ROAD MATERIALS AND PAVEMENT DESIGN
2025, VOL. 26, NO. 3, 618–657
https://doi.org/10.1080/14680629.2024.2374885

Response surface regression and machine learning models to predict the porosity and compressive strength of pervious concrete based on mix design parameters

Navaratnarajah Sathiparan [a], Sathushka Heshan Wijekoon[a], Rinduja Ravi[a], Pratheeba Jeyananthan [b] and Daniel Niruban Subramaniam [a]

[a]Department of Civil Engineering, Faculty of Engineering, University of Jaffna, Jaffna, Sri Lanka; [b]Department of Computer Engineering, Faculty of Engineering, University of Jaffna, Jaffna, Sri Lanka

ABSTRACT
This study investigates the influence of aggregate size, aggregate-to-cement ratio, and compaction effort on pervious concrete's porosity and compressive strength. It proposes using response surface methodology and machine learning techniques to predict porosity and compressive strength. Fifteen mix designs (three aggregate sizes and five aggregate-to-cement ratios) with seven compaction energy levels were employed. Experimental data analysis using various regression models revealed that the quadratic model provided the best fit for predicting both porosity and compressive strength. Machine learning models were employed to predict porosity and compressive strength more accurately. Among the models investigated, the artificial neural network achieved superior performance across all datasets (training, testing, and validation). This suggests the artificial neural network model can effectively capture the complex relationships between input and response variables. Sensitivity analysis using SHAP (SHapley Additive exPlanations) revealed that compaction energy significantly impacts both porosity and compressive, while aggregate size has the most negligible influence.

ARTICLE HISTORY
Received 11 November 2023
Accepted 25 June 2024

KEYWORDS
Pervious concrete; porosity; compressive strength; response surface regression; machine learning

Abbreviations used in this study

AC ratio	Aggregate-to-cement ratio
ANN	Artificial neural network
BTR	Boosted Tree Regression
CE	Compaction energy
f_{cp}	Compressive strength of pervious concrete
LR	Linear regression
MAE	Mean Absolute Error
ML	Machine learning
R^2	Coefficient of determination
RFR	Random Forest regression
RMSE	Root Mean Squared Error
S	Aggregate size
SHAP	SHapley Additive exPlanations

CONTACT Navaratnarajah Sathiparan nsakthiparan@yahoo.com

v_p Porosity of the pervious concrete

1. Introduction

Pervious concrete has a voided structure consisting of cement, coarse aggregate, and water. This is most relevant for use in permeable pavements, and due to its diverse pore structure, it possesses several environmental benefits (Li et al., 2023). Pavements made of pervious concrete allow water to permeate through the paved surfaces and pass into the ground, potentially reducing flash floods and erosion (Abdelhady et al., 2021; Seifeddine et al., 2023; Singh et al., 2023). Moreover, pervious concrete can absorb and dissipate heat, unlike conventional pavement materials (Wang et al., 2019). The presence of many interconnected pores makes pervious concrete suitable for sound insulation (Zhang et al., 2020). Along with enhancing all the advantages obtained from pervious concrete and reducing the demand for conventional stormwater management systems, pervious concrete is most suitable for pavement construction (Mazur & Kotwa, 2019). Porosity (v_p) and compressive strength (f_{cp}) are two critical indicators of the performance of the pervious concrete (Chandrappa & Biligiri, 2019; Radlińska et al., 2012; Sriravindrarajah et al., 2012). The v_p varies within 15–35% (Elango et al., 2021; Neithalath et al., 2010; Pradhan & Behera, 2022), whereas the f_{cp} lies within a range of 2.8 - 28 MPa (ACI-522R, 2010). These properties are dependent on aggregate size, aggregate-to-cement (AC) ratio, water-to-cement (W/C) ratio, and compaction energy (Anburuvel & Subramaniam, 2022a; Liu et al., 2024).

The aggregate-to-cement ratio is one critical factor influencing pervious concrete properties (Lyu et al., 2024). This ratio directly affects the thickness of the cement paste coating the aggregates (Seifeddine et al., 2022; Xiong et al., 2023). Numerous studies (Najah et al., 2021) have shown that a higher AC ratio leads to increased v_p, which is desirable for water infiltration but comes at the expense of lower f_{cp}. This occurs because, with a higher AC ratio, there is limited cement paste to adequately coat the aggregates and develop a strong bond. Consequently, the thin paste layer struggles to resist applied pressure, leading to premature failure of the pervious concrete (AlShareedah & Nassiri, 2021). While f_{cp} of pervious concrete can be modelled using exponential functions (Ryshkewitch, 1953; Yahia & Kabagire, 2014), it's important to note that excessively low AC ratios (below 3.0) can also be detrimental. This can be attributed to the excessive amount of cement paste, which leads to the formation of voids and a decrease in the packing density of the aggregates (Wijekoon et al., 2023). Subramaniam and Sathiparan (2022) attributed a reduction in packing density and a consequent weakening of the concrete to an excess of cement paste. This excess paste was found to create gaps within the concrete matrix.

Pervious concrete typically utilises single-sized aggregates (Xie et al., 2019; Zhong et al., 2018). However, the influence of aggregate size on f_{cp} remains a topic of debate within current research. Specific investigations show smaller uniform-sized aggregates increased the f_{cp}, enhancing the total performance of pervious concrete (Ćosić et al., 2015; Furkan Ozel et al., 2022; Yu et al., 2019b). The smaller aggregates enhance f_{cp} by increasing their contact area with the cement paste, leading to a stronger bond ((Agar-Ozbek et al., 2013; Chindaprasirt et al., 2009; Kant Sahdeo et al., 2020; Yang & Jiang, 2003; Zhong & Wille, 2016). This logic holds, but only when there's sufficient cement paste in the mix, as dictated by the AC ratio. Conversely, some studies advocate for using larger aggregates to enhance concrete strength (Huang et al., 2020b). This argument centres on the reduced requirement for cement paste to coat larger aggregate surfaces. Consequently, a thicker and potentially stronger paste layer form between the aggregates, potentially contributing to improved overall strength. This implies that pervious concrete with larger aggregates might outperform smaller ones, but only at higher AC ratios (Chindaprasirt et al., 2009). Further complicating the picture, Yu et al. (2019a) observed that larger aggregates reduce the amount of small voids within the concrete. The influence of tiny voids on overall v_p necessitates consideration of an additional factor impacting the relationship between aggregate size and compressive strength in pervious concrete. This highlights the importance of balancing the AC ratio and cement paste content to optimise aggregate size selection. Studies suggest a 9.5 mm aggregate size might be a suitable starting point for achieving this balance (Elango & Revathi, 2017). However, further research is needed to establish the ideal size based on specific project requirements (Bright Singh & Madasamy, 2022).

Effective compaction energy plays a critical role in shaping the properties of pervious concrete. It influences the arrangement of the binder-coated aggregates and the overall structure (Kevern et al., 2009; Lian & Zhuge, 2010). Compaction ensures uniform aggregate distribution and even binder coating, both essential for optimal performance. However, there's a catch: excessive compaction can clog the pores, hindering the permeability that defines pervious concrete ((Kevern et al., 2009; Lian & Zhuge, 2010; Pieralisi et al., 2016). Therefore, striking a balance during compaction is crucial. Several compaction methods exist, including static loading, standard Proctor rammer, vibration, rodding, and Marshall rammer (Kant Sahdeo et al., 2020; Lian & Zhuge, 2010; Putman & Neptune, 2011). Standard Proctor compaction is favoured for its consistent and uniform results (Putman & Neptune, 2011). The key is understanding the impact of compaction energy. As compaction increases, the aggregates pack tighter within the cement paste matrix. This translates to higher density and f_{cp}, but unfortunately, it comes at the expense of reduced v_{p}. Research suggests a 'sweet spot' exists where both porosity and f_{cp} reach a stable value. Using a standard Proctor hammer, this balance can be achieved with 30 to 45 blows of compaction (Anburuvel & Subramaniam, 2022b). The compaction energy also affects the functional qualities of pervious concrete. The packing of aggregates is adversely impacted by low AC ratios when accompanied by reduced compaction effort. The elimination of voids is hindered by insufficient compaction effort (Alimohammadi et al., 2021; Ferić et al., 2023).

Recent research has yielded substantial findings on the impact of aggregate size, aggregate-to-cement ratio, and compaction effort on the characteristics of pervious concrete (Anburuvel & Subramaniam, 2023). The authors believe that v_{p} and f_{cp} are influenced by each of these elements individually and emphasise the need to address the combined impact of these three design factors due to their interdependent nature (Cai et al., 2024; Li et al., 2021). Based on the empirical relationship between aggregate size (S in mm), aggregate-to-cement ratio (AC), compaction effect (B in number of blows), porosity (v_{p} in %) and compressive strength (f_{cp} in MPa), Wijekoon et al. (2023) proposed the Equations (1) and (2) for v_{p} and f_{cp} pervious concrete, respectively.

$$v_p = -6.93 + 0.42S + 4.85AC + (28.95 + 10.18/S - 3.20AC)e^{(-0.0102-0.0023S-0.0055AC)B} \tag{1}$$

$$f_{cp} = 10.59 - 0.22S - 0.51AC + \left(470.07 \times 0.96^{S} \times AC^{-2.31}\right)\left(1 - e^{(-0.0853-0.0007S+0.1902/AC)B}\right) \tag{2}$$

While the equation developed in the study demonstrates a satisfactory fit with the experimental data at hand, it should be noted that the equation is highly intricate and only exhibits a modest fit compared to existing literature. In contemporary engineering, surface regression models and machine learning methodologies have gained significant traction in forecasting the properties of construction materials (Ahmad et al., 2022; Feng et al., 2020; Marani & Nehdi, 2020; Song et al., 2021; Zhang et al., 2022b). Response Surface Methodology (RSM) primarily utilises the statistical regression method because of its practicality, cost-effectiveness, and relative ease of implementation. The first- and second-order polynomial equation was developed using RSM (Nooraziah & Tiagrajah, 2014; Veza et al., 2023). The term 'polynomial model' is commonly used to denote a regression model. Because the properties of pervious concrete are highly dependent on how it's mixed and influenced by many factors, machine learning can be a powerful tool to predict its characteristics (Sathiparan et al., 2024b; Sathiparan et al., 2024c). Also, machine learning models are becoming more accessible to users with open-source software while also addressing the limitations of mathematical modelling (Li et al., 2022; Sun et al., 2019; Zhang et al., 2020). Machine learning adaptation over traditional statistical methods, especially in concrete science, can model non-linear data with complex interactions and missing values (Boddy & Morris, 1999; Zhao & Du, 2016). Table 1 summarises the machine learning technique used to predict the characteristics of pervious concrete. Although some research focuses on predicting f_{cp} using the aggregate size and aggregate-to-cement ratio, compaction energy was still not considered in these studies. As published literature shows, compaction energy is a significant parameter influencing the v_{p} and f_{cp}.

Predicting pervious concrete's porosity and compressive strength through machine learning and statistical models based on mix design is a significant advancement. This allows engineers to optimise concrete performance for specific projects, reduce trial-and-error approaches, and promote sustainable construction by ensuring effective stormwater management through targeted porosity levels. The major research question is how effective the response surface regression and machine learning models are for accurately predicting the porosity and compressive strength of pervious concrete based on the chosen mix design parameters. To develop models using these techniques and evaluate their effectiveness in predicting the two key performance indicators (porosity and compressive strength) of pervious concrete based on the variables used in designing the mix (aggregate size, AC ratio and compaction energy). This study investigates using two modelling approaches to predict the porosity and compressive strength of pervious concrete. The first approach involves developing response surface models with varying levels of complexity (linear, interaction, quadratic, and cubic). The second leverages machine learning algorithms like linear regression, artificial neural networks, and decision tree regressions. To train and test these models, an experiment was conducted using fifteen mix designs in total consisting of three aggregate sizes (8.5, 15.0 and 21.5 mm) and five aggregate-to-cement ratios (3.0, 3.5, 4.0, 4.5 and 5.0) with seven-level of compaction energy (by varying compaction blow of 0, 15, 30, 45, 60, 75 and 90). The performance of all models was compared using metrics like root mean squared error (RMSE) and R^2, with lower RMSE and higher R-squared indicating better performance. Ultimately, the goal is to establish models that can accurately predict the properties of pervious concrete based on its mix design.

2. Experimental programme

2.1. Material used

The composition of pervious concrete mixture typically consists of cement, aggregates, and water. The selection of the coarse aggregate for this study was igneous granite-gneisses rock. For the preparation of pervious concrete mixes, the chosen binder material was 43-grade ordinary Portland cement (OPC), complying with IS 269 (2015) standards. The bulk density and specific gravity of the cement employed in this research were 1280 kg/m^3 and 3.15, respectively. The aggregates underwent a sieving process and were subsequently categorised into three distinct size ranges: 5–12 mm, 12–18 mm, and 18–25 mm. The quality of the aggregates was assessed and the findings are displayed in Table 2. Aggregates used was meeting the specifications outlined in BIS 383 (2016) for construction applications.

2.2. Mix design

A total of 315 cubes were produced for the experimental study. Fifteen different mix designs, including three different aggregate sizes (5–12 mm, 12–18 mm and 18–25 mm) and aggregate-to-cement ratios, were altered across a range of 3.0, 3.5, 4.0, 4.5, and 5.0. Additionally, the compaction energy was varied by applying seven different levels of blows, namely 0, 15, 30, 45, 60, 75 and 90 from a standard proctor rammer. As per the existing body of literature, the water-to-binder ratio was consistently maintained at 0.3 throughout all mix proportions to achieve a slum*p* value of zero. Table 3 displays the mix design of the pervious concrete. The necessary quantities of cement, water, and coarse aggregates were determined and measured individually using an electronic balance with a minimum precision of 0.1 mg based on the AC ratio and W/C ratio. These materials were then combined in an electrically operated concrete mixer.

Pervious concrete relies on a delicate balance between strength and permeability, achieved through the AC ratio and specific aggregate size. For a particular aggregate size in each mix design, aggregate and cement content weight operate in opposite directions as the AC ratio changes. With

Table 1. Summary of machine learning methods used to predict the properties of pervious concrete.

Ref.	Input parameters	Response parameters	Machine learning method used[a]	No. of data used
Sun et al. (2019)	• Aggregate size • Aggregate-to-cement ratio • Water-to-cement ratio	• Permeability • Compressive strength	SVR	90
Sathiparan et al. (2023a)	• Aggregate size • Aggregate-to-cement ratio • Compaction energy	• Ultrasonic pulse velocity	LR, ANN, BTR, RFR, SVR and XG boost	75
Subramaniam et al. (2023)	• Aggregate size • Aggregate-to-cement ratio • Compaction energy	• Electrical resistivity	LR, ANN, BTR, KNN, RFR and SVR	75
Huang et al. (2020a)	• Aggregate-to-cement ratio • Aggregate grading	• Permeability	RFR	69
Malami et al. (2022)	• Curing days • Rice husk ash content • Calcium carbide waste content	• Flexural strength	ANN, SVR and ANFIS	42
Sudhir Kumar et al. (2024)	• Aggregate size • Ground Granulated Blast-furnace Slag content • Curing days	• Compressive strength • Flexural strength • Permeability	ANN	36
Huang et al. (2022)	• Aggregate size • Sample thickness • clogging sands	• Permeability	SVR	84
Le et al. (2022)	• Aggregate size • Aggregate-to-cement ratio • Water-to-cement ratio • Silica fume content • Sand content	• Porosity • Compressive strength	ANN, RFR, SVR, XGN	164
Ahmad et al. (2023)	• Cement content • Gravel content • Waste glass power content • Water-to-cement ratio • Curing period	• Compressive strength	ANN	99

[a]LR: Linear Regression, ANFIS: Adaptive Neuro-Fuzzy Inference System, ANN: Artificial Neural Network, BTR: Boosted Tree regression, KNN: K-Nearest Neighbors' regression, RFR: Random Forest Regression, SVR: Support Vector Regression, and XGboost: EXtreme gradient boosting

a lower AC ratio, meaning less aggregate compared to cement, the total weight of aggregate used is lower. This allows more cement paste to fill the voids between aggregate particles, creating a stronger, denser structure. However, this comes at the cost of reduced permeability. On the other hand, a higher AC ratio necessitates more aggregate by weight. This allows for the interconnected voids that pervious concrete needed for drainage. While this increases permeability, it comes with a trade-off. Less cement is proportionally required, which can be cost-effective. However, it's crucial to remember that pervious concrete already has lower inherent strength than regular concrete. An excessively high AC ratio can jeopardise the structural integrity of the mix if not carefully balanced.

The preparation procedure for the pervious concrete mix is illustrated in Figure 1. The freshly prepared mixture underwent a rigorous mixing process to achieve a uniform consistency. The mortar mix was initially subjected to slump test as per BS 1881-102 and then transferred into 150 mm^3 molds. The mix was placed in three layers of equal height and compaction was applied to each layer using a standard protector compaction hammer that fell freely from 300 mm above the mix surface (5.08 cm diameter face and weighs 2.5 kg). Compaction energy due to a hammer blow is calculated using potential energy. The potential energy (PE) stored in the hammer due to its position at the top is calculated using Equation (1). This assumes negligible energy loss due to air resistance or other factors during the fall. Therefore, the compaction energy delivered by the hammer blow can be estimated as the potential energy calculated in Equation (3).

$$PE = n * m * g * h \tag{3}$$

where: PE is the potential energy in Joules (J), n is the number of blows, m is the mass of the hammer in kilograms (kg), g is the acceleration due to gravity (usually 9.81 m/s^2) and h is the drop height in metres (m)

Table 2. Property of the coarse aggregate.

Properties	Aggre.5–12 mm	Aggre.12–18 mm	Aggre.18–25 mm	Test standard
Bulk density (kg/m^3)	1498	1491	1480	BS-EN-1097 (2020)
Specific gravity	2.73	2.72	2.72	BS-EN-1097 (2020)
Moisture content (%)	0.40	0.35	0.32	BS-EN-1097 (2020)
Absorption (%)	1.35	1.20	1.10	BS-EN-1097 (2020)
Flakiness Index	14.8	14.7	14.2	BS-EN-933 (2017)
Aggregate Impact value – AIV (%)	25.6	25.5	25.3	BS-EN-1097 (2020)
Aggregate Crushing Value – ACV (%)	30.5	30.5	30.4	BS-EN-1097 (2020)
Elongation Index (%)	23.95	24.02	24.80	BS-EN-933 (2017)

Table 3. Mix design for 1 m^3 pervious concrete mix.

