Machine Learning-Based Prediction of Chloride Resistance in Low-Carbon Concretes Using Experimental Data

Tran Huyen Vu[1], Yuxuan Yang[2], Trinh Cao[3], Harish Srivastava[4] and Vute Sirivivatnanon [5]

[1]Postdoctoral Research Associate, University of Technology Sydney, New South Wales, Australia
[2]PhD Candidate & Research Assistant, University of Technology Sydney, New South Wales, Australia
[3] Multiplex, New South Wales, Australia
[4]Director Civil Engineering, Transport for NSW, Macquarie Park, New South Wales, Australia
[5]Professor of Civil Engineering, University of Technology Sydney, New South Wales, Australia
[6] Research Director, SmartCrete CRC

Abstract: This study presents reliable predictive models using machine learning techniques, combined with advanced functions, to predict the chloride resistance of low-carbon concretes. Chloride resistance, measured through the Rapid Migration Test Nordtest (NT) 492, is an internationally recognised durability index for assessing a concrete's ability to resist chloride penetration. The predictive models were trained and validated on a dataset of 36 experimental results, encompassing a wide range of compressive strengths and replacement quantities of supplementary cementitious materials in low-carbon concrete mixtures. The models demonstrated high accuracy, with an R-squared value of approximately 0.9, a mean square error (MSE) of 1.9, and a mean absolute error (MAE) of 0.9. In addition to NT 492, the study considered the NT 443 test, which is more onerous, and is recommended for quality control due to its capacity to assess long-term chloride resistance. This research also introduces a simplified prediction model for NT 443 based on NT 492, utilising data collected from high-quality studies in the literature. The correlation between NT 492 and NT 443 was strong, with a correlation coefficient of 0.93, and the prediction equation for NT 443 yielded an R-squared value of 0.87. These findings highlight the NT 492 test as a practical and efficient alternative to the NT 443 test, offering a simpler and more efficient approach for evaluating chloride resistance in low-carbon concretes.

Keywords: *Machine Learning, Correlation coefficient, R-squared, Chloride resistance, NT 443, NT 492*

1. INTRODUCTION

Reinforced concrete (RC) structures play a vital role in civil infrastructure but are prone to degradation over time [1–3]. Among the various factors driving this deterioration, chloride attack is particularly detrimental to the durability of RC structures, especially in marine environments [4,5]. While the penetration of chlorides alone does not harm concrete directly, the accumulation of chloride ions at the steel reinforcement level can exceed the critical threshold, causing de-passivation of the steel's protective layers and triggering corrosion[3,6–8]. This corrosion undermines the structural integrity and safety of RC structures, often necessitating early repairs and resulting in substantial economic costs. Therefore, improving resistance to chloride ingress is critical for ensuring the long-term durability of RC structures.

Among the various methods used to evaluate chloride resistance, natural exposure tests in real-world environments are often considered the most reliable [3,9]. This approach involves placing test specimens in actual service conditions of concrete structures or directly sampling existing in-situ structures for analysis. While it provides authentic and precise data, the extended exposure duration spanning years or even decades can make the results outdated and impractical. As a result, accelerated testing methods are typically preferred for assessing concrete's chloride resistance. A widely used time-saving approach is the Rapid Migration Test (NT 492) [10], which assesses chloride resistance by measuring the chloride diffusion coefficient. This coefficient, known as the "non-steady-state migration coefficient" (Dnssm), is calculated based on Fick's second law. NT 492 is currently recognised as a durability index for assessing chloride resistance in concrete cited in various international standards, including the fib Model Code 2010 [11] and BS EN 206:2013 [12], as well as by organisations such as Transport for New South Wales (Australia) [13]. Consequently, it has gained widespread adoption among researchers and regulatory bodies as a means to determine whether a concrete mix meets the required chloride resistance criteria.

In recent years, machine learning has become a powerful tool for tackling real-world challenges, providing innovative solutions across various scientific fields such as robotics, statistics, bioinformatics, computer science, and construction materials [14]. As a subset of artificial intelligence, machine learning has gained significant attention due to its ability to achieve high accuracy, efficiently handle large-scale computations, adapt to complex environments, and offer a more cost-effective alternative to traditional statistical and experimental methods. Recent studies have increasingly leveraged machine learning techniques to predict the chloride penetration resistance of concrete [15–22]. The success of machine learning models largely depends on the quality and relevance of the input data. Without accurate and well-structured data, predictions may be unreliable or misleading. It is crucial to acknowledge that different test methods, such as the 90-day salt ponding test (AASHTO [23] and ASTM C1543 [24]), bulk diffusion tests (ASTM C1556 [25]or NT 443 [26], which require a minimum of 35 days), natural exposure testing, and the Rapid Migration Test (NT 492)—all yield chloride diffusion coefficients. However, the interpretation of these coefficients varies depending on the test method used. For example, a chloride diffusion coefficient of 8×10^{-12} m²/s from NT 443 would indicate poor chloride resistance, whereas the same value from NT 492 would suggest strong resistance.

