A Non-destructive Method for the determination of Carbonation Time for Nominal Concrete Cover Depth Using Non-Linear Ensemble Prediction

Salim Idris Malami\textsuperscript{1,2}\textsuperscript{*}, Pinar Akpinar\textsuperscript{3}

\textsuperscript{1}Institute of Sustainable Built Environment, Heriot-Watt University, United Kingdom
\textsuperscript{2}Kano University of Science and Technology, Kano, Nigeria
\textsuperscript{3}Bahçeşehir Cyprus University, Nicosia, N. Cyprus, via Mersin 10, Turkey

Abstract. Carbonation, a process involving the reaction of carbon dioxide and moisture, results in the formation of powdery calcium carbonate, a critical durability issue causing reinforcement corrosion. The study analyzed carbonation data from coastal and inland buildings in the Turkish Republic of Northern Cyprus, revealing higher carbonation rates than anticipated within their lifespan. An artificial intelligence model named Support Vector Machine (SVM) was applied to predict carbonation time (T) to penetrate concrete cover of 25mm in the TRNC. Subsequently used two ensemble techniques, namely Neural Network Ensembles (NNE) and Support Vector Machine Ensembles (SVME) to enhance the performance of the prediction of T. Four performance criteria namely Correlation Coefficient (CC), Root Mean Square Error (RMSE), Correlation Coefficient ($R^2$), Mean Absolute Error (MAE) was applied to verify the modelling accuracy. The Values of $R^2$ of Ensemble techniques indicated significant increase in the performance, greater than the SVM model. This shows that using ensemble techniques is promising in getting better predictions of carbonation time (T) to penetrate concrete cover. The results obtained showed that NNE and SVME combination demonstrated the best performance under the evaluation criteria of $R^2 = 0.8721$ and $R^2 = 0.8644$ in testing phases respectively in comparison SVM-M1 to SVM-M3.

1 Introduction

Carbonation problems can significantly degrade the performance of steel-reinforced concrete buildings over time. Carbonation proceeds by Carbon dioxide (CO$_2$) in the atmosphere reacting with calcium hydroxide (Ca(OH)$_2$) to produce calcium carbonate (CaCO$_3$) \cite{25}. The reaction causes a change in the concrete microstructure, lowering the pH level in the concrete pores. Reinforcement corrosion then develops, finally resulting in significant structural damage. Carbonation rate is affected by the composition of materials, quality, strength, and ambient conditions (temperature, carbon dioxide gas (CO$_2$), and relative humidity). Due to a growth in industrial activity and population, the CO$_2$ content in the air has steadily been on the rise, and reinforced concrete structures are expected be subjected to more carbonation \cite{25}.

Ensuring high durability in concrete would imply that the building would serve satisfactorily throughout its service life and even beyond in some cases, without yielding construction and demolishing wastes that are generated from failed buildings that requires demolishing and being replaced. It is known that such construction and demolishing wastes yield in pollution of ground waters and has negative effects on soil quality and vegetation as well. Hence, ensuring elongated service life of reinforced concrete buildings would make positive contribution on the protection of the environment. Prediction of the carbonation progress in concrete, which has negative effects both on concrete and the steel bars in a building, is for sure the first step in protection of the building against this severe durability problem.

Artificial intelligence (AI) analyses human thinking and uses developed machine elements to transform that thinking into real-world circumstances. A fair degree of construction activities such as masonry work and concrete pouring, have been performed in the same way for decades now. As a result, new methodologies in construction and general civil engineering practice based on artificial intelligence and other modern techniques are required. Artificial intelligence applications have of recent demonstrated the capacity to participate critically in civil engineering and indeed other fields of technology application. More advanced uses of this technology can be found in civil engineering fields such as design of structures.

\*Corresponding author: sim2000@hw.ac.uk
and water systems, and quality control, construction project management, and risk management and construction materials' behaviour and performance \cite{10, 22, 24}. As such, it has increasingly been dawning on civil engineers that it is vital for them to improve their AI knowledge in order to open the industry more to further advances.