Mix ID	AC ratio	Cement (kg)	Aggregate 5–12 mm (kg)	Aggregate 12–18 mm (kg)	Aggregate 18–25 mm (kg)	Water (L)
S-3.0	3.0	466.8	1400.3			140.0
S-3.5	3.5	413.3	1446.5			124.0
S-4.0	4.0	370.8	1483.2			111.2
S-4.5	4.5	336.2	1513.0			100.9
S-5.0	5.0	307.6	1537.8			92.3
M-3.0	3.0	466.8		1400.3		140.0
M-3.5	3.5	413.3		1446.5		124.0
M-4.0	4.0	370.8		1483.2		111.2
M-4.5	4.5	336.2		1513.0		100.9
M-5.0	5.0	307.6		1537.8		92.3
L-3.0	3.0	466.8			1400.3	140.0
L-3.5	3.5	413.3			1446.5	124.0
L-4.0	4.0	370.8			1483.2	111.2
L-4.5	4.5	336.2			1513.0	100.9
L-5.0	5.0	307.6			1537.8	92.3

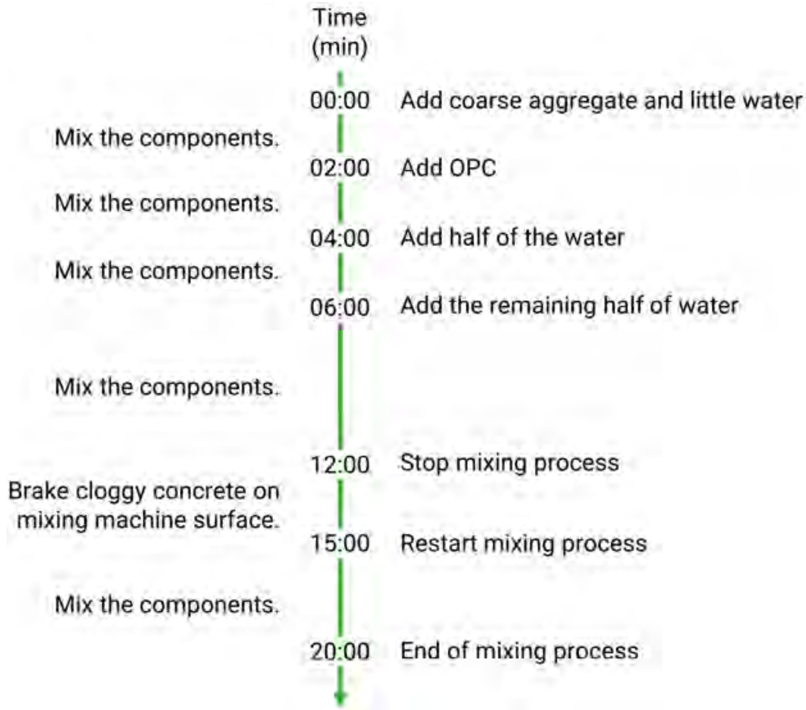

Figure 1. Preparation procedure for pervious concrete mix.

Following the third layer, an iron rod was used to level and remove the extra mix. The cast cubes were air-dried for one day and then transferred to the curing tank for 28 days.

2.3. Testing

2.3.1. Porosity

Following ASTM C1754 (2012), the porosity of each pervious concrete cube specimen was measured. This involved weighing the specimens in air and submerged in water to determine the amount of water absorbed through the pores. The test setup are presented in Figure 2(a) and 2(b). The weight difference between the dried sample and the sample dipped in water was used to compute the v_p of the pervious concrete. The formula for the v_p of pervious concrete is represented by Equation (4).

$$vp = \left(1 - \frac{w_d - w_w}{\rho_w V}\right) \tag{4}$$

where, v: total void ratio, w_d: weight of oven-dried sample [kg], w_w: weight of sample in water [kg], V: volume of the sample [m^3] and ρ_w: density of water [kg/m^3]

2.3.2. Compressive strength

The measurement of f_{cp} was conducted following the guidelines described in BS1881. Upon extraction from the curing tank after a curing duration of 28 days, the specimens were subjected to desiccation inside the laboratory environment. Subsequently, a f_{cp} assessment was conducted. The experiment utilised a Universal Testing Machine (UTM) to achieve the objective. The specimens were subjected to loading at a 1 mm/min rate, with displacement control, as depicted in Figure 2(c). The mean value of six specimens was calculated and subsequently utilised for subsequent analysis.

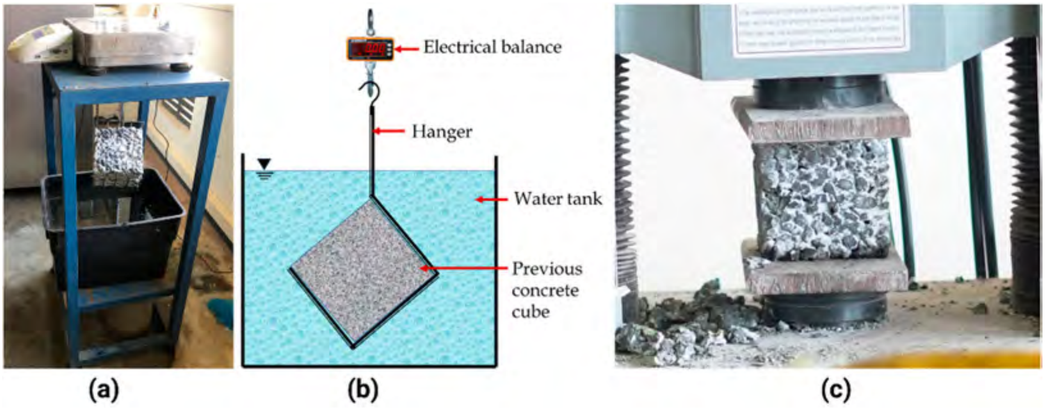

Figure 2. (a) Equipment used for porosity, (b) porosity measurement setup, and (c) compression test.

3. Statistical analysis and modelling techniques

3.1. Statistical analysis

A correlation analysis was performed on the experimental data to understand the relationships between mix design choices and the resulting properties of pervious concrete. This analysis explored the connections between three input variables and two output variables. The input variables, considered independent, were aggregate size (S: Ranging from 8.5 mm to 21.5 mm), Aggregate-to-cement ratio (AC: Varying from 3.0 to 5.0) and compaction energy (CE ranging from 0 to 673 J and equivalent to 0 to 90 number of blows). The output variables, considered dependent, were Porosity (v_p: Values ranging from 11.2% to 43.4%) and Compressive strength (f_{cp}: Corresponding to a range of 2.3 to 28.4 MPa) Pearson's correlation coefficient was used to assess the strength and direction of the relationships.

3.2. Response surface model

A response surface model is a mathematical and statistical technique that explores the relationship between one or more response variables and several explanatory variables. To effectively address prediction challenges, it is advised to minimise the complexity of the model (Veza et al., 2023). The selection of a higher-order model may be considered inappropriate when doing regression analysis. Simultaneously, a simple model, such as a linear model, often experiences a deficiency in its ability to represent the data accurately. The present study used four response surface models from the basic linear model to a more complex cubic model. The data from the present experiment was used to predict the model parameter for the best-fit formula, and then data from published literature (Table 4) was used to validate the proposed formula.

3.2.1. Linear model

Linear regression (LR) is a statistical method used to describe the relationship between a dependent variable and one or more independent variables. The process involves determining the best-fitting linear regression model that accurately represents the relationship between the dependent and independent variables, as seen in Equation (5) (Forthofer et al., 2007). Within this context, the variables a_0 to a_4 represent model parameters. In the present study, aggregate size (S), aggregate-to-cement ratio (AC), and compaction energy (CE) are independent variables, and v_p and f_{cp} are dependent variables.

$$v_p \text{ or } f_{cp} = \alpha_0 + \alpha_1(S) + \alpha_2(AC) + \alpha_3(CE) \tag{5}$$

Table 4. Validation data set collected from published literature.

	Aggregate size (mm)	AC ratio	compaction energy (J)	Porosity range (%)	Compressive strength range (MPa)
Ferić et al. (2023)	6.0–12.0	3.5–4.5	111.2–561.0	11.2–27.3	7.68–24.40
Anburuvel and Subramaniam (2022b)	17–23	5.0	0–374.0	18.6–39.7	5.95–11.31
Elango and Revathi (2017)	6.3–12.5	3.3	224.4	14.8–19.5	14.00–22.00
Joshaghani et al. (2015)	7.0–16.0	3.5–6.9	0	32.6–37.6	5.50–8.20
Torres et al. (2015)	6.3–9.5	3.6–6.5	37.4–224.4	18.0–31.3	5.20–18.60
Arcolezi (2022)	4.75–19	5.0	74.8	28.4–33.9	9.28–13.70
Elizondo-Martinez et al. (2021)	9.5	5.0	74.8–374.0	21.4–33.9	-
Dall Bello De Souza Risson et al. (2021)	19	3.3	0–374.0	16.2–33.3	7.02–15.81
Joshi and Dave (2016)	8–15	4.0	0	30.2–35.1	9.33–10.37

3.2.2. Two-factor interaction (2FI) model

The 2FI response surface method is a statistical technique employed to model the correlation between many independent variables and a single dependent variable. The approach relies on the utilisation of the hypergeometric function 2FI and is used to fit a polynomial of the second degree to the given dataset, as demonstrated in Equation (6) (Hedayat & Pesotan, 1997). The resultant model can forecast the value of the dependent variable for any amalgamation of the independent variables.

$$v_p \; or \; f_{cp} = \beta_0 + \beta_1(S) + \beta_2(AC) + \beta_3(CE) + \beta_4(S)(AC) + \beta_5(S)(CE) + \beta_6(AC)(CE) \quad (6)$$

3.2.3. Quadratic model

The regression models were developed using the response surface method, namely the quadratic model, as described by Mao (Mao, 1981) and Khuri and Cornell (Khuri & JA, 1996). The mathematical representation of the model can be expressed as Equation (7) (Tan et al., 2008):

$$Y = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} \sum_{j=1, i \geq j}^{n} b_{ij} x_i x_j \quad (7)$$

The regression coefficients, b_0, b_i and b_{ij} represent the parameters of the regression model. The input variables in the regression function are denoted as x_i and x_j. The total number of variables in the model is n. The response variable is represented as Y.

In the present study, the number of variables (n) are three, namely: aggregate size (S), aggregate to cement ratio (AC) and compaction energy (CE). The response variable is $v_{\rm p}$ and $f_{\rm cp}$. The mathematical representation of the quadratic model can be expressed as Equation (8):

$$\begin{aligned} v_p or f_{cp} = {} & b_0 + b_1(S) + b_2(AC) + b_3(CE) + b_{11}(S)^2 + b_{22}(AC)^2 \\ & + b_{33}(CE)^2 + b_{12}(S)(AC) + b_{13}(S)(CE) + b_{23}(AC)(CE) \end{aligned} \quad (8)$$

3.2.4. Cubic model

The cubic regression model is a statistical strategy that aims to identify the cubic polynomial, which is of degree 3, that provides the optimal fit for a given dataset. The general mathematical representation of the cubic model can be expressed as Equation (9) (Kumar et al., 2020):

$$Y = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} \sum_{j=1, i \geq j}^{n} b_{ij} x_i x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1, i \geq j \geq k}^{n} b_{ijk} x_i x_j x_k \quad (9)$$

The regression coefficients, b_0, b_i, b_{ij} and b_{ijk} represent the parameters of the regression model. The input variables in the regression function are denoted as x_i, x_j and x_k. The total number of variables in the model is n. The response variable is represented as Y.

In the present study, the mathematical representation of the cubic model can be expressed as Equation (10):

$$\begin{aligned} v_p or f_{cp} = {} & b_0 + b_1(S) + b_2(AC) + b_3(CE) + b_{11}(S)^2 + b_{22}(AC)^2 \\ & + b_{33}(CE)^2 + b_{12}(S)(AC) + b_{13}(S)(CE) + b_{23}(AC)(CE) \\ & + b_{111}(S)^3 + b_{222}(AC)^3 + b_{333}(CE)^3 + b_{122}(S)(AC)^2 \\ & + b_{133}(S)(CE)^2 + b_{112}(S)^2(AC) + b_{113}(S)^2(CE) \\ & + b_{223}(AC)^2(CE) + b_{233}(AC)(CE)^2 + b_{123}(S)(AC)(CE) \end{aligned} \quad (10)$$

3.3. ML modelling

Figure 3 presents a comprehensive view of the workflow diagram for machine learning modelling. An extensive collection of 315 datasets was obtained from the experimental programme to establish a robust model for predicting the properties of pervious concrete. Out of the 315 available datasets, the researcher used 210, representing two-thirds of the data, to train the model. The remaining 105 datasets, constituting one-third of the data, were reserved for testing its performance. To validate the proposed model, data collected from published literature were utilised. Specifically, the dataset consisted of 92 observations for v_p and 80 observations for f_{cp}. In the current study, four machine learning techniques, including linear regression, artificial neural network, random forest regression, and boosted decision tree model, are employed to forecast the v_p and f_{cp}. Researchers utilised several evaluation criteria to assess and compare the various models' performance. These criteria focus on a model's scientific accuracy. A key indicator is the error percentage between predicted and experimental data – a lower percentage signifies better alignment. Another metric is the coefficient of determination (R^2), which approaches one as the model explains more of the data's variation. Additionally, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) measure the difference between predicted and actual values, with lower values indicating better agreement. Furthermore, researchers considered the a20 index, which reflects the proportion of predictions falling within an acceptable 20% deviation from experimental values (Sathiparan et al., 2024c). Finally, a Taylor diagram was employed to provide a comprehensive view of how each model performs across these metrics.

3.3.1. Linear regression

Linear regression is a fundamental machine learning approach employed to predict numerical values by utilising input information. The assumption is that a linear relationship exists between the features and the target variable. The model is trained to optimise the coefficients that best fit the given data, enabling it to generate predictions for novel inputs (Miller, 2023).

3.3.2. Artificial neural networks

Artificial neural networks (ANNs) are highly parallelised systems consisting of numerous interconnected elementary processors. ANNs are widely employed as a machine learning technique in various disciplines within the sciences and engineering domains (Hedayat & Pesotan, 1997). ANNs possess two primary hyperparameters that govern the network's architecture or topology: the number of layers and the number of nodes within each hidden layer. When setting your network, it is imperative to provide explicit values for these parameters (Montesinos López et al., 2022). The optimal approach for determining the appropriate hyperparameters for your particular predictive modelling task is through methodical testing using a reliable test framework. After doing multiple iterations, it has been shown that employing a two-layer architecture with six nodes per layer yields superior performance for the models.

3.3.3. Boosted decision tree

Boosted decision trees are an ensemble learning technique that amalgamates many decision trees to enhance the model's predictive accuracy. Boosting is an iterative strategy that involves training weak models on several data subsets and merging them to form a robust model (de Zarzà et al., 2023). Ensemble learning techniques like boosted decision trees utilise decision trees as base learners. These base learners are often called 'weak models' due to their limited individual accuracy. However, the boosting algorithm leverages this by iteratively focusing on previously misclassified instances. During each iteration, the weights of data points incorrectly classified by the preceding model increase. Subsequently, a new decision tree is trained on this weighted dataset, prioritising the learning of these challenging samples. This iterative process progressively improves the ensemble's overall performance, transforming a collection of weak learners into a robust prediction model (Alsahaf et al., 2022).

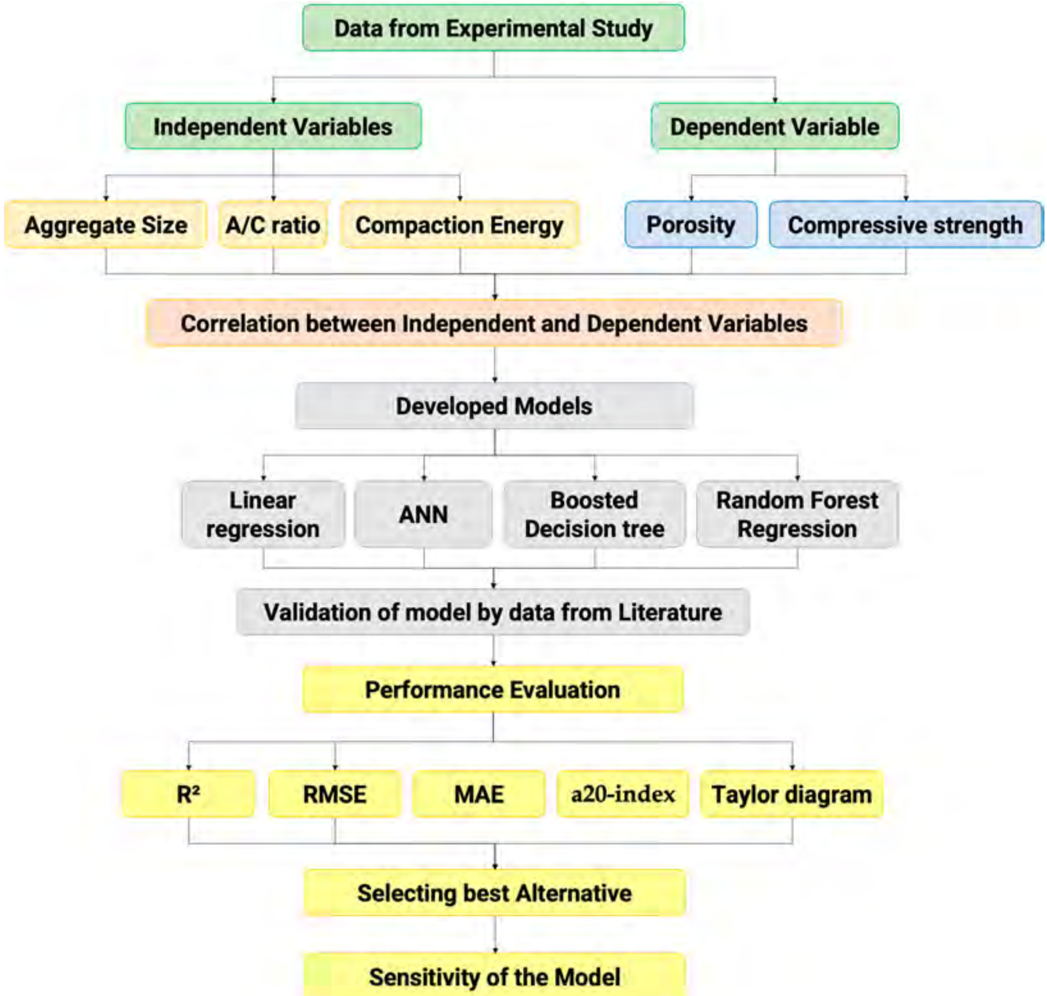

Figure 3. Outline of the ML modelling flowchart.

3.3.4. Random forest regression

Random forest regression, an ensemble learning technique, leverages the combined predictive power of multiple decision trees to enhance model precision. Similar to boosted decision trees, it utilises decision trees as weak learners. However, key distinctions exist in their training processes. Random forest trains each tree on a subset of data points and features (Khan et al., 2024). This approach, known as bagging, mitigates overfitting and improves model generalizability. The final prediction is the average of all individual tree forecasts. Conversely, boosted decision trees train sequentially, with each tree focusing on correcting the errors of its predecessor. This is achieved by assigning higher weights to misclassified samples in the training data (Durap, 2023). The final prediction aggregates all tree predictions, weighted based on their performance.

3.4. Performance indicators

Various performance indicators are employed to assess the efficacy of developed models, such as R^2, RMSE, MAE, and a20 index. In the scenario of an ideal prediction model, it is anticipated that the R^2 and a20 index values will converge towards a unity value. The a20 index quantifies the count of samples that satisfy specific requirements, specifically predicted values that fall within a 20% margin of experimental values. Equations (11) through (14) are employed for the computation of each of these criteria (Sathiparan et al., 2024b).

$$R^2 = \left(\frac{\sum_i (P_i - \bar{P})(E_i - \bar{E})}{\sqrt{\sum_i (P_i - \bar{P})^2} \sqrt{\sum_i (E_i - \bar{E})^2}} \right)^2 \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (E_i - P_i)^2}{N}} \tag{12}$$

$$MAE = \frac{\sum_{i=1}^{n} (|E_i - P_i|)}{N} \tag{13}$$

$$a20_{index} = \frac{N20}{N} \tag{14}$$

where, Pi: predicted value, *E*i: experimental value, $\bar{P}$: mean of predicted value, $\bar{E}$: mean of experimental value, *N*: total number of datasets, N_{20}: total number of predicted to the measured data of value ratio ranged from 0.8 to 1.2.

