Research Article

Prediction of Compressive Strength of Concrete in Wet-Dry Environment by BP Artificial Neural Networks

Chengyao Liang, Chunxiang Qian, Huaicheng Chen, and Wence Kang

1School of Material Science and Engineering, Southeast University, Nanjing 211189, China
2Research Institute of Green Construction Materials, Nanjing 211189, China

Correspondence should be addressed to Chunxiang Qian; cxqian@seu.edu.cn

Received 25 August 2017; Revised 1 February 2018; Accepted 13 February 2018; Published 2 April 2018

Academic Editor: Patrice Berthod

Copyright © 2018 Chengyao Liang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Engineering structure degradation in the marine environment, especially the tidal zone and splash zone, is serious. The compressive strength of concrete exposed to the wet-dry cycle is investigated in this study. Several significant influencing factors of compressive strength of concrete in the wet-dry environment are selected. Then, the database of compressive strength influencing factors is established from vast literature after a statistical analysis of those data. Backpropagation artificial neural networks (BP-ANNs) are applied to establish a multifactorial model to predict the compressive strength of concrete in the wet-dry exposure environment. Furthermore, experiments are done to verify the generalization of the BP-ANN model. This model turns out to give a high accuracy and statistical analysis to confirm some rules in marine concrete mix and exposure. In general, this model is practical to predict the concrete mechanical performance.

1. Introduction

Marine environment tends to have a negative effect on concrete structures’ performance, which has been investigated in many researches. Compressive strength is used to describe the mechanical performance of concrete. In the marine environment, especially tidal and splash zones, concrete structure degradation is quite serious and the compressive strength of that is descending with exposure age. It is investigated in either laboratory or actual marine condition that compressive strength degrades with time.

In the early 1980s, British academic Mangat and Gurusamy [1] did a research on mechanical properties of steel fiber-reinforced concrete exposed to the marine splash zone and tidal zone in the Aberdeen beach, and the exposure age is up to three years (2000 wet-dry cycles). Results indicated that melt extract fibers are suitable for marine applications. In addition, another actual marine exposure experiment was done by Kuhail and Shihada [2] in the Gaza beach for a period of seven months. And it was found that the compressive strength of concrete shows a trend of rising early but declining later. Toutanji et al. [3] focused on studying the effect of different supplementary cementitious materials on strength and durability of concrete. They found that proper mineral additives could improve the performance of concrete in the marine environment. Aye and Oguchi [4] investigated the effect of physical sulfate attack on the performance of plain and blended cement mortars. Specimens were exposed to 10% Na$_2$SO$_4$ and MgSO$_4$ solutions for 24 months. MgSO$_4$ was found to be more damaging than Na$_2$SO$_4$ considering chemical attack. However, Na$_2$SO$_4$ was more harmful than MgSO$_4$ as far as the physical attack was considered. Jiang and Niu [5] also did a research on the effect of different types of sulfate solutions on concrete performance under wet-dry cycles. Results show that the deterioration degree of concrete in magnesium sulfate solution is more severe than that in the other sulfate solutions. Chloride ions in the composite solution help decrease the deterioration rate of concrete effectively. Chen [6] established a constitutive model for concrete in wet-dry cyclic sulfate attack.

In this research, compressive strength of concrete served in the marine tidal zone and splash zone or exposed to the wet-dry cycle environment was focused on. Database is established from existing researches. An artificial neural network (ANN) is applied in this study. The ANN has been
2. Data Collection and Analysis

In order to establish a prediction model that can be widely used, the data were excerpted from large numbers of experiments carried out by researchers. 2167 sets of data were collected in all, which cover over 80 articles from China, United States, Europe, and other regions, and they are presented in Table 1 in detail.

2.1. Database Establishment. In this research, compressive strength of concrete in the wet-dry cycle exposure environment, as the target of prediction, is related to many factors, including material factors and environmental factors. Material factors include w/c, specimen sizes, initial strength, and fly ash dosage and slag dosage. Specific surface area is chosen to describe the specimen sizes. Environmental factors are various ion concentrations, including sodium, magnesium, chloride, and sulfate ions, exposure condition, and exposure age. Exposure condition contains five variables, including the wetting time, drying time, wetting temperature, drying temperature, and cycle period.

The detailed data of the factors mentioned above are excerpted from previous experiments and are functioned as the database of the artificial neural network model.