Artificial neural networks (ANN) have shown that they are valuable tools for measuring risk since they can understand collected datasets of information in construction. Neural networks have been applied for the prediction of the 28 day compressive strength of samples of ordinary concrete, high strength self-compacting concrete and other enhanced concretes such as concrete containing a high volumes of fly ash \cite{19}. In the same vein, the durability of self-compacting concrete with exposure conditions susceptible different sodium sulphate concentrations may be predicted using an AI technique (neuro-fuzzy systems). Such techniques could in addition be used to investigate carbonation induced damage in reinforced concrete \cite{10, 22}. Pattern recognition AI models play an important function today, having demonstrated convincingly that they can model and simulate random events and also non-linear hydro-meteorological time series. As a result, an AI model effectively controls the experimental time series linked to the problem using a data intelligent algorithm. Many studies using AI-models in construction projects and other and material engineering applications have shown that their predictive power can be successfully applied.

\cite{21}, for example, employed two ANN models, the radial basis function neural network (RBFNN) and the backpropagation neural network (BPN), to model and predict carbonation in pre-stressed concrete. They considered the effects of stress on the concrete carbonation process and used the accelerated carbonation experimental method. The authors discovered these ANN models could accurately predict and determine the depth of carbonation. \cite{32} developed the carbonation prediction model (CaPrM) to model and predict the progress of carbonation in reinforced concrete by making use of machine learning approaches. They combined four learning approaches and inputted 25 variables to create the model and track the carbonation process. The experimental data included 23 distinct types of concrete mixes. Integrating machine learning algorithms was found to enhance accuracy of carbonation prediction. \cite{5} investigated the feasibility of using ANNs as a non-destructive testing method of forecasting the depth of carbonation in reinforced concrete. To determine the output, 18 input factors were found. The results showed that the neural network model could accurately anticipate carbonation.

Moreover, \cite{29} created an ANN model that uses multilayer perception and a backpropagation learning method to analyze several indices of the carbonation process in concrete. The researchers looked into variables including the composition of concrete, admixtures used, $\text{CO}_2$ concentration, relative air humidity, and other mechanically related qualities. The results were encouraging, showing that ANN have a lot of potential for modelling the depth of carbonation. \cite{26} worked empirically to develop a model that used an automatic neural network search (ANS) to study the effects of different variables on carbonation depth. Their findings revealed that the depth of carbonation can be effectively regulated by using an adequate concrete mixture composition. As a result, ANS was able to confirm the experimentally acquired data with adequate precision. According to \cite{14}, ANN appears to be a very good technique for forecasting carbonation depth in reinforced concrete and indeed engineering project decision-making to determine durability. As a result, the authors constructed an ANN model so as to evaluate the carbonation depth in concrete over time in order to predict $\text{CO}_2$ dispersion. \cite{18} recently analyzed carbonation diffusion and examined samples of concrete cores collected from already built structures in Northern Cyprus' coastal and inland districts. The phenolphthalein indicator was used to examine 149 samples obtained from 39 buildings. They used Fick's Law to calculate the depth of carbonation at the conclusion of a service life of 50 years and the time for carbonation to engulf a concrete cover of 30mm. 71% of coastal samples were found satisfactory to the Eurocode performance requirements for carbonation, while just 41% of the inland samples were. \cite{4} used ANN to investigate the variables that affect the development of concrete carbonation depth. The aim of the paper was to enhance the accuracy of carbonation modelling utilizing an ANN design with 18 variables and an alternate scaled conjugate gradient (SCG). The parametric analysis yielded a promising outcome that was consistent with previous experience. Their ANN model could thus be utilised to investigate the degree of the carbonation in reinforced concrete. \cite{20} published a study on the progress of carbonation utilizing the model of the Architectural Institute of Japan (AIJ), machine learning methods, and finite element analysis. Their carbonation depth model for prediction was constructed from experimental data. It was determined that the deep learning model was the best predictor of the progress of carbonation.

The results obtained from this study will provide increased understanding on the efficiency of SVM for the prediction of carbonation problem in concrete. Hence, the correct method could be detected and could be improved according to the findings, in order to serve engineers to carry out investigation on this critical durability problem, enabling them to take precautions before severe damages occur in the buildings during their service lives.

