Indian Journal of Pure & Applied Physics
Vol. 62, April 2024, pp. 274-284
DOI: 10.56042/ijpap.v62i4.7768

Water Transmissible Pavement: A physics of Granular Sub-base Permeability through Road Dust Analysis using Machine Learning

Zishan Raza Khan[a], Syed Aqeel Ahmad[a] & Vikash Singh[a*]
Department of Civil Engineering, Integral University, Lucknow, Uttar Pradesh 226 026, India

Received 3 January 2023; accepted 5 February 2024

Amidst the rapid urbanization and heightened infrastructure demands, contemporary cities are capitalizing on every available space, converting previously permeable land into impermeable surfaces. This transition obstructs the absorption of storm water, leading to intensified runoff. To counteract this challenge and address infrastructure requirements, Low Impact Development (LID) techniques have emerged, among which permeable pavement stands out as a widely adopted solution. Serving as a transient storage facility, permeable pavements store storm water within their Granular Sub-Base (GSB) or reservoir layer, thereby diminishing the size of storm drains and contributing to the implementation of Sustainable Urban Drainage Systems (SUDS). Nevertheless, the utilization of permeable pavements is commonly recommended for walkways, parking lots, or low-volume roads due to their susceptibility to clogging. This study delves into the potential for clogging in the reservoir layer, employing Machine Learning models such as Random Forest (RF), Gradient Boosting Regressor (GBR), Light Gradient Boosting Machine (LGBM), and Extra Trees (ET). The investigation incorporates data from 200 instances with varying GSB layers, thicknesses, and combinations of road dust particle sizes. The results reveal a robust correlation ($R^2 > 0.97$) with experimental data, indicating that GSB-III demonstrates optimal clog resistance under high dust loads.

The findings suggest that GSB-V and GSB-VI may be suitable for areas with dust loads below 200 gm/month. This research provides valuable insights for the development of clog-resistant permeable pavements tailored to moderate to high-volume roads.

Keywords: Dust; Structural behaviour; Machine learning; Sustainable urban drainage system; Clog-resistant permeable pavement; GSB

1 Introduction

Given the swift pace of urbanization and the concurrent surge in impermeable surfaces, urban areas are confronted with formidable challenges in effectively managing stormwater runoff. Traditional impermeable pavements exacerbate urban flooding, water pollution, and strain on existing drainage infrastructure. To address these issues and advancement in sustainable urban development, there is a critical need for an innovative stormwater management solution to establish a Sustainable Urban Drainage System (SUDS) like any urban sustainable development efforts made by other researchers[1-5]. Permeable pavements emerge as a cost-effective and performance-efficient Low Impact Development (LID) measure for stormwater management. They operate by enabling the percolation of stormwater from a porous top layer to a sub-layer of porous aggregate reservoir[6]. These pavements are broadly categorized into three types: Porous Concrete (PC), Porous Asphalt (PA), and Pervious Interlocking Concrete Blocks (PICB)[7], each presenting distinct advantages and challenges.

PC Permeable Pavement (PC-PP) primarily composed of no-fines concrete, exhibits high porosity that facilitates stormwater infiltration. However, the presence of pores in concrete leads to low compressive strength and reduced durability. This results in ravelling over time, causing a decline in permeability[8]. Researchers have explored diverse materials, such as fly ash and polymer epoxies, to develop concrete mixes with high porosity and balanced strength[9].

PA Preamble Pavement (PA-PP), primarily consisting of coarse-grained aggregate with high porosity (20-22%) and permeability, demonstrates effectiveness in removing pollutants from stormwater[10]. However, the highly tortuous nature of pores in PA-PP leads to clogging from road dust sediments, reducing effective permeability[11-12].

PICB Permeable Pavement (PICB-PP) shares functionality with PA-PP and PC-PP. Designed for

*Corresponding author: (E-mail: vikashsinghiu96@gmail.com)

areas with minimal vehicular traffic, PICB-PP allows stormwater passage through gaps between pavers filled with poorly graded stone grit. However, choking of gaps and the bedding layer with fine particles leads to more than 80% permeability loss over time, necessitating biannual maintenance[13].The global trend of adopting permeable pavement systems as an environmentally friendly solution to manage stormwater runoff has been driven by urbanization. Cities like Berlin, Munich, and Tokyo have incorporated permeable pavements into urban planning to enhance water management and alleviate stress on traditional drainage infrastructure. However, the widespread adoption of permeable pavements has encountered common challenges related to clogging, primarily attributed to the accumulation of road dust. This challenge compromises pavement permeability, resulting in reduced effectiveness in stormwater management. Consequently, there has been a strategic shift towards developing clog-resistant permeable pavement solutions capable of withstanding the impact of road dust and maintaining optimal functionality over time.

The concept of clog-resistant permeable pavement focuses on employing different methods, with the most effective being the inclusion of 'drainage cells' or 'through holes' in the top layer to facilitate water transfer to the underlying layers, which may eventually clog with road dust sediments[14].The permeable pavement structure comprises three layers: the top layer for vehicular movement, the second bedding layer, and the third Granular Sub Base (GSB) layer, highly porous and permeable, composed of poorly graded aggregates (ASTHO 57)[15]. A novel approach is proposed to bypass the top layer and direct stormwater directly to the GSB layer, functioning as a reservoir. Despite the GSB layer's high initial permeability, the accumulation of road dust over time may reduce its permeability, jeopardizing the effectiveness of the permeable pavement. Researchers aim to study the clogging behavior of different GSB grades recommended by the Ministry of Road Transport & Highways (MoRTH), Government of India. This research utilizes Machine Learning (ML) to model the clogging behavior, leveraging ML as an effective tool for predicting future behavior[16-18].