Many studies aiming to predict chloride resistance falter at the foundational stage of data collection, often due to a limited understanding of how chloride diffusion coefficients vary across different test methods. As a result, while these studies may generate predictions, they often fail to provide meaningful insights into the actual chloride resistance of concrete. This study addresses this challenge by leveraging high-precision experimental data from 36 concrete mixes tested using NT 492 to develop machine-learning models that predict chloride resistance based on concrete mix compositions. By referencing predictions to reliable and validated experimental results, this approach enhances the accuracy and practicality of estimating chloride resistance, offering a valuable tool for optimising concrete durability.

Furthermore, all concrete samples in this study underwent a shortened curing process, remaining in a water tank for only two days after one day in the mould, rather than the conventional 28-day water curing commonly reported in chloride resistance studies. This reduced curing period was intentionally chosen to better reflect real-world site conditions rather than controlled laboratory environments, making the findings more relevant to practical construction scenarios.

Gradient Boosting Regression

Among the machine learning techniques, Gradient Boosting Regression (GBR) is a powerful one for the purpose of prediction [27]. The GBR model used in this study is a boosting ensemble method that improves predictions by combining multiple simple models, known as weak learners [28]. Typically, the construction of a GBR model involves two key steps, as described meticulously in a report by Su et al. [29]. First, weak learners are created one at a time in a sequence. Each new weak learner is trained to correct the errors made by the previous models. The goal is to adjust the new weak learners, so they focus on the areas where the overall model is making mistakes. This is done by aligning them with the negative gradient of the model's loss function, which helps improve accuracy. Second, after training multiple weak learners, they are combined using a weighted averaging method to make final predictions. This ensures that the model benefits from all the learners while reducing individual errors. However, GBR has a drawback because it is a 'greedy algorithm', meaning it keeps adding new learners to reduce errors, which can lead to overfitting if too many learners are added. To prevent this, regularisation techniques must be used to ensure the model performs well on new data.

As an ensemble learning method, GBR sequentially improves its predictions by minimising errors at each iteration, leading to high predictive accuracy. Unlike traditional regression models, GBR effectively captures the nonlinear relationships between key mix parameters, particularly the concrete proportions in this study. Additionally, GBR provides valuable insights into feature importance, which will be applied to identify the most influential factors affecting chloride resistance of various concrete mixes. Its robustness against outliers (data points that significantly deviate from the expected trend) and noisy data (random fluctuations or measurement errors) makes it well-suited for experimental studies, where variations in measurements are common.

2. RESEARCH SIGNIFICANCE

Low-carbon concrete emerges as a sustainable alternative to traditional concrete, addressing the high carbon dioxide (CO_2) emissions linked to Portland cement production [30–32]. This research defines low-carbon concrete as concrete designed to meet workability and structural performance requirements while minimising embodied carbon. As Portland cement is a significant contributor to global CO_2 emissions, incorporating supplementary cementitious materials (SCMs) such as fly ash, ground granulated blast furnace slag (GGBFS/slag), silica fume, and natural pozzolans as partial replacements for Portland cement can markedly reduce CO_2 emissions. Beyond lowering the carbon footprint, SCMs also enhance concrete's resistance to chloride penetration [33], providing a dual advantage by improving durability while promoting environmentally friendly construction practices.

This study leverages GBR to predict chloride resistance of low carbon concrete mixes, as measured by NT 492, based on mix compositions. By integrating machine learning, this approach enhances the evaluation process, enabling engineers and industry professionals to make data-driven decisions and optimise mix designs more efficiently. The adoption of ML-driven methodology advances construction materials science, offering a more intelligent and sustainable approach to assessing concrete durability, paving the way for a more efficient and AI-driven future in concrete materials science. Through advanced techniques such as feature importance evaluation, predictive modelling, and correlation analysis, this study uncovers critical insights that traditional methods often overlook. Machine learning not only identifies key factors influencing chloride resistance but also reveals complex interactions among them, providing a deeper understanding of concrete behavior. Moreover, this approach enables highly accurate predictions, significantly reducing the reliance on labor-intensive and time-consuming testing.

This study further introduces a simplified prediction model for chloride resistance using NT 443, developed based on NT 492 measurements and supported by high quality literature data. These findings will determine if NT 492 is a practical and efficient alternative to NT 443, and is capable of providing a faster and more cost-effective approach for evaluating chloride resistance in low-carbon concretes.

3. METHODOLOGY

3.1 Materials and concrete mix proportions

The binders examined in this study include ground granulated blast furnace slag (GGBFS/Slag), Class F fly ash, and ordinary Portland cement (OPC). The aggregates used in this study consist of a blend of four types: 10 mm crushed rock, 20 mm crushed rock, manufactured sand, and natural sand. For the 20 mm aggregates, the predominant particle sizes are 13.2 mm and 9.5 mm, whereas for the 10 mm aggregates, the dominant sizes are 6.7 mm and 4.75 mm. In the case of manufactured sand, most particles fall within the 2.36 mm to 1.18 mm range, while natural sand contains a significant proportion of finer particles, predominantly between 150 µm and 425 µm.

The mix designs for the evaluated 36 concrete mixes are presented in Table A in the Appendix and were produced by using the Unit Water Method.

3.2 Testing procedure

After mixing, the concrete specimens were left in the mould for one day before being demolded. Specimens designated for compressive strength testing (ASTM C39/C39M-21) were then cured in a lime water tank for an additional 27 days before testing. In contrast, specimens intended for the NT 492 test were placed in the lime water tank for only 2 days, after which they were stored in air under laboratory conditions. Upon reaching 28 days of age, these specimens were tested according to the NT 492 procedure.