While there are still notable limitations of ANN models, such as overfitting, the studied literature shows that artificial intelligence-based models, especially ANN, have captured the attention of the construction and material study disciplines. This study aims to use the potential support vector machine (SVM) computational model for accurate prediction of carbonation time (T) to penetrate concrete cover of 25mm in the Turkish Republic of Northern Cyprus (TRNC) existing buildings and also applied the ensemble techniques namely SVME and NNE to improve the model accuracy. However, to the best of our knowledge no literature available that presented SVME and NNE ensemble techniques in the prediction
of carbonation time (T) to penetrate concrete cover of 25mm in the existing buildings of TRNC. The results obtained from this study will provide increased understanding on the efficiency of SVM and ensemble techniques for the prediction of carbonation problem in concrete. Hence, the correct method could be detected and could be improved according to the findings, in order to serve engineers to carry out investigation on this critical durability problem, enabling them to take precautions before severe damages occur in the buildings during their service lives.

2. Materials and Methods
2.1 Experimental Procedures
Samples of concrete cores from 10 different reinforced concrete buildings from TRNC were analysed in this research: 5 samples were taken from Kyrenia (along the coast), while the remaining 5 were taken from Nicosia (inland). Details of the buildings used in the study such as the ages of the structures were taken from the owners. The ages of the buildings ranged from 10 up to 41 years. From each building, three cylinder-shaped core samples were obtained. The diameter of each cone was 65mm and the height, 70mm. The Turkish Chamber of Civil Engineering laboratory extracted samples in keeping with EN 12504-1 standards [31]. The core samples' compressive strength was assessed by making use of TS-EN 13791 techniques for evaluating "in-situ compressive strength in structures and precast concrete components" after which the mean of the values obtained from the samples was derived [12]. After the compression test, fractured samples were then examined using the traditional phenolphthalein indicator test to estimate carbonation depth development over the lifetime of the buildings. Although this test is routinely carried out to find the presence of carbonation in concrete core samples, no standard protocol is specified for it. Guidance manuals alone outline the method [25, 36]. One gram of the solution on phenolphthalein was first dissolved in 50 ml of ethanol. 100 ml of de-ionized water was then used to dilute the solution. A phenolphthalein indicator was then evenly sprayed on both halves of the interior surfaces of the shattered samples. When the pH if greater than 9.5, the indicator turns pink, otherwise it stays colourless. The pinkish colour indicates the presence of carbonation [8, 23, 25]. When Ca(OH)$_2$, Na(OH)$_2$, and K(OH)$_2$ are present in a concrete pore solution, the pH level is commonly between 13 and 14. It should be noted that even samples with lower pH values in the range of 10 and 12 can be affected by carbonation. For this reason, the phenolphthalein indicator test cannot give fully accurate results for the depth of carbonation [28]. Despite the fact that the real carbonation depth may be greater than that found by the phenolphthalein indicator test, this frequently used approach offers a good indication of the minimal depth of carbonation present.

![Fig. 1. Extracted cored samples from the structures](image)

3. Artificial Intelligence Models Proposed
The current study offers an SVM computational model for time (T) prediction in the TRNC, and then uses two ensemble techniques, namely Neural Network Ensembles (NNE) and Support Vector Machine Ensembles (SVME), to improve time prediction performance. Figure 2 shows the methodology followed in the research. The traditional feature extraction method is used for the simulation of T in accordance with the correlation coefficient relating the variables. Equation 1 shows the input selection and SVM combination based on sensitivity analysis.
Carbonation Time \((T)\) = \begin{cases} 
M1 = \Phi(Cd + B) \\
M2 = \Phi(Cd + B + \sigma) \\
M2 = \Phi(Cd + B + \sigma + A)
\end{cases} \tag{1}

Where:
\(M1\) to \(M3\) are the combination of the two models
\(\Phi\) is the function for each input variable
\(Cd\) is depth of carbonation
\(B\) refers to the carbonation constant
\(\sigma\) is the compressive strength
\(A\) is the age.

