Computers and Concrete, *Vol. 20, No. 6 (2017) 627-634*
DOI: https://doi.org/10.12989/cac.2017.20.6.627

Predicting compressive strength of blended cement concrete with ANNs

Uneb Gazder*[1], Omar Saeed Baghabara Al-Amoudi[2a], Saad Muhammad Saad Khan[2b] and Mohammad Maslehuddin[3c]

[1]Department of Civil Engineering, University of Bahrain, Isa Town, 32038, Bahrain
[2]Department of Civil and Environmental Engineering, King Fahd University, Dhahran, 31261, KSA
[3]Center for Engineering Research, Research Institute, King Fahd University of Petroleum & Minerals, Dhahran, KSA

(Received December 19, 2016, Revised February 16, 2017, Accepted June 19, 2017)

Abstract. Predicting the compressive strength of concrete is important to assess the load-carrying capacity of a structure. However, the use of blended cements to accrue the technical, economic and environmental benefits has increased the complexity of prediction models. Artificial Neural Networks (ANNs) have been used for predicting the compressive strength of ordinary Portland cement concrete, i.e., concrete produced without the addition of supplementary cementing materials. In this study, models to predict the compressive strength of blended cement concrete prepared with a natural pozzolan were developed using regression models and single- and 2-phase learning ANNs. Back-propagation (BP), Levenberg-Marquardt (LM) and Conjugate Gradient Descent (CGD) methods were used for training the ANNs. A 2-phase learning algorithm is proposed for the first time in this study for predictive modeling of the compressive strength of blended cement concrete. The output of these predictive models indicates that the use of a 2-phase learning algorithm will provide better results than the linear regression model or the traditional single-phase ANN models.

Keywords: blended cement concrete; compressive strength; artificial neural networks; learning algorithms

1. Introduction

The cost, time and energy involved in producing and placing concrete in the field constitutes a significant proportion of a construction budget. Furthermore, the compressive strength of concrete plays an important role in determining the volume of concrete to be used in a project. Therefore, it is important to ascertain the compressive strength of concrete to be used. Testing of concrete to determine its compressive strength is costly and requires a lot of time and effort (Khatibniah *et al.* 2016). Therefore, practitioners and researchers often rely on models to predict the specific properties of concrete with a given combination of materials and mix design (Wu *et al.* 2016).

Traditionally, concrete is made of three basic components; cement, water and aggregates. Consequently, empirical methods and formulae were commonly used to predict the compressive strength of concrete. However, it has been realized nowadays that a significant improvement in its properties can be achieved by modifying its composition, such as by using supplementary cementing materials. Further, the use of supplementary cementing materials in concrete results in technical, economic and environmental benefits (De Larrad amd Sedran 1994, Russell 1999, Santamaria *et al.* 2016) due primarily to the achievement of better durability and enhanced sustainability. Consequently, the traditional prediction models used to predict the concrete strength may no longer be valid in such cases. The complexity of such models increases with the use of supplementary cementing materials to produce binary, ternary or even quaternary cement blends.

Artificial neural network (ANN) technique has been effectively used to predict the compressive strength of concrete (Yeh 1998, Aticin 2011, Sadrmomtazi *et al.* 2013). ANN has also been used in other areas related to civil engineering, such as detection of structural damage (Feng and Bahng 1999), structural system identification (Feng and Kim 1998), concrete mix properties (Oh *et al.* 1999), traffic forecasting (Gazder and Hussain 2013).

Kasperkiewicz *et al.* (1995) used fuzzy-ARTMAP type ANN for the prediction of compressive strength of concrete. Out of 340 datasets, 200 datasets were taken to train the networks while 140 datasets were used to test the model. It was reported that the compressive strength of concrete can be accurately predicted using this type of ANN (Kasperkiewicz *et al.* 1995).

Guang and Ji-Zong (2000) predicted the 28-day compressive strength of concrete using feed-forward ANNs and reported that ANN models predicted the compressive strength with high accuracy. They used 100 data sets; 85 to train the network and 15 to test the models. The relative error between the predicted and actual compressive strength was within 5% for the first batch and 12% for the second batch.

