Prediction of Pervious Concrete Permeability and Compressive Strength Using Artificial Neural Networks

Behrooz Shirgir1*, Amir Reza Mamdoohi2, Abolfazl Hassani3

Received: 24.05.2015     Accepted: 07. 07. 2015

Abstract:
Pervious concrete is a concrete mixture prepared from cement, aggregates, water, little or no fines, and in some cases admixtures. The hydrological property of pervious concrete is the primary reason for its reappearance in construction. Much research has been conducted on plain concrete, but little attention has been paid to porous concrete, particularly to the analytical prediction modeling of its permeability. In this paper, two important aspects of pervious concrete due to permeability and compressive strength are investigated using artificial neural networks (ANN) based on laboratory data. The proposed network is intended to represent a reliable functional relationship between the input independent variables accounting for the variability of permeability and compressive strength of a porous concrete. Results of the Back Propagation model indicate that the general fit and replication of the data regarding the data points are quite fine. The R-square goodness of fit of predicted versus observed values range between 0.879 and 0.918 for the final model; higher values were observed for the permeability as compared with compressive strength and for the train data set rather than the test data set. The findings can be employed to predict these two important characteristics of pervious concrete when there are no laboratorial data available.

Keywords: Porous concrete, neural network, permeability, compressive strength.
Prediction of Pervious Concrete Permeability and Compressive Strength Using...

1. Introduction

Pervious concrete has been recommended as an alternative pavement surface to help reduce runoff as compared to traditional impervious pavements. The high void content in pervious concrete allows for storm water to flow through the pervious concrete to the sub-base below instead of running off [Ferguson, 2005; Tennis, 2004]. The advantages of pervious concrete surfaces include increased utilization of residential and commercial land and the direct replenishment of local aquifers. Pervious concrete pavements have been used in the past 20 years in lower traffic load areas such as parking lots, shoulders of roadways, airport taxiways and runways, street and local roads provided that subsoil conditions, drainage characteristics, and groundwater location are suitable. In recent decades, the use of pervious concrete for the construction of secondary roads, parking lots, driveways, walkways, and sidewalks is increasing continuously because of its various environmental benefits such as: [Nguyen et al. 2014]

1. The storm water can rapidly be filtered into soil, and the groundwater resources can be recharged.
2. The surface is air and water permeable and the soil below can be kept wet. It improves the road surface.
3. The pervious concrete pavement can absorb the noise of vehicles, which creates quiet and comfortable environment.
4. The pervious concrete pavement materials have holes that can cumulate heat. The pavement can adjust the temperature and humidity of the Earth surface and eliminates the hot island phenomenon in cities.

Hardened pervious concrete usually has a low compressive strength which in some cases (when porosity is limited to 15%) the compressive strength can increase to 28 MPa [Pala et al., 2007]. Pervious concrete can only be applied to squares, footpaths, parking lots, and paths in parks [Ghafoori and Dutta, 1995; Fukute, 1998]. High range water reducer and thickening agent are introduced in the concrete to improve its strength and workability [Raffique et al. 2012]. Research on pervious concrete is rapidly developing. However, in spite of studies conducted on the effects of aggregate gradation, there still appears to remain much to be learned about the effects its properties and gradations have on pervious concrete mixtures. Pervious concrete has been successfully utilized in the USA for over 40 years; the applications in cold climates only began around 2006 due to the pervious concrete perceived lack of freeze–thaw durability [John and Vernon, 2013]. An important aspect of the performance of pervious pavements is the permeability of pervious concrete. Defined as a measure of the ability to transmit fluid, most commonly water, a highly permeable concrete can help deposit excessive water underneath it, or on the base or sub-grade layers of a pavement system. Due to its open structure, the strength of pervious concrete is lower than regular concrete. Although permeability is considered as one of the most important characteristics of pervious concrete, the strength of the structure should not be under-estimated. Most researches on pervious concrete are based on laboratory effort and linear regression analysis, and not much research has been conducted based on heuristic methods. In this paper, a heuristic method, artificial neural networks (ANN) is developed to predict the two major properties of pervious concrete, namely permeability and compressive strength. The proposed network is intended to represent a reliable functional relationship between the input independent variables accounting for the variability of permeability and compressive strength of a pervious concrete as two output dependent variables. The main objective of this paper is, thus, to apply an ANN model trained by the valuable laboratory data to predict permeability and compressive strength of pervious concrete.

