Optimized neural network based carbonation prediction model

Woubishet Z. Taffese 1, Fahim Al-Neshawy 1, Esko Sistonen1 and Miguel Ferreira2

1 Department of Civil and Structural Engineering, Aalto University, Espoo, Finland: Phone: +358947001; Fax: +358947023826; e-mail: woubishet.taffese@aalto.fi, fahim.alneshawy@aalto.fi, esko.sistonen@aalto.fi.
2 VTT Technical Research Centre of Finland, Espoo, Finland Phone: +358401380546; e-mail: miguel.ferreira@vtt.fi.

Abstract

Concrete carbonation is one of the major causes of steel corrosion in reinforced concrete structure that can lead to shortened service life. Several carbonation prediction models including mathematical and neural network are available. The mathematical models are simplified and do not take all influential parameters of concrete materials into consideration. Most of the existing neural network based carbonation prediction models do not employ all parameters that influence the microstructural properties of the concrete. They also failed to perform certain essential steps during the model development, which in turn degrade their performance. In this work, novel neural network based carbonation prediction model is proposed. The model selects most relevant parameters, and removes irrelevant and/or redundant features from the original input data to build robust learning models. The performance evaluation of the model shows that the proposed carbonation prediction model predicts reasonably well with increased generalization ability.

Keywords: concrete carbonation, carbonation prediction, modelling, neural network

1. Introduction

Carbonation of concrete is one of the major causes of deterioration in reinforced concrete structures [1, 2]. It is a natural physicochemical process caused by the penetration of carbon dioxide from the surrounding environment into the concrete through pores in the matrix where the carbon dioxide reacts with hydrated cement and is expressed by Eq. (1). Calcium hydroxide (Ca(OH)2) in contact with carbon dioxide (CO2) forms calcium carbonate (CaCO3). This reaction neutralizes the natural protection of reinforcement steel provided by the concrete, which in turn corrosion of the steel bars will be initiated [1, 3-5]. It is a recurrent problem since about two-thirds of all structural concrete is exposed to environmental conditions that favour carbonation-induced corrosion [6, 7].

\[ Ca\,(OH)_2 + CO_2 + H_2O \rightarrow CaCO_3 + H_2O \]  

(1)

Conventionally, concrete carbonation depth at a given time in steady state conditions can reasonably be estimated using Eq. (2). This equation is based on Fick’s second law of diffusion and it is well known [1, 3].

\[ x_c(t) = k \sqrt{t} \]  

(2)

where \( x_c(t) \) is carbonation depth at the time \( t \) in [mm], \( k \) is coefficient of carbonation [mm/day^{0.5}] and \( t \) is the duration of carbonation [day].

It is generally known that coefficient of carbonation is a decisive factor in determining carbonation depth. It is mainly controlled by diffusion of CO2 into the concrete pore system. CO2 diffusion through concrete depends on several factors such as CO2 concentration, environmental condition, and concrete characteristics. Therefore, carbonation coefficient may significantly vary from one concrete element to another depending on environment and...
microstructural parameters which are linked with concrete composition and type of materials used. Coefficient of carbonation is analysed either by an accelerated carbonation test or by measuring the development of the carbonation depth from an existing concrete structure. Since carbonation is a slow process, it is usually investigated by performing accelerated test with a higher CO$_2$ concentration in a controlled environment at a defined age [8]. Then, the measured carbonation depth is used to calculate the equivalent carbonation coefficient using Eq. (2).

Developing an advanced carbonation prediction model which addresses all the governing parameters is a difficult task because concrete property is depends on many parameters that are complex to describe mathematically. In addition, it is not feasible to combine two or more mathematical models in order to simulate the carbonation process. These problems calls for the use of machine learning based models which learn the complex behaviour of carbonation of concrete from experimental data. Among several machine learning techniques, neural network is commonly used for carbonation prediction, e.g. [9-11]. However, most of the existing neural network based carbonation prediction models do not employ all the essential parameters that influence the microstructural properties of concrete. The common parameters utilized in most of the available models are cement content and water to cement ratio (w/c) to describe the concrete properties. These parameters are not sufficient to predict the concrete carbonation since it is a function of many parameters and ignoring other influential variables weakens the prediction performance of the model. An accurate prediction model can be developed if all influential parameters evaluated as a group rather than individually because imperative dependencies may be overlooked. Indeed, there are few neural network based models which consider most of the parameters but lacked to perform certain essential variable selection task during the model development. It is also known that incorporating more information regarding the concrete properties enhances the learning performance of the model. Nonetheless, some parameters or variables may be irrelevant and/or redundant for the carbonation prediction and even degrade the learning performance. Due to this, optimized neural network model which consider variable selection technique during the model development phase is essential.