In summary, the research aims to address challenges associated with clogging in permeable pavements, with a focus on innovative solutions and the application of machine learning to model and predict clogging behavior in different GSB grades under real-world conditions.

2 Research Significance

This study holds significant merit in its pioneering integration of sophisticated machine learning (ML) methodologies, incorporating nuanced model evaluation metrics such as the determination coefficient (R-squared), root mean square error (RMSE), mean absolute error (MAE), and mean squared error (MSE). Harnessing the capabilities of advanced ensemble methods including Gradient Boosting Regressor (GBR), Random Forests (RF), Extra Tree (ET), and Light Gradient Boosting Machine (LGBM), the research endeavours to unravel the intricate interdependencies among different grades of Granular Sub-Base (GSB) layers, diverse thicknesses of layers, varying compositions of clogging materials by percentage, and their collective impact on the permeability of permeable pavement.

Through the application of these sophisticated ML techniques, this research not only contributes to the establishment of a robust and comprehensive framework for the accurate prediction and optimization of permeability across diverse layers but also serves as a foundational step towards a more data-driven and sustainable approach in realizing the concept of an innovative stormwater management solution, ultimately leading to the establishment of a Sustainable Urban Drainage System. The intricate insights derived from this study have the potential to revolutionize the understanding and implementation of permeable pavement systems, fostering advancements in urban water management practices.

3 Materials and Method

The investigation encompasses Lucknow, situated at Latitude 26.846695 and Longitude 80.946167, serving as the capital of Uttar Pradesh in northern India, with an estimated population exceeding 3.9 million in the year 2023[3]. Lucknow experiences an annual rainfall average of 981mm, predominantly concentrated within the Monsoon season from June to September, where more than 850mm occurs. The remaining precipitation is distributed across the winter months, spanning from December to January. The prevailing wind direction in Lucknow predominantly moves from the West to the North, exhibiting higher speeds during the daytime, ranging between 14 to 18 kmph, and slower speeds at night, averaging 4 to 7 kmph. The climatic conditions in Lucknow feature hot and humid summers, marked by average maximum temperatures ranging from 43°C to 46^0C in May-June, and cold, dry winters with average minimum temperatures ranging from 6°to 10°C.

A major factor influencing the clogging of permeable pavement in Lucknow is road dust. This study involves the collection and analysis of road sweepings from various locations in Lucknow City. The grain size distribution curve (Fig. 1) indicates a Coefficient of Uniformity (Cu) of 3.27 and a Coefficient of Curvature (Cc) equal to 1.00. The fineness on the 0.075 mm IS sieve is measured at 10.44%, falling within the 5-12% range. However, the condition Cu>6 and 1≤Cc ≤3 is not met. Consequently, the road sweeping is classified as SP-SM, denoting Poorly Graded Sand Containing Silt (IS: 1498-1970)[19]. The results signify that the permeability of this SP-SM soil ranges between 2.55E-05 m/s to 5.35E-04 m/s, categorizing it as having average permeability. Additionally, the findings suggest that the permeable pavement surrounding this soil is prone to a high dust load and is susceptible to clogging if not adequately maintained.

The grain size analysis reveals that over 50% of the road sweepings consist of particles finer than 0.180mm, falling into the classification of road dust. These finer particles are susceptible to being entrained and transported by stormwater runoff, posing a potential risk of clogging within the reservoir layer of permeable pavement, specifically the Granular Sub-Base (GSB), unless the pavement is meticulously designed to mitigate such issues. In the scope of this study, three distinct grain size ranges (Fig. 2) have been delineated: 0.18mm to 0.15mm, 0.15mm to 0.075mm, and less than 0.075mm. The investigation aims to conduct cyclic clogging experiments by evaluating the permeability of the GSB layer, considering variations in thickness, specifically 100mm, 150mm, and 200mm. This experimental approach seeks to elucidate the impact of different grain size ranges on the cyclic clogging behavior of the GSB layer at varying thicknesses, providing insights into the dynamic interaction between road dust particles and permeable pavement structures.

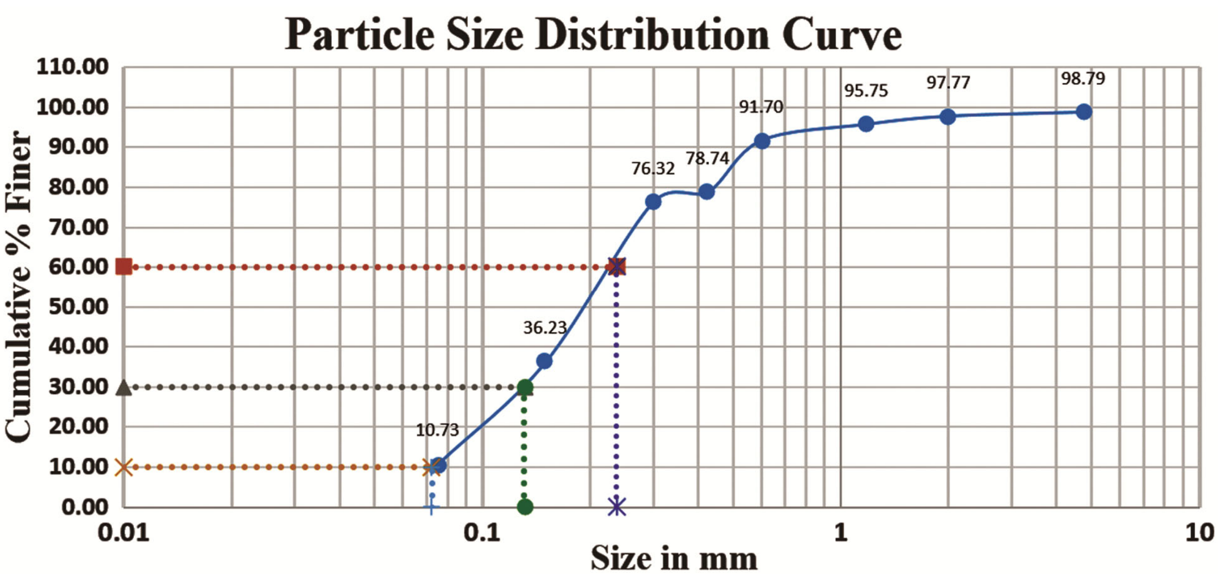

Fig. 1 — Grain size distribution curve of the road sweepings.