The data used in the research were split in two: 70\% for training and 30\% for testing. Model performance was validated by making use of the the k-fold cross validation method, which is one of the feasible techniques for achieving un-biased model’s prediction performance with limited sets of data.

\[X_i = \frac{x_u - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \] \tag{2}

where:
\(X_i\) is the normalized quantity,
\(x_u\) is the unnormalized quantity,
\(x_{\text{min}}\) is the minimum value of the data set
\(x_{\text{max}}\) is the maximum value of the data set

3.1 Support Vector Machine (SVM)

[35] created and implemented the Support Vector Machine. SVM provides solutions to problems involving classification, pattern recognition, regression and prediction. The SVM includes a model driven by data with the concept of a learning machine. The most important use case of SVM is the minimisation of structural risk and statistical learning theory. The qualities that differentiate the SVM from ANNs are a curtailment in error and complexity as well as an improvement in the performance ability of the network. The SVM kernel function can be divided into two: linear support vector regression (L-SVR) and non-linear support vector regression (N-SVR). It has been used in various technical domains, such as rainfall prediction, compressive strength prediction and prediction of flexural strength [3, 15, 24]. SVR is a layer-based SVM that includes kernel function being weighted on the input variable and a sum total of kernel outputs that is function-weighted. When employing the SVM, a linear regression line was first fitted onto the data, with the non-linear kernel catching a...
non-linear trend. The data set given is \( \{(x_i, d_i)\}_{i=1}^{N} \) (where \( x_i \) is the input vector, \( N \) refers to the number of patterns of data and \( d_i \) is the actual value). The SVM function is shown in Equation 3.

\[
y = f(x) = w \varphi(x_i) + b
\]  

(3)

Feature spaces indicated by \( \varphi(x_i) \), vector \( x \) mapped the non-linear. The regression parameters \( b \) and \( w \) can be obtained by “mapping positive values for the slack parameters of \( \xi^* \) and \( \xi \) as well as in Equation 4, which is the minimized objective function” [37].

\[
\text{Minimize: } \frac{1}{2} \|w\|^2 + C(\sum_{i} \xi_i + \xi_i^*)
\]  

(4)

Subject to: 

\[
\begin{align*}
& w_i \varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i^* \\
& d_i - w_i \varphi(x_i) + b_i \leq \varepsilon + \xi_i \\
& \xi_i, \xi_i^* \geq 0, \quad i=1,2,...,N
\end{align*}
\]  

(5)

Where:

\[
\frac{1}{2} \|w\|^2 \text{ refer to the vector norm weights}
\]

\( C \) refers to the regularized constant.

A generalised concept model structure of the SVM is shown in Fig.4. Lagrange multiplier parameters here are signified as \( \alpha_i \) and \( \alpha_i^* \). The Vector \( w \) in Eq. (4) is determined sequel to the solving of the optimisation problem.

\[
w^* = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \varphi(x_i)
\]  

(6)

It follows that the complete form of the SVM modifies Eq. 5.

\[
f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) k(x, x_i) + b
\]  

(7)

Where

\( k(x, x_i) \) represents kernel function

\( b \) represents bias term.

The Gaussian Radial Basis Function is the kernel function that is used by most researchers. Its form is:

\[
k(x, x_2) = \exp(-\gamma \|x_1 - x_2\|^2)
\]  

(8)

### 3.2 Ensemble Techniques

Ensemble learning is a sort of machine learning, which is a procedure that combines numerous predictors to improve the performance of a model. It consists of a trained base classifier, the decisions of which are merged to produce fresh outcomes [7]. Engineering, computer science and medicine are just some of the diverse applications of ensemble learning. It enhances the accuracy of a singular model by putting together the results of different individual models. This process of mixing several different individual models is ensemble modeling. Data heterogeneity and features are another rationale for using this technique. Any model's performance accuracy is influenced by data variations such as normalcy, size, and linearity. Furthermore, multiple studies have demonstrated that performance outcomes for different models might vary even for the same sort of data set, implying that the ensemble approach should be used. Ensembles can be classified into simple, weighted, neural ensembles, and SVM ensembles [11,30]. Both studies cited indicated that artificial neural ensembles perform best with regards to accuracy owing to their nonlinear nature. For this reason, this paper uses a nonlinear neural ensemble (NNE) and support vector machine ensembles (SVME) to enhance the accuracy of the predictions of single models.