In this work, optimized neural network based carbonation prediction model is developed using experimental data. The model selects most relevant parameters, and removes irrelevant and/or redundant variables from the original input data to build robust learning models.

2. **Neural network**

A neural network is a computational network which consists of partially or fully interconnected simple processing units called artificial neurons [12]. They have a capability of capturing nonlinear and complex underlying characteristics of any physical process with a high degree of accuracy. Neural network can be categorized in a number of different ways depending on their architecture which is intimately linked with the learning algorithm used to train the network. Multilayer feedforward architecture and backpropagation training procedure is broadly used for nonlinear regression problems [12, 13] and due to this it is suitable for modelling carbonation process. This architecture usually has three or more layers. The first and the last layers are input and output layers whereas the intermediate layers (hidden layers) which aids in performing useful computations before directing the input into the output layer. The architectural graph in Figure 1 illustrates the layout of a multilayer feedforward neural network with a single hidden layer. The network can be viewed as a flexible nonlinear parametric function from a set of inputs, $x_i$, to a set of outputs, $y_m$. First linear combinations of the weighted inputs are formed, and then transformed using a
nonlinear activation function $\phi(.)$ [13], Eq. (3). The intermediate variables are then linearly combined to produce the outputs as in Eq. (4). The outputs of these first-layer neurons are multiplied by the layer of the interconnection weights that connect them to the next layer of neurons. This process continues until the output nodes compute their outputs. For neurons in the same layer, the same activation functions are used. Various forms of activation functions can be defined depending on the characteristics of applications.

$$z_j = \phi\left(\sum_i w_{ji}^{(1)} x_i\right)$$  \hspace{1cm} (3)$$

$$y_m = \sum_j w_{mj}^{(2)} z_j$$  \hspace{1cm} (4)$$

where $w_{ji}^{(1)}$ and $w_{mj}^{(2)}$ are the synaptic weights of the network of the network which are initially set to random values, and then adjusted during training by back propagation using the response data.

Neural network is one of the commonly used machine learning based methods in concrete research. To mention few, prediction of: chloride penetration in concrete [14], chloride threshold of pitting corrosion [15], hygrothermal forecasting in in thick-walled concrete [16] and corrosion currents of reinforced concrete [17].

3. Neural network modelling process

In this section, the development steps of the proposed carbonation prediction model are presented. As any machine learning based model, the major development phases consist of data, training, validation and testing. The workflow of this carbonation prediction model is illustrated in Figure 2. The red rectangular boxes represent the main processes or tasks of the modelling procedure. The major inputs into the process are shown in blue rectangular boxes. The experimental data, training, validation, and testing phases of the model are discussed in the following subsections.
3.1 Experimental data

Experimental data used in this work for the development of carbonation prediction model were prepared for Finnish Duralnt-project. This project was carried out in cooperation with Aalto University and VTT Technical Research Centre of Finland. In Finnish Duralnt-project concrete specimen from 23 different mix proportion, which represents mainly prevailing common industrial mixes in Finland, were prepared. The data consists of concrete mixture ingredients and fresh and hardened properties of 46 specimens. Accelerated carbonation tests for half of the concrete specimens are performed by exposing them in a climatic control chamber for 28 and the remaining half for 56 days. These tests were performed by applying CO\textsubscript{2} of 1\% in a controlled environment (temperature 21°C and relative humidity 60\%) in accordance with EN 13295. Carbonation front depth readings, from all sides, in a freshly broken surface of 100 x 100 mm\textsuperscript{2} were made. The depth was determined by spraying a pH indicator solution of phenolphthalein. The arithmetic mean of the four sides carbonation depth measurements of every specimen was considered as the representative value. Carbonation front of two groups of concrete specimens after 56 days in the accelerated carbonation chamber is illustrated in Figure 3. Surface areas with a pink colour indicate the pH is above 9 and are non-carbonated part. The carbonated parts of the specimens are the area where the colour of the concrete is unchanged.