3.3 Evaluation of model accuracy of ML model for predicting NT492

Assessing the accuracy and performance of a machine learning (ML) model is essential to ensure its reliability and effectiveness. The selection of an appropriate evaluation metric depends on the nature of the problem whether it is a classification or regression task, as well as the characteristics of the dataset. In regression problems, commonly used metrics include Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2), as they provide key insights into predictive accuracy.

The R^2 metric (see Eq.1) reflects how well the model captures relationships between predictors and the target variable. A higher R^2 value, approaching 1, indicates that the model accounts for most of the data's variability. However, relying solely on R^2 can be misleading, as it is sensitive to overfitting and does not convey the actual magnitude of errors or the influence of outliers.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad \text{(Eq.1)}$$

where $SS_{res} = \sum\left(y_{true} - y_{pred}\right)^2$ is the residual sum of squares and $SS_{tot} = \sum(y_{true} - \bar{y})^2$ =is the total sum of squares.

MAE (see Eq.2) and MSE (see Eq.3) serve as complementary metrics to R^2 by offering a more detailed perspective on error distribution. MAE calculates the average absolute deviation between predicted and actual values, providing an intuitive measure of accuracy that is less affected by large errors. In contrast, MSE measures how far the predicted values are from the actual values by taking the average of the squared differences between them. Using both helps detect bias in error distribution and ensures a well-rounded understanding of model performance.

$$MAE = \frac{\sum|y_{true} - y_{pred}|}{n} \quad \text{(Eq.2)}$$

$$MSE = \frac{\sum\left(y_{true} - y_{pred}\right)^2}{n} \quad \text{(Eq.3)}$$

where: y_{true} is the actual NT492 values, y_{pred} is the predicted NT492 values, and n is the total number of observations.

By incorporating these diverse evaluation metrics, a more comprehensive understanding of model performance can be achieved, balancing goodness-of-fit with error minimisation.

3.3 Boosting algorithm for decision trees applied in this study

Gradient Boosting can be implemented with various types of base learners; in this study, Decision Trees are utilised. These trees are built following the boosting methodology, where each new tree corrects the errors made by the previous trees. The algorithm of boosting of decision trees applied in this study is based on Friedman's (2002) [28].

Algorithm 1: Gradient_TreeBoost

$F_0(\mathbf{x}) = \arg\min_\gamma \sum_{i-1}^{N} \Psi(y_i, \gamma)$.

For $m = 1$ to M do:

$\quad \tilde{y}_{im} = -\left[\frac{\partial\Psi(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)}\right]_{F(\mathbf{x})=F_{m-1}(\mathbf{x})}, \ i = 1, N$

$\quad \{R_{lm}\}_1^L = L -$ terminal node $tree(\{\tilde{y}_{im}, \mathbf{x}_i\}_1^N)$

$\quad \gamma_{lm} = \arg\min_\gamma \sum_{\mathbf{x}_i \in R_{lm}} \Psi(y_i, F_{m-1}(\mathbf{x}_i) + \gamma)$

$\quad F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \nu \cdot \gamma_{lm} 1(\mathbf{x} \in R_{lm})$

endFor.

Figure 1. Algorithm of boosting of decision trees [28]

The key steps in the boosting decision trees are as follows:

i. **Initialise model**: The first step is to initialise the model with a constant function that minimises the loss function across all training data:

$F_0(x) = \underset{\gamma}{arg\ min} \sum_{i=1}^{N} \psi(y_i, \gamma)$ for i = 1 to N, where γ is a constant that best fits the training data.

ii. **Iterative Boosting Process**:
For each boosting iteration m (from 1 to M):

- Compute pseudo-residuals, in which pseudo-residuals represent the gradient of the loss function concerning the current model's predictions. These residuals indicate how much correction is needed at each training point. Instead of simply using the residuals (y_i−F(x_i)), Gradient Boosting computes a more general correction by using the gradient of the loss function**:**

$$\tilde{y}_{im} = -\left[\frac{\partial\Psi(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$$

In this study, the loss function is measured using Mean Squared Error (MSE), presented as Ψ(y,F(x))=(y-(F(x)). Taking the derivative for F(x), then $\frac{\partial L}{\partial F(x)} = -2\left(y - F(x)\right)$, thus pseudo-residuals become $\tilde{y}_{im}$= -2(y-F(x)). The result is negated, ensuring that the next model step moves in the opposite direction of the gradient to minimise the loss.

- Train a Regression Tree on Residuals
  A new regression tree is trained using these pseudo-residuals as the target variable, instead of the actual y values. The input remains the original feature set x_i, but the target values (or predicted values) are now the computed residuals $\tilde{y}_{im}$. The final tree consists of L terminal nodes (leaves) and each leaf contains a subset of training samples that share similar pseudo-residuals. A constant value is assigned to each leaf, which is the average of the pseudo-residuals within that region

- Compute best terminal node predictions:

$$\gamma_{lm} = arg\ \min_{\gamma} \sum_{x_i \epsilon R_{lm}} \psi(y_{i,F_{m-1}}(x_i) + \gamma)$$

  After the tree is built, the best prediction value (denoted as γ_{lm}) for each region is determined, then this value is chosen to minimise the overall loss in that region. Mathematically, γ_{lm} is found by adjusting it until the error in that region is as small as possible.