2. Artificial neural Network

ANN as generalization of mathematical models is derived primarily in analogy to biological nervous systems. A first wave of interest in neural networks emerged after the introduction of simplified neurons by McCulloch and Pitts [McCulloch and Pitts, 1990]. Neurons or nodes are the basic processing elements of neural networks. In a simplified mathematical model of a neuron, the effects of synapses are exerted by connection weights that regulate the effect of related input signals, and the nonlinear characteristic exhibited by neurons is represented by a transfer function. The neuron output is computed as the weighted sum of the input signals, transformed by the transfer function. The learning ability of a neuron is attained by adjusting the...
weights according to the chosen learning algorithm. Consisting of many simple processing units (neurons) with dense parallel interconnections, an ANN can provide meaningful answers even when the data to be processed include errors or are incomplete and can process information extremely rapidly when applied to solve real world problems [Lippmann, 1987]. The first studies on ANN started in 1943, when McCulloch and Pitts defined artificial neurons for the first time to develop a cell model. Together with the developments in computer technology, the use of artificial neural networks has become more efficient after 1980 [Anderson and Brown, 1983; Hopfield, 1982; Lee, 2003]. In 1958, Frank Rosenblatt devised a machine called perceptron that operated much the same way as human mind. An ANN is quite simple and small in size when compared to the human brain, and due to its similarity to the human brain, it has some powerful characteristics in knowledge and information processing. Hence, an ANN can be a powerful tool for engineering applications [Kewalramani and Gupta, 2006]. In recent years, ANN (single layer or multilayer) has been applied to various civil engineering problems ranging from the detection of structural damage, structural system identification, modeling of material behavior, structural optimization, structural control, ground water monitoring, and prediction of experimental studies, to concrete mix proportions [Adhikary and Mutsuyoshi, 2006].

2.1 Architecture of neural network

An ANN, basically, consists of three types of neuron layers: input, hidden, and output layers. In feed-forward networks, the signal flow is one way only: from input to output units, where the data processing can extend over multiple (layers of) units, but no feedback connections are present. Recurrent or feedback networks contain feedback connections allowing for signal flow in both directions. There are several other neural network architectures (Elman network, adaptive resonance theory maps, competitive networks, etc.), depending on application properties and requirements. The multilayer perceptron is the most widely used type of neural network, which is both simple and based on solid mathematical calculation. The input layer consists of as many neurons as the number of input variables of the problem, and the output layer, as many as the desired number of values computed from the inputs [Wu and Lim, 1993]. Hidden layers may contain a large number of hidden processing units depending on the complexity of the phenomenon being modeled; however, all problems which can be solved by a perceptron can also be solved with only one hidden layer [Adhikary and Mutsuyoshi, 2006]. The process of a model based on neural network involves five main parts: (a) data acquisition, analysis, and problem representation; (b) determining of architecture; (c) determining the learning process; (d) network training; and (e) network testing for generalization evaluation.

The weighted sums of the input elements \((\text{net})_j\) are calculated using Eq. (1)

\[
(\text{net})_j = \sum_{i=1}^{n} w_{ij} x_i + b
\]

Where \((\text{net})_j\) is the sum of weight of the \(j^{th}\) neuron for the input received from the preceding layer with \(n\) neurons, \(w_{ij}\) is the weight between the \(j^{th}\) neuron in the preceding layer, and \(x_i\) is the output of the \(i^{th}\) neuron in the preceding layer. Weights are values that express the effect of an input set or another process element in the previous layer on this process element; sum function is a function that calculates the effect of inputs and weights totally on this process element, and \(b\) is the bias and is used to model the threshold [Hola and K. Schabowicz, 2005; Schaefer, 2006; Topcu and Saridemir, 2008]. Activation function is a function that processes the net input obtained from sum function and determines the cell output. The most common activation functions are ramp, sigmoid, and Gaussian function. In general, for multilayer receptive models \((f(.))\), sigmoid function is used. Figure 1 shows a typical neural network with input, sum function, sigmoid activation function, and output. The neuron output \((\text{out})_j\) is calculated employing Eq. (2) with a sigmoid function as follows [Adhikary and Mutsuyoshi, 2006; Wu and Lim, 1993; Schaefer, 2006]:

\[
(\text{out})_j = f(\text{net})_j = \frac{1}{1 + e^{-\alpha(\text{net})_j}}
\]

Where \(\alpha\) is a constant used to control the slope of the semi linear region. The sigmoid nonlinearity is activated in every layer except in the input layer [Hola and K. Schabowicz, 2005; Topcu and Saridemir, 2008; Hagan, 1996]. The connection weights and bias values are initially chosen as random numbers and then fixed by the
results of the training process. Many alternative training processes are available; however, a few are used widely like back propagation scheme. The goal of any training algorithm is to minimize the root mean square error (RMSE), defined as the difference between predicted outputs of the model and observation outputs (used in the training dataset).