0.18mm-0.15mm 0.15mm-0.075mm <0.075mm

Fig. 2 — Different particle sizes of road dust.

Following the stipulations outlined by the Ministry of Road Transport and Highways (MoRTH)[23], Government of India, in the year 2019, the Granular Sub-Base (GSB) material is systematically classified into six distinct grades. For this investigation, particular attention is directed towards the examination of GSB grades III, V, and VI, as delineated in Table 1. This selective focus aligns with the specified grading system established by MoRTH, facilitating a detailed scrutiny of the specified GSB grades with regard to their inherent characteristics and suitability for designated applications within the realm of road construction and infrastructure development.

A clogging experiment was executed to scrutinize the permeability characteristics of coarse aggregate when subjected to muddy water, thereby inducing progressive clogging across consecutive test cycles to emulate the phenomenon of stormwater transporting road dust. Furthermore, the outcomes of the laboratory permeability assessment test will undergo comparative analysis with the predictions derived from a Machine Learning (ML) model, providing a comprehensive validation of the numerical model's efficacy in predicting permeability reduction behaviors. The laboratory test employs the constant head method based on Darcy's law as shown in Fig. 3.

The investigation focuses on GSB-III, V, and VI grades, in accordance with Clause 9.6.1 and Table 9.3 of the Ministry of Road Transport and Highways (MoRTH) specifications from 2019. These grades represent the intended reservoir layer within a

Table 1 — Adopted GSB Grade on the basis of recommended of Table 9.3 of MoRTH guidelines

IS Sieve Designation	Percent by Weight Passing the IS Sieve					
	Grade III		Grade V		Grade VI	
	Recommended	Adopted	Recommended	Adopted	Recommended	Adopted
75.0mm	-	-	100	100	-	-
53.0mm	100	100	80-100	100	100	100
26.5mm	55-75	65	55-90	70	75-100	75
9.50mm	-	-	35-65	50	55-75	65
4.75mm	10-30	20	25-50	30	30-55	35
2.36mm	-		10-20	15	10-25	18
0.85mm	-		2-10	3	-	-
0.425mm	-		0-5	0	0-8	0
0.075mm	<5	0	-	-	0-3	0

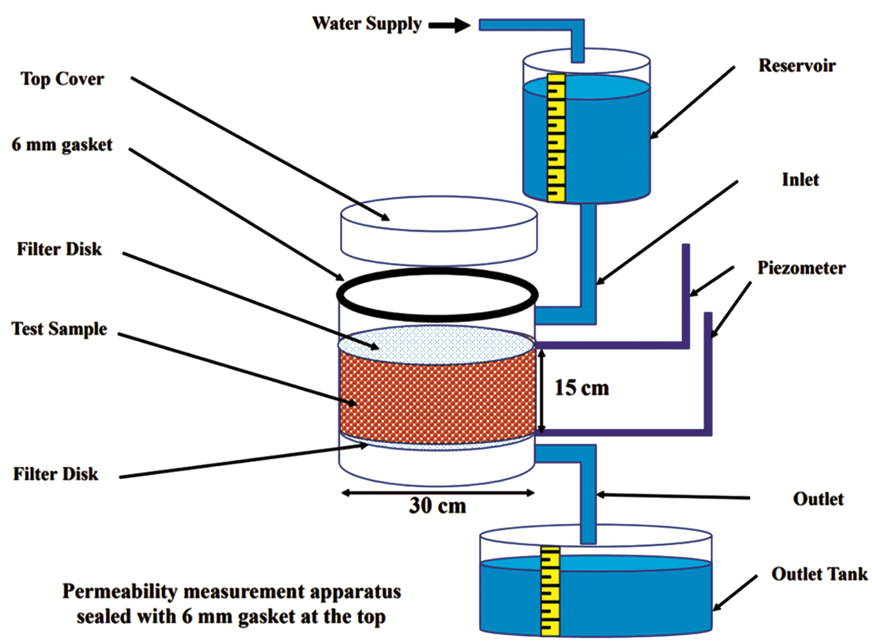

Fig. 3 — Different components of the permeability measurement apparatus with specifications.

permeable pavement structure. The test specimen is positioned within a permeability cell and compacted to attain a minimum bulk density of 1.95gm/cm^3, adhering to MoRTH guidelines. A porous filter disk is placed atop the sample to ensure uniform flow distribution. The permeability cell is hermetically sealed with a lid equipped with a 6mm gasket to eliminate any potential leakage, and the sample undergoes a 24-hour saturation process with water to eliminate entrapped air within the aggregate pores.

The experimental apparatus includes a cylindrical cell, piezometer tubes, filter disk, measuring tank, inlet and outlet pipes, all interlinked with a constant head reservoir containing water (Fig. 4). Subsequently, three sets of samples, each with thicknesses of 100mm, 150mm, and 200mm, are subjected to cyclic sediment loading using diverse combinations. This meticulous experimental design aims to replicate real-world conditions, facilitating a comprehensive understanding of how different sediment loads impact the permeability of the GSB layer under varying thicknesses, contributing valuable insights to the study of clogging behavior in permeable pavements.