A total of 15 variables, 14 describes the specimen’s mix proportion and one represent carbonation period, are selected for the modelling purpose. List of all variables are: cement type, water to binding ratio (w/b), cement, blast-furnace slag (BFS), fly ash (FA), total effective water, total aggregate, aggregate < 0.125mm, aggregate < 0.25mm, aggregate < 4mm, product name of plasticizer, plasticizer, product name of air-entraining agent, air-entraining agent and carbonation period. A total of six types of cements, according to the
classification of EN 197-1, are utilized. These are portland cement (CEM I 42,5 N-SR, CEM I 52,5 N and CEM I 52,5 R), portland limestone cement (CEM II/A-LL 42,5 R), portland composite cement (CEM II/A-M(S-LL) 42,5 N) and portland slag cement (CEM II/B-S 42,5 N). Partially replaced portland limestone cement, CEM II/A-LL 42,5 R, with pulverized blast-furnace slag and fly ash are also used in the experiment. The water to binder ratios (w/b) ranges from 0.40 to 0.60. Each concrete mix employs one plasticizer type from three producers, VB-Parmix, Glenium G 51 or Teho-Parmix. Every concrete mix types also employ an air-entraining agent (named either Ilma-Parmix or Mischöl).

Figure 3. Carbonation fronts of two groups of concrete specimens after exposure of 56 days

3.1.1 Data types

A total of 15 variables which represent the concrete mixture ingredients and the carbonation period are assigned as inputs. The target variable is the accelerated carbonation depths which are measured after exposing the non-carbonated concrete specimens in a climatic control chamber for 28 and 56 days. The input variables consist of continuous and nominal data types. Continuous variables are real numbers, such as results of quantitative measurements (e.g. cement content and total aggregate). Nominal variables are non-numeric and descriptive data types (e.g. types of cement, product name of plasticizers and air-entraining agents).

3.1.2 Data encoding

Neural network do not have the ability to work with nominal variables. An exclusive binary representation of all the non-numeric variables such as binder types, product name of plasticizers and air-entraining is necessary by assigning a single neural unit to each value to be represented. This method is often known as 1-of-n encoding and it is the most widely used approach. To represent a given value, setting of all the units’ activations to zero is a must except for the single unit which represents the value to be encoded. This unit has its activation set to one. All the encoded non-numeric variables are listed in Table 1.
Table 1. 1-of-N encoding process for non-numeric variables

<table>
<thead>
<tr>
<th>Binder materials</th>
<th>Encoded output</th>
<th>Product names of plasticizers and air-entraining agents</th>
<th>Nominal input variables</th>
<th>Encoded output</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEM I 42.5 N – SR</td>
<td>&quot;100000000&quot;</td>
<td>Plasticizers</td>
<td>Glenium G 51</td>
<td>&quot;100&quot;</td>
</tr>
<tr>
<td>CEM I 52.5 N</td>
<td>&quot;010000000&quot;</td>
<td></td>
<td>Teho-Parmix</td>
<td>&quot;010&quot;</td>
</tr>
<tr>
<td>CEM I 52.5 R</td>
<td>&quot;001000000&quot;</td>
<td></td>
<td>VB-Parmix</td>
<td>&quot;001&quot;</td>
</tr>
<tr>
<td>CEM II/A-LL 42.5 R</td>
<td>&quot;000100000&quot;</td>
<td></td>
<td>Ilma-Parmix</td>
<td>&quot;10&quot;</td>
</tr>
<tr>
<td>CEM II/A-M(S-LL) 42.5 N</td>
<td>&quot;000010000&quot;</td>
<td></td>
<td>Mischöl</td>
<td>&quot;01&quot;</td>
</tr>
<tr>
<td>CEM II/B-S 42.5 N</td>
<td>&quot;000001000&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEM II/A-LL 42.5 R &amp; BFS KJ400</td>
<td>&quot;000000100&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEM II/A-LL 42.5 R &amp; FA - Fineness N, Class A</td>
<td>&quot;00000001&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.1.3 Data normalization