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - O_i)^2}$$

where $P_i$ and $O_i$ are the predicted and observed values, respectively. $N$ is the total number of data points in validation.

2.2 Back Propagation network

The back propagation (BP) learning is an iterative search process adjusting the weights from output layer back to input layer in each run so that no further improvement in RMSE value (or other parameter) is found. The BP algorithm calculates the error, employs it to adjust the weights first in the output layer, and then distributes it backward from the output to hidden and input nodes (Figure 2). The steepest gradient descent principle is employed to direct the change in weight towards the negative of the error gradient, reaching an output close to one of its extreme values (usually 0 or 1). Eq. (4) presents this mathematically as:

$$\Delta w_n = \alpha \Delta w_{n-1} - \eta \frac{\partial E}{\partial w}$$

Where, $w$ is the weight between any two nodes; $\Delta w_n$ and $\Delta w_{n-1}$ are the changes in this weight at $n$ and $n-1$ iteration; $\alpha$ is the momentum factor; $\eta$ is the learning rate; and $E$ is the error term.

The final connection weights are kept fixed, and new input patterns are introduced to the network to produce the corresponding output compatible with the internal representation of the input/output mapping.

3. Experimental program and Data Collection

3.1 Overview of the Sample Data

The scope of this research is to obtain data by conducting laboratory tests on pervious concrete samples, using 13 different aggregate gradations from an aggregate source (Shahryar Akamshen), utilizing eight different water to cement (W/C) ratios. The sample data used for modeling includes 105 records. A data set was designed in the form of pairs of vectors, and the associated permeability and compressive strength as target vectors.

The input variables were: (i) fine aggregate (FA); (ii) porosity (Po); (iii) water-to-cement ratio (W/C); (iv) coefficient of uniformity ($C_u$); and (v) the maximum specific gravity ($G_{max}$). The boundary range of inputs and output of records including minimum, maximum, mean, and standard deviation values are listed in Table 1. Increasing the number of training samples provides more information on the shape of the solution surface(s) and thus, increases the potential level of accuracy that can be achieved by the network. Having too few data samples leads to poor generalization by the network. An optimal data set for training would be the one that fully represents the modeling domain and has the minimum number of repetitive samples (i.e. identical inputs with different outputs) in training. Thus, about 70% of the 105 records were randomly chosen for model calibration (training), 15% for validation, and the rest (15%) were kept for model test. The training data set is used to compute the gradient and update weights and biases. The error in the validation sets is monitored during the training process to prevent over-fitting of training data.
The error in the test sets is not used during training, but it is used to compare different models.

Table 1. Boundary range of inputs and output of records

<table>
<thead>
<tr>
<th>Mix components</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input (independent variables)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine aggregate (FA)</td>
<td>0</td>
<td>53</td>
<td>15.04</td>
<td>14.76</td>
<td>%</td>
</tr>
<tr>
<td>Porosity (Po)</td>
<td>16.4</td>
<td>30.8</td>
<td>24.92</td>
<td>2.65</td>
<td>%</td>
</tr>
<tr>
<td>W/C</td>
<td>0.27</td>
<td>34</td>
<td>0.31</td>
<td>0.03</td>
<td>-</td>
</tr>
<tr>
<td>Coefficient of uniformity (C_u)</td>
<td>1.2</td>
<td>3.28</td>
<td>2.28</td>
<td>0.73</td>
<td>-</td>
</tr>
<tr>
<td>G_{mm}</td>
<td>2500</td>
<td>2780</td>
<td>2630</td>
<td>80</td>
<td>Kg/m³</td>
</tr>
<tr>
<td><strong>Output (dependent variables)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permeability (P)</td>
<td>0.33</td>
<td>13</td>
<td>4.23</td>
<td>2.47</td>
<td>mm/sec</td>
</tr>
<tr>
<td>Compressive Strength (CS)</td>
<td>6.55</td>
<td>23.56</td>
<td>13.55</td>
<td>3.05</td>
<td>MPa</td>
</tr>
</tbody>
</table>