Kim *et al.* (2004) used several concrete mix proportions from two ready mix concrete plants. The dataset from the

*Corresponding author, Ph.D.
E-mail: unebgazdar@gmail.com
[a]Ph.D. Professor
[b]M.Sc.
[c]Ph.D.

http://www.techno-press.org/?journal=cac&subpage=8
ISSN: 1598-8198 (Print), 1598-818X (Online)

first plant contained 10 parameters while the dataset from the other contained nine parameters. In total, ANN was trained on 98 datasets to predict the compressive strength of concrete. It was reported that the maximum error between the predicted and tested results was 3.9%.

Aitcin (2011) used regression model and ANN for predicting the compressive strength of concrete. The models were tested on 28 specimens that were cured from 3 to 180 days. He reported that ANN models can capture the complex relationship between the compressive strength and other variables that may not be the case for classical regression models.

Awwad (2004) developed a forecasting model for compressive strength of concrete using relevance vector machine (RVM) that is a type of Bayesian regression model used for training generalized linear models through a probabilistic approach. Other variations of this method include the support vector machine (SVM) technique. The data were modeled as a chaotic system in order to capture the behavior of the complete population for forecasting. Ninety data sets were used to train and test the predicted models. It was reported that these models accurately predicted the compressive strength of concrete cured for up to 28 days.

Bingol *et al.* (2013) used ANNs to model the compressive strength of light- and semi-lightweight concrete. Three parameters were taken as input. It was reported that the ANNs can predict the compressive strength with adequate accuracy.

Chou and Pham (2013) used ANN as well as ensemble models to predict the compressive strength of high performance concrete. Ensembles are referred to those models in which more than one model is synergistically combined to produce better results. They also developed predictive models using ensemble of chi-squared automatic interaction detector, support vector machine and classification and regression trees with ANNs. They constructed ensembles using the above techniques in different combinations with the maximum of three techniques used in an ensemble. It was reported that an ensemble exhibited better predictive nature and the difference between predictions of the ensembles and ANNs was not significant.

Khan *et al.* (2013) used multi-layer feed-forward ANNs to predict the compressive strength of concrete. Out of 55 datasets, 19 were used to test the ANNs. They also compared the results of ANNs with other theoretical models and reported that the prediction from ANNs was closest to the experimental results.

Sadrmomtazi *et al.* (2013) employed regression, ANNs and adaptive network-based fuzzy inference system (ANFIS), for predicting the compressive strength of lightweight concrete. ANFIS is also an ensemble, which is developed with the combination of ANN and fuzzy inference systems. They reported that ANNs with two hidden layers performed better than ANFIS or polynomial regression models.

2. Research significance

Literature review indicates that single ANNs are relatively accurate in predicting the compressive strength of concrete. However, ensemble learning has also been tried successfully to improve the predictive performance of ANNs. However, this approach can be computationally complicated due to the selection of appropriate techniques and methods of combining their results. Moreover, it also requires the knowledge of more than one technique.

In this paper, a simplified and effective approach is proposed to enhance the accuracy of ANNs for predicting the compressive strength of concrete blended with natural pozzolan (NP). The effectiveness of ANNs in comparison with regression models to predict the compressive strength of blended cement concrete was also assessed. Further, different learning algorithms for ANN were tested and compared for their performance in predicting the compressive strength of blended cement concrete. Last but not the least, a 2-phase learning algorithm is proposed for developing an ANN model. A comparison of the models was made using a 5-fold cross validation approach to corroborate the results of this study.

3. Materials and methodology

3.1 Experimental program

The experimental program was designed to use NP as a supplementary cementing material in concrete. Ordinary Portland cement (OPC) was replaced with 5-60% of NP, to determine the optimum dosage of NP. To enhance the reactivity of the used NP, 1 to 7% silica fume (SF) or 1 to 25% hydrated lime (HL) were added to produce blended cement, namely SF-OPC-NP or HL-OPC-NP. The compressive strength of concrete specimens was evaluated after a curing period of 7 to 180 days. Out of 65 specimens, 80% of the datasets, selected randomly in each validation, were used to train the ANN while the remaining 20% were used to test the model. Table 1 summarizes the details of the input data and the compressive strength values. The scope of this study was limited to the determination of effect of the curing period and the variation in the binder type (cementitious materials content) on the compressive strength. Therefore, only these variables were used for developing the prediction models. However, further research with more variables can be carried out using the same approach for studying the relationship between other parameters. Table 2 describes the variables used in developing the prediction models for compressive strength of concrete.