The R^2 coefficient and the a20 index exhibit a numerical scale ranging from zero to one, wherein one value is regarded as the optimal outcome. The RMSE and MAE values encompass the entire numerical range, ranging from zero to infinity. It is advisable to maintain low values, with zero being the optimal outcome (Sathiparan & Jeyananthan, 2024).

In addition, a statistical analysis was conducted for the response surface models to assess the relationships between the independent variables (e.g. aggregate size, AC ratio and compaction energy) and the dependent variable (e.g. porosity and compressive strength) of the pervious concrete mixes. This analysis involved two key steps:

- Model Sum of Squares: This test will evaluate the variation in the dependent variable explained by the fitted model compared to the total variation. A higher model sum of squares indicates a better fit of the model to the data.
- Model Summary Statistics: This will provide information about the model's overall performance, including parameters like R^2 (coefficient of determination). The R^2 represents the proportion of variance in the dependent variable explained by the model. Additionally, the model summary statistics will likely include F-statistic and *p*-value, which can be used to assess the significance of the model and individual predictor variables.

These statistical tests are used to understand the strength of the relationships between the variables and the overall effectiveness of the model in predicting the properties of pervious concrete.

3.5. Validation of the models

During the verification phase of validating v_p and f_{cp} predicted by the model, the analysis utilised data gathered from existing literature sources. Ninety-two data points for v_p and eighty data points for f_{cp} were obtained through a comprehensive review of nine published literature sources. This process began by collecting data on pervious concrete through an extensive search for relevant research articles. Scopus and Google Scholar databases were utilised to identify articles containing the terms 'pervious concrete,' 'compressive strength,' 'aggregate size,' 'aggregate to cement ratio,' and 'compaction' within their titles, abstracts, or keywords. Following the initial search, a screening process was implemented to eliminate duplicate articles and ensure the data originated from unique studies. The data were compiled and presented in Table 4, focusing on the variables of aggregate size, A/C ratio, and compaction energy. The study encompassed diverse aggregate sizes ranging from 4.75 mm to

23 mm, A/C ratios ranging from 3.3 to 7.0, and compaction efforts ranging from 0 to 561 J. The v_p and f_{cp} exhibited a range of values, with v_p ranging from 11.2% to 39.7% and f_{cp} ranging from 5.2 MPa to 24.4 MPa.

3.6. Sensitivity analysis

Machine learning models, particularly those with complex structures, can sometimes exhibit opaque behaviour, making it difficult to understand how they arrive at their predictions (Shah et al., 2022). We employed the SHAP (SHapley Additive exPlanations) analysis method to address this issue. This technique helps us evaluate the significance and impact of individual input parameters on the predicted compressive strength of pervious concrete (Quan Tran et al., 2022; Zhang et al., 2022a). This study uses SHAP analysis to gain insights into the behaviour of our high-performing machine learning model for predicting compressive strength in pervious concrete. Analysing the SHA*P* values can identify which input factors have the most significant influence on the model's predictions.

4. Experimental results

Figures 4 and 5 illustrate the discrepancy of v_p and f_{cp} with AC ratios and compaction energy for various aggregate sizes, respectively. Figure 4 reveals a consistent increasing trend for v_p as the AC ratio increased. This relationship has been observed in numerous studies (Chandrappa & Biligiri, 2016; Debnath et al., 2023; Sata & Chindaprasirt, 2020). Singh et al. (2020) highlights the established concept of A/C impacting porosity. With a higher A/C, there's less paste to fill the gaps, leading to increased void content. Through their experimental results, Le et al. (2022) demonstrates the positive correlation between A/C ratio and porosity. Luo et al. (2023) explore the impact of A/C (indirectly through binder content) on porosity, concluding that a higher A/C leads to higher porosity. This pattern is primarily attributed to the fact that as the AC ratio increases, there is insufficient paste to coat the aggregates adequately. Consequently, the aggregate mass per unit volume becomes higher, while the cement mass is compromised (Chindaprasirt et al., 2008; Khankhaje et al., 2024). This imbalance results in a greater void volume, leading to an increased v_p. Furthermore, when the aggregate size increases, there is a corresponding increase in v_p, however, this is not gradual (Fu et al., 2014). Several studies have observed this trend, such as Huang et al. (2020b) and Luo et al. (2023), which found a direct correlation between the size of coarse aggregates and porosity in pervious concrete mixes. This phenomenon occurs because smaller aggregates provide a larger contact area, effectively reducing the voids within the mixture. For a given AC ratio, v_p declines with increased compaction effort (Ibrahim et al., 2014). This is because compaction leads to enhanced aggregate packing density, thereby reducing the overall pore volume within the material. Chandrappa and Biligiri (2016) found an apparent decrease in porosity with increasing compaction effort, highlighting the trade-off between achieving the desired strength and maintaining sufficient permeability. Debnath et al. (2023) observed a similar trend, where higher compaction resulted in denser pervious concrete with lower porosity.

This experiment investigated the effects of aggregate size, water-cement ratio, and compaction energy on porosity. The minimum porosity, measured at 11.2%, was observed for a combination of 8.5 mm aggregate size, water-cement ratio of 3, and compaction energy of 448.8 J (equivalent to 60 blows). In contrast, the maximum porosity of 43.4% was obtained with a larger aggregate size of 21.5 mm, a higher water-cement ratio of 4.5, and no compaction applied.

As shown in Figure 5, a reduction in f_{cp} was observed as AC ratio increased. Several studies have documented this relationship, as Furkan Ozel et al. (2022), Xu et al. (2023) and Shilar et al. (2022) found that a higher A/C ratio resulted in lower compressive strength for pervious concrete mixes. This outcome stemmed from the content of the cement paste and aggregates present in the mix. Specifically, a low AC ratio allowed higher cement paste to envelop the aggregates and occupy the interstitial spaces, forming a dense and stronger pervious concrete matrix. However, as the AC ratio increased, insufficient cement paste became available to comprehensively coat the aggregates (Yahia & Kabagire, 2014). This

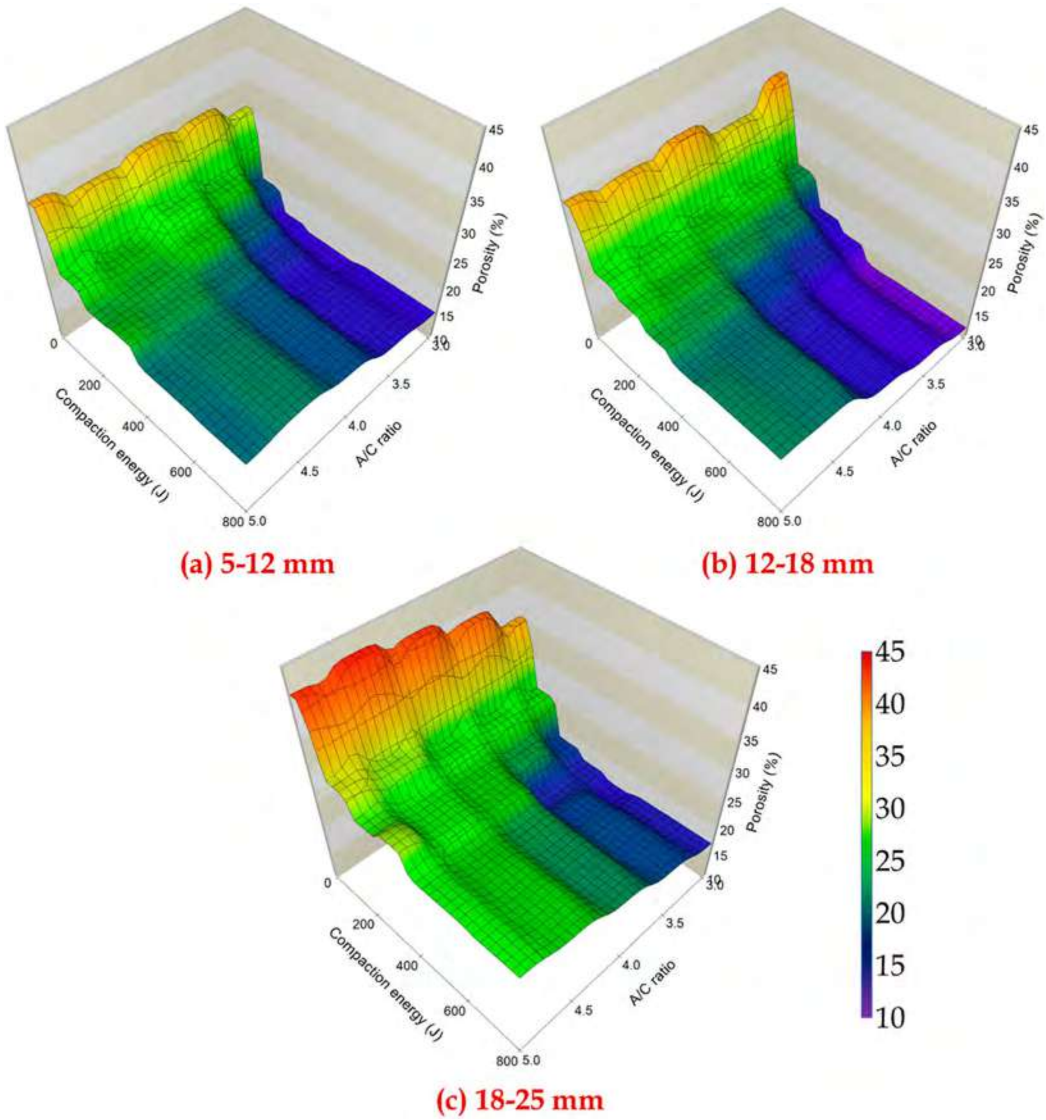

Figure 4. Variation of porosity of the pervious concrete with compaction energy and A/C ratio for the aggregate size of (a) 5–12 mm, (b) 12–18 mm, and (c) 18–25 mm.

deficiency in paste coverage translated into reduced structural integrity. As aggregate content or size increases within a reasonable range, the resistance to cracking generally increases. Larger aggregate particles can distribute the load over a larger area than smaller ones. This reduces stress concentration points within the concrete matrix (Ferić et al., 2023). A stronger interfacial transition zone (ITZ) is crucial as The ITZ is the zone between the cement paste and aggregate. Its strength is vital for load transfer between them (Li et al., 2012). The key is understanding that increasing aggregate content or size isn't always beneficial. An excessive amount of aggregate can lead to a very stiff mix with less cement paste. This can result in a poorly bonded matrix, creating weak spots and potentially increasing overall cracking (Li et al., 2021). Increasing aggregate content or size won't necessarily improve crack resistance if the ITZ is weak (due to poor mix design or improper curing). Larger cracks might propagate more easily through a weak ITZ.

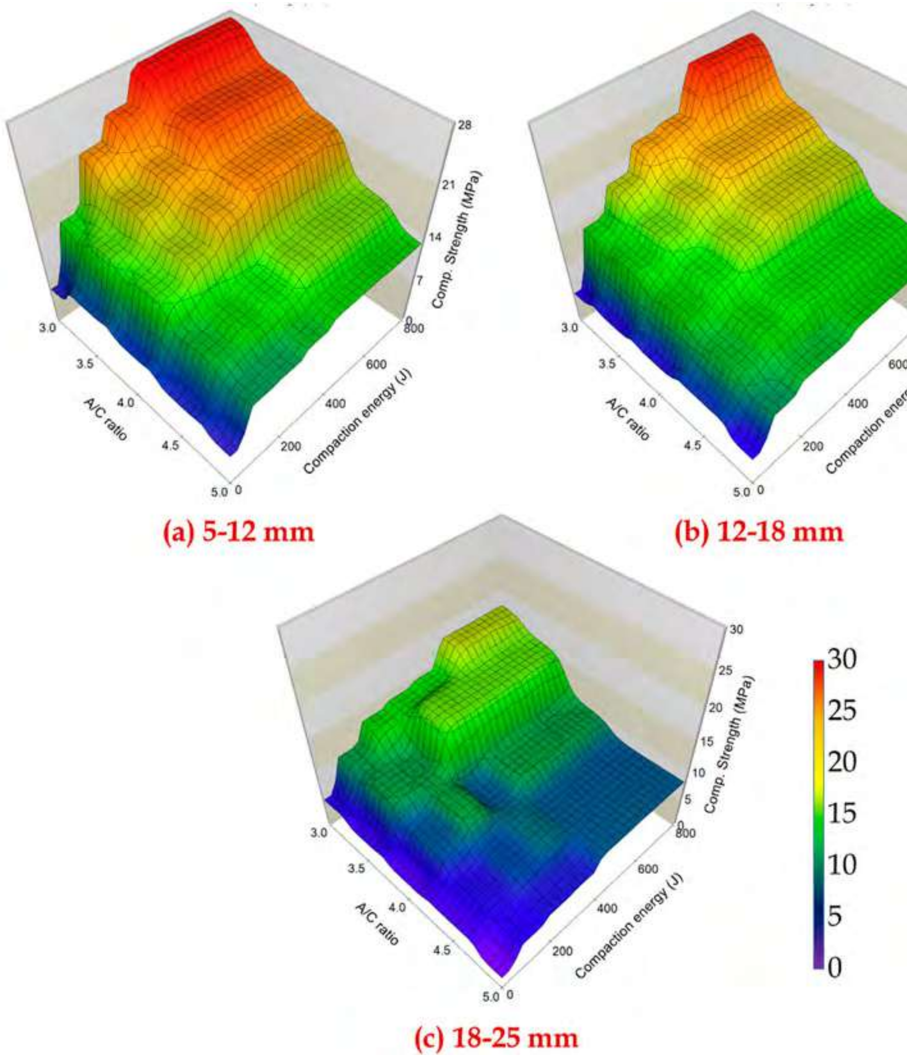

Figure 5. Variation of compressive strength of the pervious concrete with compaction energy and A/C ratio for aggregate size of (a) 5–12 mm, (b) 12–18 mm, and (c) 18–25 mm.

An intriguing observation was made when smaller aggregate sizes were employed with higher AC ratios. In such cases, a slightly enhanced strength was evident. This phenomenon was attributed to the increased contact area between the cement paste and smaller aggregates, promoting superior interfacial adhesion and improved material properties. As the AC ratio continued to increase, the disparities in strength between pervious concrete compositions featuring smaller and larger aggregates diminished. The decrease in the difference in strength between concrete with smaller and larger aggregates was because the smaller aggregates needed more cement paste to cover their increased surface area (Jebli et al., 2018; Yap et al., 2018). However, this was challenging to achieve with higher AC ratios, which made the concrete weaker. While maintaining a consistent AC ratio, the f_{cp} exhibited an increment with increasing compaction energy. Several studies support this observation as Sahdeo et al. (2021), Debnath et al. and Ferić et al. (2023) demonstrate a positive correlation between compaction energy and compressive strength in pervious concrete mixes. The enhancement in aggregate packing density brought about by increasing the number of blows compaction strengthened the mixture's capacity to endure elevated compressive forces. With low AC ratios and lesser compaction, it

Table 5. Summary of performance indicators for response surface models.

		Test				Valid			
		R^2	RMSE	MAE	a20	R^2	RMSE	MAE	a20
v_p	Linear	0.68	3.79	2.96	0.82	0.77	3.88	3.17	0.70
	SF1	0.69	3.73	2.95	0.79	0.75	4.01	3.34	0.71
	Quadratic	0.89	2.23	1.73	0.95	0.85	3.04	2.46	0.84
	Cubic	0.94	1.63	1.28	0.98	0.02	777.88	601.09	0.01
f_{cp}	Linear	0.75	3.02	2.36	0.67	0.74	2.54	1.95	0.69
	SF1	0.81	2.64	2.11	0.73	0.78	2.08	1.65	0.76
	Quadratic	0.90	1.91	1.40	0.81	0.72	2.57	2.05	0.61
	Cubic	0.93	1.58	1.11	0.88	0.05	868.74	661.55	0.01

forms too much cement paste, which creates empty spaces and makes it difficult for the aggregates to fit together correctly. However, weak compaction cannot remove these empty spaces (Ferić et al., 2023). However, as with the increment of the compaction energy, the aggregates moved around in the cement paste and came together more tightly, which led to significant differences in strength for concrete with low AC ratios at different compaction levels.

This experiment investigated the influence of aggregate size, water-cement ratio, and compaction energy on compressive strength. The highest strength, reaching 28.4 MPa, occurred with a combination of 8.5 mm aggregate, water-cement ratio of 3, and strong compaction (561 J or 75 blows). Conversely, the lowest strength of 2.3 MPa resulted from larger aggregates (21.5 mm), higher water-cement ratio (5), and no compaction applied.

5. Statistical analysis

Correlation between input (independent) variables and output (dependent) variables within the context of pervious concrete mix design from present experimental results was investigated and illustrated in Figure 6. Upon conducting statistical analysis, it was observed that the Pearson Correlation revealed a relatively weak correlation among the input parameters (S, AC, and CE), indicating a high degree of independence between them. Conversely, there was a more substantial correlation between the output variables, specifically between v_p and f_{cp}. As anticipated, the independent parameters showed lesser dependence, while the dependent variables exhibited dependencies on both independent and dependent variables. Furthermore, compaction energy displayed a moderate correlation with both v_p and f_{cp}, whereas the AC Ratio demonstrated a relatively weaker correlation with these variables.

6. Response surface models

Three key elements that influence the v_p and f_{cp} are the aggregate size (*S*), the aggregate-to-cement ratio (*AC*), and the compaction energy (*CE*). Therefore, in the current study, the independent variables selected were *S*, *AC*, and *CE*, whereas the dependent variables were v_p and f_{cp}. Table 5 present the model performance indicators, which are calculated using Equations (11)–(14).