2.2. Database Analysis

2.2.1. Statistical Analysis. Before establishing the model, a statistical analysis has been done on descriptive statistics of the database [7]. Statistical analysis is a multiplex method that can be used to describe the regularity and distribution of the sample database, and the fluctuation of database can be seen according to some statistics such as variance, standard deviation, skewness, and kurtosis [20]. Table 2 demonstrates some statistics of 1078 sets of data of factors influencing the compressive strength of concrete in the wet-dry environment. And those data were selected randomly from 2169 sets of data.

Table 2 mainly demonstrates six statistics of influencing factors. They are, respectively, standard deviation, variance, skewness, kurtosis, minimum, and maximum. Standard deviation is used to reflect the fluctuation of a series of data, and the instability of a statistic is measured by its variance [21]. In environmental factors, data of four types of solution ion concentrations share the similar discrete degree, so do the dosage of fly ash and slag in material factors. And compared to other factors, especially exposure age, variance and standard deviation of those six factors’ data are much lower because almost the choice of ion concentration refers to seawater ion concentrations and the dosage of fly ash and slag is mainly concentrated at 0.3, which is considered as the optimum content. Nevertheless, the exposure age ranges from 2 days to several years, contributing to the high variance and standard deviation of exposure age. On top of this, the degree of dispersion and variation of initial strength is close to that of final strength.

Skewness and kurtosis are used to characterize the distribution of data [22, 23]. Normal distribution is the most common distribution. Skewness refers to the frequency distribution of asymmetric degree of skew direction. Deviation between normal distributions is often reflected by the coefficient of skewness. When the coefficient of skewness is higher than 0, the oblique direction of the distribution is right (positive); on the contrary, the oblique distribution is left (negative). Right oblique direction distribution has a thin
and long tail on the left, which means that the data mainly focused on the small figures range. And taking a general observation of the data distribution of the factors, all of them are right oblique direction distributed. The coefficient of kurtosis indicates the shape of data distribution. It can be divided into spire distribution, standard distribution, and flat distribution. When the coefficient of kurtosis is higher than 3, it belongs to spire distribution; conversely, it belongs to flat distribution. Throughout the statistics given above, it can be concluded that except for exposure age, the other factors’ data are within normal limit. The coefficient of skewness of variables is all above 0, and it means that all the variables’ distribution directions show a positive oblique trend. Nevertheless, according to the results of coefficient of kurtosis, the four variables, respectively, fly ash dosage, slag dosage, sodium ions, and sulfate ions, show the spire distribution and the five variables, respectively, specific surface area, initial strength, magnesium ions, chloride ions, and final strength, show the flat distribution. And due to the particularity of the distribution of exposure age, it belongs to neither spire nor flat distribution. And initial strength shares a similar distribution with the final strength. Figures 1 and 2 describe the distribution of initial strength and final strength. Both of them are close to normal distribution.

In Table 3, the strength range is divided into 7 intervals and the distribution of values of initial strength and final strength is demonstrated. It can be clearly seen that the samples whose values of initial strength are between 25 and 50 MPa account for the largest proportion. So are the final strength of samples. Moreover, there is a decrease in the numbers of final strength samples of high-strength concrete compared with that of initial strength samples of high-strength concrete.

2.2.2. Correlation Analysis. On the basis of statistical analysis, correlation between final strength and various influencing factors is made in this research. The Pearson correlation coefficient, Kendall coefficient, and Spearman coefficient are commonly used. Nevertheless, the Pearson correlation coefficient is used in linear relationships [24]. The Kendall coefficient and Spearman coefficient belong to nonparametric statistics [25]. In this study, Kendall and Spearman coefficients are used for a correlation analysis. The analysis results are demonstrated in Table 4.

The correlation coefficients calculated above are based on two different formulas, both of which can be adopted more widely in nonlinear relationships. The results are in absolute value. From the analysis results, it can be clearly seen that w/c and initial strength have a significant effect on the final strength. However, fly ash dosage is more effective than slag dosage. It is quite a complex relationship between mineral admixtures and compressive strength in the marine environment, and mineral admixtures are considered to have a positive influence on concrete properties. The difference between these two correlations is attributed to the fact that data collected focused more on fly ash dosage. Besides, seawater also plays a significant role in the evolution of marine environment conditions.
Nevertheless, the correlation coefficients calculated seem hard to reflect that. Taking an observation of data collected, ion data are based on marine hydrographic data and almost have no change. Hence, correlation coefficient calculation formulas are built on data-changing trends [26]. Consequently, ion concentration seems to have a subtle effect.

In general, the correlation analysis results can be used as a simple reference for construction of artificial neural networks.