Figure 5 provides an overview of the general processes involved in the development of the Machine Learning (ML) models, with a focus on their implementation and execution in the research

4 Results and discussion

Within the framework of this investigation, the acquired dataset underwent a meticulous pre-processing phase to ensure its appropriateness for subsequent analysis and modelling. A pivotal step in this pre-processing involved addressing missing values through the application of imputation techniques, wherein absent data points were estimated or filled based on available information. This step holds significance as the presence of missing data could introduce biases or inaccuracies in the subsequent analytical procedures. The methodical assembly of data in this investigation involved the compilation of a comprehensive dataset encompassing 180 data points. The dataset incorporates crucial input parameters, including diverse types of Granular Sub-Base (GSB) grade layers, varying thicknesses of GSB layers, and different combinations of clogging materials denoted as S_1 (0.18mm-0.15mm), S_2

Fig. 4 — Laboratory test for permeability measurement under cyclic clogging with road dust.

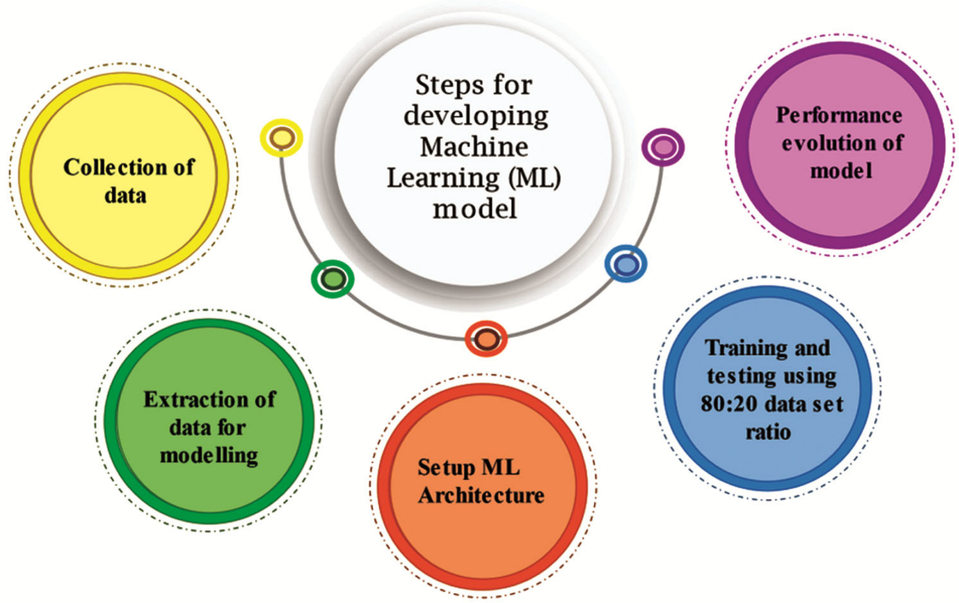

Fig. 5 — Methodology Steps for ML Model Development in the Current Study.

(0.15mm-0.075mm), and S_3 (<0.075mm), represented as percentages. Out of the total dataset, 70% of the data points were specifically allocated for the training phase of the machine learning (ML) models. Simultaneously, the remaining 30% of the dataset was reserved for rigorous testing and comprehensive evaluation of the ML models' performance. This meticulous approach ensures a robust and representative training set for the models, facilitating their adeptness in generalizing patterns and behaviors to effectively predict and optimize permeability across diverse scenarios within the scope of permeable pavement layers.

Additionally, the dataset underwent scaling procedures to standardize its distribution, aiming for a mean of 0 and a standard deviation of 1. Data scaling is a standard practice in the realm of data preprocessing, particularly when confronted with variables of disparate scales. This normalization process proves essential for certain machine learning algorithms, ensuring equitable influence from each variable in the subsequent analytical and modelling processes.

The implementation of the current study hinged on a synergistic utilization of various software tools, with Matlab and Python serving as the primary coding and data analysis platforms. Both languages are widely adopted in scientific computing and machine learning domains owing to their extensive libraries and tools tailored for data manipulation, analysis, and modelling. The integration of these software tools played a pivotal role in expediting the study's execution and extracting meaningful insights from the dataset. Each tool was strategically employed throughout the research workflow, facilitating efficient data analysis, modelling, and visualization, thereby bolstering the overall success of the study and enhancing the reliability of its findings.

Table 2 — Descriptive statistics of the input and output parameters

Variables	Thickness of Pavement (mm)	S_1 (%)	S_2 (%)	S_3 (%)	Permeability (m/day)
Count	187	187	187	187	187
Mean	156.7	35.4	34.8	29.8	387.4
SD	40.8	37.4	37.6	35.3	128.0
Min	100.0	0.0	0.0	0.0	233.9
Q1	100.0	0.0	0.0	0.0	264.3
Median	150.0	33.3	33.3	0.0	436.9
Q3	200.0	50.0	50.0	50.0	471.1
Max	200.0	100.0	100.0	100.0	585.6
Skewness	-0.3	0.6	0.6	0.9	0.2
Kurtosis	-1.5	-0.9	-0.9	-0.5	-1.5

4.1 Descriptive statistics of selected study parameters

Descriptive statistics reveal fundamental patterns in data through measures like mean, median, and mode, aiding in the identification of central tendencies and potential outliers. In the current study forecasting GSB grade layers' permeability, Table 2 presents concise numerical summaries, facilitating insights into key dataset characteristics. Researchers utilize these statistics to discern central tendencies and identify outliers, vital for accurate prediction models and further investigation into anomalies or errors.

