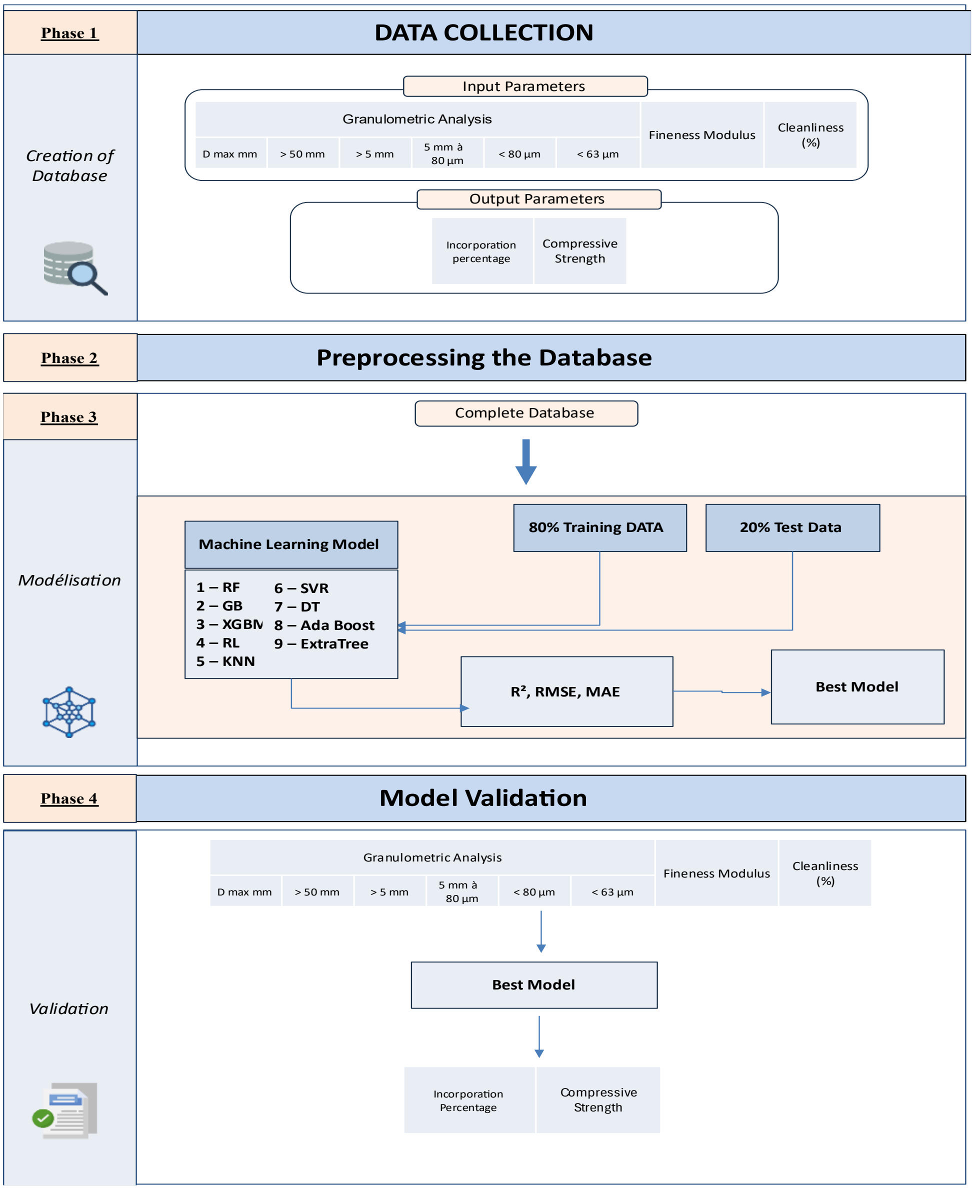

Figure 1. Data processing workflow.

2.1. Description of data base

The database employed in this study comprises 105 samples of dune sand and marine dredged sediments. The measured parameters include granulometric characteristics, sediment incorporation percentage, sediment cleanliness, and the 28-day compressive strength (CS) of the resulting concrete. **Table 1** summarizes the descriptive statistics of the key parameters, revealing a broad range of variations, especially in sediment incorporation percentages and compressive strength. This variability provides a rich and diverse dataset, essential for robust model training and accurate predictions.

Table 1. Database description.