3.2 Prediction models

Linear and quadratic regression models were developed to predict the compressive strength of concrete for the experimental regime employed in this research. It was followed by the development of ANN models. Fig. 1 shows the methodology of this research. The performance of ANN models was trained by using different learning algorithms, such as back-propagation (BP), Levenberg-Marquardt (LM) and Conjugate-Gradient Descent (CGD) (Davoodi and

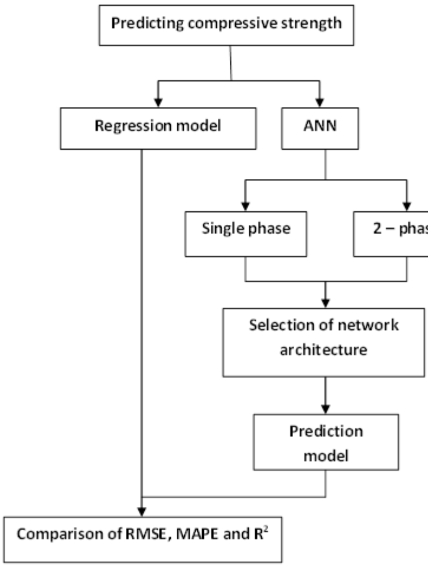

Fig. 1 Modeling methodology

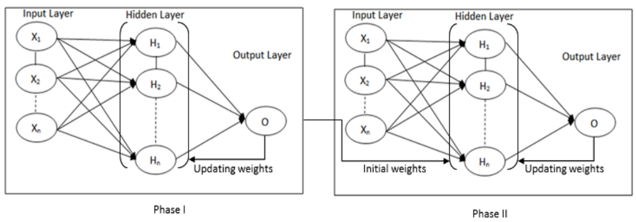

Fig. 2 2-Phase ANN framework used in the study

Khanteymoori 2010). In addition, a 2-phase learning algorithm was proposed for ANN model using different combinations of the above-mentioned algorithms. For example, ANN was trained using BP algorithm and then the trained network was further optimized by employing LM algorithm. An illustration of this approach is given in Fig. 2.

ANNs consists of a number of processing units, referred to as neurons or perceptrons. The neurons can be classified into three categories; input, hidden and output neurons. These categories are also known as layers. An input layer processes the input vector to the next layer with randomly generated weights and bias (Englebrecht 2007). It has the same number of neurons as the input variables. Each neuron in the hidden layer performs a two-step process; firstly it takes the weighted sum of input vector and secondly, it applies a pre-defined activation function on this sum value (Zurada 1992). The final value, after applying the activation function, is forwarded to the succeeding layer (either output or hidden). The process is depicted in phase I of Fig. 2. The number of hidden layers and the number of neurons in each hidden layer can be selected arbitrarily. However, these parameters affect the accuracy of the models greatly; hence, this criterion is used for determining their optimum values. The output layer consists of the number of neurons that are equal to the number of output variables. Each neuron in the output layer receives the value from the previous layer and performs the same procedure as the hidden layer to give the final result of the model (Jain *et al.* 1996).

The training algorithm of an ANN model is the process by which the network minimizes the error to give the best

Table 1 Description of experimental regime

Concrete mix composition	Compressive strength (MPa)				
	7 Days	14 Days	28 Days	90 Days	180 Days
100% OPC	49.68	53.25	60.98	65.41	69.56
95% OPC + 5% NP	47.93	52.11	60.54	68.98	72.32
90% OPC + 10% NP	45.68	51.93	58.53	70.11	76.6
85% OPC + 15% NP	36.96	44.31	48.36	63.82	69.89
80 % OPC +20% NP	35.28	41.18	51.91	59.67	68.23
60% OPC + 40% NP	16.68	20.73	25.19	35.16	47.31
40% OPC + 60% NP	9.21	16.01	19.72	30.73	42.47
84% OPC + 15% NP + 1% SF	48.23	54.01	63.22	68.59	73.08
82% OPC + 15% NP + 3% SF	50	55.63	65.28	71.13	75.89
80% OPC + 15% NP + 5% SF	53.04	56.11	68.47	74.13	79.78
78% OPC + 15% NP + 7% SF	54.42	59.81	69.83	75.76	81.22
84% OPC + 15% NP + 1% HL	-	-	47.6	-	-
82% OPC + 15% NP + 3% HL	-	-	51.8	-	-
80% OPC + 15% NP + 5% HL	-	-	57	-	-
78% OPC + 15% NP + 7% HL	-	-	67.8	-	-
76% OPC + 15% NP + 9% HL	-	-	62.8	-	-
75% OPC + 15% NP + 10% HL	-	-	60.6	-	-
70% OPC + 15% NP + 15% HL	-	-	49.8	-	-
65% OPC + 15% NP + 20% HL	-	-	47	-	-
60% OPC + 15% NP + 25% HL	-	-	39.8	-	-
Minimum compressive strength (MPa)	9.21	16.01	19.72	30.73	42.47
Maximum compressive strength (MPa)	54.42	59.81	69.83	75.76	81.22
Average compressive strength (MPa)	40.65	45.92	53.81	62.13	68.76
Cementitious material content (kg/m^3)	360				
*w/b	0.42				
[a]FA/CM	1:2				
[b]CA/CM	1:5				