Normalization of the inputs and the target variables is a standard practice before processing them in a neural network. It puts different variables on a common scale and is very important especially where the inputs are generally on a wide different scale. Since Matlab neural network toolbox is used for modelling, it automatically normalizes both the input and target variables. In this way, the network output always falls into a normalized range [-1, 1]. The network output can then be reverse transformed back into the units of the original target data. So, there is no need for a separate normalization process. The normalization algorithm carried out by Matlab is described in Eq. (5).

\[ y = \left( y_{max} - y_{min} \right) \cdot \frac{x - x_{min}}{x_{max} - x_{min}} + y_{min} \]  

(5)

where: \( y \) is the network output, \( y_{max} \) is the maximum value of the network output interval, (+1), \( y_{min} \) the minimum value of the network output interval, (-1), \( x \) is the original inputs or target variables, \( x_{max} \) is the maximum value for variable \( x \), and \( x_{min} \) is the minimum value for variable \( x \).

3.2 Training

The graphical representation of the proposed neural network model to predict carbonation is principally the same with Figure 1. The model has three layers: an input layer, a hidden layer and an output layer. All the variables, except the accelerated carbonation depth, are assigned as input neurons. The training dataset holds 75% of the dataset. The optimal number of neurons in the hidden layer is determined based on the generalization error after executing several trainings. The measured accelerated carbonation depth is assigned as output neuron. The applied learning algorithm is Levenberg-Marquardt. It is the fastest backpropagation procedure that updates weight and bias values in the negative gradient direction. The activation functions selected for the hidden layer is tan-sigmoid transfer function so that the model can learn the nonlinear relationships between input and target variables. Linear transfer activation function is applied in the output layer.
3.3 Validation and testing

The validation and testing represent 15% and 10% of the total dataset, respectively. Validation dataset is used to measure the network generalization, and to halt training when the generalization stops improving. Test dataset are used to measure network performance during and after training. The following three statistical measures are used to evaluate the performance of the developed carbonation prediction model.

The mean absolute error (MAE), also called the absolute loss, is an average of the absolute residuals/errors (the difference between the predicted and the actual value) and measured in the same units as the data. It is mathematically expressed by Eq. (6).

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |Y_i - \hat{Y}_i|
\]  

(6)

The mean square error (MSE) is the mean of the squared difference between the target and its predicted value. It is the most widely employed loss function and described in Eq. (7).

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2
\]  

(7)

The root mean square error (RMSE) is simply the square root of the MSE, Eq. (8).

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}
\]  

(8)

where \(\hat{Y}_i\) is the predicted output value, \(Y_i\) is the measured target value, and \(N\) is the number of observations.

The RMSE is sometimes preferable than the MSE because understanding error values of the MSE is difficult due to a squaring effect, particularly, if the target value represents quantities in units of measurements. RMSE retains original units as also MAE. All these indicators measure the spread between the monitored and the predicted output from the network.

3.4 Variable selection

Variable selection is a technique to select most relevant variables subset, and removes irrelevant and/or redundant variables according to some criterion from the original input variables to build robust learning models [18]. Indeed, having more information about the concrete properties, the performance of the model is expected to be better. It is also equally true that some variables may be irrelevant for carbonation prediction which ultimately degrades the learning performance of the model. Therefore, pertinent variables have to be selected after the initial validation test is carried out.

Many feature selection methods have been proposed in the literature. In the proposed model, the most common feature selection method called sequential feature selection is applied. This method has two components: a ranking criterion and a sequential search algorithm [19]. The criterion is formulated as an objective function that finds/ranks input variables which are most appropriate for carbonation prediction. The later component adds or removes features from the input variables while evaluating the criterion. Mean squared error is set as criterion in the adopted feature selection method. Frequentlly selected input varaibles by sequential feature selection method is presented in Table 2. It can be seen that some of the parameters such as plasticizers, air and aggregate contents are identified as influential carbonation predictors for this specific dataset. These parameters were overlooked in many existing analytical models.
Table 2. Influential input variables identified by sequential feature selection method