Paramete	N	Mean	Standard Deviation	Min	Max
Dmax (mm)	105	2,64	1,58	1,00	8,00
> 50 mm (%)	105	0,00	0,00	0,00	0,00
> 5 mm (%)	105	0,32	1,20	0,00	7,30
5 mm to 80 µm (%)	105	94,10	7,67	52,40	98,50
< 80 µm (%)	105	5,58	7,07	1,50	47,60
< 63 µm (%)	105	4,92	7,25	1,00	47,60
Fineness modulus	105	1,40	0,37	1,13	3,40
Cleanliness (%)	105	70,80	11,49	24,00	78,00
Incorporation percentage (%)	105	91,70	23,68	10,00	100,00
28-day Compressive Strength (MPa)	105	53,10	3,40	46,10	60,30

2.2. Statistical data analysis

In this study, a representative sample of 105 specimens was analyzed to characterize the granulometric and mechanical properties of the material. **Table 1 s**ummarizes the key descriptive statistics for each measured parameter:

The particle size distribution of the studied materials reveals key characteristics for concrete formulation. The maximum aggregate size (Dmax) has an average of 2.64 mm, with a standard deviation of 1.58 mm, indicating notable variability between samples, ranging from 1.00 mm to 8.00 mm. No particles larger than 50 mm were detected (0%), which is consistent with expectations for this type of material. Particles larger than 5 mm are scarcely present, with an average of 0.32% (standard deviation of 1.20%), although some samples may contain up to 7.30%. The majority of particles fall within the 5 mm to 80 µm range, representing an average of 94.10% of the material (standard deviation of 7.67%), a favorable characteristic for concrete compactness[9].

In contrast, the fractions below 80 µm and 63 µm show respective averages of 5.58% and 4.92%, but with high variability (standard deviations of 7.07% and 7.25%), which may influence the density and strength of the concrete. The fineness modulus, reflecting the distribution of particle sizes, has an average of 1.40 (standard deviation of 0.37), with a range of 1.13 to 3.40, indicating moderate dispersion essential for mixture cohesion. The sediment cleanliness, measured as a percentage, has an average of 70.80% (standard deviation of 11.49%), with values varying from 24% to 78%, which can impact the quality of the concrete. The incorporation rate shows a high average of 91.70% (standard deviation of 23.68%), covering a wide range from 10% to 100%, allowing for the evaluation of the impact of different incorporation levels on concrete performance.

Finally, the 28-day compressive strength (RC 28d) has an average of 53.10 MPa (standard deviation of 3.40 MPa), with values ranging from 46.10 MPa to 60.30 MPa, demonstrating homogeneous and satisfactory mechanical performance despite the variability of other parameters[10].

2.3. Synthesis and perspectives

Overall, the results demonstrate a high degree of homogeneity for most granulometric and mechanical parameters, with minimal variations observed across the majority of the samples. However, the presence of a few extreme values, particularly in the fine fractions and incorporation percentages, suggests that variations in the process or specific anomalies may account for these discrepancies. These findings highlight the need for further research to optimize quality control and manufacturing conditions, ensuring consistent and optimal performance of the material under investigation[11].

2.4. Modeling methodology

The methodology adopted in this study, as presented in **Figure 1**, comprises four main stages:

Data collection and preparation

The data originate from experimental tests conducted on dune sands, marine dredged sediments, and concretes containing these materials. They include granulometric characteristics, cleanliness (sand equivalent), fineness modulus, the percentage of sand incorporation in concrete, and the 28-day compressive strength (RC). These data form the database necessary for constructing and evaluating predictive models[12].
The compressive strength tests were conducted on cylindrical specimens of 16×32 cm, cured under water at 20°C for 28 days. Data normalization was systematically applied before training all models. Hyperparameter tuning was conducted for each algorithm using grid search and cross-validation.

Data preprocessing

This step involves normalizing the variables, handling missing values through imputation, and removing outliers to improve the quality of the inputs for the models.

Model training: Machine learning algorithms

Machine learning models

To predict the optimal incorporation percentage of dredged sediments and the compressive strength (RC) of concrete, nine machine learning algorithms were tested and compared in this study (**Table 2**). These algorithms, selected for their regression efficiency and ability to handle complex multivariate datasets, include:

- **Ensemble methods** : RandomForestRegressor, XGBoost, LightGBM, Extra Trees[13]
- **Linear approaches** : Linear Regression[14]
- **Similarity- and tree-based techniques** : SVR, KNN, Decision Tree, AdaBoost[15]

To ensure a thorough comparison, each model was trained on 80% of the data and tested on the remaining 20%. Performance was evaluated using specific metrics, enabling the identification of the most suitable algorithm for each of the two prediction objectives.

Table 2. Machine learning models.

Model	Description
Random Forest (RF)	Ensemble method based on the construction of multiple decision trees. It enhances robustness and accuracy by aggregating the predictions from each tree..
XGBoost	An optimized variant of Gradient Boosting, it leverages first- and second-order derivatives to better fit the loss function. XGBoost is renowned for its computational speed and ability to efficiently handle missing values.
LightGBM	Developed by Microsoft, LightGBM builds its decision trees "leaf-wise" rather than "level-wise," improving accuracy and reducing computation time, especially for large datasets.
Régression Linéaire (RL)	A basic model that establishes a linear relationship between the explanatory variables and the target variable. It is simple and easy to interpret but may be insufficient for capturing nonlinear relationships.
SVR (Support Vector Regression)	A variant of the Support Vector Machine method, it aims to find a regression function by minimizing the error beyond a certain threshold. SVR is particularly suited for complex problems but may require precise tuning of its hyperparameters.