*w/b=Water-to-binder ratio, [a]FA/CM=Fine aggregate-to-cementitious materials ratio, [b]CA/CM=Coarse aggregate-to-cementitious materials ratio

Table 2 Description of variables used in modelling

Variable	Description
C	Compressive strength of concrete (MPa)
P.OPC	% of OPC
P.NP	% of NP
P.SF	% of SF
P.HL	% of HL
CP	Curing period (days)

possible performance (Bingöl *et al.* 2013). A variety of training algorithms are available, however, BP is the most widely used training algorithm for the prediction problems (Hong-Guang and Ji-Zong 2000, Kim *et al.* 2004, Bingöl *et al.* 2013, Khan *et al.* 2013, Chou *et al.* 2011, Cheng *et al.* 2012, Hacene *et al.* 2014).

In consideration of different 10 m height wind speed v10 and the power law exponent index α results shown in

Table 2, the representative upstream typhoon wind fields at different directions used as the input data for training ANN model are determined, which is shown in Tables 1-2.

In BP algorithm, the outputs are calculated and processed forward while the errors are calculated, by taking difference of the model results with the actual outputs, which is propagated backward.

LM is another algorithm that is found to be useful for the prediction problems. It is represented by the following equation

$$\boldsymbol{x_{k+1} = x_k - [J^T(x_k)J(x_k) + \mu_k I]^{-1} J^T(x_k)\vartheta(x_k)} \tag{1}$$

Where J is the Jacobian matrix, μ_K is the Marquardt parameter, I is the unit matrix and x_k is iteration k. This function is designed for minimizing the sums of squares for non-linear functions (Bingöl *et al.* 2013).

CGD is another algorithm that has been often used for the prediction problems. It is considered an advanced search method. Its basic concept is that the search is done in one direction along with the weight space. The previous direction of search is weighted to the minimum with the new direction, in order to find the optimal step size in the new direction. It can be represented by the following equation

$$\boldsymbol{S_{ij}^{new} = -\nabla J^{new} + \alpha S^{old}} \tag{2}$$

Where S is the direction vector, J is the performance surface, and α is a trade-off parameter between the two directions (Vlahogianni *et al.* 2005).

All the models in this study were tested using k-fold cross validation technique, which is a popular approach for prediction modeling (Erdal 2013). In this technique, the sample is divided into 'k' parts, each part is used as a test sample in different runs of model development and the remaining parts are used as training samples for that run. The accuracy of the model is reported as an average value of 'k' runs of model testing. This technique is used to minimize any biasness in the model accuracy due to sampling errors (Cheng *et al.* 2012). For this study, the value of 'k' is set to be 5, which is similar to the work done by Martins and Camões (2013).

Mean absolute percentage error (MAPE), root mean square error (RMSE) and R2 (also known as co-efficient of determination) were estimated as the performance measures for each model. These parameters have been commonly used in studies related to the prediction of compressive strength of concrete (Chou *et al.* 2011, Cheng *et al.* 2012, Khan *et al.* 2013, Martins and Camões 2013). These parameters were estimated as follows

$$\boldsymbol{MAPE = \frac{\left\{\sum_{i=1}^{n}\left[\frac{|y_i - Y_i|}{Y_i}\right] X100\right\}}{n}} \tag{3}$$

$$\boldsymbol{RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - Y_i)^2}{n}}} \tag{4}$$