3.2 Testing Program

As PC contains much higher hydraulic permeability than ordinary concrete, conventional methods which are used to evaluate the permeability of normal concrete cannot be directly applied. Therefore, based on [Neithalath, Weiss, and Olek, 2006], a falling head permeability test device was mounted to estimate the permeability of the cylindrical samples (Figure 3). To prevent unsaturated overflow during the test, drainage pipe was 10 mm overhead compared from the top of the specimen. A graduated acrylic cylinder of 400 mm long was fixed to the top of the specimen assembly and clamped tightly to prevent water leakage. Water level was monitored during the test by reading the level in the graduated cylinder. To remove the air trapped in the specimen and ensure that the specimen was completely saturated, it was conditioned via allowing the water to drain out through the pipe until it was leveled in the graduated cylinder and the drain pipe. Having the valve closed, the graduated cylinder was filled with water.

After opening the valve, the time (t) required for the water level to fall from an initial head of 350 mm (h₁) to a final level of 50 mm (h₂) was measured in seconds. This procedure was repeated three times, and the average t was used for calculation purposes. Average coefficient of permeability (k) was figured out using Eq. 1, which follows Darcy’s law and assumes laminar flow [Shirgir, Hassani, and Khodadadi, 2011].

\[ k = \frac{aL}{A_t} \ln \left( \frac{h_1}{h_2} \right) \]  

where, \( a = \) the average cross-section of the graduated cylinder
\( A_t = \) the average cross-section of the test sample
\( L = \) the length of the specimen.

The gradation specifications used followed the guidelines of ASTM C 33. These aggregate gradations fit the specifications of the single-sized ½ inch, #67-I, #67, A12.5-C, A12.5, single-sized ¾ inch, A9.5-C, A9.5, #78-C, #78-C, and single sized #89-C, #89-F and #89. Uniformity coefficient (C_u) dominated the effects of aggregate gradation where this parameter is defined as:

\[ CU = \frac{D_{60}}{D_{10}} \]  

where,
\( D_{60} = \) the diameter of aggregate corresponding to 60% finer
\( D_{10} = \) the diameter of aggregate corresponding to 10% finer

The maximum specific gravity test, ASTM D 2041 (Rice Test), although specified for bituminous paving mixtures according to the ASTM standard, was deemed valid since pervious concrete is a relatively new concept for which standard specifics are still being tested and developed. Also, loose matured cement coated pervious concrete aggregate is similar to, in form, loose pervious asphalt or a bituminous paving mixture which is essentially loose aggregate coated with binder which compares to the loose aggregate coated with cement paste. The loose cement coated aggregate was oven-dried with a minimum mass of 1500 g (each was placed in two metal bowls). The sample weights were taken, and water with a temperature of approximately 25 °C was added to the bowls to completely cover the samples. The two bowls were placed on the mechanical agitation device, and the air trapped in the sample...
was removed and later the bowl and sample submerged weight were recorded as documented in the ASTM D 2041 standard, Section 9.4 through 9.5.1. Calculations for the maximum specific gravity were conducted based on the bowls used under water determination:

\[ G_{\text{mm}} = \frac{A}{A-(C-B)} \]  

(7)

where,

- \( G_{\text{mm}} \) = maximum specific gravity of the mixture
- \( A \) = mass of the dry sample in air
- \( B \) = mass of bowl under water
- \( C \) = mass of bowl and sample under water

3.3 Pre-processing of Data

The data was pre-processed by Principal Component Analysis (PCA) technique to identify and delete the insignificant variables; however, the PCA results showed that all five input factors (listed in Table 1) are significant. It is generally recommended to normalize the input and output data before presenting them to the network. In this research, a linear normalization as Eq. (8) was used to limit the variation of the variables to the interval (0, 1):

\[ S = \frac{(V - V_{\text{min}})}{(V_{\text{max}} - V_{\text{min}})} \]  

(8)

Where, \( S \) is the normalized value of the variable; \( V \), \( V_{\text{min}} \), and \( V_{\text{max}} \) are variable, minimum, and maximum values, respectively.