- Update model: Once the optimal leaf values γ_{lm} are determined, the model is updated to incorporate the new tree's corrections. The updated model is expressed as:

$$F_m(x) = F_{m-1}(x) + v \sum_{l=1}^{L} \gamma_{lm}\, 1(x \in R_{lm})$$

  where: $F_{m-1}(x)$ is the previous iteration's model, v is the learning rate, controlling the step size for each update, γ_{lm} is the constant prediction for each leaf, and $1(x \in R_{lm})$ is an indicator function, meaning the update only applies to data points that belong to region R_{lm}. This gradually improves predictions by correcting errors step by step.

iii. **Repeat Until Convergence**: The boosting process continues for m iterations, with each new tree refining the model by correcting the errors left by previous trees. As the iterations progress, the model becomes more accurate, gradually minimising the loss function. The final output is not just a single tree but an ensemble of all the trees, each contributing to the overall prediction.

4. RESULTS AND DISCUSSIONS

4.1 Results of 28-day compressive strength, experimental chloride resistance as per NT 492, and performance of GBR in predicting chloride resistance

The compressive strength at 28 days (com28) of 36 concrete mixes ranges from 34 to 83 MPa, is detailed in Table 1. Meanwhile, the experimentally measured chloride resistance according to NT 492 (NT492) varies from 4.5 to 38.1 $10^{-12}m^2/s$. The input parameters, including OPC cement (Ce), Fly ash (FA), Slag (Slag), Combined aggregate (Aggregate) amounts, and water to binder ratio (w/b) for each mix, are also shown in Table 3.

The Gradient Boosting Regressor (GBR) achieved exceptional predictive performance, achieving high R^2 values for both the training set (0.9999) and the test set (0.9775). The values indicate that GBR can explain a large proportion of the variance in the data, even on unseen test data, thus showcasing excellent generalisation capabilities.

Furthermore, the model exhibited low error metrics, with a test Mean Absolute Error (MAE) of 0.9568 and a test Mean Squared Error (MSE) of 1.989, confirming its high predictive accuracy with minimal error. This combination of strong performance metrics and robust generalisation ability establishes GBR as a highly suitable model for predicting chloride resistance in concrete mixes.

Table 1. Results of compressive strength, experimental and predicted chloride resistance

Order	Ce	FA	Slag	Aggregate	w/b	com28	Experimental NT492	Predicted NT492	Set
1	338	0	0	1874	0.57	49.6	37.9	37.8	Train
2	255	84	0	1857	0.57	45.6	38.1	38.0	Train
3	205	137	0	1851	0.57	36.7	30.5	31.3	Test
4	171	171	0	1842	0.58	34	29.9	30.1	Train
5	220	0	118	1864	0.57	43.7	21.0	21.0	Train
6	167	0	167	1856	0.57	42.4	17.2	17.1	Train
7	118	0	218	1859	0.57	37	12.3	12.4	Train
8	135	68	135	1855	0.55	44.3	14.1	14.1	Train
9	101	101	135	1857	0.54	37.6	13.5	13.5	Train
10	386	0	0	1812	0.5	54.1	27.4	27.2	Train
11	294	98	0	1816	0.5	54	26.9	27.0	Train
12	234	156	0	1789	0.5	48.8	29.7	29.6	Train
13	199	199	0	1803	0.5	44	28.7	27.9	Test

14	254	0	137	1821	0.5	55.3	12.3	12.3	Train
15	195	0	195	1819	0.5	47.4	9.6	9.6	Train
16	136	0	253	1815	0.5	41.4	9.5	8.8	Test
17	159	79	159	1829	0.44	51.9	7.6	7.6	Train
18	119	119	159	1820	0.47	47.1	10.6	10.6	Train
19	461	0	0	1767	0.42	70.7	13.0	16.6	Test
20	346	117	0	1746	0.41	66.5	14.4	14.4	Train
21	280	186	0	1741	0.42	60	15.4	15.5	Train
22	233	233	0	1727	0.42	58.5	20.6	20.4	Train
23	297	0	161	1749	0.42	62.7	7.1	7.0	Train
24	233	0	233	1754	0.42	58.8	6.5	6.5	Train
25	163	0	304	1751	0.41	47.4	6.5	5.9	Test
26	190	95	190	1759	0.4	58.6	7.1	7.1	Train
27	141	141	187	1735	0.39	55.6	7.4	7.4	Train
28	574	0	0	1671	0.34	83.3	10.1	10.5	Train
29	432	144	0	1641	0.34	78.6	11.1	10.5	Test
30	348	232	0	1625	0.34	74.5	9.4	9.2	Train
31	290	290	0	1607	0.34	72.9	11.2	11.2	Train
32	372	0	200	1650	0.34	72.7	5.8	5.9	Test
33	286	0	286	1638	0.34	67.3	5.3	5.3	Train
34	199	0	369	1623	0.34	56.4	4.5	4.5	Train
35	228	114	228	1612	0.34	61.9	6.1	6.1	Train
36	171	171	228	1607	0.34	61.8	5.9	6.4	Test

The visualisation of predicted vs. experimental NT492 is presented in Figures 3a and 3b. In Figure 3a, the dashed black line (Y = X) represents the ideal case where predictions perfectly match actual values, while the orange best-fit regression line provides insight into the model's overall performance. A high R^2 value of 0.9872 indicates that the GBR model effectively captures the variance in the data, demonstrating strong predictive accuracy. The majority of data points align closely with the Y = X line, signifying minimal error in the model's predictions. However, slight deviations can be observed at higher NT492 values, suggesting potential room for fine-tuning to enhance prediction consistency.