Table 3 Regression models

Type of model	Model	R^2	RMSE (MPa)	MAPE (%)
Linear	C = 56.76 – 0.73P.NP + 2.21P.SF + 0.14CP	0.84	7.16	13
Quadratic	C = -492.434 + 16.319 OPC – 6.152 P.NP + 0.888 P.SF + 0.355 CP – 0.109 P.OPC^2 + 0.109 P.NP^2 – 0.001 CP^2	0.92	29.59	38

$$\boldsymbol{R^2 = 1 - \frac{SSE}{SST}} \tag{5}$$

Where: y_i is the output from the model, Y_i is the actual output and n is the number of observations in the test sample. *SSE* is the sum of square of errors and *SST* is the total sum of squares for the model (Chou *et al.* 2011, Devore 2012). It is to be noted that these measures have unique characteristics and may or may not conform to one another. R^2 focuses on determining how much variability of the data is captured by the model by taking the ratio of error variance (SSE) and total variance (SST). RMSE depends upon the deviation of the predicted value from the observed value, but it also magnifies the deviations through squaring them, and hence, large deviations become larger while small (fractional) deviations become smaller. Lastly, MAPE also illustrates the deviation between the predicted and observed values but this term is also sensitive to the magnitude of the observed values as the deviation is expressed as the percentage of observed value.

4. Results

4.1 Regression models

As stated earlier, the validity of traditional models becomes doubtful in case of blended cement concretes. Hence, regression models were developed for the prediction of compressive strength of natural pozzolan-based blended cement concrete using the variables stated in Table 2. The results of the regression models are shown in Table 3.

It can be observed from the data in Table 3 that the linear regression model shows better accuracy than the quadratic model. Hence, variables having statistically significant coefficients at 95% confidence level were selected for the model. Moreover, the R-square value for the linear model is, although less than quadratic, it is reasonable. Therefore, the linear regression model would be employed in this study for further comparison with ANN models. The possible reason for having high R2 value could be due to the fact that this parameter indicates the amount of variance of the observed data captured by the model.

The linear regression model shows that the quantity of NP and SF influences the compressive strength of concrete along with the time of curing. The negative coefficient for NP may be ascribed to the inclusion of other and more efficient cementing materials. However, further studies can be conducted to explore the effect of inclusion of NP as a cementing material on other properties of concrete.

It can also be observed that among the variables with statistically significant coefficients, the quantity of SF has the greatest influence on strength. On the other hand, curing

Table 4 Number of neurons and hidden neurons for ANN models

Learning Algorithm	Number of Hidden Neurons	Number of Iterations (Phase I)	Number of Iterations (Phase II)
BP	3	100	N/A
LM	4	200	N/A
CGD	1	100	N/A
BP-LM	2	100	100
BP-CGD	3	100	100
LM-CGD	3	100	200
LM-BP	1	100	100
CGD-BP	1	100	100
CGD-LM	4	100	200

period has the least impact (in terms of coefficient) on the compressive strength of concrete, which is probably due to the inclusion of NP in the concrete mixtures. Hence, it can be realized that the cementing material has a higher impact on the compressive strength of blended cement concrete.

It should also be noted at this point that the reason for insignificant effect of P.OPC could be attributed to the fact that it is correlated with the other cementing materials, i.e. it decreases with an increase in the quantity of the supplementary cementing material.

4.2 ANN Models

The number of hidden neurons and number of iterations for the training algorithm are important parameters for the performance of ANN models. The optimum numbers of these factors were determined by using different combinations and comparing the performance of ANN model of these combinations in terms of RMSE, MAPE, and R2. The combinations giving the best values of these parameters for specific leaning algorithms are given in Table 4. It should be noted that the number of iterations for phase II applies only to the models in which the proposed 2-phase learning process has been used by employing two leaning algorithms in turn.

It can be observed from the data in Table 4 that 100 iterations were sufficient to give the best performance for the given dataset in all cases, except when LM, LM-CGD and CGD-LM algorithms were employed. The maximum number of iterations required was 200 that were employed for the LM algorithm and phase II of LM-CGD combination. The maximum number of hidden neurons was required for LM algorithm which was 4 compared with only one hidden neuron that was needed to be employed for CGD algorithm. Hence, it can be said that ANN models developed with CGD algorithm have a simplified structure and less processing time is required for this particular dataset.

The performance of ANN models, in terms of RMSE, MAPE and R2, with different learning algorithms and their combinations is presented in Tables 5 and 6. It can be observed from the data in Table 5 that LM algorithm has the lowest error parameters among the models with a single learning algorithm, while BP algorithm has the highest error values in this regard.