### 3.3 Evaluation criteria

After the models development (SVM, NNE and SVME), four statistical indicators were used for the prediction accuracies; correlation coefficient (R), Determinacy coefficient (DC), mean absolute error (MAE) and root mean square error (RMSE) are obtained using the following:
I. Correlation Coefficient

\[ CC = \frac{\sum_{i=1}^{N}(T_{\text{com},i} - \bar{T}_{\text{com}})(T_{\text{pre},i} - \bar{T}_{\text{pre}})}{\left(\sum_{i=1}^{N}(T_{\text{com},i} - \bar{T}_{\text{com}})^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{N}(T_{\text{pre},i} - \bar{T}_{\text{pre}})^2\right)^{\frac{1}{2}}} \]  

(9)

II. Determinacy coefficient

\[ DC = 1 - \frac{\sum_{i=1}^{N}(T_{\text{com},i} - T_{\text{pre},i})^2}{\sum_{i=1}^{N}(T_{\text{com},i} - \bar{T}_{\text{com}})^2} \]  

(10)

III. Mean absolute error

\[ MAE = \frac{1}{N} \sum_{i=1}^{N} |T_{\text{pre},i} - T_{\text{com},i}| \]  

(11)

IV. Root mean square error

\[ RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (T_{\text{com},i} - T_{\text{pre},i})^2} \]  

(12)

Where:

T is the carbonation time for 25mm concrete cover

\( T_{\text{pre},i} \) is the predicted time

\( T_{\text{com},i} \) is the computed time

\( \bar{T}_{\text{pre}} \) is the mean predicted time

\( \bar{T}_{\text{com}} \) is the mean computed time

N refers to the number of data points.

Desired results are high DC and DC as well as low values of MAE and RMSE. The models will be ranked in order of their performance under these criteria. Models having the greatest values of DC and CC as well as the least values of MAE and RMSE are chosen for better T prediction in the research area.

4. Results and Discussion

4.1 Depth of carbonation and compressive strength

The averages of both compressive strength and depth of carbonation of the core samples was presented in Figure 3. The concrete core samples used in this study have different exposure conditions and ages.

![Fig. 3. The carbonation time and carbonation depth of the structures](image)