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>No.</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Water to binder ratio (w/b)</td>
<td>5</td>
<td>Cement content</td>
</tr>
<tr>
<td>2</td>
<td>Plasticizer content</td>
<td>6</td>
<td>Total aggregate</td>
</tr>
<tr>
<td>3</td>
<td>Carbonation period</td>
<td>7</td>
<td>Aggregate &lt; 0.250mm</td>
</tr>
<tr>
<td>4</td>
<td>Air content</td>
<td>8</td>
<td>Aggregate &lt; 0.125mm</td>
</tr>
</tbody>
</table>

4. Results and discussions

The training, validation and testing performance of the developed carbonation prediction model is presented in this section. The performance of the developed model using the whole dataset is shown as a regression plot in Figure 4a. The correlation coefficient (R-values) between measured and predicted carbonation depth is 0.98269. Such a high correlation coefficient indicates that the developed model predicts the carbonation depths well during the training phases. The measured and predicted carbonation depth with the predicted error is illustrated in Figure 4b. It can be clearly observed that the prediction error, the difference between the predicted and the measured value, is small with mean absolute error of 0.31076.

![Regression plot](image)

Figure 4. Predicted vs measured carbonation depth a) regression plot b) scatter plot

The prediction accuracy of the developed carbonation prediction model is analysed statistically using the test dataset. These dataset are unseen previously during the training phase. The average of five test results of the model before and after incorporating variable selection method are presented in Table 3. The MAE, MSE and RMSE of the model after considering variable selection method for the test dataset are 0.41421, 0.72356 and 0.85062, respectively. The lower the value of the error statistics (MSE, RMSE, and MAE) is the better the prediction accuracy of the model. As it is clearly seen in Table 3, all the statistical measures are small for the model which incorporates variable selection method compared to the model which does not include variable selection. This proves that the employed variable selection enhances the prediction performance of the carbonation prediction model considerably. Generally, from all the error metrics presented in the same table, it can be concluded that the generalization ability and the accuracy of the developed carbonation prediction model is reasonably high. This confirms the suitability of the proposed model for predicting the accelerated carbonation depth.
Table 3. Statistical performance evaluation of the developed carbonation prediction model

<table>
<thead>
<tr>
<th>Statistical performance indicators</th>
<th>Without variable selection</th>
<th>With variable selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean absolute error, MAE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.54356</td>
<td>0.30681</td>
</tr>
<tr>
<td>Testing</td>
<td>0.70269</td>
<td>0.41421</td>
</tr>
<tr>
<td>Mean square error, MSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.38253</td>
<td>0.13815</td>
</tr>
<tr>
<td>Testing</td>
<td>0.95624</td>
<td>0.72356</td>
</tr>
<tr>
<td>Root mean square error, RMSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.61849</td>
<td>0.37168</td>
</tr>
<tr>
<td>Testing</td>
<td>0.97787</td>
<td>0.85062</td>
</tr>
</tbody>
</table>

The complex nature of the carbonation process in combination with a wide variety of concrete mixtures makes reaching at an equation which fully represent the carbonation process challenging. As presented in the results section, neural network based model is able to predict the accelerated carbonation depth with reasonably high accuracy without assuming a predetermined equation as a model. In addition, unlike the conventional models, the proposed model is able to select influential parameters which govern the carbonation process from the whole dataset. This allows evaluating all governing parameters as a group rather than individually and thus ensures the reliability of the prediction since imperative dependencies are not overlooked.

5. Conclusions

In this work, an optimized neural network based carbonation prediction model was developed and presented. The model consists of multilayer feedforward architecture and it was trained using the Levenberg-Marquardt backpropagation training algorithm using a dataset obtained from the Finnish DuraInt-project. It employed a total of 15 variables as inputs. The target variable is the accelerated carbonation depths measured after exposing the specimens for 28 and 56 days in a chamber. The statistical performance indicators, MAE, MSE, and RMSE were applied to measure the prediction accuracy. The test performance of the presented model demonstrated that the model predicts the accelerated carbonation depth with minimal error. In addition, the incorporated variable selection method identified influential input variables and thus enhances the prediction performance of the model. Hence, the results demonstrate the applicability of the model in concrete service life management system.

References