Model	Description
KNN (K-Nearest Neighbors)	Similarity-based algorithm: the prediction for a new sample is derived from the values of its k nearest neighbors. KNN is easy to understand but can be sensitive to noise and the choice of distance metric used.
Decision Tree (DT)	A basic algorithm that partitions the variable space using hierarchical decision rules. While intuitive, it tends to overfit unless some form of regularization or depth limitation is applied.
AdaBoost	A sequential boosting technique where each new estimator focuses on the samples misclassified (or poorly predicted) by the previous estimators. AdaBoost is effective for reducing bias but can be sensitive to noise.
Extra Trees (ET)	An ensemble method similar to Random Forest, with the difference that the selection of split points in the trees is random. This can lead to greater tree diversity and, potentially, better generalization.

Table 2. (*Continued*)

Statistical indicators

To evaluate the performance of the models in predicting the incorporation percentage and compressive strength (RC), several statistical indicators (performance metrics) were used: the coefficient of determination (R^2), the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE) (**Table 3**).

These complementary indicators provide an assessment of the overall performance of the models in terms of accuracy and explanatory capacity.

Table 3. Performance metrics

Performance Metric	Formula	Interpretation
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$ [16]	An R^2 value close to 1 indicates that the model effectively explains the variability in the data.
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$ [17]	A lower RMSE indicates more accurate predictions.
Mean Absolute Error (MAE)	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert y_i - \hat{y}_i\rvert$ [18]	The lower the MAE, the better the model's accuracy.
Mean Absolute Percentage Error (MAPE)	$MAPE = \frac{100}{N}\sum_{i=1}^{N}\frac{\lvert y_i - \hat{y}_i\rvert}{\lvert y_i\rvert}$ [16]	The lower the MAPE, the more accurate the model is.

Prediction models

Two distinct prediction models were developed and trained using the RandomForestRegressor algorithm (**Table 4**). The comprehensive analysis and modeling process, designed to optimize the incorporation of marine sediments in concrete as outlined in **Figure 2**, begins with a detailed examination of granulometric characteristics, including grain size and distribution, an assessment of cleanliness, and data normalization, while addressing outliers to ensure the data's reliability.

The first model predicts the optimal percentage of marine sediment incorporation, using granulometric characteristics and cleanliness as input variables. The second model, referred to as the RC model, predicts compressive strength by incorporating both the sediment incorporation percentage and the previously mentioned parameters.

These models undergo validation through experimental testing, followed by fine-tuning to enhance their predictive accuracy. They are ultimately validated for practical use, ensuring that the results are not only scientifically sound but also applicable in real-world scenarios. This approach blends scientific rigor with industrial relevance, offering a robust methodology for optimizing material performance in concrete production.

Table 4. Training process of the models used.

Model	Objective	Explanatory Variables (features)	Target Variable	Data Split	Algorithm and Parameters
Incorporation Percentage Model	Predict the incorporation percentage	Granulometric characteristics + cleanliness	Incorporation percentage	80% training, 20% test	RandomForestRegressor, XGBoost, LightGBM, Linear Regression, SVR, KNN, Decision Tree, AdaBoost, Extra Trees
Compressive Strength (RC) Model	Predict compressive strength (RC)	Granulometry + cleanliness + incorporation percentage	Compressive strength (RC)	80% training, 20% test	