### 3. Construction of Artificial Neural Networks

Artificial neural network (ANN) model is a prediction model that has been widely applied in many fields. Similar to human brains, which respond to external stimuli through connections and exchanging information among ten billions neurons, the artificial neural network model is able to deal with the message inputted and realize result prediction [27]. One important training rule of the ANN is the delta rule, which is based on the idea of gradient descent. The delta rule is applied to determine the fraction of difference between the target and output [28].

The backpropagation (BP) neural network model is adopted in compressive strength in the research. The BP algorithm underpins the delta rule to neural nets with hidden nodes. The whole operation is demonstrated in Figure 3.

The operation can be divided into six steps [29]. Step 1: input training factors. Fifteen influencing factors are inputted into the model; Step 2: hidden nodes calculate the output. This is a quite complex process that detailed the calculated algorithm is invisible; Step 3: output nodes calculate outputs. Step 4: compare the outputs with targets and figure out the difference; Step 5: adjust the model parameters on the basis of the training rule using the results of Step 4; Step 6: input testing factors, output prediction. The whole operation is demonstrated in Figure 4.

### Table 3: Sample numbers of initial and final strength distribution intervals.

<table>
<thead>
<tr>
<th>Strength intervals (MPa)</th>
<th>0~25</th>
<th>25~50</th>
<th>50~75</th>
<th>75~100</th>
<th>100~125</th>
<th>125~150</th>
<th>&gt;150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample numbers of initial strength</td>
<td>43</td>
<td>621</td>
<td>361</td>
<td>40</td>
<td>7</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Sample numbers of final strength</td>
<td>83</td>
<td>588</td>
<td>405</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 4: Kendall and Spearman correlation coefficients of influencing factors.

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Specific surface area</th>
<th>Initial strength</th>
<th>FA dosage</th>
<th>Slag dosage</th>
<th>Sodium ions</th>
<th>Magnesium ions</th>
<th>Chloride ions</th>
<th>Sulfate ions</th>
<th>Exposure age</th>
<th>w/c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendall</td>
<td>0.01</td>
<td>0.406**</td>
<td>0.172**</td>
<td>0.04</td>
<td>0.00</td>
<td>0.180**</td>
<td>0.01</td>
<td>0.00</td>
<td>0.078**</td>
<td>0.483**</td>
</tr>
<tr>
<td>Spearman</td>
<td>0.00</td>
<td>0.577**</td>
<td>0.226**</td>
<td>0.05</td>
<td>0.00</td>
<td>0.243**</td>
<td>0.02</td>
<td>0.00</td>
<td>0.126**</td>
<td>0.627**</td>
</tr>
</tbody>
</table>

**This factor has a quite high relevance with the final strength.

![Figure 3: Backpropagation neural network abstract.](image)

![Figure 4: Comparison between six networks.](image)
and calculate every $d(n)$ in hidden nodes in this step; Step 6: the results of Step 5 are used to carry on a second training until the error is suitably small.

3.1. Determination of Input Neurons and Output Neurons and Pattern Classification. As shown in Figure 3, fifteen inputs are given and they represent fifteen influencing parameters. 2169 sets of data are collected for the establishment of the BP neural network model. The datasets are categorized into three parts: training set, validation set, and testing set, respectively. There are various pattern classification methods. For example, Chen et al. [30] used 86% dataset for training and 14% dataset for testing to establish an ANN model to predict the strength of concrete. Hence, the conventional and widely recognized division ratio is 50% training set, 25% validation set, and 25% testing set [31]. Prechelt has done a research on some benchmark rules and problems of the neural network model establishment and come up with a basic pattern classification. The division ruler is following the convention in the literature of using half of the images for training and half for testing. And he proved that this pattern classification method can well avoid the overfitting of the model [32]. Hence, in this research, 50% of the data are used as the training set, 25% are used as the testing set, and the rest 25% are used as the validation set to predict the compressive strength of concrete that is exposed to the wet-dry environment.

3.2. Model Establishment. The training and testing program is written and put into MATLAB. The inner network structure is discussed. Ash used the DNC method to select the proper node parameters [33]. DNC means dynamic node creation. DNC sequentially adds nodes one at a time to the hidden layer(s) of the network until the desired approximation accuracy is achieved. There are six networks and they are discussed in Figure 4. Detailed structure information is given below, and Figure 4 presents the training results.

Network 1: 15-25-10-1
Network 2: 15-15-5-1
Network 3: 15-30-15-1
Network 4: 15-30-1
Network 5: 15-25-1
Network 6: 15-20-1

From Figure 4, Network 3, 15-30-15-1, turns out to have the lowest mean square error and is applied into the BP neural network model. Because the mean squared error tends to be a plateau after 3000 epochs, the epochs are set as 3000 epochs in the parameter setting.