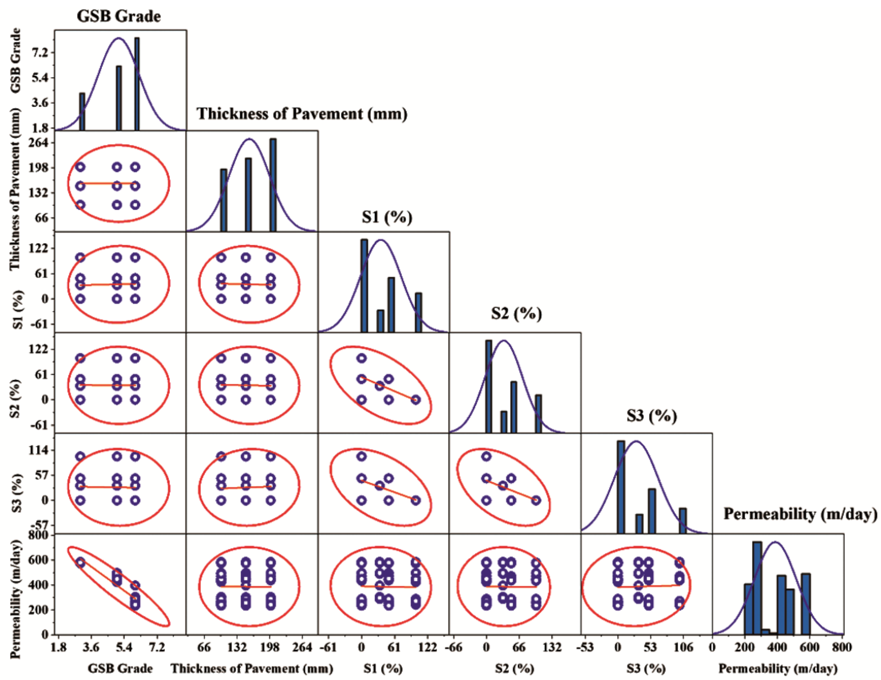

Fig. 6 — Combination of Scatter plots and histogram of permeability of various input parameters with respect to output parameters.

4.2 Scatter Plots and Histogram

Scatter plots, a key visual tool in data analysis, illustrate relationships between two numeric variables in a Cartesian coordinate system. In the current study, Fig. 6 displays scatter plots revealing patterns in the relationship between input parameters and the permeability of different GSB grade layers. Clear linear trends in these plots signify significant impacts of specific input parameters on resulting permeability values. This visual analysis aids in informed decision-making and optimizing concrete mix designs for desired permeability in engineering applications. The study's use of scatter plots enhances data interpretability, facilitating meaningful conclusions about critical factors influencing concrete performance. Recognition of moderate linear relationships in some plots emphasizes the importance of assessing variable contributions to prediction accuracy, considering their independence and information value.

A histogram, a vital graphical tool, visually represents the frequency distribution of both continuous and discrete data. In Fig. 6 of the current study, histograms depict normalized values of input and output parameters, offering insights into data distribution patterns and outliers. These visualizations aid in assessing parameter adequacy, identifying data trends, and detecting anomalies. The information gleaned guides researchers in understanding parameter significance and its impact on research outcomes. Histograms contribute to robust data exploration, revealing dataset structure, central tendencies, and outliers for enhanced decision-making and research reliability. In summary, histograms are crucial for intuitive data analysis, providing valuable insights and supporting the study's success.

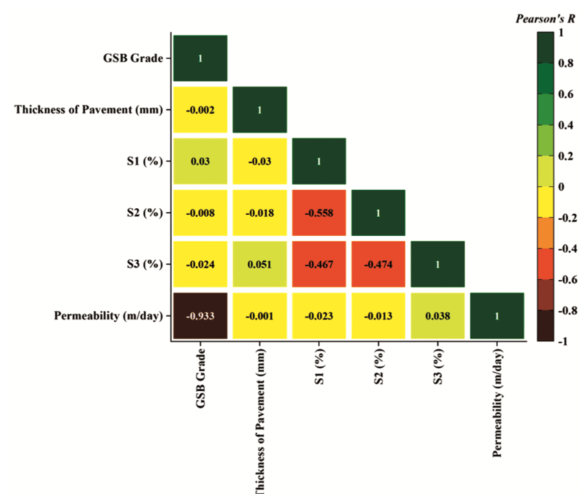

Fig. 7 — Pearson's R showing the current study's correlation matrix between input parameters and output parameters.

4 Correlation Matrix

A correlation matrix, a tabular display of correlation coefficients between variables, is pivotal in uncovering interrelationships, detecting multi-collinearity, and pinpointing outliers. In the ongoing study, Fig. 7 illustrates the correlation matrix, revealing associations among different input and output variables of the current study. The coefficients in the matrix quantify positive or negative associations, aiding in understanding dependencies. Discrepancies between correlation values may highlight evaluation issues, impacting analysis efficiency. The matrix provides valuable insights into variable interrelationships, informing modelling decisions and identifying patterns or issues within the dataset. In summary, the correlation matrix is a potent tool for analyzing relationships, guiding variable selection, and ensuring robust analyses in research investigations.