Table 5 Performance of ANN models with single phase learning algorithm

Learning Algorithm	RMSE (MPa)	MAPE (%)	R^2
BP	10.09	14.52	0.72
LM	5.88	9.22	0.88
CGD	8.02	12.79	0.70

Table 6 Performance of ANN models with 2-phase learning algorithm

Learning Algorithm	Arrangement	RMSE (MPa)	MAPE (%)	R^2
BP-LM	Forward	6.37	11.09	0.85
	Backward	6.87	14.31	0.84
BP-CGD	Forward	5.08	9.95	0.91
	Backward	6.59	11.92	0.85
LM-CGD	Forward	4.34	6.95	0.94
	Backward	4.76	9.15	0.92

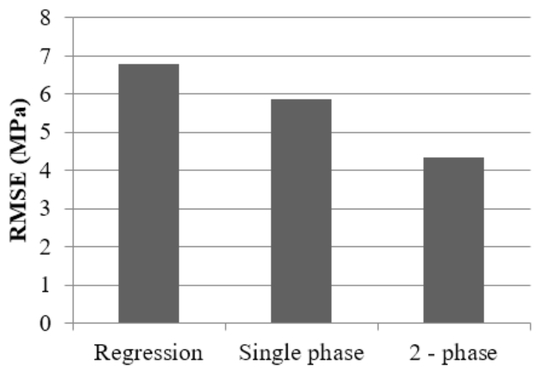

Fig. 3 Comparison of RMSE of best models

Almost all the combinations performed better than single learning algorithm as shown by the lower values of RMSE and MAPE and higher values of R2, as shown in Table 6. The data presented in Table 6 also show that combinations employing CGD algorithm in phase II perform better than other combinations in terms of the above-mentioned performance measures. The best performance is achieved when LM and CGD algorithms are employed in combination using forward order. This could be attributed to the fact that both of these algorithms are advanced search techniques for finding solutions to non-linear problems. LM algorithm is based upon minimizing sums of squares while CGD algorithm is based upon searching solution in a weighted space scenario. Hence, these algorithms complement each other, thereby resulting in higher accuracy models.

5. Discussion

Figs. 3 through 5 compare the performance of best models in terms of RMSE, MAPE and R^2. It is evident from these data that the average difference between single phase ANN model and regression model is approximately 1 MPa

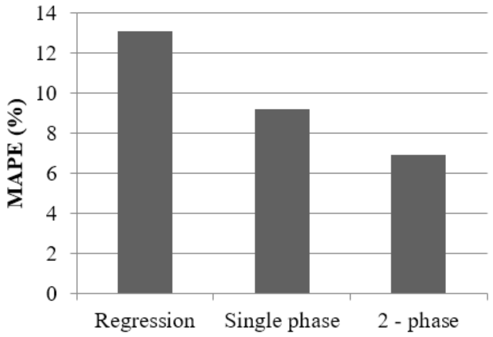

Fig. 4 Comparison of MAPE of best models

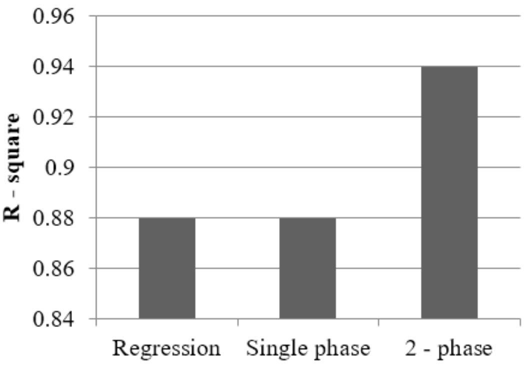

Fig. 5 Comparison of R-square of best models

for RMSE (Fig. 3) and 4% for MAPE (Fig. 4). The average R^2 is the same for both models (Fig. 5). Therefore, it can be said that the performance of regression model and single-phase ANN is comparable with one another. Moreover, the performance of 2-phase learning algorithm was approximately 6 MPa for RMSE (Fig. 3), 6% for MAPE (Fig. 4), which were the lowest among all types of models, and 0.94 for R^2 (Fig. 5) which was the highest. Therefore, it can be concluded that the 2-phase learning algorithms were the best model among all types of models used in this study.