In this model, trainlm is a network training function that updates weight and bias values according to the Levenberg–Marquardt optimization. trainlm is often the fastest backpropagation algorithm in the toolbox and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms. And the detailed trainlm training parameters are given as follows:

- `net.trainParam.show` equals 10
- `net.trainParam.lr` equals 0.05
- `net.trainParam.goal` equals $10^{-4}$
- `net.trainParam.epochs` equals 3000
- `net.trainParam.max_fail` equals 7

In this model, performance goal $10^{-4}$ is set as the threshold value. And the best validation performance is reached at the 49th epoch. The maximum validation failure is set as 7; thus, the model is calculated to the 56th epoch before ending. And the validation error reached the setting value. Significant parameters are demonstrated in Figure 5. After the 56th epoch run, the gradient reaches 0.0044514. And the decrease factor Mu is 0.0001 in this research. Validation vectors are used to stop training early if the network performance on the validation vectors fails to
improve or remains the same for max_fail epochs in a row. Test vectors are used as a further check that the network is generalizing well but does not have any effect on training.

3.3. Prediction Results. After the establishment of this prediction model, the prediction results are demonstrated in Figure 6.

In the training set, the correlation coefficient $R$ is 0.9684 (Figure 6), and the training model is quite appropriate. Then, the validation set is used to optimize the training set until it reaches the setting error value. And the final correlation coefficient of the validation set is 0.92702. After the prediction model establishment, testing set data is to test the performance of the neural network model. The correlation coefficient of that is 0.94778. The general prediction is listed in Figure 7 ($R = 0.962$). And this neural network model achieves a good prediction of 2169 sets of data.

In addition, the three error indexes, RMSE, MAPE, and MSE, of this model, respectively, are calculated in Table 5. The MAE of training data, validation data, and testing data was all less than 4.05 MPa, the MAPE of training data, validation data, and testing data was all less than 5.5%, and
the RMSE of training data, validation data, and testing data was all less than 4.5 MPa. These results indicate that this model has an excellent testing performance in predicting compressive strength of concrete exposed to the wet-dry environment.

3.4. Comparison with the Linear Regression Model. Relationship between multiple input variables (explanatory variables) and one output variable (a response variable) also can be expressed by the linear regression model [34–36]. Linear regression model is a statistical model determined as follows:

\[ Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n + \beta, \]

where \( Y \) is the output value and responds to compressive strength of concrete in this research; \( X_i \) represents different influencing factors, such as \( w/c \) and initial strength; \( \alpha_i \) are the weighing coefficients of different influencing factors; and \( \beta \) is the error-modifying coefficient. Enter, stepwise, forward, and backward are four regression methods for ordinary least squares estimation.

A linear model of compressive strength responding to various influencing factors is established through SPSS 22.0.

A stepwise method is chosen, and the fitting curve of outputs and targets of the linear model is shown in Figure 8.

The fitting result is obviously worse than that of the BP neural networks. And the MAE, RMSE, and MAPE are calculated in Table 6. From the comparison with the prediction results of the BP neural model, it is concluded that the BP neural model has a much better prediction performance.

4. Experiment Validation

In order to verify the generalization ability of this prediction model, a set of experiments were done.

4.1. Experiment Work and Data Collection

4.1.1. Raw Materials. Concrete specimens with three different strength grades were cast in the experiments. Three strength grades are, respectively, C30, C50, and C80. P52.5 Portland cement coming from Huaxin cement factory was chosen. Fine aggregates are fluvial sands and have a modulus of fineness of 2.7. The sediment content of the fluvial sands is 1%. Coarse aggregates are 5–25 continuously graded limestones and the sediment content is 0. A water-reducing agent with a solid content of 30% is polycarboxylate high-performance water-reducing admixture, manufactured by Subote New Materials Co., Ltd.

4.1.2. Mix Proportions and Specimen Preparation. Nine 100 mm × 100 mm × 100 mm cubes were cast to test the
mechanical properties, and three specimens were prepared at every strength grade level. Table 7 shows the mix design of concrete specimens.

Water-reducing admixture dosage is adjusted according to the rheological properties of fresh concrete in trial mix, and the final mix design was determined. The specimens are curing according to GB/T50081-2002. Curing temperature is 20 ± 2 °C, and relative humidity is 95%. After curing for 90 days, cement hydration is almost completed, and the strength development reaches a plateau.