4.4 Assessment of errors

4.4.1 Mean absolute error

For model evaluations, the Mean Absolute Error (MAE) serves as a performance metric, gauging the average absolute difference between actual and predicted values in a dataset. This metric is crucial for assessing the accuracy of Machine Learning (ML) models, providing insights into the alignment of model predictions with actual values.

$$MAE = \frac{1}{n}\sum_{i=1}^{N}|P_i - O_i| \quad (1)$$

Where, P_i = predicted value, O_i = true value and n = total number of data points

4.4.2 Root mean squared error

For assessing model performance with large error values, Root Mean Squared Error (RMSE) is a suitable metric. RMSE, derived from the standard deviation of prediction errors, is a variation of Mean Squared Error (MSE), incorporating the square root for accuracy determination (Zhang et al., 2021). RMSE offers a balanced representation in scenarios where substantial errors can impact model performance significantly.

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^{N}(P_i - O_i)^2}{n}\right)} \quad (2)$$

Where, P_i = predicted value, O_i = true value and n = total number of data points

4.4.3 Correlation Coefficient (R^2)

The coefficient of correlation, denoted as R^2, serves as a statistical metric gauging the alignment between data points and the fitted regression line. It quantifies the proportion of variation in the dependent variable predictable from the independent variable(s) within a regression model. A higher R^2 value signals a more favorable model fit to the reference data, indicative of enhanced predictive capability.

In this study, Fig. 8 visually presents the relationship between predicted and experimental permeability for each machine learning (ML) model under consideration. These visual representations offer insights into the efficacy of the ML models in approximating actual permeability values, facilitating a comparative evaluation against experimental data. Scrutinizing these visualizations provides valuable perspectives on the ML models' performance in predicting the permeability of different GSB layers of varying thicknesses.

The analysis of Fig. 9 unveils a data distribution during training and testing and found it in inter quartile range The mean value increases during testing. The elevated coefficients of determination (R^2) for these models strongly support their exceptional fitness and predictive accuracy concerning the acquired data. Notably, Extra Tree (ET), Random Forest (RF), Light GBM (LGBM), and Gradient Boosting Regression (GBR) models exhibit an almost linear relationship between test and forecasted permeability values. During numerical validation, GBR, RF, and ET models exhibit slightly higher scatter and dispersion between the diagonal

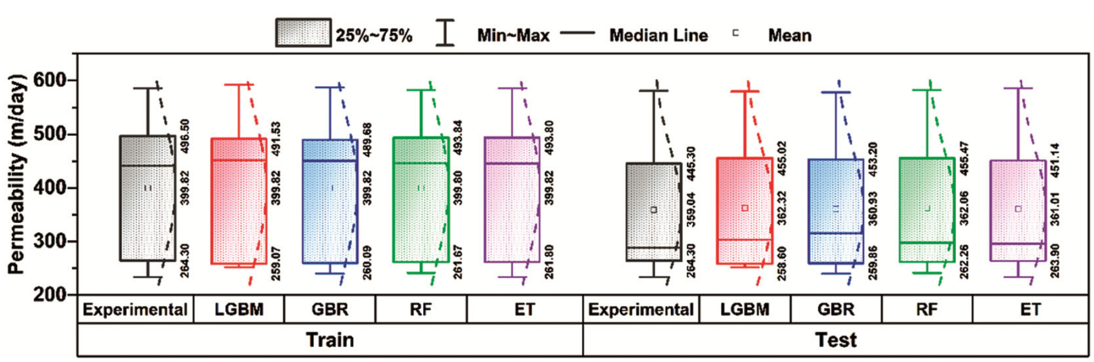

Fig. 8 — Comparison of data set during training and testing.

and test data points compared to LGBM during training as shown in Fig. 9. However, during testing, RF emerges as the best-performing model compared to others. The detailed performance of all models during training and testing is presented in Tables 3.

For the provided models (LGBM, GBR, RF, and ET) during training: LGBM has the lowest MAE (13.67) and RMSE (18.32), indicating relatively smaller errors. RF and ET have similar performance in terms of R^2, RMSLE, and MAPE, with slightly higher errors compared to LGBM.GBR shows competitive performance but has slightly higher errors compared to LGBM. During the testing phase: Random Forest (RF) performs exceptionally well with the lowest MAE, MSE, RMSE, and the highest R^2 (0.99), indicating accurate predictions and good model fit. Extra Trees (ET) also demonstrates strong performance, particularly with the lowest MAPE (0.03), indicating precise percentage-wise predictions. Gradient Boosting Regress or (GBR) and Light Gradient Boosting Machine