Constant values for carbonation wire calculated as “B” as shown in table 1 using current carbonation depth and current ages. Using these “B” values, time for the carbonation to reach 25mm nominal cover depth were computed. The predicted times for carbonation to get to 25mm nominal cover depth are shown in fig. 3 and table 1. From the fig. 3 it is observed that half of the structures (1, 3, 8, 9 and 10) already carbonated more than 25mm at the age between 7.4 years and 33.8 years before reaching 50 years of construction (service life). From the predictions, it can be seen that structure 4 and 6 have critical carbonation depths of 25mm at the age of 1.9 and 4 respectively. This is showing that the structures in Cyprus Island are facing potential damages due to carbonation problems.
<table>
<thead>
<tr>
<th>Structure No./ Location</th>
<th>Age (years)</th>
<th>Compressive strength $\sigma$(MPa)</th>
<th>Current Carbonation depth (mm)</th>
<th>Carbonation Constant $B$(mm/yr$^{0.5}$)</th>
<th>Expected carbonation time for 25mm nominal cover $T$(years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. 1 (Inland)</td>
<td>18.00</td>
<td>18.93</td>
<td>39.00</td>
<td>0.92</td>
<td>7.40</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>19.30</td>
<td>38.96</td>
<td>0.92</td>
<td>7.41</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>18.56</td>
<td>39.04</td>
<td>0.92</td>
<td>7.38</td>
</tr>
<tr>
<td>No. 2 (Inland)</td>
<td>31.00</td>
<td>14.63</td>
<td>5.00</td>
<td>0.09</td>
<td>775.00</td>
</tr>
<tr>
<td></td>
<td>31.00</td>
<td>14.75</td>
<td>10.50</td>
<td>0.19</td>
<td>175.74</td>
</tr>
<tr>
<td></td>
<td>31.00</td>
<td>14.36</td>
<td>7.76</td>
<td>0.14</td>
<td>321.75</td>
</tr>
<tr>
<td>No. 3 (Inland)</td>
<td>18.00</td>
<td>22.65</td>
<td>47.00</td>
<td>1.11</td>
<td>5.09</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>40.40</td>
<td>24.30</td>
<td>0.57</td>
<td>19.05</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>24.29</td>
<td>27.00</td>
<td>0.64</td>
<td>15.43</td>
</tr>
<tr>
<td>No. 4 (Inland)</td>
<td>18.00</td>
<td>22.65</td>
<td>24.00</td>
<td>0.57</td>
<td>19.53</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>34.79</td>
<td>23.70</td>
<td>0.56</td>
<td>20.03</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>28.72</td>
<td>23.84</td>
<td>0.56</td>
<td>19.79</td>
</tr>
<tr>
<td>No. 5 (Inland)</td>
<td>10.00</td>
<td>21.08</td>
<td>26.20</td>
<td>0.83</td>
<td>9.10</td>
</tr>
<tr>
<td></td>
<td>10.00</td>
<td>36.47</td>
<td>4.00</td>
<td>0.13</td>
<td>390.63</td>
</tr>
<tr>
<td></td>
<td>10.00</td>
<td>27.01</td>
<td>24.16</td>
<td>0.76</td>
<td>10.71</td>
</tr>
<tr>
<td>No. 6 (Coastal)</td>
<td>21.00</td>
<td>33.95</td>
<td>24.67</td>
<td>0.54</td>
<td>21.57</td>
</tr>
<tr>
<td></td>
<td>21.00</td>
<td>17.49</td>
<td>10.67</td>
<td>0.23</td>
<td>115.28</td>
</tr>
<tr>
<td></td>
<td>21.00</td>
<td>19.08</td>
<td>32.83</td>
<td>0.72</td>
<td>12.18</td>
</tr>
<tr>
<td>No. 7 (Coastal)</td>
<td>10.00</td>
<td>26.92</td>
<td>5.20</td>
<td>0.16</td>
<td>231.14</td>
</tr>
<tr>
<td></td>
<td>10.00</td>
<td>20.21</td>
<td>19.99</td>
<td>0.63</td>
<td>15.64</td>
</tr>
<tr>
<td></td>
<td>10.00</td>
<td>31.92</td>
<td>5.80</td>
<td>0.18</td>
<td>185.79</td>
</tr>
<tr>
<td>No. 8 (Coastal)</td>
<td>40.00</td>
<td>17.62</td>
<td>32.50</td>
<td>0.51</td>
<td>23.67</td>
</tr>
<tr>
<td></td>
<td>40.00</td>
<td>21.66</td>
<td>30.34</td>
<td>0.48</td>
<td>27.16</td>
</tr>
<tr>
<td></td>
<td>40.00</td>
<td>25.40</td>
<td>18.33</td>
<td>0.29</td>
<td>74.41</td>
</tr>
<tr>
<td>No. 9 (Coastal)</td>
<td>41.00</td>
<td>26.69</td>
<td>28.84</td>
<td>0.45</td>
<td>30.81</td>
</tr>
<tr>
<td></td>
<td>41.00</td>
<td>42.79</td>
<td>42.17</td>
<td>0.66</td>
<td>14.41</td>
</tr>
<tr>
<td></td>
<td>41.00</td>
<td>23.07</td>
<td>20.17</td>
<td>0.32</td>
<td>62.99</td>
</tr>
<tr>
<td>No. 10 (Coastal)</td>
<td>41.00</td>
<td>26.08</td>
<td>44.67</td>
<td>0.70</td>
<td>12.84</td>
</tr>
<tr>
<td></td>
<td>41.00</td>
<td>15.12</td>
<td>35.57</td>
<td>0.56</td>
<td>20.25</td>
</tr>
<tr>
<td></td>
<td>41.00</td>
<td>26.45</td>
<td>31.34</td>
<td>0.49</td>
<td>26.09</td>
</tr>
</tbody>
</table>
4.2 Results of machine learning models