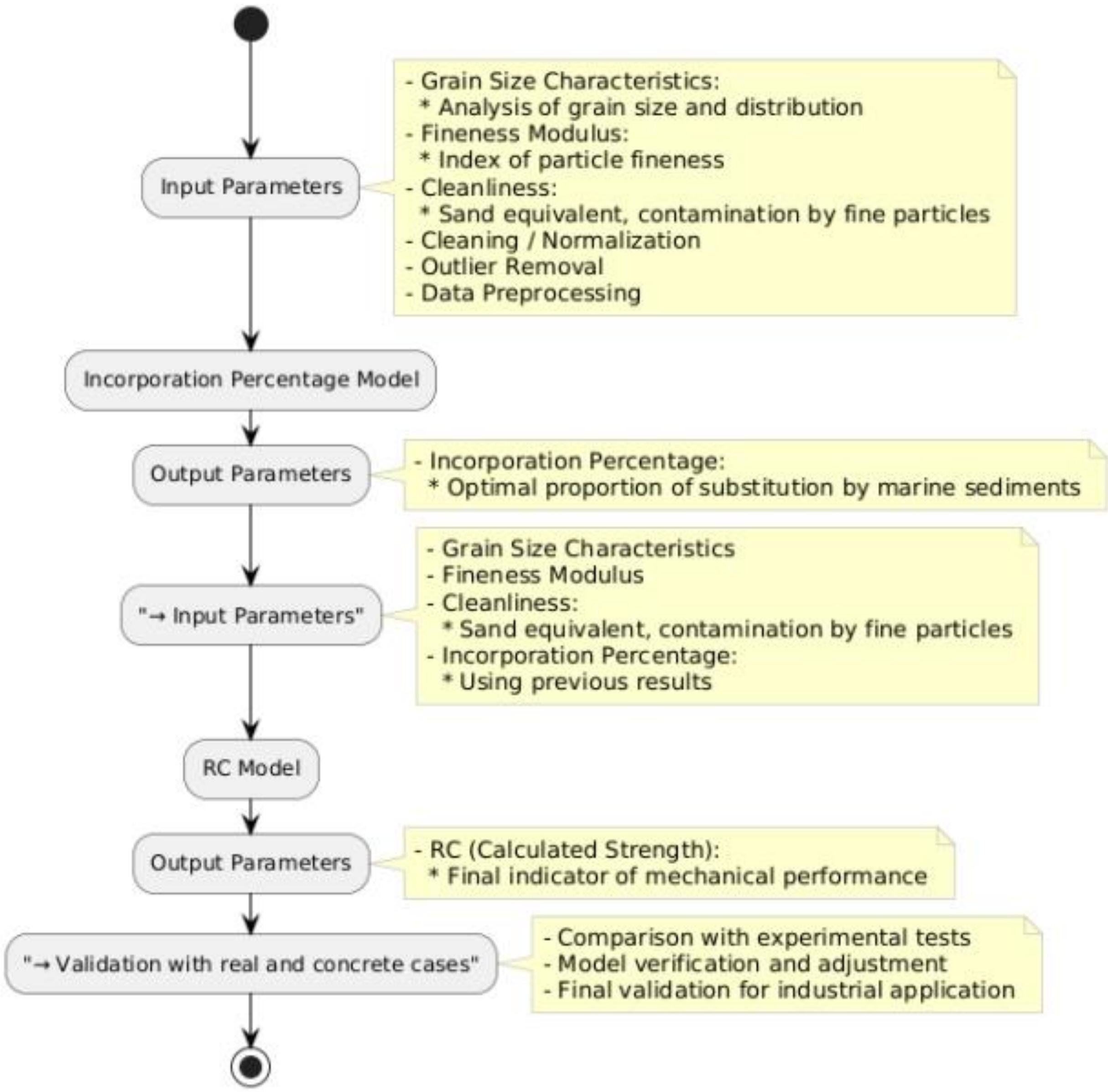

Figure 2. Prediction process for incorporation percentage and compressive strength.

Validation and evaluation

The generated models were applied to new real-world datasets (practical cases) to evaluate their performance in operational scenarios and confirm their generalization capabilities. This approach aims to ensure the robustness of predictions while providing practical recommendations for formulating concrete that incorporates dredged sediments.

Each model is assessed using the following metrics: **RMSE** and **R²**.

3. Results and discussion

This section presents the results obtained from the tested models, along with an analysis of their performance in predicting the incorporation percentage and the compressive strength (RC) of concrete.

3.1. Model performance: Comparative analysis of models

The results presented in **Table 5** reflect a rigorous comparative evaluation of nine machine learning algorithms applied to predicting the optimal incorporation percentage of dredged sediments and the compressive strength (RC) of concrete, based on performance indicators. It is observed that ensemble methods, particularly **Random Forest (RF)** and **Extra Trees (ET)**, demonstrate remarkable performance with very low RMSE and MAE values and a coefficient of determination (R^2) close to 1, indicating an excellent ability to capture data variability. Although linear regression, despite its simplicity, provides competitive results, some models such as **LightGBM** and **SVR** show difficulties, suggesting that their parameterization requires finer optimization for this type of data. These findings highlight the advantage of ensemble approaches for modeling complex phenomena and provide valuable insights for selecting predictive models in materials engineering.

The poor performance of LightGBM and SVR may be due to overfitting or suboptimal hyperparameters, highlighting their sensitivity to data characteristics. Feature importance in the sensitivity analysis was calculated using Gini importance derived from the Random Forest model.

Table 5. Model performances on the test set.