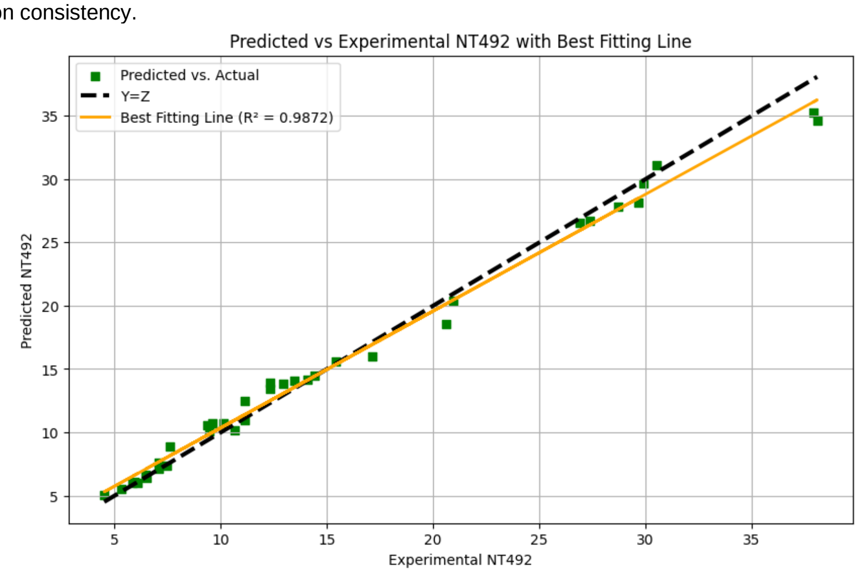

Figure 3. Predicted vs. Experimental NT492

The residuals plot (Figure 4) visualises the differences between the predicted and experimental NT492 values, offering insight into the model's accuracy and potential biases. The X-axis represents the experimental NT492 values, while the Y-axis represents the residuals (Predicted - Experimental values). The dashed black line at 0 serves as a reference, where perfect predictions would result in all residuals being equal to zero. A random distribution of residuals around the zero line indicates most residuals are small, suggesting that the model makes relatively accurate predictions across different NT492 values. However, at higher NT492 values (above 30), the residuals show larger negative deviations, indicating that the model tends to underestimate NT492 in extreme cases. While the residuals are not showing a strong pattern, the spread at higher values suggests that further fine-tuning or feature engineering might improve the model's predictions in extreme cases.

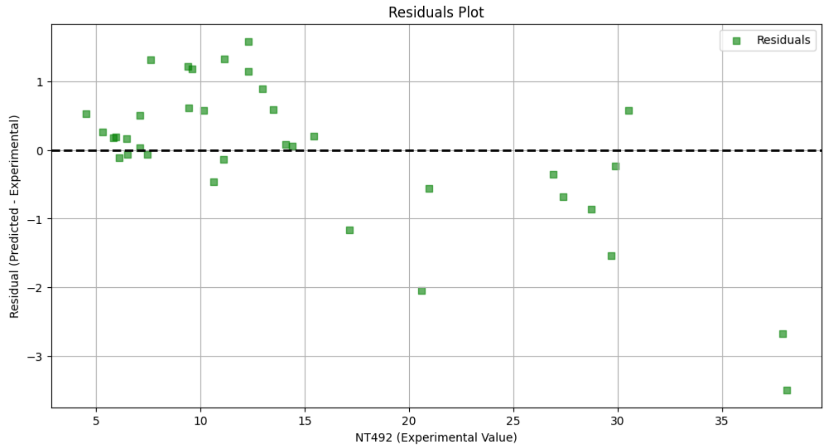

Figure 4. The residuals plot

Overall, these visualisations support the hypothesis that that the GBR model is highly effective in predicting NT492 values, showing excellent generalisation capabilities across the dataset.

4.2 Analysis of Feature Importance

In decision tree-based models like GBR, 'feature importance' is typically derived based on how often and how significantly a feature contributes to the decision-making process in the trees. In other words, feature importance, as presented in Figure 5 provides deep insights into how various combinations of factors influence the chloride resistance of concrete, as measured by NT 492. The water-to-binder ratio (w/b) is identified as the most influential factor in predicting NT492. This aligns with concrete durability principles, as a lower w/b ratio generally leads to denser concrete with reduced permeability, thereby improving resistance to chloride ingress. Additionally, slag and slag w/b interactions show high importance, suggesting that the replacement of cement with slag plays a crucial role in enhancing chloride resistance. Slag contributes to durability by refining the pore structure and reducing permeability, making concrete more resistant to aggressive environments. The aggregate w/b interaction and slag aggregate interaction also play a significant role, indicating that the combined effects of binder chemistry and aggregate proportions influence NT492. This suggests that optimising both binder composition and aggregate selection can improve concrete resistance to chloride penetration. These findings highlight key mix design strategies for optimising chloride resistance.