This part involved the application data from laboratory analysis of extracted concrete core samples obtained from concrete buildings in North Cyprus, in which some from Kyrenia (coastal area) and other from (Nicosia) inland area. In this data driving models (SVM, NNE and SVME), the data input for the experiment consists of the following: Age (A) (Year), carbonation depth (Cd) (mm), carbonation constant (B) (mm/yr0.5) and Compressive strength (σ) (MPa). Time to carbonate 25mm concrete cover (T) (Years) is the output variable. This is presented in Table 1. Prior to the AI modelling, data portioning and portioning of the targeted and input variables were conducted. The normalization was used in creating common scale and decreasing the redundancy of the data set, by this, the integrity of the data was increased. The basic statistical information of the data set was presented in Table 2. This basic statistical information were utilized in many AI based researches [1, 17, 27, 33]. Looking in to table 2, we can say that the data from the experiments is reliable in AI modelling and analysis due to its low skewness. Data used for AI was partitioned into two: training and testing data with an external validation due to the limited data set as reported by [6, 9, 33]. Most dominant and suitable input variables combination were determined by sensitivity analysis using correlation matrix as in Fig. 4. The matrix represents the linear relationship between all variables.

Table 2. Statistical parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>A(years)</th>
<th>σ(MPa)</th>
<th>Cd (mm)</th>
<th>B(mm)</th>
<th>T (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>24.80</td>
<td>24.43</td>
<td>24.92</td>
<td>0.53</td>
<td>89.28</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>11.97</td>
<td>7.59</td>
<td>12.44</td>
<td>0.27</td>
<td>162.52</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-1.51</td>
<td>0.07</td>
<td>-0.81</td>
<td>-0.64</td>
<td>10.86</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.31</td>
<td>0.79</td>
<td>-0.16</td>
<td>0.10</td>
<td>3.08</td>
</tr>
<tr>
<td>Minimum</td>
<td>10.00</td>
<td>14.36</td>
<td>4.00</td>
<td>0.09</td>
<td>5.09</td>
</tr>
<tr>
<td>Maximum</td>
<td>41.00</td>
<td>42.79</td>
<td>47.00</td>
<td>1.11</td>
<td>775.00</td>
</tr>
</tbody>
</table>

Model accuracy was analysed using the statistical indexes (MSE, R2, CC and RMSE), which included both goodness of the fit and error criteria. Considering Table 3, it can be seen that the results are showing that all the criteria were obtained by model combinations M2 and M3 except combination M1. With regards to R2, a radar diagram was used. This is an
especially important step to compare the overall performance of the models. This is shown clearly on Fig. 5 encompassing the performance of the SVM-M1, SVM-M2, SVM-M3, NNE and SVM models. The NNE model has the best prediction performance of $T$ of all the models, yielding $R^2 = 0.5637$ in training and 0.8721 in testing as shown on Table 3. Quantifiably, NNE improved the prediction accuracy of SVM by 0.88%, SVM-M3 by 8.49%, SVM-M2 by 14.31% and SVM-M1 by 57.57% respectively. The general better fitness of the best-performing models is indicative of their capacities to handle non-linearity between the $T$ variables. For the NNE and SVME, the values of $R^2$ showed marginal increase in the performance than the other models. This shows that NNE and SVME techniques are promising in the increase in the prediction performance of carbonation time $T$. It is worth reiterating here how reliable ensemble AI models have proven to be not only in engineering, but in science in general [20, 38, 39].