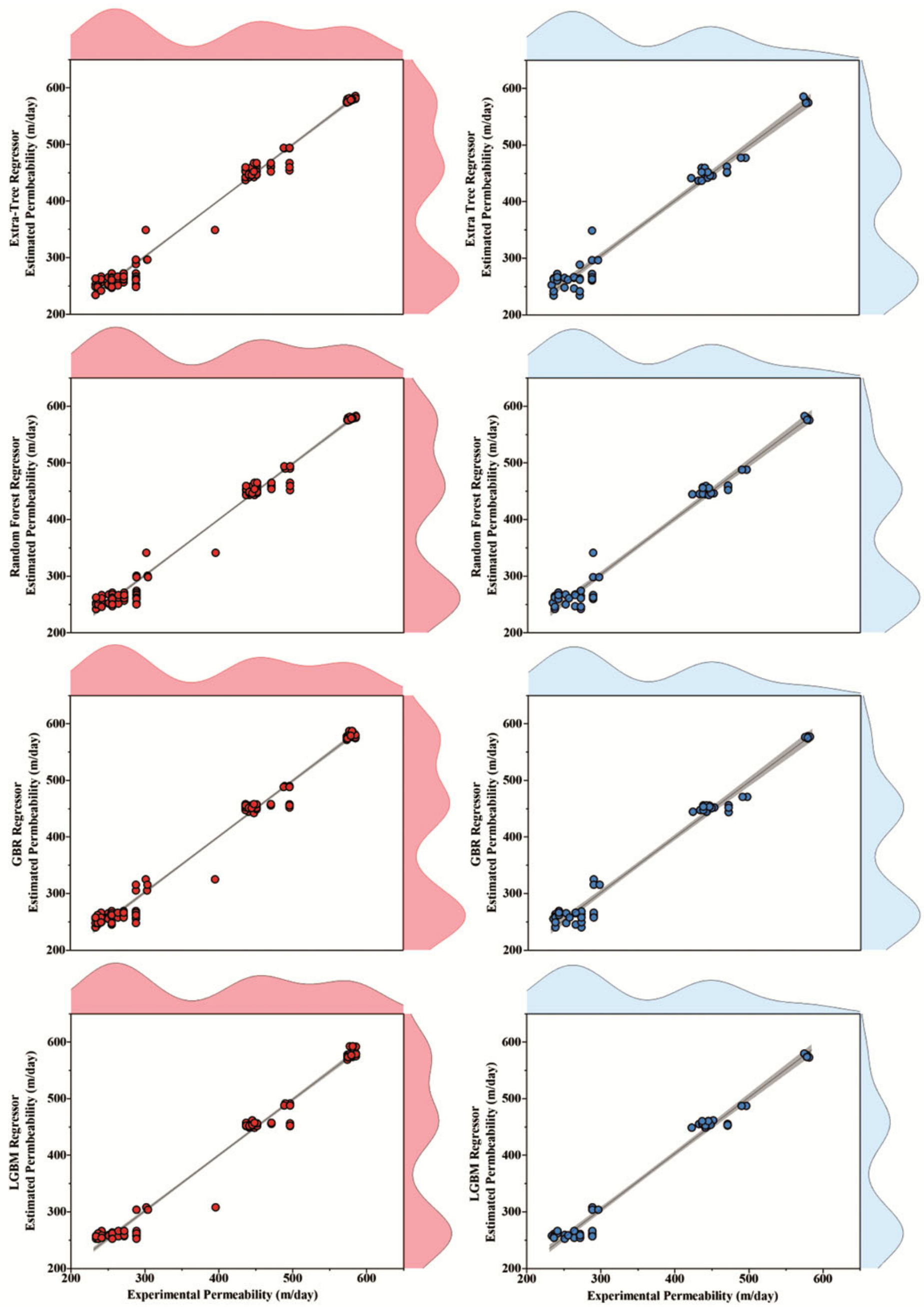

Fig. 9 — Scatter diagram showing R^2 during training (left) and testing (right).

Table 3 — Performance of different ML models during training and testing (right)

	Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE
Training	LGBM	13.67	364.67	18.32	0.98	0.06	0.04
	GBR	14.83	426.59	19.99	0.97	0.06	0.05
	RF	14.84	460.06	20.63	0.97	0.07	0.05
	ET	15.60	536.30	22.12	0.97	0.07	0.05
	Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE
Test	LGBM	12.79	284.01	16.85	0.98	0.05	0.04
	GBR	12.59	274.48	16.57	0.98	0.05	0.04
	RF	11.16	241.07	15.53	0.99	0.05	0.04
	ET	10.82	247.11	15.72	0.98	0.05	0.03

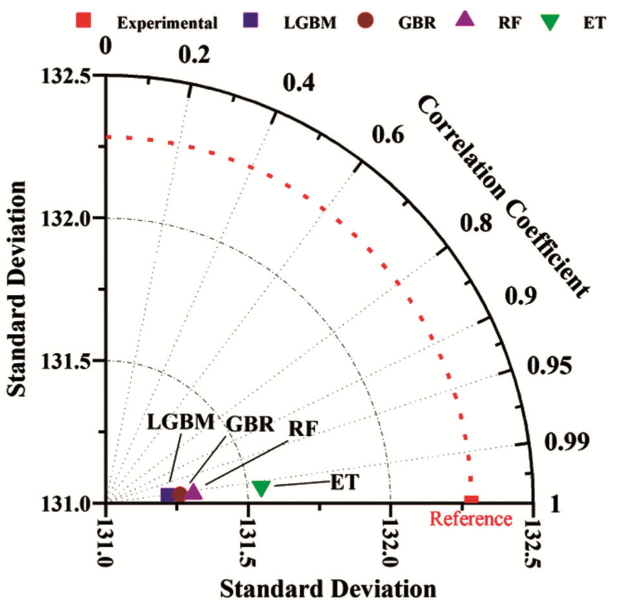

Fig. 10 — Performance of ML models for permeability prediction using Taylor's diagram.

(LGBM) show competitive results, with R^2 values of 0.98 and low error metrics, signifying accurate predictions and good overall performance.

In summary, these metrics provide insights into how well each model is capturing the patterns in the training data, with lower values generally indicating better performance.

5 Taylor's diagram

The Taylor diagram serves as a graphical tool for assessing the concordance between observed patterns or combinations of observed patterns and reference data. It quantifies similarity by evaluating correlation, Root Mean Square Error (RMSE) difference, and standard deviations of changes in each pattern. Taylor diagrams are particularly valuable for a comprehensive examination of complex model characteristics and the comparative evaluation of distinct models. In this investigation, Fig. 10 illustrates Taylor's diagram for all considered models. Analysis of Fig. 10 reveals that the Gradient Boosting Regression (GBR) and Light GBM (LGBM) models exhibit remarkable performance, closely aligning with the reference data compared to other models such as Extra Tree (ET) and Random Forest (RF). Models with higher correlation values and lower RMSE are deemed favorable in Taylor's diagram, indicating stronger agreement with the reference data.