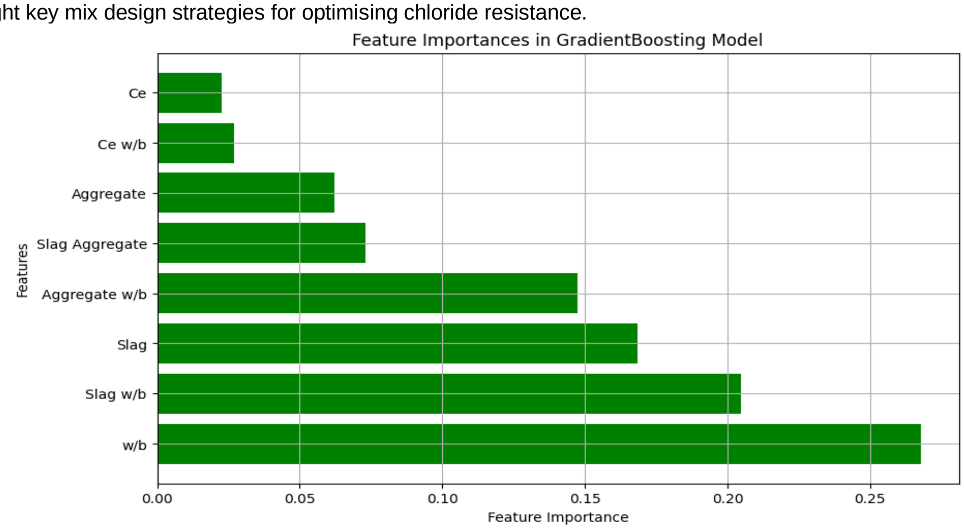

Figure 5. Feature importance

4.3 Correlation of chloride resistance measured by NT 492 and NT443

Data was gathered from high-quality studies in the literature, as outlined in Table 2. The chloride diffusion coefficient, measured using NT443 and NT492 for the same concrete mixes, exhibited a strong correlation based on linear regression analysis. The model achieved a train R^2 of 0.8718, test R^2 of 0.8204, and overall R^2 of 0.8688, demonstrating its reliability. Additionally, the Mean Absolute Error (MAE) values for the training and test sets were 1.4365 and 0.9886, respectively. These findings, along with the linear regression visualisation, are presented in Figure 6, accompanied by the formula provided below:

$$y=0.8587x - 0.1696$$

where y is chloride resistance measured by NT 492 and x is chloride resistance measured by NT 443

Table 2. Chloride diffusion coefficient measured by NT443 and NT492 of the same concrete mixes

ID	Chloride diffusion coefficient (10^{-12} m^2/s)		Ref	ID	Chloride diffusion coefficient (10^{-12} m^2/s)		Ref	ID	Chloride diffusion coefficient (10^{-12} m^2/s)		Ref
	NT 443	NT 492			NT 443	NT 492			NT 443	NT 492	
1	3.23	2.85	[34]	15	8.79	6.43	[35]	29	6.42	4.82	[35]
2	2.85	1.75		16	15.32	11.14		30	6.75	4.99	
3	1.91	0.9		17	24.85	20.91		31	13.2	9.1615	[36]
4	1.81	0.2		18	25.26	26.72		32	4.59	6.3015	
5	7.99	4.76		19	3.91	3.24		33	5.92	3.934	
6	13.48	6.02		20	7.01	5.71		34	11.46	5.231	
7	6.59	7.036	[37]	21	7.86	7.93					
8	4.061	4.309		22	8.97	9.51					
9	1.803	2.217		23	4.85	4.63					
10	1.701	1.569		24	5.03	5.76					
11	1.553	2.314		25	5.88	6.3					
12	1.058	0.744		26	6.4	6.96					
13	2.283	2.898		27	4.32	3.9					
14	3	3.466		28	6.08	5.39					