In response surface methodology, mathematical equations were developed to describe the relationship between independent variables (aggregate gate size, AC ratio and compaction energy) and a desired outcome (porosity and compressive strength). This equation is built using a specific technique called data fitting. Data fitting techniques analyse the present experimental dataset to identify a mathematical equation that best captures the observed trends. Once a candidate equation is obtained through data fitting, it’s crucial to validate its accuracy and generalizability. This validation step involves using data collected from published literature. In this research, predictions made by equations were compared with those reported in other published studies (provided in Table 4). This

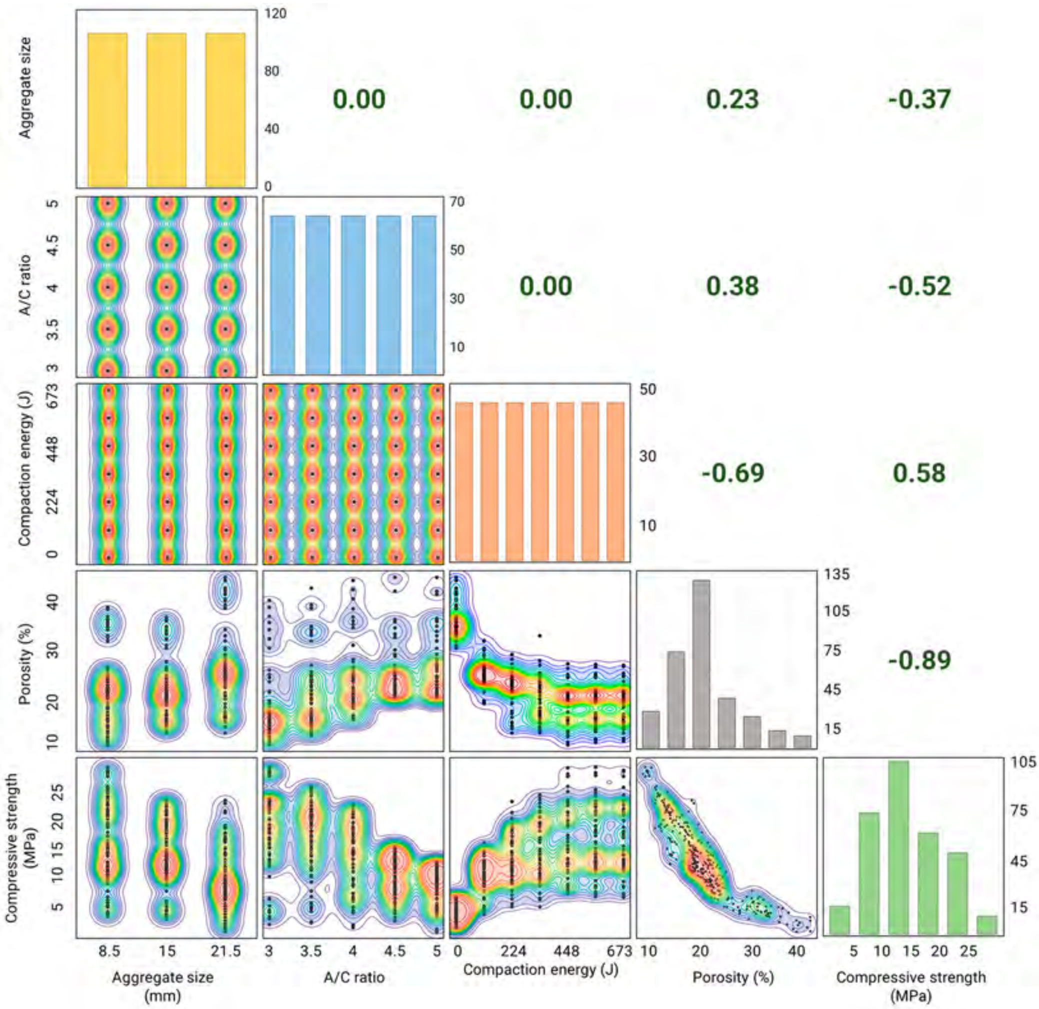

Figure 6. Data distribution and Contour maps of porosity and compressive strength of pervious concrete from experimental datasets.

two-pronged approach helps ensure that the developed equation is not just an artefact of the specific experiments conducted but reflects a more generalisable relationship.

Two distinct tests were performed to assess the accuracy of the model in predicting v_p and f_{cp}: the sequential model sum of squares and the model summary statistics. The results are presented in Tables 6–8. The model's fitness was assessed using various metrics, including R^2, RMSE, MAE, and the a20 index. A model is considered better when the values of R^2 and the a20 index approach unity. In a similar vein, a model is considered to be better when it exhibits lower values of root mean squared error (RMSE) and mean absolute error (MAE) (Sathiparan & Jeyananthan, 2023).

6.1. Linear model

The linear equation used to forecast the empirical relationship between the response variable and the independent variables in the experimental data was denoted as Equation (15) and Equation (16) for v_p and f_{cp}, respectively. The examination of the relationship revealed a positive correlation between

Table 6. ANOVA analysis for the linear model.

	Source	Sum of Squares	df	Mean Square	F-value	*p*-value
v_p	Model	9667.58	3	3222.53	221.15	< 0.0001
	(S)	763.09	1	763.09	52.37	<0.0001
	(AC)	2098.84	1	2098.84	144.03	<0.0001
	(CE)	6805.65	1	6805.65	467.04	<0.0001
	Residual	4531.85	311	14.57		
	Lack of Fit	4215.43	101	41.74	27.7	<0.0001
	Pure Error	316.42	210	1.51		
	Cor Total	14199.42	314			
f_{cp}	Model	8508.94	3	2836.31	307.47	<0.0001
	(S)	1562.49	1	1562.49	169.38	<0.0001
	(AC)	3123.21	1	3123.21	338.57	<0.0001
	(CE)	3823.24	1	3823.24	414.46	<0.0001
	Residual	2868.87	311	9.22		
	Lack of Fit	2440.28	101	24.16	11.84	<0.0001
	Pure Error	428.59	210	2.04		
	Cor Total	11377.81	314			

Table 7. ANOVA analysis for 2F1 model.

	Source	Sum of Squares	df	Mean Square	F-value	*p*-value
v_p	Model	9816	6	1636	114.95	< 0.0001
	(S)	2.07	1	2.07	0.1457	0.7029
	(AC)	15.46	1	15.46	1.09	0.2981
	(CE)	522.72	1	522.72	36.73	<0.0001
	(S)(AC)	35.68	1	35.68	2.51	0.1144
	(AC)(CE)	110.22	1	110.22	7.74	0.0057
	(S)(CE)	2.52	1	2.52	0.1772	0.6741
	Residual	4383.42	308	14.23		
	Lack of Fit	4067	98	41.5	27.54	<0.0001
	Pure Error	316.42	210	1.51		
	Cor Total	14199.42	314			
f_{cp}	Model	9189.16	6	1531.53	215.53	<0.0001
	(S)	71.07	1	71.07	10	0.0017
	(AC)	126.89	1	126.89	17.86	<0.0001
	(CE)	1213.06	1	1213.06	170.71	<0.0001
	(S)(AC)	27.92	1	27.92	3.93	0.0484
	(AC)(CE)	485.65	1	485.65	68.34	<0.0001
	(S)(CE)	166.66	1	166.66	23.45	<0.0001
	Residual	2188.65	308	7.11		
	Lack of Fit	1760.06	98	17.96	8.8	<0.0001
	Pure Error	428.59	210	2.04		
	Cor Total	11377.81	314			

the variables *S* and *AC* with v_p, while a negative correlation was observed with f_{cp}.

$$v_p = 10.81 + 0.29(S) + 3.65(AC) - 0.02(CE) \tag{15}$$

$$f_{cp} == 32.74 - 0.42(S) - 4.45(AC) + 1.55E - 02(CE) \tag{16}$$

The findings of the analysis of variance (ANOVA) for the linear equation are displayed in Table 6. The analysis of variance (ANOVA) indicates that the equation adequately represents the connection between the response variable and the significant variables. The coefficient term's significance increases when the value of F increases and the value of p decreases (Amini et al., 2008; Kalavathy et al., 2009). The *p*-value was less than 0.05, suggesting that the model has a statistical significance of more than 95% confidence (Di Leo & Sardanelli, 2020).

For the v_p and f_{cp}, the ANOVA results indicated that the F-value for the model was 221.15 and 307.47, respectively, implying that most of the variation in the response could be explained by the

Table 8. ANOVA analysis for the quadratic model.

	Source	Sum of Squares	df	Mean Square	F-value	*p*-value
v_p	Model	12631.56	9	1403.51	273.03	< 0.0001
	(S)	110.62	1	110.62	21.52	< 0.0001
	(AC)	81.96	1	81.96	15.94	<0.0001
	(CE)	1969.42	1	1969.42	383.12	<0.0001
	$(S)^2$	144.85	1	144.85	28.18	<0.0001
	$(AC)^2$	69.77	1	69.77	13.57	0.0003
	$(CE)^2$	2600.94	1	2600.94	505.97	<0.0001
	(S)(AC)	35.68	1	35.68	6.94	0.0089
	(AC)(CE)	110.22	1	110.22	21.44	<0.0001
	(S)(CE)	2.52	1	2.52	0.4907	0.4842
	Residual	1567.86	305	5.14		
	Lack of Fit	1251.44	95	13.17	8.74	<0.0001
	Pure Error	316.42	210	1.51		
	Cor Total	14199.42	314			
f_{cp}	Model	10228.46	9	1136.5	301.59	<0.0001
	(S)	0.8268	1	0.8268	0.2194	0.6398
	(AC)	0.074	1	0.074	0.0196	0.8886
	(CE)	2069.55	1	2069.55	549.19	<0.0001
	$(S)^2$	54.78	1	54.78	14.54	0.0002
	$(AC)^2$	9.44	1	9.44	2.51	0.1145
	$(CE)^2$	975.07	1	975.07	258.75	<0.0001
	(S)(AC)	27.92	1	27.92	7.41	0.0069
	(AC)(CE)	485.65	1	485.65	128.87	<0.0001
	(S)(CE)	166.66	1	166.66	44.23	<0.0001
	Residual	1149.35	305	3.77		
	Lack of Fit	720.76	95	7.59	3.72	<0.0001
	Pure Error	428.59	210	2.04		
	Cor Total	11377.81	314			

regression equation and that the model was significant. Furthermore, the probability $p < 0.0001$ for all input parameters also suggested that the model was significant.

The model summary statistics shown in Table 5 show that the R^2 for the v_p and f_{cp} was in moderate accuracy. As illustrated in Figure 7(a) and 7(b), the data were further examined to assess the association between the experimental and predicted v_p and f_{cp}, respectively. The scatter plot illustrates that the data points exhibit a moderate distribution with a straight line. Based on the present experimental data, the R^2 values for v_p and f_{cp} prediction are 0.68 and 0.75, respectively. Similarly, when validated with published literature data, the R^2 values for v_p and f_{cp} predictions are 0.77 and 0.74, respectively. These R^2 values indicate a moderate relationship between the experimental and predicted values of the response. Additionally, the findings suggest that 81.9% and 66.7% of the data fell within a 20% margin of error for v_p and f_{cp}, respectively. The analysis indicates that the chosen linear model was insufficient in accurately predicting the response variables based on the experimental data.

6.2. Two-factor interaction (2FI) model

The 2FI equation used to forecast the empirical relationship between the response variable and the independent variables in the experimental data was denoted as Equation (17) and Equation (18) for v_p and f_{cp}, respectively.

$$v_p = 21.6 - 0.09(S) + 1.05(AC) - 0.04(CE) + 0.09(S)(AC) + 7.51E - 05(S)(CE) + 3.73E - 03(AC)(CE) \tag{17}$$

$$f_{cp} == 23.88 - 0.53(S) - 3.01(AC) + 0.06(CE) + 0.08(S)(AC) - 6.11E - 04(S)(CE) - 0.01(AC)(CE) \tag{18}$$

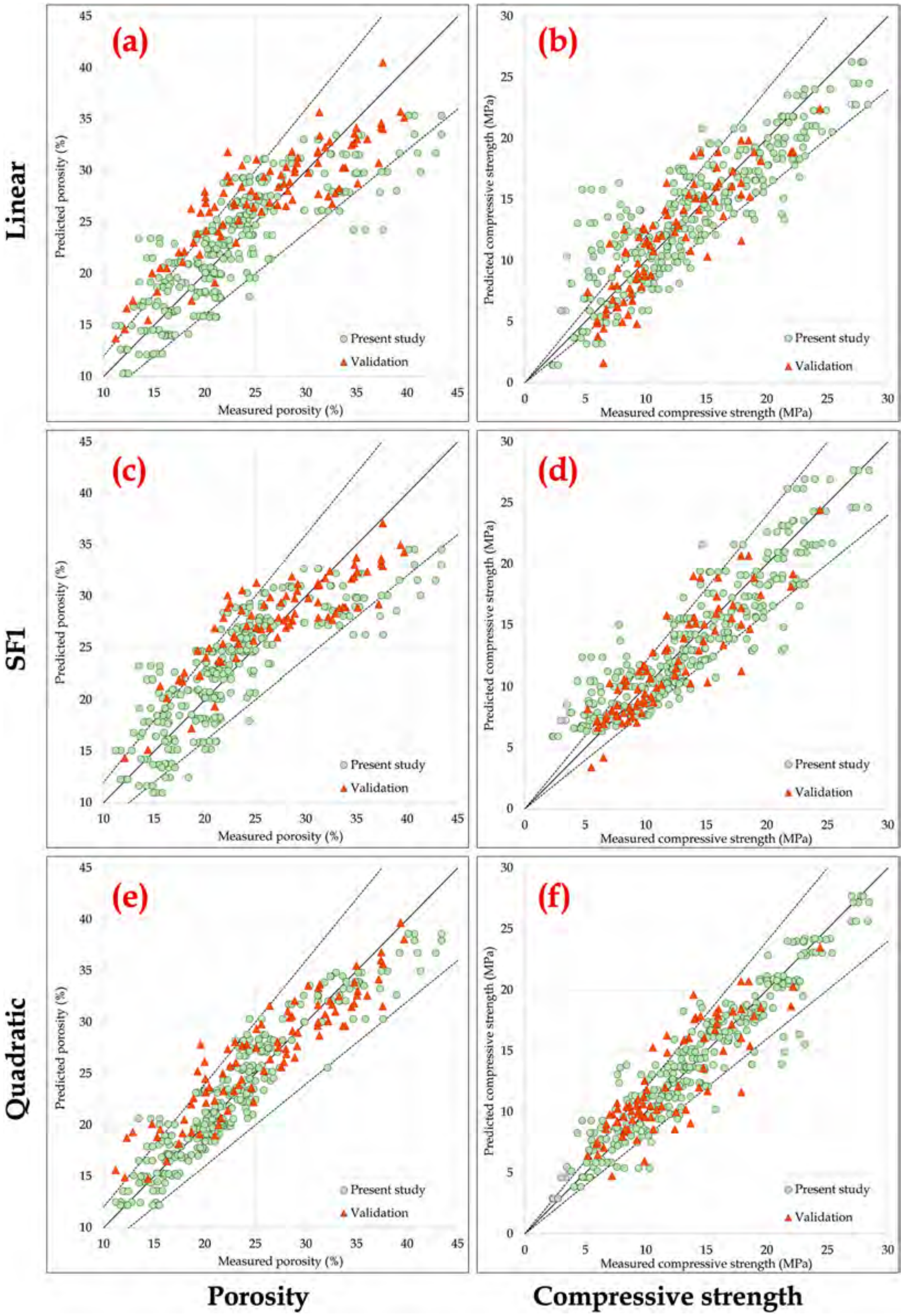

Figure 7. Predicted vs. measured porosity and compressive strength for different response surface models

The outcomes of the analysis of variance (ANOVA) for the 2FI equation are displayed in Table 7. The analysis of variance (ANOVA) indicates that the equation adequately represents the connection between the response variable and the significant variables. The ANOVA analysis revealed that the F-values for the model were 114.95 and 215.33, respectively, indicating that a substantial portion of the variability in the response variables, namely v_p and f_{cp}, could be accounted for by the 2FI equation. These findings suggest that the model holds statistical significance. In addition, the probability value of $p < 0.0001$ further reinstates the statistical significance of the model. The significance of v_p is highly influenced by the *(CE)*, whereas the interception of parameters of the *(AC)(CE)* was shown to be significant. The additional model terms indicated by their *p*-values of more than 0.05 did not exhibit substantial influence. Regarding f_{cp}, the parameters *(AC)*, *(CE)*, *(AC)(CE)*, and *(S)(CE)* exhibit a high level of significance. The remaining model components shown in Table 7 with *p*-values below 0.05 were also significant factors.

The model summary statistics shown in Table 5 show that the R^2 for the v_p and f_{cp} was in moderate accuracy. As illustrated in Figure 7(c) and 7(d), the data were further examined to assess the association between the experimental and predicted v_p and f_{cp}, respectively. The scatter plot illustrates that the data points exhibit a moderate distribution with a straight line. Based on the present experimental data, the R^2 values for v_p and f_{cp} prediction are 0.69 and 0.81, respectively. Similarly, when validated with published literature data, the R^2 values for v_p and f_{cp} prediction are 0.75 and 078, respectively. Although these values are slightly better than the linear model, these R^2 values indicate a moderate relationship between the experimental and predicted response values. Additionally, the findings suggest that 72.7% of the data fell within a 20% margin of error for v_p and f_{cp}, respectively. The analysis shows that the chosen 2FI model was insufficient in accurately predicting the response variables based on the experimental data.

6.3. Quadratic model

The quadratic equation used to forecast the empirical relationship between the response variable and the independent variables in the experimental data was denoted as Equation (19) and Equation (20) for v_p and f_{cp}, respectively.

$$\begin{aligned} v_p = {} & 15.01 - 1.11(S) + 10.05(AC) - 8.11E-02(CE) \\ & + 3.40E-02(S)^2 - 1.12(AC)^2 + 6.56E-05(CE)^2 \\ & + 8.97E-02(S)(AC) + 7.51E-05(S)(CE) \\ & + 3.73E-03(AC)(CE) \end{aligned} \tag{19}$$

$$\begin{aligned} f_{cp} = {} & 10.81 + 0.10(S) + 0.30(AC) + 8.31E-02(CE) \\ & - 2.09E-02(S)^2 - 0.41(AC)^2 - 4.03E-05(CE)^2 \\ & + 7.93E-02(S)(AC) - 6.11E-04(S)(CE) \\ & - 7.83E-03(AC)(CE) \end{aligned} \tag{20}$$

The outcomes of the analysis of variance (ANOVA) for the quadratic equation are displayed in Table 8 The analysis of variance (ANOVA) indicates that the equation adequately represents the connection between the response variable and the significant variables. The ANOVA analysis revealed that the F-values for the v_p and f_{cp}, prediction model were 273.95 and 301.59, respectively, indicating that a substantial portion of the variability in the response variables, namely v_p and f_{cp}, could be accounted for by the quadratic equation. These findings suggest that the model holds statistical significance. In addition, the probability value of $p < 0.0001$ further reinstates the statistical significance of the model. The significance of v_p is highly influenced by all the parameters except $(S)^2$ and *(S)(CE)* whereas these two parameters are also shown to be significant. Regarding f_{cp}, the parameters *(CE)*, $(CE)^2$, *(AC)(CE)*, and

$(S)(CE)$ exhibit a high level of significance. Meanwhile, the parameters of $(S)^2$ and $(S)(AC)$ were shown to be significant. The additional model terms indicated by their p-values more than 0.05 in Table 8 did not exhibit significant influence.

The model summary statistics shown in Table 5 show that the R^2 for the v_p and f_{cp} was good accuracy. The association between the experimental and predicted v_p and f_{cp} is illustrated in Figure 7(e) and 7(f), respectively. The scatter plot illustrates that the data points exhibit a good distribution with a straight line. Based on the present experimental data, the R^2 values for v_p and f_{cp} prediction are 0.89 and 0.90, respectively. Similarly, when validated with published literature data, the R^2 values for v_p and f_{cp} prediction are 0.85 and 0.72, respectively. These values are better than both linear and 2F1 models. Additionally, the findings indicate that only 95.2% and 80.6% of the data fell within a 20% margin of error for v_p and f_{cp}, respectively. It is even reduced to 83.7% and 61.2% of the data, which fell within a 20% margin of error for v_p and f_{cp}, respectively, for validated from published literature data. The analysis indicates that the chosen quadratic model was sufficient in accurately predicting the response variables based on the experimental data.