4.1.3. Details of Wet-Dry Exposure. Cube specimens are exposed to a wet-dry environment after 90-day standard curing. The wet-dry cycle system is in accordance with GB/T50082-2009 but a little different from that. In this system, the specimens were immersed in solution for 16 hours, where the solution temperature is 20 °C. Then, the specimens were put into the oven where the baking temperature is 50 °C for 8 hours. The whole process is treated as a complete cycle, and one wet-dry cycle takes one day.

4.1.4. Testing Procedure. After mixing and standard curing, all hardened specimens from each group were tested to estimate the compressive strength. In order to achieve nondestructive testing of concrete, the ultrasonic method was chosen to measure compressive strength of concrete. Compressive strength at 90 days is regarded as initial strength. After 5, 10, 15, 20, and 25 wet-dry cycles, compressive strength was tested through the ultrasonic method.

4.1.5. Experimental Results and Discussion. Figure 9 shows the compressive strength of cube specimens after wet-dry exposure. An assemble list including experiment results and influencing factors is shown in Table 8, and lists A–O are 15 influencing parameters of final compressive strength.

4.2. Consistency between Experiment Results and Prediction Results. The experimental parameters were brought into the input neurons of this model to test the difference between prediction values and actual values. Figure 10 shows the prediction value and actual value. These white dots are prediction values and red asterisks are actual values. It turns out to be a good compatibility. In order to assess the performance of this BP-ANN model, three statistical indicators are calculated. Table 9 shows the detailed values and error percentages. A maximum error percentage is 3.93%, and all the errors are within acceptable range. The data fitting is ideal. Hence, this model could be used for the wet-dry environment concrete compressive strength prediction.

The MAE of the prediction is 0.6356 MPa, the RMSE is 0.8144 MPa, and the average error percentage is 1.09%. From the calculation results of MAE and RMSE, it can be concluded that this model is quite accurate for predicting the strength.

Table 7: Mix design of raw materials.

<table>
<thead>
<tr>
<th>Grades</th>
<th>Cement (kg/m³)</th>
<th>Silica fume</th>
<th>Water (kg/m³)</th>
<th>Sand (kg/m³)</th>
<th>Coarse aggregate (kg/m³)</th>
<th>Gravel (kg/m³)</th>
<th>Water reducer</th>
<th>w/c</th>
</tr>
</thead>
<tbody>
<tr>
<td>C30</td>
<td>375</td>
<td>—</td>
<td>200</td>
<td>750</td>
<td>675</td>
<td>450</td>
<td>—</td>
<td>0.53</td>
</tr>
<tr>
<td>C50</td>
<td>450</td>
<td>—</td>
<td>136.1</td>
<td>675</td>
<td>675</td>
<td>450</td>
<td>2.0</td>
<td>0.35</td>
</tr>
<tr>
<td>C80</td>
<td>514</td>
<td>57</td>
<td>141.65</td>
<td>684</td>
<td>615.6</td>
<td>410.4</td>
<td>9.5</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Figure 9: Compressive strength of concrete exposed to cyclic numbers.
From the errors of the prediction values and true values in Table 10, the strength can be accurately predicted by this BP neural network model. However, because this BP neural network model is a four-layer structure, it may cost a long time to do the calculation. Besides, the influencing factors may not be considered totally, and this will have an effect on the prediction results. Last but not least, weighing coefficients could be calculated by some new method, such as Grey relational theory, in advance. It could shorten the running time and simplify the model structure. And this method is under exploration.

5. Conclusions

The BP artificial neural network was developed for compressive strength of concrete in wet-dry environment prediction. The following conclusions are obtained in this study:

(1) Data collected to establish a prediction model are relatively representative in this research. However, data of some particular factors need more experiments to supplement, such as chloride ion and magnesium ion concentrations and slag dosage.
And new data are continuously added into the database.

(2) Through multiple attempts, a four-hidden layer architecture BP artificial neural network is more appropriate to predict a multifactorial target because the combined effect of fifteen factors is quite complex. It turns out that the model is appropriate to predict concrete compressive strength in the wet-dry environment.

(3) The generalization of this model is validated by experiments. Although long-term strength prediction accuracy is not available, errors of short-term experimental values and prediction values are subtle. The BP artificial neural network is practical for concrete compressive strength prediction.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was financially supported by the National Program on Key Basic Research Project of China (Grant no. 2015CB6551002). The authors would like to express their thanks to the Key Basic Research Project of China for finance grant.

References