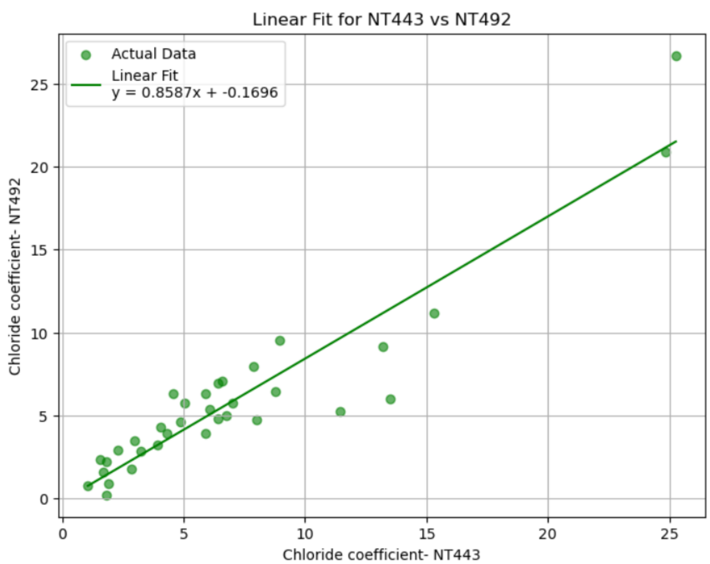

Figure 6. The linear fit for NT 443 and NT 492

5. Conclusions

This study successfully developed and validated machine learning-based predictive models for assessing the chloride resistance of low-carbon concretes using experimental data. By leveraging Gradient Boosting Regression (GBR), the model demonstrated high predictive accuracy, achieving an R^2 value of approximately 0.9, along with low Mean Squared Error (MSE) and Mean Absolute Error (MAE) values. These results highlight the efficacy of machine learning in accurately predicting chloride resistance, making it a valuable tool for optimising concrete mix designs and improving durability assessments.

A key contribution to this research is the introduction of a simplified prediction model for NT 443, developed based on NT 492 measurements. The strong correlation ($R^2 = 0.87$) between NT 443 and NT 492 confirms that NT 492 can serve as a reliable alternative to NT 443, offering a quicker and more cost-effective method for evaluating chloride resistance in low-carbon concretes. This has significant implications for construction practices, enabling engineers and industry professionals to conduct efficient durability assessments without extensive long-term testing.

Furthermore, the study highlights the importance of feature selection in chloride resistance predictions, with water-to-binder ratio (w/b) and slag content identified as the most influential parameters. The findings reinforce the need for a strategic approach to mix design to enhance concrete durability while maintaining sustainability goals.

In conclusion, this research demonstrates the potential of ML - driven approaches in concrete materials science, paving the way for data-driven decision-making in infrastructure development. The integration of machine learning techniques not only improves prediction accuracy but also contributes to a more efficient, cost-effective, and environmentally friendly approach to assessing and optimising low-carbon concrete durability.

6. ACKNOWLEDGEMENTS

This research is funded through the 'SmartCrete' CRC (https://smartcretecrc.com.au) with the support of Cement Concrete and Aggregates Australia (https://www.ccaa.com.au) and Transport for New South Wales (https://www.transport.nsw.gov.au). The authors are grateful for the support of the three organisations and the University of Technology Sydney (UTS).