Modèle	RMSE	R^2	MAE
Random Forest (RF)	0.147	0.997	0.084
XGBoost	0.218	0.994	0.111
LightGBM	1.578	0.661	0.953
Régression Linéaire (RL)	0.229	0.993	0.167
SVR	1.979	0.481	0.919
KNN	0.424	0.976	0.231
Decision Tree (DT)	0.173	0.993	0.130
AdaBoost	0.229	0.993	0.112
Extra Trees (ET)	0.189	0.994	0.108

3.2. Analysis of predictions

The comparison between the actual and predicted values of the incorporation percentage and compressive strength (RC) for the Random Forest model shows a strong proximity of the points to the identity line y=x (**Figures 3** and **4**), demonstrating the accuracy of the predictions with minimal deviations. This performance confirms the model's ability to generalize effectively to new data, although Gradient Boosting and XGBoost models also exhibit good, albeit slightly lower, performance compared to Random Forest. The results illustrated in **Figures 3** and **4** highlight the relevance of machine learning in predicting the mechanical properties of concretes incorporating dredged sediments, effectively capturing experimental variability. Notably, the deviations for compressive strength remain relatively low, with most samples showing less than 5% variation (as depicted in **Figure 1**). Similarly, the predictions for the incorporation percentage, shown in **Figure 2**, align closely with the experimental values, allowing for the precise identification of the optimal incorporation rate. These results highlight the robustness of machine learning in processing complex datasets and generating reliable predictions, even in practical, operational settings.

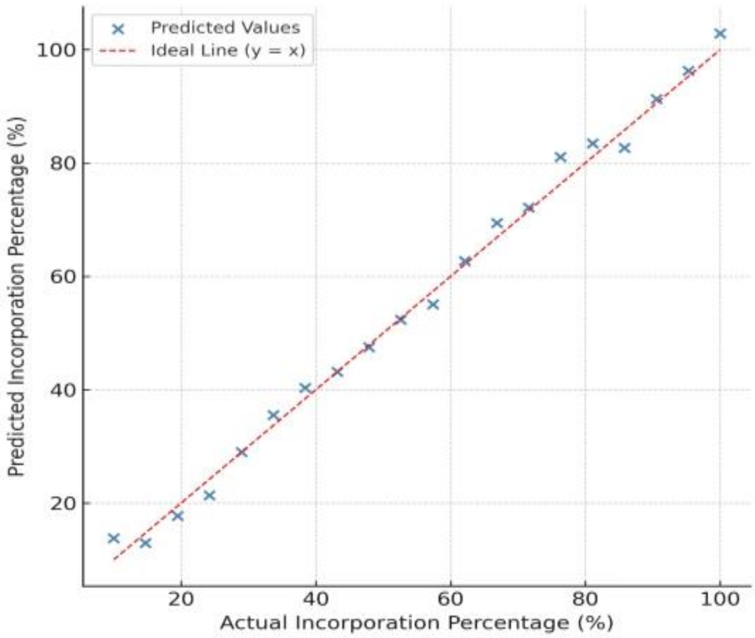

Figure 3. Comparison between actual and predicted values of the incorporation percentage for the Random Forest model.

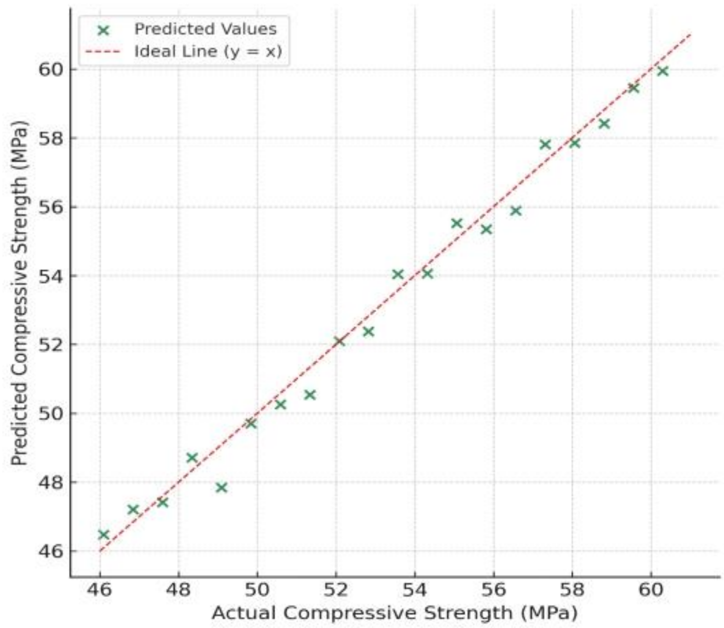

Figure 4. Comparison between actual and predicted values of compressive strength for the Random Forest model.

3.3. Sensitivity analysis

The sensitivity analysis (**Figure 5**) was conducted to determine the input variables that have the greatest influence on the compressive strength of concrete. Initially, it appears that the incorporation percentage, sediment cleanliness, and maximum aggregate diameter are the most critical parameters: variations in these factors can significantly affect the material's mechanical performance. This observation provides engineers with an important lever for optimizing the formulation while promoting the valorization of dredged sediments.