6.4. Cubic model

From the Cubic model summary statistics shown in Table 5, it can be seen that the R^2 for the v_p and f_{cp} was in excellent accuracy for the present experimental data. Despite achieving an excellent fit with the present experimental data, the model suffers when the model is applied to published literature data. Both R^2 and a20 indexes showed very low values when the model was applied to published literature data. It is due to the cubic model capturing too much random noise from the limited data, leading to a high coefficient on the training data but poor performance on unseen data. This reduces the model's generalizability. The passage mentions that the model might implicitly account for missing variables like gradation. However, not explicitly including them could lead to a lower coefficient in the cubic model. The analysis indicates that the chosen cubic model was not recommended for predicting the response variables.

7. Performance of ML models

7.1. Porosity prediction

The outcomes of the training, testing, and validation datasets for the machine learning models used in predicting the v_p of pervious concrete are displayed in Table 9. Each performance indicator is calculated using Equations (11) to (14) from the predicted values for the training, test, and validation datasets. Figure 8(a) to (d) depict the graphical depiction of the outcomes obtained from the training, testing, and validation datasets for the LR, ANN, RFR, and BDT models, respectively. Except for the linear regression model, it is evident that the discrepancy between the predicted values of various models and the actual values typically falls within a 20% margin for the prediction models, indicating a high level of prediction accuracy. This suggests that regression prediction models are effective in serving experimental purposes.

Table 9. Summary of performance indicators for prediction of porosity based on machine learning models.

	Train				Test				Validation			
	R^2	RMSE	MAE	a20	R^2	RMSE	MAE	a20	R^2	RMSE	MAE	a20
LR	0.66	3.84	3.00	82.86	0.72	3.71	2.99	73.33	0.78	3.60	2.96	81.52
ANN	0.94	1.57	1.16	99.05	0.94	1.86	1.24	97.14	0.85	3.05	2.43	89.13
RFR	0.96	1.28	0.94	99.52	0.97	1.30	0.95	99.05	0.80	3.33	2.65	88.04
BDT	0.96	1.23	0.94	100.00	0.96	1.38	0.99	99.05	0.79	3.39	2.74	86.96

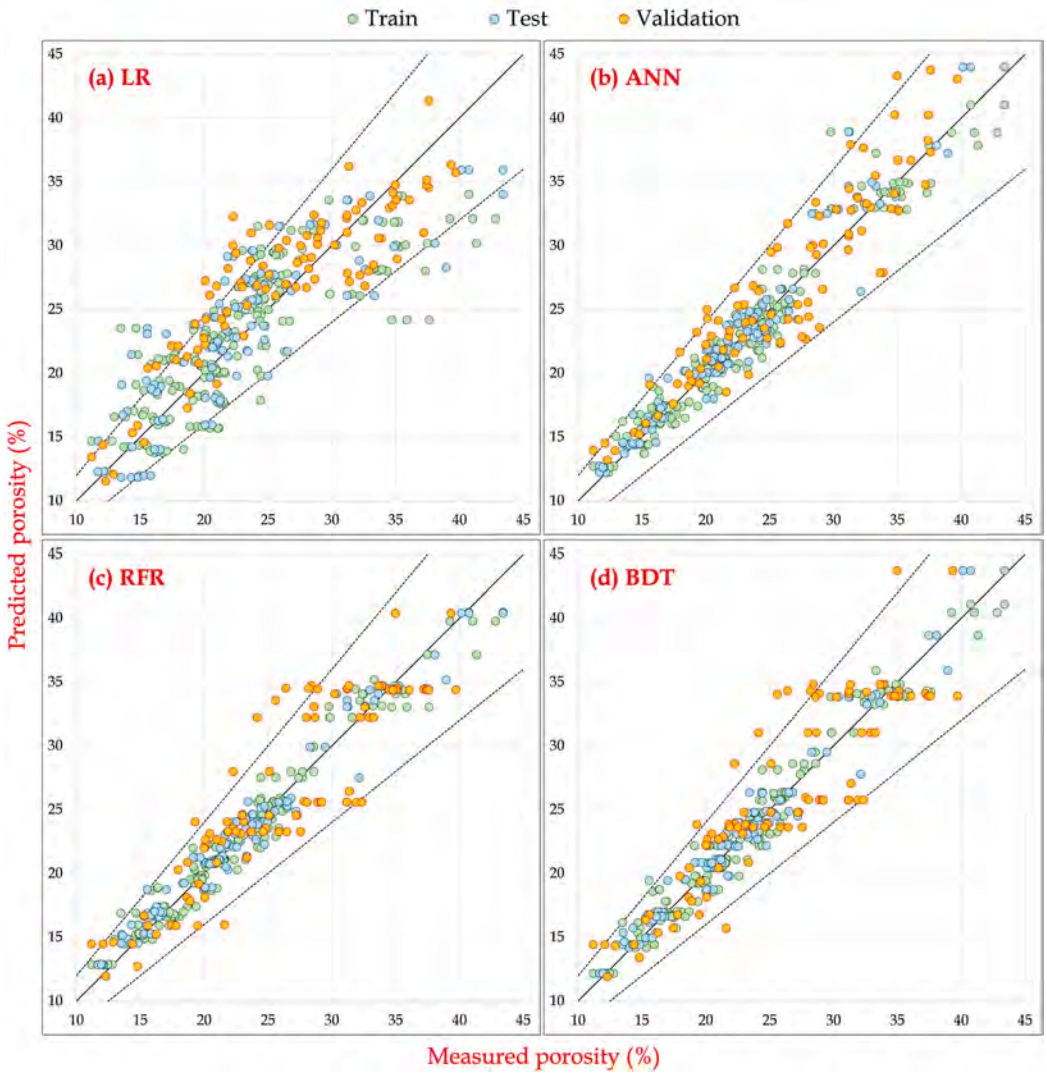

Figure 8. Predicted vs. measured porosity for different machine learning models.

The initial model examined in this analysis is linear regression. The underlying assumption of this model is that the predictors are independent of each other and that the residuals exhibit homoscedasticity and follow a normal distribution (Starbuck, 2023). The utilisation of linear regression's performance serves as a fundamental reference point for evaluating the efficacy of other models and ascertaining their suitability for the given dataset. Additionally, it can serve as a means of verification alongside a mathematical model, wherein both employ the same algorithm (Maulud & Abdulazeez, 2020). The model exhibits an R^2 value of 0.66, indicating a moderate level of explained variance. The RMSE is 3.84%, representing the average magnitude of the residuals. The MAE is determined to be 3.00%, indicating the average absolute difference between predicted and observed values. Lastly, the a20-index was measured at 0.83, which implies significance in the model's accuracy in predicting the value within the 20% error range of the observed values. Likewise, values were recorded for both test and validation datasets. It is worth mentioning that the R^2 value is relatively low, suggesting that a linear model can only explain a moderate amount of variability in the data. There exist two potential factors contributing to the suboptimal performance of this model. One factor contributing to this phenomenon is the presence of unmeasured predictor factors within the dataset. As a result, the model lacks essential information and exhibits suboptimal performance. Another potential explanation is that the data does not adhere to the linear assumption of the model, which posits that the predictors have a linear relationship with the output variable (Daoud, 2017; James et al., 2013). The aforementioned alternatives are subjected to additional evaluation through scatter plots, as shown in Figure 8(a). In this scenario, many outlier points exert a more substantial influence on the regression analysis. The findings of this study demonstrate that the linear model exhibits suboptimal predictive performance, suggesting the presence of unmeasured variables necessary for accurate v_p prediction. Additionally, more advanced modelling techniques may be required to improve predictive accuracy. However, it is justifiable to explore the utilisation of other models to ascertain whether enhanced performance may be attained.

The results presented in Figure 8(b) indicate that the ANN exhibits superior performance in predicting v_p, as observed throughout the study's training and testing datasets. Compared to linear regression, the utilisation of ANN for training, testing, and validating data reduces the RMSE by 59.1%, 49.9%, and 15.3%, respectively. Based on the results, it can be inferred that the ANN model exhibits significantly superior performance to the LR model. Moreover, it is worth noting that over 97% of the data falls within a 20% error margin for both the train and test datasets. For the validation dataset, approximately 89% of the data falls within a 20% margin of error.

The RFR model mitigates the inherent instability of basic regression trees by employing an ensemble of trees that leverage bootstrap aggregation and random variable selection (Genuer et al., 2017). The performance indicators suggest the random forest method effectively predicts v_p using the given dataset. The model exhibits the second lowest RMSE and MAE, with the highest R^2 and a20 value (1.28%, 0.94%, 0.96, and 1.00, respectively). The observed outcome could be attributed to tree-based approaches' capacity to acquire inconsistent variable importance within the dataset. Put differently, it is worth noting that each tree when trained on a specific subset of the data, has the potential to acquire a distinct set of weights that determine the importance of variables (Papadopoulos et al., 2018). The random forest algorithm, when applied collectively, exhibits improved predictive capabilities for the target variable (Schonlau & Zou, 2020). However, upon evaluating the data obtained from published literature, it was seen that the model's performance was comparatively lower than that of the ANN model. As a result, the model exhibits a reduction in the RMSE for training, testing, and validating by 66.7%, 65.0%, and 7.54%, respectively, compared to the LR model. Additionally, it demonstrates an enhancement in the RMSE for the ANN training and testing phases, with improvements of 18.5% and 30.1%, respectively. Nevertheless, the mean validating RMSE exhibits a decrease of 9.2% in comparison to the validation RMSE of the ANN model. This suggests the potential presence of overfitting in the random forest model.

Based on the performance indicators, the BDT approach is considered the most effective for prediction in the training dataset. However, BDT emerges as the second and third most effective method when it comes to testing and validating data collection. As a result, the model exhibits a reduction in the RMSE for training, testing, and validating by 67.9%, 62.4%, and 5.7%, respectively, compared to the linear model. However, the RMSE was shown to be lower in the validation process compared to both the ANN and RFR models. Despite the use of a nested cross-validation procedure, it is evident that the BDT model exhibits a modest degree of overfitting, as indicated by the more excellent value of the RMSE seen during the validation phase in comparison to the RMSE values obtained during the training and testing phases.

Overall, while the RFR and BDT models outperformed the ANN model in both the training and testing phases, the ANN model demonstrated superior validation performance using published literature data.

The prediction error for each machine-learning model studied is shown in Figure 9. This error is determined by calculating the difference between the predicted v_p and the actual v_p. The RFR, BDT, and ANN models have comparable error distributions in predicting v_p with 14.9%, 15.3%, and 15.2%, respectively. The linear regression model demonstrates a broader spread of errors in the prediction of

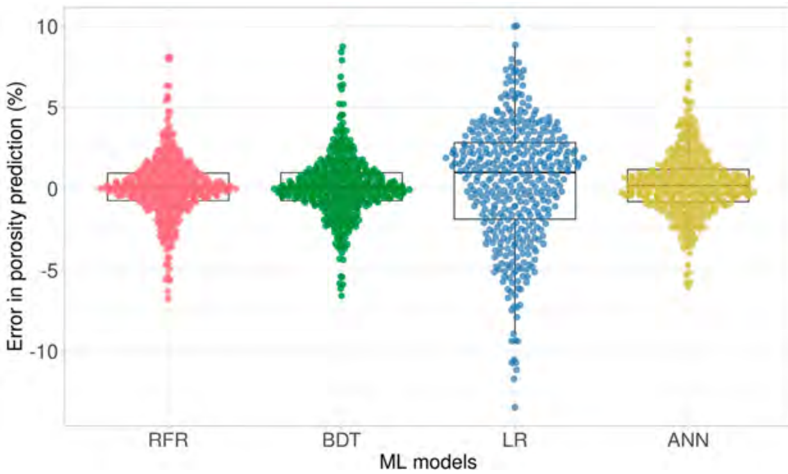

Figure 9. Error distribution in predicting porosity for different machine learning models.

v_p, with an extent of 23.5%. The RFR and BDT models exhibit symmetrical errors, with average errors of 0.10% and 0.13%, respectively. In contrast, the LR and ANN models had average errors of 0.36% and 0.31%, respectively. In addition, it is worth noting that, except for the LR model, all other models have positive skewness. The ANN model has a skewness value of 0.60, the highest among the models considered. Conversely, the LR model shows a skewness of −0.49. The use of many statistical and graphical approaches can significantly enhance the evaluation of prediction models. This suggests that using a diverse range of statistical indicators and graphical representations to evaluate the effectiveness of prediction models might provide a more comprehensive study.

Figure 10 illustrates the Taylor diagram, which serves as a visual tool for assessing the predictive capabilities of machine learning models. The Taylor diagram, a statistical instrument, offers a visual framework for evaluating and comparing several models. The provided visual representation depicts the level of concordance between each model and the reference data, quantified by correlation, standard deviation, and RMSE measurements.

The diagram can depict the relative competency of each model to a reference point (Taylor, 2001). The Taylor diagram typically includes a reference point representing the 'perfect' model. This point would correlate 1, an RMSE of 0, and a standard deviation equal to the standard deviation of the measured data. The closer a model's point on the diagram is to the reference point, the better its overall performance in replicating the observed patterns and variability. The degree of closeness between the pentagram and the reference location (measured point in the Taylor diagram) directly correlates with the model's accuracy in predicting v_p. Among all machine learning models, the BDT and RFR models have the best accuracy when applied to both the training and testing datasets. Conversely, ANN models display the highest level of accuracy for the validation dataset. Based on the above criteria, the ML models may be arranged in descending order of performance: ANN exhibits the best performance, followed by BDT, RFR, and LR with the lowest performance. The findings indicate a robust association with the performance indicator values that were previously stated.

7.2. Compressive strength prediction

The outcomes of the training, testing, and validation datasets for the machine learning models used in predicting the f_{cp} are displayed in Table 10. Each performance indicator is calculated using Equations (11) to (14) from the predicted values for the training, test, and validation datasets. Figure 11(a) to (d) depict the graphical depiction of the outcomes obtained from the training, testing, and validation datasets for the LR, ANN, RFR, and BDT models, respectively.

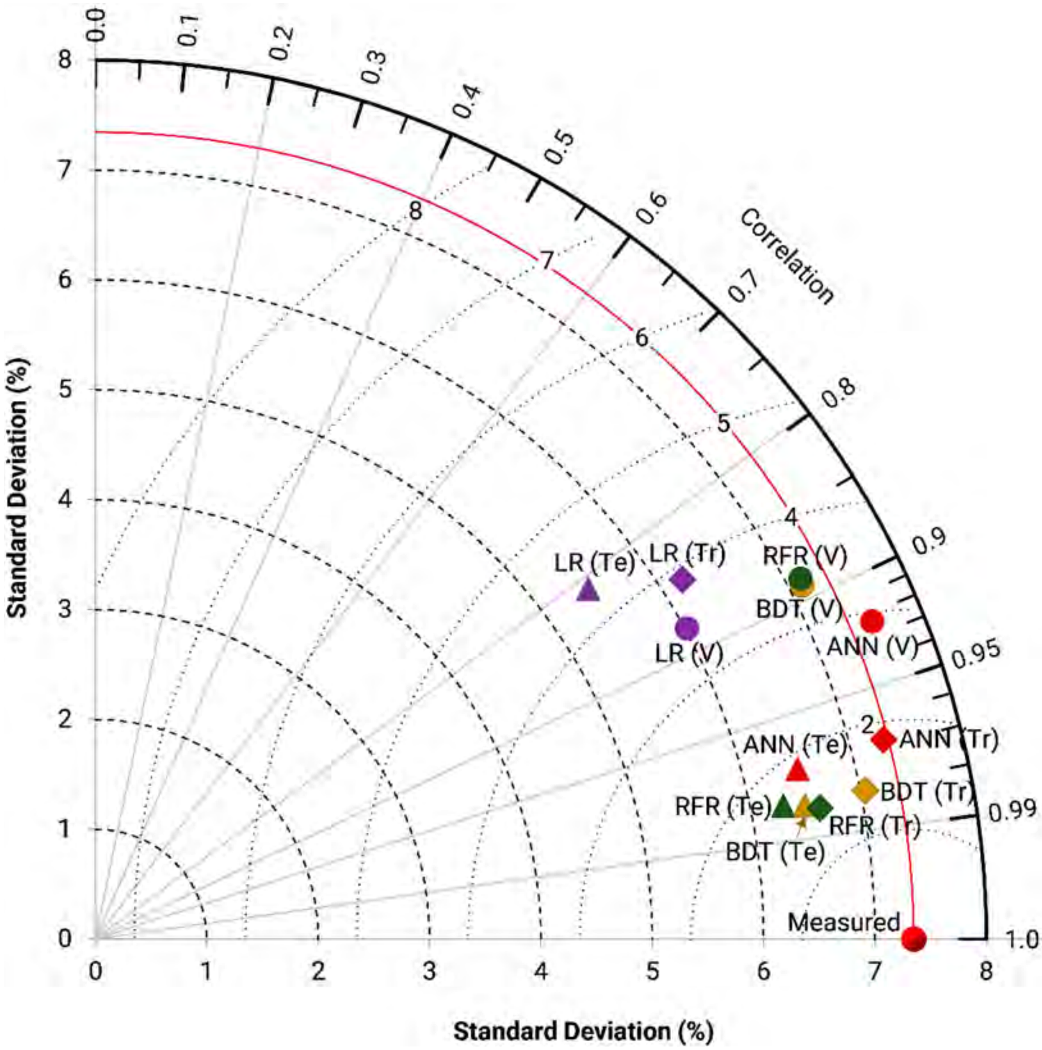

Figure 10. Taylor diagram for the prediction of porosity.

Table 10. Summary of performance indicators for prediction of compressive strength based on machine learning models.