The results presented in Fig. 10 for the machine learning (ML) models align with earlier findings in relevant studies[20-21]. These results suggest that the GBR and LGBM models are well-suited for predicting permeability through different Granular Sub Base (GSB) layers, showcasing higher concordance with reference data and lower RMSE differences compared to other models.

6 Conclusion

Urbanization and increased infrastructure demands necessitate the transformation of permeable surfaces into impermeable ones in urban areas, impeding the natural absorption of stormwater and inducing intensified runoff. Mitigating these challenges, Low Impact Development (LID) techniques, exemplified by permeable pavements, are employed. Permeable pavements function as transient stormwater storage units, diminishing the dimensions of storm drains and contributing to the establishment of Sustainable Urban Drainage Systems (SUDS). Similar works have been carried out by different researchers[23-24].

The presented regression models (LGBM, GBR, RF, and ET) undergo training on a designated dataset, with subsequent evaluation on an independent testing dataset. Notably, Random Forest (RF) and Extra Trees (ET) exhibit commendable performance across both training and testing phases, manifesting lower errors and superior goodness of fit as indicated by the R^2 metric. LGBM and GBR also demonstrate satisfactory performance, albeit RF and ET outperform them across multiple metrics during the testing phase.

Specifically, during training, LGBM attains the lowest Mean Absolute Error (MAE) of 13.67 and Root Mean Squared Error (RMSE) of 18.32, indicative of relatively diminutive errors. RF and ET showcase comparable performance concerning R^2, Root Mean Squared Logarithmic Error (RMSLE), and Mean Absolute Percentage Error (MAPE), albeit with slightly elevated errors compared to LGBM. GBR exhibits competitive performance, albeit with marginally higher errors compared to LGBM.

During the testing phase, RF excels, achieving the lowest MAE, Mean Squared Error (MSE), RMSE, and the highest R^2 (0.99), reflecting precision in predictions and an exemplary model fit. ET similarly displays robust performance, particularly with the lowest MAPE (0.03), signifying accurate percentage-wise predictions. GBR and LGBM exhibit competitive results, featuring R2 values of 0.98 and low error metrics, denoting precise predictions and overall model efficacy. Furthermore, the permeability prediction performance of machine learning models, assessed using Taylor's diagram, is elucidated in the broader context.

Acknowledgements

The authors acknowledge the support of Integral University, Lucknow for granting seed money (IUL/IIRC/SMP/2021/003) through Integral Information & Research Centre for carrying out the research. The manuscript communication no. is IU/R&D/2024-MCN0002320.

References

1 Khan M K & Singh B P, *IOSR J Mech Civil Eng*, 12 (2015) 101.
2 Abuamer I M, Sadat M & Tampère C M, Int Conf Comput Character Tech Eng Sci, CCTES, (2018) 127.
3 Mumtaz N, Izhar T, Pandey G & Labhasetwar P K, *Elsevier*, Woodhead Publishing, (2022) 259.
4 Banerjee R, Srivastava J B & Gupta N, *Int J Inter Des Manuf*, (2023) 1.
5 D'Ambrosio R & Longobardi A, *Sustain Cities Soc*, 98 (2023) 104856.
6 Putri F K, Hidayah E & Ma'ruf M F, *Water Sci Technol*, (2023).
7 Madrazo-Uribeetxebarria E, Antín M G, Eguilegor G A & Andrés-Doménech I, *Construct Build Mater*, 387 (2023) 131587.
8 Razzaghmanesh M & Beecham S, *Water*, 10 (2018) 337.
9 Sathiparan N, Jeyananthan P & Subramaniam D N, *Asian J Civil Eng*, (2023). https://doi.org/10.1007/s42107-023-00790-3
10 Hammes G &Thives L P, *Water*, 15 (2023) 2962.
11 Hommel J, Coltman E & Class H, *Trans Porous Med,* 124 (2018) 589.
12 Kia A, Imperial College London, (2018).
13 Lamberti L, Piccilli D G A, Minetto B, Perius C F & Tassi R, *Ciência e Natura*, 45 (2023) e84089.
14 Scholz M & Grabowiecki P, *Build Environ*, 42 (2007) 3830.
15 ASSTHO 57 reference Stormwater Management Guidebook, Center for Watershed Protection-U.S.A-2013 Stormwater Management Guidebook, District Department of the Environment Watershed Protection Division, District of Columbia, USA, (2013).
16 Bano S, Mumtaz N, Bano F & Ahmad S A, *Asian J Civil Eng*, (2023). https://doi.org/10.1007/s42107-023-00872-2
17 Ahmad S A, Rafiq S K, Hilmi H D M, *et al. Asian J Civil Eng*, (2023). https://doi.org/10.1007/s42107-023-00811-1
18 Wijekoon S H, Shajeefpiranath T, Subramaniam D N, *et al., Asian J Civil Eng*, (2023). https://doi.org/10.1007/s42107-023-00757-4.
19 IS1498:1970, Classification and Identification of Soils for General Engineering Purposes (FIRST REVISION)
20 Pandey M, Jamei M, Ahmadianfar I, Karbasi M, Lodhi A S & Chu X, *J Hydrol*, 2021 (2021) 127330.
21 Yesiloglu-Gultekin N & Gokceoglu C, *J Non-destruct Eval*, 41 (2022) 10.
22 Minstry of Roads Transport and Highways (MoRTH), Specifications for Road and Bridges works, Fifth Revision, Indian Road Congress, New Delhi, (2013).
23 Nivedya M K & Mallick R B, *Int J Pavem Eng*, 21 (2020) 1057.
24 Radfar, Ata, Rockaway T D & Ehsaei A, *Watershed Manag*, 2015 (2015) 149.