To further refine the study, a complementary analysis was conducted by deliberately excluding cleanliness and incorporation percentage. The results indicate that granulometric fractions greater than 63 µm and 80 µm remain highly influential, followed by the fraction ranging from 5 mm to 80 µm. Conversely, the variables D_{max}, particles smaller than 5 mm, and particles smaller than 50 mm exhibit a more limited impact. This distribution underscores the importance of controlling the proportion of fine and medium particles, particularly those exceeding 63 µm and 80 µm, to enhance the concrete's resistance.

In summary, these two complementary approaches confirm, on the one hand, the crucial role of cleanliness and the incorporation percentage, and on the other hand, the need to precisely control the particle size distribution to maximize mechanical performance. Engineers can thus more effectively target their formulation choices, balancing the optimization of compressive strength with the reasoned integration of dredged sediments[19].

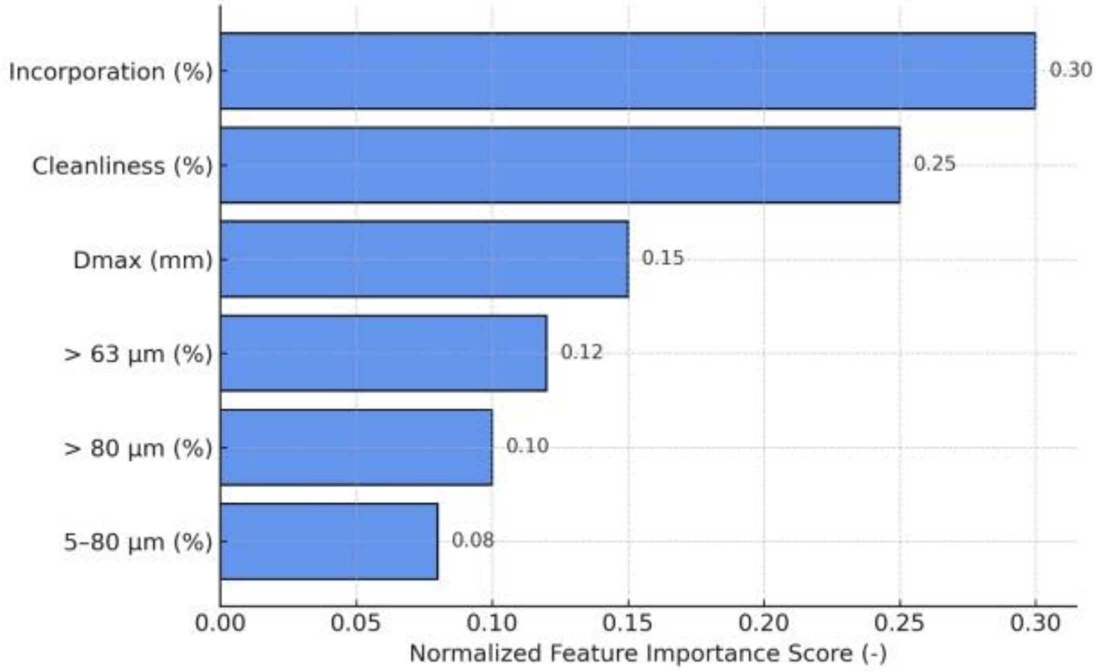

Figure 5. Importance of input variables based on Random Forest model. Variables are ranked by their impact on the compressive strength prediction.

The results obtained confirm that ensemble learning models, particularly Random Forest and XGBoost, provide highly accurate predictions for concrete properties, which aligns with findings from other studies. For instance, Wang & al. (2022)[12] demonstrated that Random Forest and XGBoost were the most effective algorithms for predicting the compressive strength of concrete containing industrial by-products such as fly ash and slag, with R2R^2R2 values exceeding 0.95 474747. Similarly, Tran & al. (2021)[20] used artificial neural networks (ANNs) to predict the mechanical performance of stabilized dredged sediments, achieving an accuracy above 90% 404040. The results of this study further validate the reliability of AI-based models, particularly ensemble learning techniques, in optimizing concrete formulations with unconventional materials.

However, some studies have highlighted the limitations of AI models when applied to small datasets. Segovia & al. (2025)[7] reported that while Random Forest and XGBoost performed well, their accuracy decreased when applied to datasets with high variability in sediment properties, suggesting that larger and more diverse training datasets are necessary to improve generalization 323232. In contrast, the present study benefited from a relatively extensive dataset, allowing for robust training and validation, which contributed to the high predictive accuracy of the models.

One of the key contributions of this study is the identification of the most influential granulometric fractions on compressive strength. The results indicate that fractions greater than 63 µm and 80 µm play a critical role in enhancing concrete performance, while particles smaller than 50 µm have a less significant impact. This observation aligns with previous research emphasizing the importance of optimizing the particle size distribution to improve mechanical properties.