![Fig. 5. Radar plots for $R^2$ of (a) SVM (b) all models (c) SVM-M3, NNE and SVME](image1)

![Fig. 6. Scatter plots for all the models](image2)
Table 3: Results of data-driven models

<table>
<thead>
<tr>
<th>Models</th>
<th>Training</th>
<th>R²</th>
<th>MSE</th>
<th>R</th>
<th>RMSE</th>
<th>R²</th>
<th>MSE</th>
<th>R</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-M1</td>
<td>0.2535</td>
<td>0.0432</td>
<td>0.5034</td>
<td>0.2079</td>
<td>0.3700</td>
<td>0.0004</td>
<td>0.6082</td>
<td>0.0209</td>
<td></td>
</tr>
<tr>
<td>SVM-M2</td>
<td>0.5786</td>
<td>0.0244</td>
<td>0.7607</td>
<td>0.1562</td>
<td>0.7473</td>
<td>0.0002</td>
<td>0.8645</td>
<td>0.0132</td>
<td></td>
</tr>
<tr>
<td>SVM-M3</td>
<td>0.5548</td>
<td>0.0258</td>
<td>0.7449</td>
<td>0.1606</td>
<td>0.7980</td>
<td>0.0001</td>
<td>0.9476</td>
<td>0.0084</td>
<td></td>
</tr>
<tr>
<td>NNE</td>
<td>0.5637</td>
<td>0.0253</td>
<td>0.7508</td>
<td>0.1590</td>
<td>0.8721</td>
<td>0.0001</td>
<td>0.9338</td>
<td>0.0094</td>
<td></td>
</tr>
<tr>
<td>SVME</td>
<td>0.5594</td>
<td>0.0255</td>
<td>0.7479</td>
<td>0.1597</td>
<td>0.8644</td>
<td>0.0001</td>
<td>0.9297</td>
<td>0.0097</td>
<td></td>
</tr>
</tbody>
</table>

It is common practice for time series plots to be used in civil and material engineering as in [2,34]. Another importance of time series is that it helps in understanding data set and in interpreting the exact meaning of a data set [7,13,16]. Further investigation of the predictive accuracy of the models can be seen visually from a time series plot as depicted on Fig. 7. It should note be that NNE and SVME displayed a good correspondence with similar patterns relating to the observed values of T.

In general, the conditions of AI testing and the nature of experimental data employed in this study and in other studies within the literature are not directly comparable, hence direct comparison of the accuracies would not be possible. However, the accuracies obtained in the other studies are presented here to provide a general view of the current status of the studies. [23] provided and reported the ability of SVM, MLR, ELM and ANFIS models in the prediction of carbonation depth. [21] reported that, ANN models could accurately predict and determine the depth of carbonation. Similarly, [4] reported that ANN model has the capability of investigating the variables that affect the development of concrete carbonation depth.

5. Conclusion
Considering both inland and coastal buildings in North Cyprus, the buildings have a greater carbonation rate than that ordinarily expected within their life span.
The results showed all the AI models (SVM, NNE and SVME) performing reasonably well at the training stage and with R²-values greater than 0.7 at the testing stage (except SVM-M1 with R² of 0.37). The combination of NNE and SVME displayed the best performance under the evaluation criteria of R², with R² = 0.8721 and R² = 0.8644 in testing phases respectively in comparison to models SVM-M1 to SVM-M3. The results indicated that NNE increased the prediction accuracy of SVME by 0.88%, SVM-M3 by 8.49%, SVM-M2 by 14.31% and SVM-M1 by 57.57% respectively. It can be concluded that the results of the proposed NNE and SVME models using $C_0$, $B$, $\sigma$ and $A$ as input parameters demonstrate good performance in predicting the carbonation time $T$ with R² greater than 7.0. This has shown that the proposed techniques are promising in the increase in the prediction performance of carbonation time $T$. The results are also amenable to the suggestion that other models such as hybrid models, emerging algorithms and optimization methods could be further explored to enhance model performance. There are key factors to consider in future studies to improve the ensemble models presented in this paper. Larger datasets and more input variables should be used. Exposure conditions and curing conditions are some of the input variables that could be used.

Acknowledgements
- The Civil Engineering Laboratory of Near East University, Cyprus and the Laboratory of Turkish Cypriot Chamber of Civil Engineers gave copious support throughout this study. The authors acknowledge the support and are grateful.
- Kano State Government Scholarship Bord for their support.

Conflict of Interest
The authors declared no conflict of interest in the manuscript.

References