	Train				Test				Validation			
	R^2	RMSE	MAE	a20	R^2	RMSE	MAE	a20	R^2	RMSE	MAE	a20
LR	0.76	3.05	2.39	65.71	0.73	2.96	2.30	68.57	0.67	3.31	2.24	66.25
ANN	0.97	1.12	0.78	93.33	0.93	1.56	1.05	88.57	0.84	2.05	1.63	63.75
RFR	0.95	1.44	1.01	87.62	0.95	1.24	0.86	91.43	0.64	3.08	2.67	40.00
BDT	0.93	1.64	1.09	83.33	0.94	1.28	0.90	91.43	0.71	2.67	2.13	51.25

The outputs of the LR model for predicting the f_{cp} are shown in Figure 11(a). The LR model demonstrated moderate accuracy in estimating f_{cp}; however, some divergence was seen between the actual and predicted values. The RMSE values obtained for the training dataset (3.05 MPa), test dataset (2.96 MPa), and validation dataset (3.31 MPa) indicate that the LR approach used for estimating f_{cp} exhibits worse performance when compared to other machine learning models. Finally, the a20-index was calculated to be 0.66 for the evenly distributed training dataset. This indicates that the model's

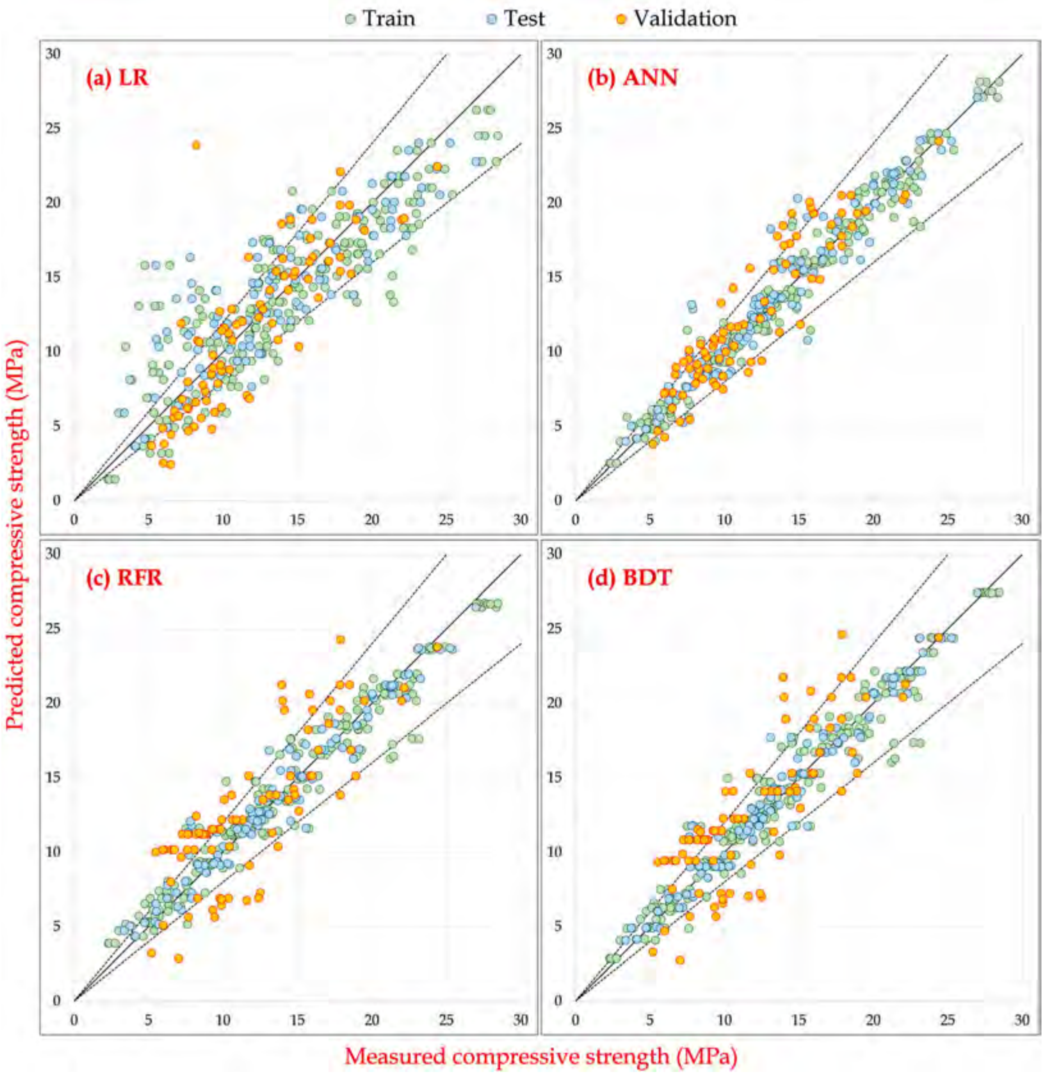

Figure 11. Predicted vs. measured compressive strength for different machine learning models.

inability to predict values within a 20% error range accurately was only 66% of the experimental values. According to the data shown in Figure 12, the error values exhibited a range of −14.02 to 15.69 MPa, with an average value of 0.21 MPa. Furthermore, an assessment was conducted on the percentage variance of errors, revealing that 90% of the error measurements were within the range of −4.33 to 4.74 MPa. The examination of errors revealed that the linear regression model did not accurately predict f_{cp}.

The results of the ANN model in predicting the f_{cp} are shown in Figure 11(b). When comparing the LR model, it was seen that the ANN technique yielded more accurate results and exhibited a more minor discrepancy between the actual and predicted values. Compared to other models, the ANN model has a higher level of precision, as seen by its RMSE scores of 1.12, 1.56, and 2.05 MPa for the training, testing, and validation datasets, respectively. While the ANN model improved the percentage of predicted values falling within ±20%, reaching 93.3% for the training dataset and 88.6% for the test dataset, it saw a decrease to 64% for the validation dataset. The ANN model produced a range of

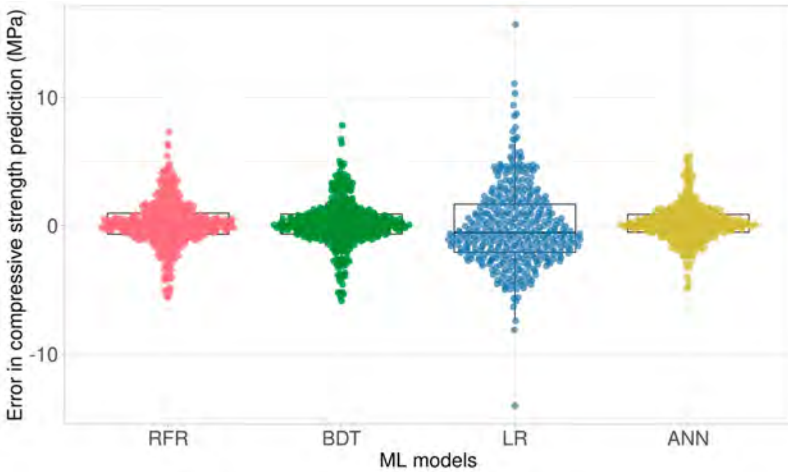

Figure 12. Error distribution in predicting compressive strength for different machine learning models.

error values from −4.79 to 5.41 MPa, with a mean error value of 0.21 MPa. The error values exhibited a distribution whereby 90% fell within the range of −1.83 to 2.91 MPa, representing the lowest values across all the models. Based on the analysis, it was determined that the ANN model had a lower standard deviation of errors, indicating a higher level of accuracy compared to the alternative models. ANNs can acquire knowledge from data without explicit rules or equations. ANNs are well-suited for addressing complicated and nonlinear situations that are challenging to represent using conventional methodologies. Consequently, the ANN model demonstrates enhanced accuracy in comparison to traditional approaches. Noisy, incomplete, or inconsistent data may be effectively managed by using many techniques, including regularisation, dropout, and data augmentation.

The results of the RFR and BDT models in predicting the f_{cp} are shown in Figure 11(c) and Figure 11(d), respectively. Both models provide very similar results. Although BDT and RFR are not the same, both work the same because they are ensemble methods that use multiple decision trees to make predictions. Both models show better accuracy than the LR model but lower accuracy than the ANN model.

The prediction error for each machine-learning model studied is shown in Figure 12. The linear regression model demonstrates a broader spread of errors in predicting v_p. In addition, it is worth noting that all the models have a positive skewness. The LR model has a skewness value of 0.65, the highest among the models considered. Conversely, the RFR model shows the lowest skewness of 0.16. This suggests that using a diverse range of statistical indicators and graphical representations to evaluate the effectiveness of prediction models might provide a more comprehensive study.

Figure 13 illustrates the Taylor diagram for predicting f_{cp} for each machine-learning model. The degree of closeness between the pentagram and the reference location (measured point in the Taylor diagram) directly correlates with the model's accuracy in predicting f_{cp}. Among all machine learning models, the ANN models have the best precision when considering all the datasets. Based on the above criteria, the ML models may be arranged in descending order of performance: ANN exhibits the best performance, followed by BDT, RFR, and LR with the lowest performance. The findings indicate a robust association with the performance indicator values that were previously stated.

7.3. Sensitivity analysis

Non-linear and complicated models like ANN often exhibit a black-box nature due to their intricate structure and behaviour (Shah et al., 2022). The utilisation of SHAP (SHapley Additive exPlanations)

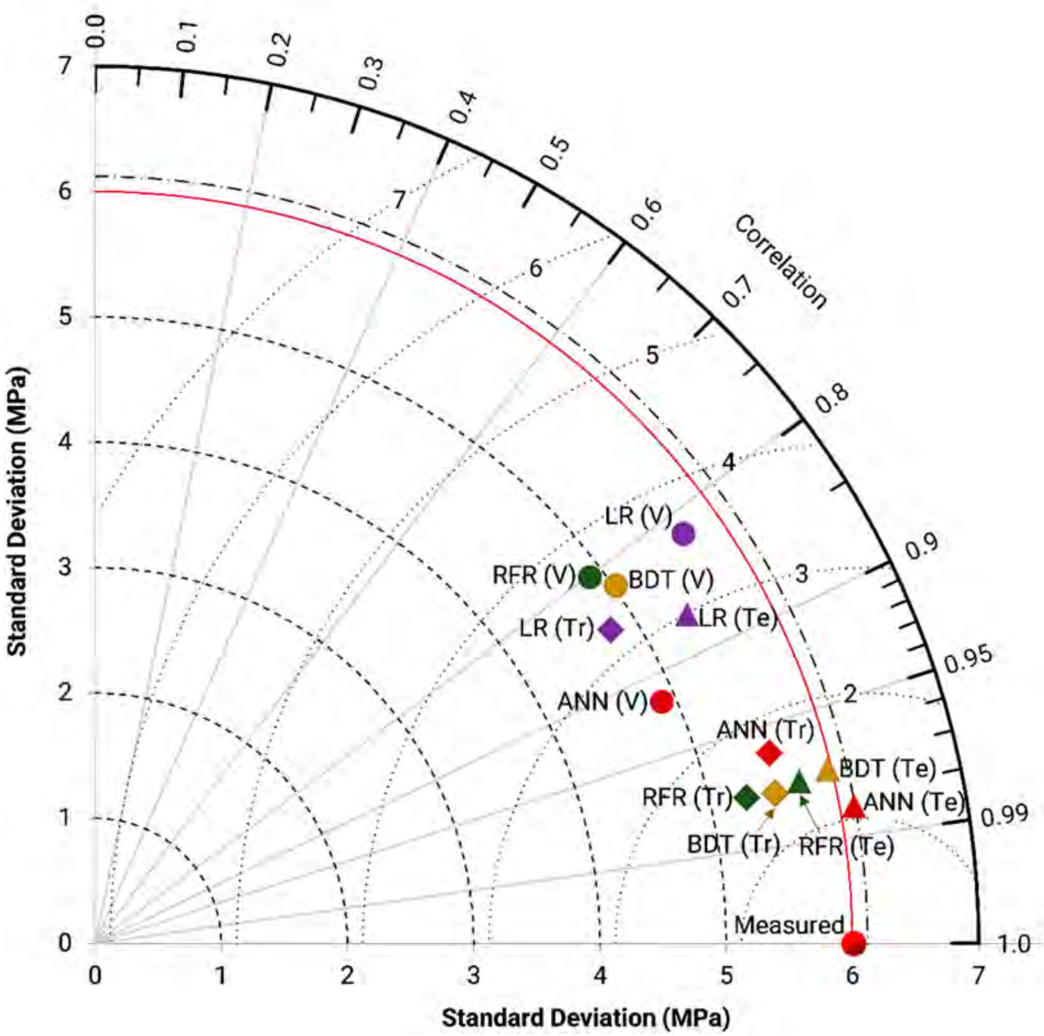

Figure 13. Taylor diagram for the prediction of compressive strength.

is very advantageous in examining intricate machine-learning models encompassing diverse aspects (Quan Tran et al., 2022; Zhang et al., 2022a). The ANN was shown to provide superior results in predicting the v_p and f_{cp}. Consequently, the outcomes obtained from the ANN were used for further analysis utilising SHAP methodology. The SHAP approach was created as a means to examine predictions generated by machine learning models via the use of Shapely values. The primary principle behind SHAP is to calculate the Shapley values for each attribute of the given sample, which requires interpretation. Each Shapley value represents the corresponding feature's impact on the prediction.

Figure 14(a) and (b) show the average SHA*P* values corresponding to the forecasts of v_p and f_{cp}, respectively, as determined by the ANN model. Based on the findings, it is evident that the compaction energy exhibits the highest SHA*P* value, indicating its significant effect on the prediction of both v_p and f_{cp}. This finding suggests that the level of compaction energy plays a pivotal role in accurately predicting the properties of pervious concrete. In contrast, it was shown that the aggregate size exhibited the lowest SHA*P* value, suggesting a lack of substantial impact on the prediction of v_p and f_{cp}.

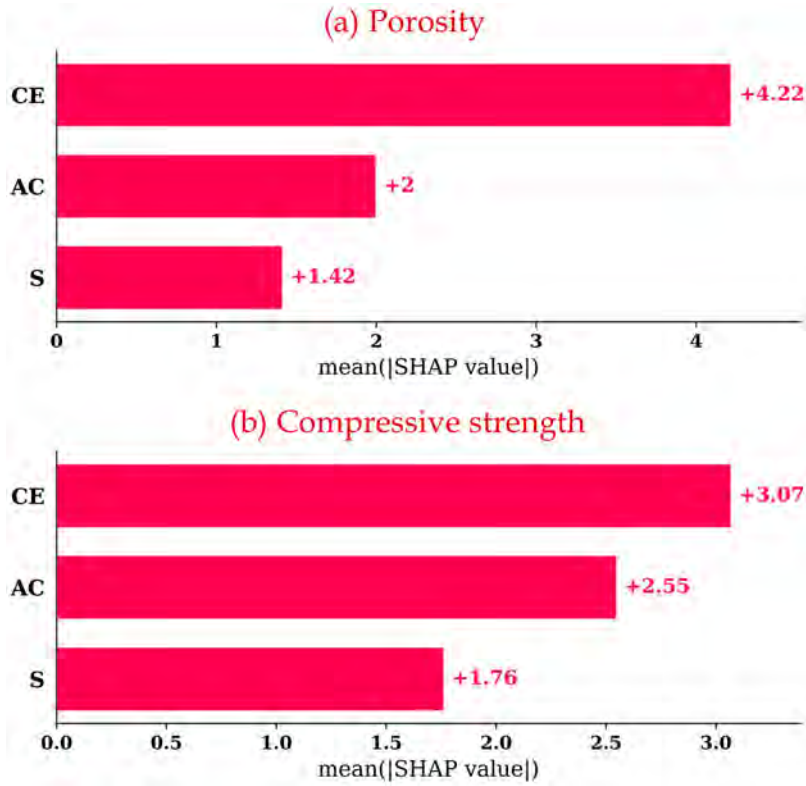

Figure 14. Mean SHA*P* values (a) porosity and (b) compressive strength.

The SHAP summary plots depicting the ANN's predictions of v_p and f_{cp} are shown in Figure 15(a) and (b), respectively. The colour gradient represents the spectrum of feature values, while the x-axis displays either the SHA*P* value or the feature's contribution toward the anticipated result. To v_p, it is found that a SHA*P* value of 10, situated on the extreme right side (positive), indicates that a decrease in compaction energy (from 112.2 J to zero compaction) might result in a 10% increase in v_p compared to the average value. Conversely, a SHA*P* value of −7 on the far-left end (negative) indicates that a significant increase in compaction energy (from 112.2 J to 673.2 J) might reduce v_p by 7% below the mean value. The correlation between v_p and compaction energy is readily comprehensible since prior research has shown that compaction energy has a substantial and direct impact on the arrangement of the mixture (Anburuvel & Subramaniam, 2022b). One notable result of the SHAP study is that aggregate size less influences predicting v_p. The aforementioned phenomenon was consistent with f_{cp}, with the compaction energy emerging as the primary input variable that substantially influences the prediction of f_{cp}. The outcomes from the SHAP analysis indicate that using the game theory approach for calculating SHA*P* values might enhance the understanding of the recommended hybrid machine learning models. Additionally, these findings demonstrate that the prediction accuracies of the models are both reasonable and satisfactory.

8. Comparison with previous works

Table 11 provides a comprehensive overview of existing research on prediction models for porosity and compressive strength of pervious concrete. These models leverage both response surface methodology and various machine learning techniques. While some studies compared multiple techniques (e.g. Le et al. (2022) and Sathiparan et al. (2023b) found XGB to be the most accurate), others focused on specific algorithms (e.g. Adewumi et al. (2016) used SVR and Sudhir Kumar et al. (2024)

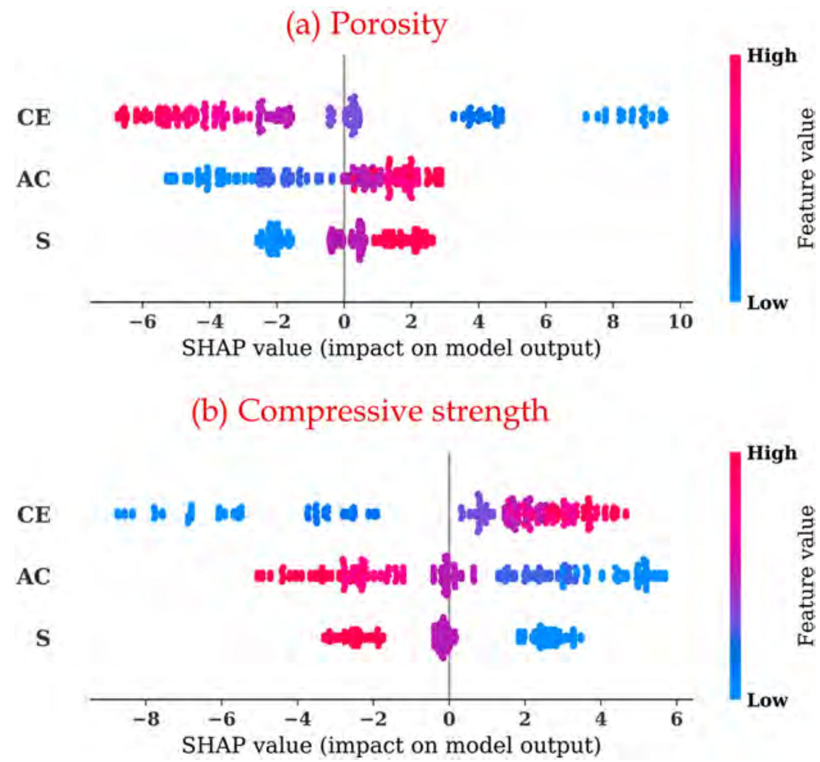

Figure 15. Summary of SHA*P* values (a) porosity and (b) compressive strength.

used ANN for prediction of compressive strength. Most of the studies accurately predict porosity and compressive strength. Common limitation identified in these studies is the use of limited datasets and machine learning techniques. Significantly, the prediction was used for conversational previous concrete (without supplementary cementing materials). This present study addresses these limitations by employing a comprehensive dataset and exploring a wider range of response surface methodologies and machine learning methods. The proposed Artificial Neural Network (ANN) model achieved superior porosity and compressive strength prediction accuracy, with an R^2 value of 0.94 and 0.97, respectively.

9. Limitations of the proposed models

The equation of the response surface model and the machine learning models used in this study agree with the experimental data. However, it is essential to note that this model only considers three specific design factors: aggregate size, A/C ratio, and compaction energy. Several factors, such as the water-to-cement ratio, aggregate grading, fine content, and the use of superplasticizers, among others, may influence the features of pervious concrete. Therefore, a model that includes these factors should be checked and corrected to reflect the model considered. The constants included in the formulated models may encompass the influences of these supplementary components. Nevertheless, developing a single equation that includes all these elements is not desirable since it may not be deemed appropriate. In contrast, the machine learning model might potentially derive advantages by including all these characteristics. Furthermore, the performance of the machine learning model might potentially be enhanced by the incorporation of additional data. Hence, it is vital to maintain a comprehensive dataset that includes mixed design parameters and characteristics of previous concrete.