7. REFERENCES

[1] M.R. Sherman, D.B. McDonald, D.W. Pfeifer, Durability aspects of precast prestressed concrete part 2: Chloride permeability study, PCI Journal 41 (1996) 76–95. https://doi.org/10.15554/pcij.07011996.76.95.
[2] D.E. Angulo Ramirez, G.R. Meira, M. Quattrone, V.M. John, A review on reinforcement corrosion propagation in carbonated concrete – Influence of material and environmental characteristics, Cem Concr Compos 140 (2023). https://doi.org/10.1016/j.cemconcomp.2023.105085.
[3] A. Neville, Chloride attack of reinforced concrete: an overview, Mater Struct 28 (1995) 63–70. https://doi.org/10.1007/BF02473172.
[4] O.E. Gjorv, O. Vennesland, Diffusion of Chloride Ions from Seawater into Concrete, Cem Concr Res 9 (1979) 229–238.
[5] CCAA, Chloride Resistance of Concrete. Cement Concrete & Aggregates Australia, Cement Concrete & Aggregates Australia, 2009.
[6] T. Huyen Vu, L.C. Dang, G. Kang, V. Sirivivatnanon, Chloride induced corrosion of steel reinforcement in alkali activated slag concretes: A critical review, Case Studies in Construction Materials 16 (2022). https://doi.org/10.1016/j.cscm.2022.e01112.
[7] Q. Yuan, C. Shi, G. De Schutter, K. Audenaert, D. Deng, Chloride binding of cement-based materials subjected to external chloride environment - A review, Constr Build Mater 23 (2009) 1–13. https://doi.org/10.1016/j.conbuildmat.2008.02.004.
[8] V. Sirivivatnanon, R. Khatri, Testing and Specifying Chloride Resistance of Concrete, 8th Austroads Bridge Conference (2011) 472–487.
[9] M. Castellote, C. Andrade, C. Alonso, Measurement of the steady and non-steady-state chloride diffusion coefficients in a migration test by means of monitoring the conductivity in the anolyte chamber. Comparison with natural diffusion tests, Cem Concr Res 31 (2001) 1411–1420. https://doi.org/10.1016/S0008-8846(01)00562-2.
[10] Nordtest, Chloride migration coefficient from non-steady-state migration experiments- NT Build 492, (1999) 1–8.
[11] fib, Model code 2010.fib Model Code for Concrete Structures, Fédération internationale du béton, 2010.
[12] BSI, Concrete – Specification, performance, production and conformity- BS EN 206:2013, British Standards Institution, 2013.
[13] Transport for NSW, QA Specification B80- Concrete Work for Bridges, 2021.
[14] C. Sammut, G.I. Webb, Encyclopedia of Machine Learning and Data Mining Second Edition, Springer Nature, 2017.
[15] A.R. Boğa, M. Öztürk, I.B. Topçu, Using ANN and ANFIS to predict the mechanical and chloride permeability properties of concrete containing GGBFS and CNI, Compos B Eng 45 (2013) 688–696. https://doi.org/10.1016/j.compositesb.2012.05.054.
[16] M. Marks, M.A. Glinicki, K. Gibas, Prediction of the chloride resistance of concrete modified with high calcium fly ash using machine learning, Materials 8 (2015) 8714–8727. https://doi.org/10.3390/ma8125483.
[17] M. Marks, M.A. Glinicki, K. Gibas, Prediction of the chloride resistance of concrete modified with high calcium fly ash using machine learning, Materials 8 (2015) 8714–8727. https://doi.org/10.3390/ma8125483.
[18] W.Z. Taffese, L. Espinosa-Leal, Prediction of chloride resistance level of concrete using machine learning for durability and service life assessment of building structures, Journal of Building Engineering 60 (2022). https://doi.org/10.1016/j.jobe.2022.105146.
[19] W. Liu, G. Liu, X. Zhu, Applicability of machine learning algorithms in predicting chloride diffusion in concrete: Modelling, evaluation, and feature analysis, Case Studies in Construction Materials 21 (2024). https://doi.org/10.1016/j.cscm.2024.e03573.
[20] W.Z. Taffese, L. Espinosa-Leal, Multitarget regression models for predicting compressive strength and chloride resistance of concrete, Journal of Building Engineering 72 (2023). https://doi.org/10.1016/j.jobe.2023.106523.
[21] L. Li, L. Su, B. Guo, R. Cai, X. Wang, T. Zhang, Prediction and prevention of concrete chloride penetration: machine learning and MICP techniques, Front Mater 11 (2024). https://doi.org/10.3389/fmats.2024.1445547.
[22] Q. feng Liu, M.F. Iqbal, J. Yang, X. yang Lu, P. Zhang, M. Rauf, Prediction of chloride diffusivity in concrete using artificial neural network: Modelling and performance evaluation, Constr Build Mater 268 (2021). https://doi.org/10.1016/j.conbuildmat.2020.121082.
[23] ASHTO, Standard method of test for resistance of concrete to chloride ion penetration - ASHTO T 259-02., 2021.
[24] ASTM, Standard Test Method for Determining the Penetration of Chloride Ion into Concrete by Ponding - ASTM C1543-02, 2009. www.astm.org.
[25] ASTM, Standard Test Method for Determining the Apparent Chloride Diffusion Coefficient of Cementitious Mixtures by Bulk Diffusion - ASTM C1556-04, 2009. www.astm.org,.
[26] Nordtest, Concrete, hardened: Accelerated chloride penetration- NT BUILD 443, (1995). www.nordtest.org.
[27] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artif Intell Rev 54 (2021) 1937–1967. https://doi.org/10.1007/s10462-020-09896-5.
[28] J.H. Friedman, Stochastic gradient boosting, 2002. www.elsevier.com/locate/csda.
[29] T. Su, H. Li, Y. An, A BIM and machine learning integration framework for automated property valuation, Journal of Building Engineering 44 (2021). https://doi.org/10.1016/j.jobe.2021.102636.
[30] M.R. Ahmad, A. Fernàndez-Jimenez, B. Chen, Z. Leng, J.G. Dai, Low-carbon cementitious materials: Scale-up potential, environmental impact and barriers, Constr Build Mater 455 (2024). https://doi.org/10.1016/j.conbuildmat.2024.139087.
[31] F. Althoey, W.S. Ansari, M. Sufian, A.F. Deifalla, Advancements in low-carbon concrete as a construction material for the sustainable built environment, Developments in the Built Environment 16 (2023). https://doi.org/10.1016/j.dibe.2023.100284.
[32] M. Schneider, The cement industry on the way to a low-carbon future, Cem Concr Res 124 (2019). https://doi.org/10.1016/j.cemconres.2019.105792.
[33] T.H. Vu, Y. Yang, L.C. Dang, V. Sirivivatnanon, Comparative analysis of chloride and acid resistance in one-part geopolymer and low-carbon concrete, Magazine of Concrete Research 0 (n.d.) 1–26. https://doi.org/10.1680/jmacr.24.00355.
[34] F.K.S. Junior, G.B. Wally, F.R. Teixeira, F.C. Magalhães, Experimental assessment of accelerated test methods for determining chloride diffusion coefficient in concrete, Revista IBRACON de Estruturas e Materiais 14 (2021). https://doi.org/10.1590/S1983-41952021000400007.
[35] R.W. Shiu, C.C. Yang, Evaluation of migration characteristics of opc and slag concrete from the rapid chloride migration test, Journal of Marine Science and Technology (Taiwan) 28 (2020) 69–79. https://doi.org/10.6119/JMST.202004_28(2).0001.
[36] M. Maes, E. Gruyaert, N. De Belie, Resistance of concrete with blast-furnace slag against chlorides, investigated by comparing chloride profiles after migration and diffusion, Materials and Structures/Materiaux et Constructions 46 (2013) 89–103. https://doi.org/10.1617/s11527-012-9885-3.
[37] Hou Hao-bo, ZHANG Guo-zhi, Assessment on Chloride Contaminated Resistance of Concrete with Non-steady-state Migration Method, Journal of Wuhan University of Technology Mater. Sci. Ed (2004).