For example, Achour & al. (2019)[21] found that dredged marine sediments could be incorporated into concrete up to 12.5% without significant loss of strength, but exceeding this threshold led to decreased durability due to increased porosity 333333. Similarly, Soleimani & al. (2023)[4] observed that the nature of the sediments played a crucial role: sandy sediments allowed for incorporation levels of up to 30–40%, whereas clay-rich sediments negatively affected concrete strength beyond 10–15% 232323. The findings of this study reinforce these conclusions by showing that controlling the granulometric composition of sediments is essential to maximizing compressive strength while maintaining durability.

Additionally, the sensitivity analysis conducted in this study highlights the significant role of sediment cleanliness and maximum aggregate diameter in determining concrete performance. These results are consistent with Alloul. (2023)[22], who demonstrated that impurities and high fine content in dredged sediments could negatively affect mechanical properties unless properly treated 363636. Although this study focuses on mechanical performance, it is essential to consider the environmental impact of using dredged sediments. Future work should assess potential leaching behavior and chemical risks through extraction tests to ensure environmental safety.

This underscores the importance of pre-treatment processes such as washing and screening to remove undesirable fines and improve sediment quality before incorporation into concrete.

4. Conclusion

This study successfully demonstrates the potential of machine learning algorithms in predicting the optimal incorporation percentage of marine dredged sediments in concrete and estimating the compressive strength (RC) of the resulting formulations. By evaluating a variety of machine learning models, including Random Forest (RF), XGBoost, LightGBM, Linear Regression (LR), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Decision Tree (DT), AdaBoost, and Extra Trees (ET), the study provides valuable insights into the complex relationships between the granulometric characteristics of sediments, the percentage of sediment incorporation, and the mechanical performance of concrete. The Random Forest (RF) model, in particular, demonstrated exceptional performance, outperforming the other models in terms of both accuracy and robustness. Achieving an R^2 of 0.98, a root mean square error (RMSE) of 0.15 MPa, and a mean absolute error (MAE) of 0.12 MPa, the RF model highlights its ability to capture the nonlinear interactions between the various input variables, such as granulometric characteristics, sediment cleanliness, and maximum aggregate diameter. These findings are critical, as they underscore the value of machine learning techniques in addressing complex problems that are traditionally difficult to model through conventional methods.

The sensitivity analysis conducted in this study revealed that key parameters, including the incorporation percentage, sediment cleanliness, and maximum aggregate diameter, exert the greatest influence on the compressive strength of the concrete. This information is vital for optimizing the performance of concrete mixtures, as it provides a clearer understanding of the factors that need to be carefully controlled in order to achieve desired material properties. The identification of these critical parameters paves the way for future optimization efforts, ensuring that the incorporation of marine dredged sediments into concrete could be fine-tuned to produce high-performance, sustainable materials.

Furthermore, the use of machine learning models provides a powerful tool for predicting the behavior of concrete mixtures in real-world scenarios, enabling faster and more efficient development of optimized formulations. The ability to predict both the incorporation percentage and the compressive strength with high accuracy is a significant step forward in the field of sustainable construction materials, particularly in the context of utilizing waste products such as marine dredged sediments. This approach not only addresses the challenge of reducing the environmental impact of dredging activities but also contributes to the broader goals of circular economy principles by valorizing materials otherwise considered waste.

Looking ahead, several avenues for further research remain. Expanding the dataset to include a wider range of sediment types, particularly from different geographic regions or with varying mineral compositions, could enhance the generalizability of the models. Additionally, investigating other important properties of concrete, such as permeability and chemical durability, would provide a more comprehensive understanding of the material's long-term performance and sustainability. The experimental validation of the optimized formulations is also crucial in confirming the practical relevance and industrial applicability of the results. However, the relatively small dataset (105 samples) may present risks of overfitting, limiting the model's generalization capabilities. Additionally, the variability of sediments from different regions may affect the model's transferability. Future efforts should aim at expanding the dataset and validating the model in diverse geographic and environmental conditions, by testing the predictions in real-world conditions, it will be possible to ensure that the optimized concrete formulations not only perform well in laboratory settings but also meet the rigorous standards required for construction use.

In conclusion, this study illustrates the promising potential of machine learning in optimizing the use of marine dredged sediments in concrete production. By accurately predicting the optimal incorporation percentage and compressive strength, it provides a robust and scalable methodology for developing sustainable construction materials. The insights gained from this research could support future development of environmentally friendly concrete formulations, supporting the transition to a more sustainable construction industry. Ultimately, this work contributes to both the scientific understanding of material behavior and the practical implementation of sustainable practices in the built environment.

Conflict of interest

The authors declare no conflict of interest.