Table 11. Comparison of present study with published literature on prediction models (Response surface methodology and machine learning) for porosity and compressive strength of pervious concrete.

Properties	Ref	Input parameters[b]	No of datasets	Modelling algorithms used[c]	Best model	Accuracy
Porosity	Adewumi et al. (2016)	S, Cement, W/C, Agg.	24	SVR	–	$R^2 = 0.8697$
	Le et al. (2022)	Agg.S, A/C, W/C, Sand, SF	164	ANN, RFR, SVM, XGB	XGB	$R^2 = 0.88$, RMSE $= 3.2\%$
	Almasaeid and Salman (2022)[a]	A/C, W/C, Density	369	MLR, ANN	ANN	$R^2 = 0.886$
	Present study	Agg.S, A/C, Compaction energy	315	RSM, ANN, RFR, BDT, LR	ANN	$R^2 = 0.94$ RMSE $= 1.57\%$
Compressive strength	Sonebi and Bassuoni (2013)	Agg.S, W/C, Cement, Agg.	9	RSM	Quadratic	$R^2 = 0.87$
	Adewumi et al. (2016)	Agg.S, Cement, W/C, Agg.	24	SVR	–	$R^2 = 0.9806$ RMSE $= 0.3186$ MPa
	Hari and Mini (2023)	Cement, Agg., W/C, SF	45	RSM	Quadratic	$R^2 = 0.9973$
	Sun et al. (2019)	Agg.S, A/C, W/C	90	SVR	–	$R^2 = 0.9997$ RMSE $= 0.353$ MPa
	Sudhir Kumar et al. (2024)	Agg.S, Ground Granulated Blast-furnace Slag content, Curing days	36	ANN	–	$R^2 = 0.82$
	Le et al. (2022)	Agg.S, A/C, W/C, Sand, SF	164	ANN, RFR, SVR, XGB	XGB	$R^2 = 0.98$ RMSE $= 1.44$ MPa
	Sathiparan et al. (2023b)[a]	A/C, FA/C, W/C, Agg.S, Curing period	437	LR, ANN, BDT, RF, SVR, XGB	XGB	$R^2 = 0.98$ RMSE $= 1.62$ MPa
	Sathiparan et al. (2024a)	Cement, SF, Sand, Agg., W/C, Admixture, Agg.S, Curing period	222	LR, ANN, BDT, KNN, RF, SVR, XGB	XGB	$R^2 = 0.98$ RMSE $= 0.24$ MPa
	Present study	Agg.S, A/C, Compaction energy	315	RSM, ANN, RFR, BDT, LR	ANN	$R^2 = 0.97$ RMSE $= 1.22$ MPa

[a] Performance indicator for all datasets, for others it is for test datasets.
[b] Agg.S – Aggregate size, cement – cement content, Agg – aggregate content, W/C – water to cement ratio, A/C – aggregate to cement ratio, Sand – sand content, SF – silica fume content, FA – fly ash content
[c] ANN – Artificial Neural Network, BDT – Boosted Decision Tree Regression, KNN – K-Nearest Neighbour, MLR – Multiple Linear Regression, NLR – Non-Linear Regression, LR – Linear regression, RFR – Random Forest Regression, SVR – Support Vector Regression, XGB – Extreme Gradient Boosting, RSM – Response surface methodology

10. Conclusions

The framework for forecasting the porosity and compressive strength using response surface regression and machine learning models is proposed in the present study. There were 315 experimental specimens (data points) to train and test the models and validate with data from published literature. The following conclusions may be derived from the findings:

- This experiment showed that Porosity increases with higher AC ratio and larger aggregate size while decreasing with increased compaction energy. Compressive Strength is reduced with higher AC ratios but showed a slight increase with smaller aggregate sizes at lower AC ratios. Increased compaction energy led to higher compressive strength.
- The optimal combination for minimum porosity (11.2%) was achieved with smaller aggregates (8.5 mm), lower AC ratio (3), and high compaction energy (448.8 J). Conversely, the maximum porosity (43.4%) resulted from larger aggregates (21.5 mm), higher AC ratio (4.5), and no compaction. The highest value (28.4 MPa) for compressive strength was obtained with the same combination as the minimum porosity. The lowest strength (2.3 MPa) occurred with larger aggregates, higher AC ratio, and no compaction.
- The statistical analysis demonstrates that the output variables porosity and compressive strength have a stronger correlation. This aligns with the understanding that porosity can influence the strength of pervious concrete. Furthermore, the analysis revealed a moderate positive correlation between compaction energy (CE) and both porosity and compressive strength. The aggregate-cement ratio showed a weaker correlation with both output variables than compaction energy.
- The quadratic model provided the most accurate predictions for both porosity and compressive strength in four response surface models. This model exhibited high R^2 values (> 0.85), indicating a strong correlation between the predicted and experimental values. A significant portion of the data points fall within a 20% margin of error for both porosity and compressive strength. However, it's important to note that when validated with published literature data, the accuracy of the quadratic model decreased. This suggests that the model may not be universally applicable, and further research might be required for broader generalizability.
- The results revealed that all models except LR exhibited high levels of prediction accuracy, with discrepancies between predicted and actual values typically falling within a 20% margin for the training and testing datasets. Also, machine learning models (ANN, BDT and RFR) performed much better than response surface models. ANN demonstrated the strongest overall performance among these models, achieving the highest R^2 values and lowest RMSE across all datasets. However, when evaluating data from published literature (validation set), BDT displayed the most accurate predictions.
- The SHAP analysis revealed that compaction energy significantly impacts both porosity and compressive strength predictions. This aligns with existing knowledge that compaction energy plays a crucial role in the structure of pervious concrete. Conversely, aggregate size exhibited minimal influence on the predicted values.

11. Future research direction

The present study incorporated three key factors: compaction energy, aggregate-to-cement ratio, and aggregate size. While the machine learning models showed good agreement with experimental results for porosity and compressive strength, they were limited by the data range. Aggregate size ranged from 5 to 25 mm, A/C ratio from 3 to 5, and compaction from 0 to 90 blows. Future research should explore a broader spectrum of these parameters, as real-world applications may involve values outside this range. Other factors like aggregate gradation, fine content, and water-to-cement ratio influence porosity and compressive strength (Wijekoon et al., 2023). The models might implicitly account for these missing variables, but including them explicitly could improve accuracy. Expanding the data with a wider range of mix parameters, porosity, and compressive strength values would further enhance the model's performance.

While traditional machine learning techniques like linear regression, ANN, random forest regression, and boosted decision tree regression used here have merits, they also have limitations. Despite their success in predicting porosity and compressive strength, there's room for improvement. These methods may struggle with complex, non-linear relationships in data due to their limited ability to learn complex patterns and reliance on pre-defined rules. Feature engineering, the manual process of extracting relevant information, can be time-consuming and might miss crucial details. Deep learning, with its sophisticated architecture, offers a powerful alternative. Deep learning excels at automatically extracting features and capturing non-linear patterns directly from raw data, eliminating the need for manual feature engineering. This allows for more flexible and accurate models, especially with high-dimensional or complex data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Navaratnarajah Sathiparan http://orcid.org/0000-0001-8570-0580
Pratheeba Jeyananthan http://orcid.org/0000-0003-0476-3964
Daniel Niruban Subramaniam http://orcid.org/0000-0002-1023-1617

References

Abdelhady, A., Hui, L., & Zhang, H. (2021). Comprehensive study to accurately predict the water permeability of pervious concrete using constant head method. *Construction and Building Materials*, *308*, 125046. https://doi.org/10.1016/j.conbuildmat.2021.125046

ACI-522R. (2010). *Report on pervious concrete, technical documents*. American Concrete Institute.

Adewumi, A. A., Owolabi, T. O., Alade, I. O., & Olatunji, S. O. (2016). Estimation of physical, mechanical and hydrological properties of permeable concrete using computational intelligence approach. *Applied Soft Computing*, *42*, 342–350. https://doi.org/10.1016/j.asoc.2016.02.009

Agar-Ozbek, A. S., Weerheijm, J., Schlangen, E., & van Breugel, K. (2013). Investigating porous concrete with improved strength: Testing at different scales. *Construction and Building Materials*, *41*, 480–490. https://doi.org/10.1016/j.conbuildmat.2012.12.040

Ahmad, A., Ahmad, W., Aslam, F., & Joyklad, P. (2022). Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. *Case Studies in Construction Materials*, *16*, e00840.

Ahmad, S. A., Rafiq, S. K., Hilmi, H. D. M., & Ahmed, H. U. (2023). Mathematical modeling techniques to predict the compressive strength of pervious concrete modified with waste glass powders. *Asian Journal of Civil Engineering*, *25*, 773–785.

Alimohammadi, V., Maghfouri, M., Nourmohammadi, D., Azarsa, P., Gupta, R., & Saberian, M. (2021). Stormwater runoff treatment using pervious concrete modified with various nanomaterials: A comprehensive review, sustainability.

Almasaeid, H. H., & Salman, D. G. (2022). Application of artificial neural network to predict the properties of permeable concrete. *Civil Engineering and Architecture*, *10*, 2305–2305. https://doi.org/10.13189/cea.2022.100605

Alsahaf, A., Petkov, N., Shenoy, V., & Azzopardi, G. (2022). A framework for feature selection through boosting. *Expert Systems with Applications*, *187*, 115895. https://doi.org/10.1016/j.eswa.2021.115895

AlShareedah, O., & Nassiri, S. (2021). Pervious concrete mixture optimization, physical, and mechanical properties and pavement design: A review. *Journal of Cleaner Production*, *288*, 125095. https://doi.org/10.1016/j.jclepro.2020.125095

Amini, M., Younesi, H., Bahramifar, N., Lorestani, A. A. Z., Ghorbani, F., Daneshi, A., & Sharifzadeh, M. (2008). Application of response surface methodology for optimization of lead biosorption in an aqueous solution by Aspergillus niger. *Journal of Hazardous Materials*, *154*, 694–702. https://doi.org/10.1016/j.jhazmat.2007.10.114

Anburuvel, A., & Subramaniam, D. (2022a). Investigation of the effects of compaction on compressive strength and porosity characteristics of pervious concrete. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*, 036119812210872. https://doi.org/10.1177/03611981221087236

Anburuvel, A., & Subramaniam, D. N. (2022b). Influence of aggregate gradation and compaction on compressive strength and porosity characteristics of pervious concrete. *International Journal of Pavement Engineering*, 1–14.

Anburuvel, A., & Subramaniam, D. N. (2023). Influence of aggregate gradation and compaction on compressive strength and porosity characteristics of pervious concrete. *International Journal of Pavement Engineering*, *24*, 2055022. https://doi.org/10.1080/10298436.2022.2055022

Arcolezi, K. (2022). *Influence of aggregate sizes and packing combination on the properties of pervious concrete*. São Paulo State University (UNESP).

ASTM-C1754/C1754M-12. (2012). *Standard test method for density and void content of hardened pervious concrete*. ASTM International.

BIS-269. (2015). *Ordinary Portland cement - Specification*. Bureau of Indian Standards.

BIS-383. (2016). *Coarse and fine aggregate for concrete - Specification*. Bureau of Indian Standards.

Boddy, L., & Morris, C. W. (1999). Artificial neural networks for pattern recognition. In A. H. Fielding (Ed.), *Machine learning methods for ecological applications* (pp. 37–87). Springer US.

Bright Singh, S., & Madasamy, M. (2022). Investigation of aggregate size effects on properties of basalt and carbon fibre-reinforced pervious concrete. *Road Materials and Pavement Design*, *23*, 1305–1328. https://doi.org/10.1080/14680629.2021.1886158

BS-EN-1097. (2020). *Tests for mechanical and physical properties of aggregates*. British Standards Institution (BSI).

BS-EN-933. (2017). *Tests for geometrical properties of aggregates*. British Standards Institution (BSI).

Cai, J., Shen, C., Ye, M., Huang, S., He, J., & Cui, D. (2024). Influencing Factors of Porosity and Strength of Plant-Growing Concrete, Materials.

Chandrappa, A. K., & Biligiri, K. P. (2016). Comprehensive investigation of permeability characteristics of pervious concrete: A hydrodynamic approach. *Construction and Building Materials*, *123*, 627–637. https://doi.org/10.1016/j.conbuildmat.2016.07.035

Chandrappa, A. K., & Biligiri, K. P. (2019). Effect of pore structure on fatigue of pervious concrete. *Road Materials and Pavement Design*, *20*, 1525–1547. https://doi.org/10.1080/14680629.2018.1464500

Chindaprasirt, P., Hatanaka, S., Chareerat, T., Mishima, N., & Yuasa, Y. (2008). Cement paste characteristics and porous concrete properties. *Construction and Building Materials*, *22*, 894–901. https://doi.org/10.1016/j.conbuildmat.2006.12.007

Chindaprasirt, P., Hatanaka, S., Mishima, N., Yuasa, Y., & Chareerat, T. (2009). Effects of binder strength and aggregate size on the compressive strength and void ratio of porous concrete. *International Journal of Minerals, Metallurgy and Materials*, *16*, 714–719.

Ćosić, K., Korat, L., Ducman, V., & Netinger, I. (2015). Influence of aggregate type and size on properties of pervious concrete. *Construction and Building Materials*, *78*. https://doi.org/10.1016/j.conbuildmat.2014.12.073

Dall Bello De Souza Risson, K., Sandoval, G. F. B., Cofani Pinto, F. S., Camargo, M., Campos De Moura, A., & Martins Toralles, B. (2021). Molding procedure for pervious concrete specimens by density control. *Case Studies in Construction Materials*, *15*, e00619.

Daoud, J. I. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series*, *949*, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Debnath, B., Sarkar, P. P., & Deb, P. (2023). Effects of compaction on the properties of pervious concrete made with local aggregates. *Proceedings of the Institution of Civil Engineers - Construction Materials*, *177*(3), 147–158. https://doi.org/10.1680/jcoma.23.00024

de Zarzà, I., de Curtò, J., Hernández-Orallo, E., & Calafate, C. T. (2023). Cascading and Ensemble Techniques in Deep Learning, Electronics.

Di Leo, G., & Sardanelli, F. (2020). Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. *European Radiology Experimental*, *4*, 18. https://doi.org/10.1186/s41747-020-0145-y

Durap, A. (2023). A comparative analysis of machine learning algorithms for predicting wave runup. *Anthropocene Coasts*, *6*, 17. https://doi.org/10.1007/s44218-023-00033-7

Elango, K. S., Gopi, R., Saravanakumar, R., Rajeshkumar, V., Vivek, D., & Raman, S. V. (2021). Properties of pervious concrete – A state of the art review. *Materials Today: Proceedings*, *45*, 2422–2425. https://doi.org/10.1016/j.matpr.2020.10.839

Elango, K. S., & Revathi, V. (2017). Fal-G binder pervious concrete. *Construction and Building Materials*, *140*, 91–99. https://doi.org/10.1016/j.conbuildmat.2017.02.086

Elizondo-Martinez, E.-J., Ossa-Lopez, A., & Rodriguez-Hernandez, J. (2021). Evaluation of the effect of different compaction methods on porous concrete pavements: Correlation with strength and permeability. *Journal of Materials in Civil Engineering*, *33*, 04021184. https://doi.org/10.1061/(ASCE)MT.1943-5533.0003791

Feng, D.-C., Liu, Z.-T., Wang, X.-D., Chen, Y., Chang, J.-Q., Wei, D.-F., & 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

Ferić, K., Sathish Kumar, V., Romić, A., & Gotovac, H. (2023). Effect of aggregate size and compaction on the strength and hydraulic properties of pervious concrete. *Sustainability*, *15*(2), 1146.

Forthofer, R. N., Lee, E. S., & Hernandez, M. (2007). 13 - Linear regression. In R. N. Forthofer, E. S. Lee, & M. Hernandez (Eds.), *Biostatistics (Second edition)* (pp. 349–386). Academic Press.

Fu, T. C., Yeih, W., Chang, J. J., & Huang, R. (2014). The influence of aggregate size and binder material on the properties of pervious concrete. *Advances in Materials Science and Engineering*, *2014*, 963971.

Furkan Ozel, B., Sakallı, Ş., & Şahin, Y. (2022). The effects of aggregate and fiber characteristics on the properties of pervious concrete. *Construction and Building Materials*, *356*, 129294.
Genuer, R., Poggi, J.-M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). Random forests for big data. *Big Data Research*, *9*, 28–46. https://doi.org/10.1016/j.bdr.2017.07.003
Hari, R., & Mini, K. M. (2023). Mechanical and durability properties of basalt-steel wool hybrid fibre reinforced pervious concrete – A Box Behnken approach. *Journal of Building Engineering*, *70*, 106307. https://doi.org/10.1016/j.jobe.2023.106307
Hedayat, A. S., & Pesotan, H. (1997). Designs for two-level factorial experiments with linear models containing main effects and selected two-factor interactions. *Journal of Statistical Planning and Inference*, *64*, 109–124. https://doi.org/10.1016/S0378-3758(96)00209-1
Huang, J., Duan, T., Zhang, Y., Liu, J., Zhang, J., & Lei, Y. (2020a). Predicting the permeability of pervious concrete based on the beetle antennae search algorithm and random forest model. *Advances in Civil Engineering*, *2020*, 8863181.
Huang, J., Luo, Z., & Khan, M. B. E. (2020b). Impact of aggregate type and size and mineral admixtures on the properties of pervious concrete: An experimental investigation. *Construction and Building Materials*, *265*, 120759. https://doi.org/10.1016/j.conbuildmat.2020.120759
Huang, J., Zhang, J., & Gao, Y. (2022). Evaluating the clogging behavior of pervious concrete (PC) using the machine learning techniques. *Computer Modeling in Engineering & Sciences*, *130*, 805–821.
Ibrahim, A., Mahmoud, E., Yamin, M., & Patibandla, V. C. (2014). Experimental study on Portland cement pervious concrete mechanical and hydrological properties. *Construction and Building Materials*, *50*, 524–529. https://doi.org/10.1016/j.conbuildmat.2013.09.022
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). Linear regression. In G. James, D. Witten, T. Hastie, & R. Tibshirani (Eds.), *An introduction to statistical learning: with applications in R* (pp. 59–126). Springer New York.
Jebli, M., Jamin, F., Malachanne, E., Garcia-Diaz, E., & El Youssoufi, M. S. (2018). Experimental characterization of mechanical properties of the cement-aggregate interface in concrete. *Construction and Building Materials*, *161*, 16–25. https://doi.org/10.1016/j.conbuildmat.2017.11.100
Joshaghani, A., Ramezanianpour, A. A., Ataei, O., & Golroo, A. (2015). Optimizing pervious concrete pavement mixture design by using the Taguchi method. *Construction and Building Materials*, *101*, 317–325. https://doi.org/10.1016/j.conbuildmat.2015.10.094
Joshi, T., & Dave, U. (2016). Evaluation of strength, permeability and void ratio of pervious concrete with changing w/c ratio and aggregate size. *International Journal of Civil Engineering and Technology*, *7*, 276–284.
Kalavathy, M. H., Regupathi, I., Pillai, M. G., & Miranda, L. R. (2009). Modelling, analysis and optimization of adsorption parameters for H3PO4 activated rubber wood sawdust using response surface methodology (RSM). *Colloids and Surfaces B: Biointerfaces*, *70*, 35–45. https://doi.org/10.1016/j.colsurfb.2008.12.007
Kant Sahdeo, S., Ransinchung R.N, G. D., Rahul, K. L., & Debbarma, S. (2020). Effect of mix proportion on the structural and functional properties of pervious concrete paving mixtures. *Construction and Building Materials*, *255*, 119260. https://doi.org/10.1016/j.conbuildmat.2020.119260
Kevern, J. T., Schaefer, V. R., & Wang, K. (2009). Evaluation of pervious concrete workability using gyratory compaction. *Journal of Materials in Civil Engineering*, *21*, 764–770. https://doi.org/10.1061/(ASCE)0899-1561(2009)21:12(764)
Khan, A. A., Chaudhari, O., & Chandra, R. (2024). A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. *Expert Systems with Applications*, *244*, 122778. https://doi.org/10.1016/j.eswa.2023.122778
Khankhaje, E., Kim, T., Jang, H., Kim, C.-S., Kim, J., & Rafieizonooz, M. (2024). A review of utilization of industrial waste materials as cement replacement in pervious concrete: An alternative approach to sustainable pervious concrete production. *Heliyon*, *10*, e26188. https://doi.org/10.1016/j.heliyon.2024.e26188
Khuri, A., & JA, C. (1996). *Response surfaces: Designs and analyses*. CRC Press.
Kumar, R. R., Mukhopadhya, T., Pandey, K. M., & Dey, S. (2020). Chapter 5 - Prediction capability of polynomial neural network for uncertain buckling behavior of sandwich plates. In P. Samui, D. Tien Bui, S. Chakraborty, & R. C. Deo (Eds.), *Handbook of probabilistic models* (pp. 131–140). Butterworth-Heinemann.
Le, B.-A., Vu, V.-H., Seo, S.-Y., Tran, B.-V., Nguyen-Sy, T., Le, M.-C., & Vu, T.-S. (2022). Predicting the compressive strength and the effective porosity of pervious concrete using machine learning methods. *KSCE Journal of Civil Engineering*, *26*, 4664–4679. https://doi.org/10.1007/s12205-022-1918-z
Li, L. G., Feng, J.-J., Zhu, J., Chu, S.-H., & Kwan, A. K. H. (2021). Pervious concrete: Effects of porosity on permeability and strength. *Magazine of Concrete Research*, *73*, 69–79. https://doi.org/10.1680/jmacr.19.00194
Li, W., Xiao, J., Sun, Z., Kawashima, S., & Shah, S. P. (2012). Interfacial transition zones in recycled aggregate concrete with different mixing approaches. *Construction and Building Materials*, *35*, 1045–1055. https://doi.org/10.1016/j.conbuildmat.2012.06.022
Li, H., Yang, J., Yu, X., Zhang, Y., & Zhang, L. (2023). Permeability prediction of pervious concrete based on mix proportions and pore characteristics. *Construction and Building Materials*, *395*, 132247. https://doi.org/10.1016/j.conbuildmat.2023.132247
Li, Z., Yoon, J., Zhang, R., Rajabipour, F., Iii, W., Dabo, I., & Radlińska, A. (2022). Machine learning in concrete science: Applications, challenges, and best practices. *NPJ Computational Materials*, *8*, 127.