3.4 Model Construction and Performance Evaluation

There are no fixed guidelines to determine the best architecture of the network, and hence this has to be done by the trial-and-error method. The network ability to separate the data is affected by the number of hidden neurons. A large number of hidden neurons will ensure correct learning, and the network is able to correctly predict the data it has been trained on, but its performance on new data and its ability to generalize are compromised. The number of hidden neurons is a crucial decision. BP networks with one and two hidden layers and varying number of neurons were developed. A parametric study was carried out by changing the number of neurons in the hidden layers to test the stability of the network. Figure 4 shows the performance of the BP networks with various numbers of neurons in one hidden layer. Based on RMSE (Root Mean Square Error), it can be observed that the network with seven neurons (BP 5-7-2) in one hidden layer results in a stable and optimum network.

A further test on whether additional second hidden layer could improve the network performance was carried out, in which the number of seven neurons in the first hidden layer was fixed and various numbers of neurons in the second hidden layer were used. BP network having the structure 5-7-3-2 (seven neurons in the first and three neurons in the second hidden layer) produced the best results; however, still not as good as the one-layer network with seven neurons presented in Figure 5.

4. Results and Discussion

The same mixtures were tested for bulk density, compressive strength, and permeability coefficient in a previous study [Gesoglu et al. 2014]. The results of compressive strength showed that the pervious concrete was observed when waste rubber was replaced with single sized natural coarse aggregate [Gesoglu et al., 2014]. Figures 6 and 7 show the scatter plots of predicted versus observed values for the BP network depicted in Fig-
Figure 5 for the train and the test data set, respectively, for permeability (P) and compressive strength (CS) of porous concrete. As observed, the general fit and replication of the data regarding the data points are quite fine. For further analysis of the results, Table 2 reveals the performance measures of the BP networks. The R-square goodness of fit measures of predicted versus observed values range between 0.879 and 0.918 for the final model of this study (5-7-2); higher values were obtained for P as compared with CS and for the train data set as compared with the test data set. Interestingly, not much decrease is observed in this measure (R-square) for the test data (0.006 and 0.043, for P and CS, respectively). The RMSE measures follow more or less the same pattern, having the best (least) value for P and train data set. The difference between the train and test data is quite modest (0.143 and 0.107 for P and CS, respectively). Relative error measures (maximum, mean, and minimum) are also reported in the table. The range of relative error is larger for P and for train data; however, the mean is, quite expectedly, higher for the test data as compared to the train data. Regarding the increase in mean relative error from the train data, P is much better modeled and predicted (less than 5% increase from the train data) than CS (more than 50%). Table 2 compares the results of the two network structures (5-7-2 and 5-7-3-2) discussed in section 3-4 based on the performance measures. The one-layer network is revealed to be generally preferable to the two-layer network, despite the additional hidden layer. Figures 8 and 9 show the predicted and observed values for the BP network depicted along with their error, by data point. Figure 8 depicts P for both the train data set and the test data set, whereas Figure 9 depicts this for CS of pervious concrete. The absolute error shaded in these figures on the horizontal axis varies between -2 and 2, indicating rather small error for both the train and test data, despite the observed variations (in P and CS). It can, thus, be concluded that the neural network model proposed in this paper is generally a robust model with high predictive power which can be used to predict P and CS of pervious concrete based on five input variables.
5. Conclusion and Recommendations
An important aspect of the performance of pervious pavements is the permeability of pervious concrete. Also, due to its open structure, the strength of pervious concrete is lower than regular concrete. Although permeability is considered as one of the most important characteristics of pervious concrete, the compressive strength of the structure should not be under-estimated. An artificial neural network (ANN) was developed to predict these two major properties for pervious concrete. The proposed network intended to represent a reliable functional relationship between the input independent variables accounting for the variability of permeability and compressive strength of pervious concrete, based on the valuable laboratory data, gathered exclusively for this purpose. Results of the Back Propagation model indicate that the overall fit and reproduction of the observed laboratory data by the proposed model are quite fine. The R-square (goodness-of-fit measure) of predicted versus observed values range between 0.879 and 0.918 for the final model; higher values were obtained for the permeability, as compared with compressive strength, and for the train data set as compared with the test data set. The findings of the current study can be employed to predict the two important characteristics (permeability and compressive strength) of pervious concrete without the need to expensive and time consuming laboratory gathered data. This present research is limited by the number of data observations, due to time consuming and expensive nature of laboratory data experiments. Thus, for further research, it is suggested to utilize more extensive and quantitative data-sets considering admixtures, time duration, cement type, and a large number of test specimens.
6. References


Prediction of Pervious Concrete Permeability and Compressive Strength Using...