References

1. Benhaddou K, Souileh A, Mabrouk A, Ouadif L, Abdelhak S, Baba K et al. Sustainable valorization of marine dredged sediments from Jebha port as a partial sand replacement in eco-friendly concrete. Journal of Ecological Engineering. 2025;26(6).
2. Safhi AEM, Mejjad N, El FadilI H, Bortali M (2024) Dredged materials in Morocco: Current practices, policies, and roadmap for sustainable management. Case Studies in Construction Materials 20:e03045. https://doi.org/10.1016/j.cscm.2024.e03045
3. Zhao Z, Benzerzour M, Abriak N-E, et al (2018) Use of uncontaminated marine sediments in mortar and concrete by partial substitution of cement. Cement and Concrete Composites 93:155–162. https://doi.org/10.1016/j.cemconcomp.2018.07.010
4. Soleimani T, Hayek M, Junqua G, et al (2023) Environmental, economic and experimental assessment of the valorization of dredged sediment through sand substitution in concrete. Science of The Total Environment 858:159980. https://doi.org/10.1016/j.scitotenv.2022.159980
5. Chou J-S, Lin J-W (2020) Risk-Informed Prediction of Dredging Project Duration Using Stochastic Machine Learning. Water 12:1643. https://doi.org/10.3390/w12061643
6. Safhi AEM, Dabiri H, Soliman A, Khayat KH (2024) Prediction of self-consolidating concrete properties using XGBoost machine learning algorithm: Rheological properties. Powder Technology 438:119623. https://doi.org/10.1016/j.powtec.2024.119623
7. Segovia JA, Toaquiza JF, Llanos JR, Rivas DR (2023) Meteorological Variables Forecasting System Using Machine Learning and Open-Source Software. Electronics 12:1007. https://doi.org/10.3390/electronics12041007
8. Jianing P, Bai K, Solangi YA, et al (2024) Examining the role of digitalization and technological innovation in promoting sustainable natural resource exploitation. Resources Policy 92:105036. https://doi.org/10.1016/j.resourpol.2024.105036
9. Cui Y, Wang L, Liu J, et al (2022) Impact of particle size of fly ash on the early compressive strength of concrete: Experimental investigation and modelling. Construction and Building Materials 323:126444. https://doi.org/10.1016/j.conbuildmat.2022.126444
10. Zhou A, Chen J, Li K, et al (2024) Developing green and economical low-alkalinity seawater sea sand concrete via innovative processing underground sediment. Journal of Cleaner Production 443:140927. https://doi.org/10.1016/j.jclepro.2024.140927

11. Afolayan A, Mildner M, Hotěk P, et al (2024) Characterization of Mineralogical and Mechanical Parameters of Alkali-Activated Materials Based on Water Sediments Activated by Potassium Silicate. Buildings 14:3077. https://doi.org/10.3390/buildings14103077
12. Hongwei Wang RZ, Wang D (2023) Predicting the compaction parameters of solidified dredged fine sediments with statistical approach. Marine Georesources & Geotechnology 41:195–210. https://doi.org/10.1080/1064119X.2021.2023827
13. Barhrhouj A, Ananou B, Ouladsine M (2025) Assessing the Impact of Temporal Data Aggregation on the Reliability of Predictive Machine Learning Models. In: Julian V, Camacho D, Yin H, et al (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2024. Springer Nature Switzerland, Cham, pp 481–492
14. James G, Witten D, Hastie T, et al (2023) Linear Regression. In: An Introduction to Statistical Learning. Springer International Publishing, Cham, pp 69–134
15. Kang M-C, Yoo D-Y, Gupta R (2021) Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Construction and Building Materials 266:121117. https://doi.org/10.1016/j.conbuildmat.2020.121117
16. Chicco D, Warrens MJ, Jurman G (2021) The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science 7:e623. https://doi.org/10.7717/peerj-cs.623
17. Hodson TO (2022) Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci Model Dev 15:5481–5487. https://doi.org/10.5194/gmd-15-5481-2022
18. Robeson SM, Willmott CJ (2023) Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. PLoS ONE 18:e0279774. https://doi.org/10.1371/journal.pone.0279774
19. Ali R, Muayad M, Mohammed AS, Asteris PG (2023) Analysis and prediction of the effect of Nanosilica on the compressive strength of concrete with different mix proportions and specimen sizes using various numerical approaches. Structural Concrete 24:4161–4184. https://doi.org/10.1002/suco.202200718
20. Benayad A, Bikri S, Hindi Z, et al (2024) Transition toward Sustainability in the Moroccan Food System: Drivers, Outcomes, and Challenges. World 5:627–644. https://doi.org/10.3390/world5030032
21. Achour R, Zentar R, Abriak N-E, et al (2019) Durability study of concrete incorporating dredged sediments. Case Studies in Construction Materials 11:e00244. https://doi.org/10.1016/j.cscm.2019.e00244
22. Alloul A High-performance geopolymer and low-carbon cementitious binders using flash-calcined dredged sediments and excavated clays