Lian, C., & Zhuge, Y. (2010). Optimum mix design of enhanced permeable concrete – An experimental investigation. *Construction and Building Materials*, *24*, 2664–2671. https://doi.org/10.1016/j.conbuildmat.2010.04.057

Liu, W., Zhang, H., Chen, Z., Tao, J., Zhang, X., & Zhao, B. (2024). Performance of pervious concrete influenced by typical properties of recycled concrete aggregate and suggestions for practical use. *Road Materials and Pavement Design*, *25*, 1090–1111. https://doi.org/10.1080/14680629.2023.2238082

Luo, Y., Lv, Y., Wang, D., Jiang, Z., & Xue, G. (2023). The influence of coarse aggregate gradation on the mechanical properties, durability, and plantability of geopolymer pervious concrete. *Construction and Building Materials*, *382*, 131246. https://doi.org/10.1016/j.conbuildmat.2023.131246

Lyu, Q., Dai, P., & Chen, A. (2024). Correlations among physical properties of pervious concrete with different aggregate sizes and mix proportions. *Road Materials and Pavement Design*, 1–25. https://doi.org/10.1080/14680629.2024.2331193

Malami, S. I., Musa, A. A., Haruna, S. I., Aliyu, U. U., Usman, A. G., Abdurrahman, M. I., Bashir, A., & Abba, S. I. (2022). Implementation of soft-computing models for prediction of flexural strength of pervious concrete hybridized with rice husk ash and calcium carbide waste. *Modeling Earth Systems and Environment*, *8*, 1933–1947. https://doi.org/10.1007/s40808-021-01195-4

Mao, S. (1981). Regression analysis and experiment design. Huadong Normal University, Shanghai, China, p. 375.

Marani, A., & Nehdi, M. L. (2020). Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Construction and Building Materials*, *265*, 120286. https://doi.org/10.1016/j.conbuildmat.2020.120286

Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*, *1*, 140–147. https://doi.org/10.38094/jastt1457

Mazur, B., & Kotwa, A. (2019). Porous concrete as an Anti–Aquaplaning building material. *IOP Conference Series: Materials Science and Engineering*, *471*, 032027. https://doi.org/10.1088/1757-899X/471/3/032027

Miller, D. D. (2023). Chapter 7 - The calculated uncertainty of scientific discovery: From maths to deep maths. In S. G. Krantz, A. S. R. Srinivasa Rao, & C. R. Rao (Eds.), *Handbook of statistics* (pp. 203–226). Elsevier.

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Fundamentals of artificial neural networks and deep learning. In O. A. Montesinos López, A. Montesinos López, & J. Crossa (Eds.), *Multivariate statistical machine learning methods for genomic prediction* (pp. 379–425). Springer International Publishing.

Najah, N. S. H., Saloma, S., Hanafiah, H., Nurjannah, S. A., Muliawan, S., & Eric, E. (2021). Compressive strength, permeability, and porosity analysis of pervious concrete by variation of A/C without fine aggregate. *AIP Conference Proceedings*, *2339*, 020253. https://doi.org/10.1063/5.0044591

Neithalath, N., Sumanasooriya, M. S., & Deo, O. (2010). Characterizing pore volume, sizes, and connectivity in pervious concretes for permeability prediction. *Materials Characterization*, *61*, 802–813. https://doi.org/10.1016/j.matchar.2010.05.004

Nooraziah, A., & Tiagrajah, V. J. (2014). A study on regression model using response surface methodology. *Applied Mechanics and Materials*, *666*, 235–239. https://doi.org/10.4028/www.scientific.net/AMM.666.235

Papadopoulos, S., Azar, E., Woon, W.-L., & Kontokosta, C. E. (2018). Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. *Journal of Building Performance Simulation*, *11*, 322–332. https://doi.org/10.1080/19401493.2017.1354919

Pieralisi, R., Cavalaro, S. H. P., & Aguado, A. (2016). Discrete element modelling of the fresh state behavior of pervious concrete. *Cement and Concrete Research*, *90*, 6–18. https://doi.org/10.1016/j.cemconres.2016.09.010

Pradhan, S. K., & Behera, N. (2022). Performance assessment of pervious concrete road on strength and permeability by using silica fume. *Materials Today: Proceedings*, *60*, 559–568. https://doi.org/10.1016/j.matpr.2022.02.018

Putman, B. J., & Neptune, A. I. (2011). Comparison of test specimen preparation techniques for pervious concrete pavements. *Construction and Building Materials*, *25*, 3480–3485. https://doi.org/10.1016/j.conbuildmat.2011.03.039

Quan Tran, V., Quoc Dang, V., & Si Ho, L. (2022). Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach. *Construction and Building Materials*, *323*, 126578. https://doi.org/10.1016/j.conbuildmat.2022.126578

Radlińska, A., Welker, A., Greising, K., Campbell, B., & Littlewood, D. (2012). Long-term field performance of pervious concrete pavement. *Advances in Civil Engineering*, *2012*, 380795.

Ryshkewitch, E. (1953). Compression strength of porous sintered Alumina and Zirconia. *Journal of the American Ceramic Society*, *36*, 65–68. https://doi.org/10.1111/j.1151-2916.1953.tb12837.x

Sahdeo, S. K., Chandrappa, A., & Biligiri, K. P. (2021). Effect of compaction type and compaction efforts on structural and functional properties of pervious concrete. *Transportation in Developing Economies*, *7*, 19. https://doi.org/10.1007/s40890-021-00129-0

Sata, V., & Chindaprasirt, P. (2020). 19 - Use of construction and demolition waste (CDW) for alkali-activated or geopolymer concrete. In F. Pacheco-Torgal, Y. Ding, F. Colangelo, R. Tuladhar, & A. Koutamanis (Eds.), *Advances in construction and demolition waste recycling* (pp. 385–403). Woodhead Publishing.

Sathiparan, N., & Jeyananthan, P. (2023). Predicting compressive strength of cement-stabilized earth blocks using machine learning models incorporating cement content, ultrasonic pulse velocity, and electrical resistivity. *Nondestructive Testing and Evaluation*, 1–25. https://doi.org/10.1080/10589759.2023.2240940

Sathiparan, N., & Jeyananthan, P. (2024). Predicting compressive strength of quarry waste-based geopolymer mortar using machine learning algorithms incorporating mix design and ultrasonic pulse velocity. *Nondestructive Testing and Evaluation*, 1–24. https://doi.org/10.1080/10589759.2024.2304257

Sathiparan, N., Jeyananthan, P., & Subramaniam, D. N. (2023a). Effect of aggregate size, aggregate to cement ratio and compaction energy on ultrasonic pulse velocity of pervious concrete: Prediction by an analytical model and machine learning techniques. *Asian Journal of Civil Engineering*, *25*, 495–509.

Sathiparan, N., Jeyananthan, P., & Subramaniam, D. N. (2023b). Prediction of compressive strength of fly ash blended pervious concrete: A machine learning approach. *International Journal of Pavement Engineering*, *24*, 2287146. https://doi.org/10.1080/10298436.2023.2287146

Sathiparan, N., Jeyananthan, P., & Subramaniam, D. N. (2024a). Silica fume as a supplementary cementitious material in pervious concrete: Prediction of compressive strength through a machine learning approach. *Asian Journal of Civil Engineering*, *25*, 2963–2977.

Sathiparan, N., Jeyananthan, P., & Subramaniam, D. N. (2024b). Surface response regression and machine learning techniques to predict the characteristics of pervious concrete using non-destructive measurement: Ultrasonic pulse velocity and electrical resistivity. *Measurement*, *225*, 114006. https://doi.org/10.1016/j.measurement.2023.114006

Sathiparan, N., Wijekoon, S. H., Jeyananthan, P., & Subramaniam, D. N. (2024c). Soft computing to predict the porosity and permeability of pervious concrete based on mix design and ultrasonic pulse velocity. *International Journal of Pavement Engineering*, *25*, 2337916. https://doi.org/10.1080/10298436.2024.2337916

Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, *20*, 3–29. https://doi.org/10.1177/1536867X20909688

Seifeddine, K., Amziane, S., & Toussaint, E. (2022). State of the art on the mechanical properties of pervious concrete. *European Journal of Environmental and Civil Engineering*, *26*, 7727–7755. https://doi.org/10.1080/19648189.2021.2008511

Seifeddine, K., Amziane, S., & Toussaint, E. (2023). State of the art on the hydraulic properties of pervious concrete. *Road Materials and Pavement Design*, *24*, 2561–2596. https://doi.org/10.1080/14680629.2022.2164332

Shah, S. F. A., Chen, B., Zahid, M., & Ahmad, M. R. (2022). Compressive strength prediction of one-part alkali activated material enabled by interpretable machine learning. *Construction and Building Materials*, *360*, 129534. https://doi.org/10.1016/j.conbuildmat.2022.129534

Shilar, F. A., Ganachari, S. V., & Patil, V. B. (2022). Advancement of nano-based construction materials-A review. *Construction and Building Materials*, *359*, 129535. https://doi.org/10.1016/j.conbuildmat.2022.129535

Singh, A., Sampath, P. V., & Biligiri, K. P. (2020). A review of sustainable pervious concrete systems: Emphasis on clogging, material characterization, and environmental aspects. *Construction and Building Materials*, *261*, 120491. https://doi.org/10.1016/j.conbuildmat.2020.120491

Singh, A., Sampath, P. V., & Biligiri, K. P. (2023). Field performance monitoring of pervious concrete pavements. *Road Materials and Pavement Design*, *24*, 3013–3028. https://doi.org/10.1080/14680629.2023.2176164

Sonebi, M., & Bassuoni, M. T. (2013). Investigating the effect of mixture design parameters on pervious concrete by statistical modelling. *Construction and Building Materials*, *38*, 147–154. https://doi.org/10.1016/j.conbuildmat.2012.07.044

Song, H., Ahmad, A., Farooq, F., Ostrowski, K. A., Maślak, M., Czarnecki, S., & Aslam, F. (2021). Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. *Construction and Building Materials*, *308*, 125021. https://doi.org/10.1016/j.conbuildmat.2021.125021

Sriravindrarajah, R., Wang, N. D. H., & Ervin, L. J. W. (2012). Mix design for pervious recycled aggregate concrete. *International Journal of Concrete Structures and Materials*, *6*, 239–246. https://doi.org/10.1007/s40069-012-0024-x

Starbuck, C. (2023). Linear regression. In C. Starbuck (Ed.), *The fundamentals of people analytics: With applications in R* (pp. 181–206). Springer International Publishing.

Subramaniam, D. N., Jeyananthan, P., & Sathiparan, N. (2023). Soft computing techniques to predict the electrical resistivity of pervious concrete. *Asian Journal of Civil Engineering*, *25*, 711–722.

Subramaniam, D. N., & Sathiparan, N. (2022). Comparative study of fly ash and rice husk ash as cement replacement in pervious concrete: Mechanical characteristics and sustainability analysis. *International Journal of Pavement Engineering*, 1–18.

Sudhir Kumar, B., Srikanth, K., & Eeshwar, T. (2024). Implementation of soft computing techniques in forecasting compressive strength and permeability of pervious concrete blended with ground granulated blast-furnace slag. *Journal of Soft Computing in Civil Engineering*, *8*, 19–45.

Sun, J., Zhang, J., Gu, Y., Huang, Y., Sun, Y., & Ma, G. (2019). Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression. *Construction and Building Materials*, *207*, 440–449. https://doi.org/10.1016/j.conbuildmat.2019.02.117

Tan, I. A. W., Ahmad, A. L., & Hameed, B. H. (2008). Optimization of preparation conditions for activated carbons from coconut husk using response surface methodology. *Chemical Engineering Journal*, *137*, 462–470. https://doi.org/10.1016/j.cej.2007.04.031

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*, 7183–7192. https://doi.org/10.1029/2000JD900719

Torres, A., Hu, J., & Ramos, A. (2015). The effect of the cementitious paste thickness on the performance of pervious concrete. *Construction and Building Materials*, *95*, 850–859. https://doi.org/10.1016/j.conbuildmat.2015.07.187

Veza, I., Spraggon, M., Fattah, I. M. R., & Idris, M. (2023). Response surface methodology (RSM) for optimizing engine performance and emissions fueled with biofuel: Review of RSM for sustainability energy transition. *Results in Engineering*, *18*, 101213. https://doi.org/10.1016/j.rineng.2023.101213

Wang, H., Li, H., Liang, X., Zhou, H., Xie, N., & Dai, Z. (2019). Investigation on the mechanical properties and environmental impacts of pervious concrete containing fly ash based on the cement-aggregate ratio. *Construction and Building Materials*, *202*, 387–395. https://doi.org/10.1016/j.conbuildmat.2019.01.044

Wijekoon, S., Shajeefpiranath, T., Subramaniam, D., & Sathiparan, N. (2023). A mathematical model to predict the porosity and compressive strength of pervious concrete based on the aggregate size, aggregate-to-cement ratio and compaction effort. *Asian Journal of Civil Engineering*, 1–13.

Xie, N., Akin, M., & Shi, X. (2019). Permeable concrete pavements: A review of environmental benefits and durability. *Journal of Cleaner Production*, *210*, 1605–1621. https://doi.org/10.1016/j.jclepro.2018.11.134

Xiong, B., Gao, H., Chen, J., Lu, X., Tian, B., Chen, B., & Liu, W. (2023). Study on influencing factors of paste coating thickness of pervious concrete, buildings.

Xu, L., Ding, X., Niu, L., Huang, Z., & Sun, S. (2023). Experimental study on the mechanical properties and influencing factors of glass fiber-reinforced permeable concrete, materials.

Yahia, A., & Kabagire, K. D. (2014). New approach to proportion pervious concrete. *Construction and Building Materials*, *62*, 38–46. https://doi.org/10.1016/j.conbuildmat.2014.03.025

Yang, J., & Jiang, G. (2003). Experimental study on properties of pervious concrete pavement materials. *Cement and Concrete Research*, *33*, 381–386. https://doi.org/10.1016/S0008-8846(02)00966-3

Yap, S. P., Chen, P. Z. C., Goh, Y., Ibrahim, H. A., Mo, K. H., & Yuen, C. W. (2018). Characterization of pervious concrete with blended natural aggregate and recycled concrete aggregates. *Journal of Cleaner Production*, *181*, 155–165. https://doi.org/10.1016/j.jclepro.2018.01.205

Yu, F., Sun, D., Hu, M., & Wang, J. (2019a). Study on the pores characteristics and permeability simulation of pervious concrete based on 2D/3D CT images. *Construction and Building Materials*, *200*, 687–702. https://doi.org/10.1016/j.conbuildmat.2018.12.135

Yu, F., Sun, D., Wang, J., & Hu, M. (2019b). Influence of aggregate size on compressive strength of pervious concrete. *Construction and Building Materials*, *209*, 463–475. https://doi.org/10.1016/j.conbuildmat.2019.03.140

Zhang, J., Li, D., & Wang, Y. (2020). Predicting uniaxial compressive strength of oil palm shell concrete using a hybrid artificial intelligence model. *Journal of Building Engineering*, *30*, 101282. https://doi.org/10.1016/j.jobe.2020.101282

Zhang, L. V., Marani, A., & Nehdi, M. L. (2022b). Chemistry-informed machine learning prediction of compressive strength for alkali-activated materials. *Construction and Building Materials*, *316*, 126103. https://doi.org/10.1016/j.conbuildmat.2021.126103

Zhang, J., Niu, W., Yang, Y., Hou, D., & Dong, B. (2022a). Machine learning prediction models for compressive strength of calcined sludge-cement composites. *Construction and Building Materials*, *346*, 128442. https://doi.org/10.1016/j.conbuildmat.2022.128442

Zhao, W., & Du, S. (2016). Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, *54*, 4544–4554. https://doi.org/10.1109/TGRS.2016.2543748

Zhong, R., Leng, Z., & Poon, C.-S. (2018). Research and application of pervious concrete as a sustainable pavement material: A state-of-the-art and state-of-the-practice review. *Construction and Building Materials*, *183*, 544–553. https://doi.org/10.1016/j.conbuildmat.2018.06.131

Zhong, R., & Wille, K. (2016). Compression response of normal and high strength pervious concrete. *Construction and Building Materials*, *109*, 177–187. https://doi.org/10.1016/j.conbuildmat.2016.01.051