Table 7 presents an overall comparison of the sets of RMSE, MAPE and R^2; obtained using both the single- and 2-phase learning approaches to further reinforce the superior performance of 2-phase learning algorithm over the single-phase learning algorithm. It can be clearly seen from the data in Table 7 that the average performance parameters are better for the 2-phase learning algorithm. Moreover, the difference in highest and lowest values (range) is approximately 5 MPa and 2 MPa for single- and 2-phase learning algorithms, respectively. In addition, the difference between the highest and lowest values in terms of R^2 is 0.18 and 0.10 for single and 2-phase algorithms, respectively. These results indicate that the 2-phase approach is much more consistent in terms of accuracy than the single-phase learning algorithm. A significant variation in the MAPE could be ascribed to the fact that it is more sensitive to the value of observation itself. However, other parameters, such as R^2 and RMSE, mainly focus on the deviation between the observed and predicted values. Chou and Pham (2013) conducted a study to compare the ensemble models with ANN and linear regression. The performance measures for their ANN and regression models

Table 7 Comparison between single- and 2-phase algorithms

Parameter	Single-phase algorithm	2-phase algorithm
Average RMSE (MPa)	5.67	8.00
Range of RMSE (MPa)	5.88-10.09	4.34-6.87
Average MAPE (%)	10.56	12.18
Range of MAPE (%)	9.22-14.52	6.95-14.31
Average R^2	0.77	0.89
Range of R^2	0.70-0.88	0.84-0.94

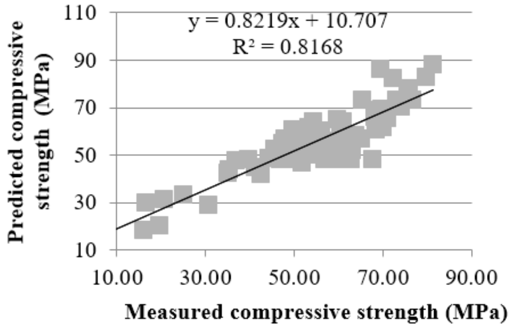

Fig. 6 Measured vs. predicted compressive strength values for regression model

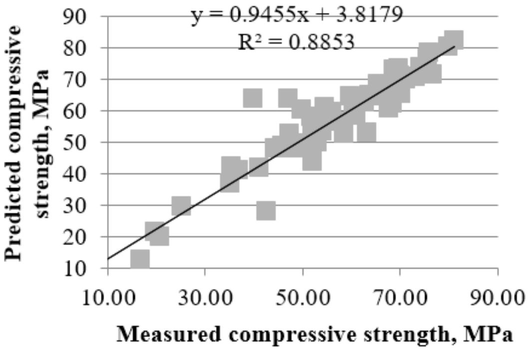

Fig. 7 Measured vs. predicted compressive strength values for best single-phase ANN model

are highly comparable to the data shown in Figs. 3 through 5. They reported approximately an average increase of 0.02 in the R^2 and approximately average reduction of 5% in error values with the use of ensemble models. The 2-phase learning algorithm employed in the present study resulted in an average increase of 0.06 in R^2 value and approximately 6% reduction in the error values compared to the regression model and 2% reduction in the single-phase ANN models. Hence, it can be concluded that the proposed 2-phase algorithm can give better improvement over regression models compared to the ensembles.

Moreover, Erdal *et al.* (2013) used ensemble of different ANN models to enhance their prediction performance. It was reported that the use of ensemble resulted in R^2 value of approximately 0.94 which is comparable to those achieved in the current study utilizing the 2-phase algorithm. Further, an RMSE of 4.5 MPa was obtained by using ensembles that are somewhat similar to the best 2-phase algorithm used in this study.

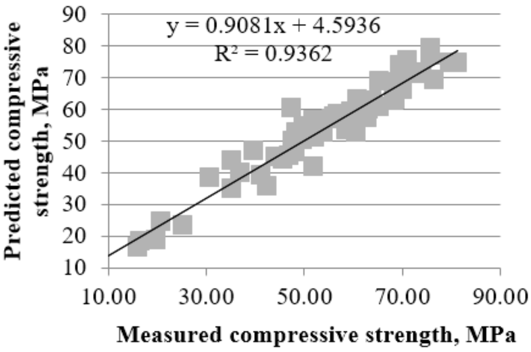

Fig. 8 Observed vs. predicted values for best 2-phase ANN model

Figs. 6 through 8 present the observed and predicted values for the best models using regression and single- and 2-phase ANNs. The points in these figures should fit more closely to the best fit line if the model predictions confirm with the observed values. In other words, these figures are a graphical representation of the R^2 values. It can be noted from these figures that the values predicted by the 2-phase ANN are closest to the best-fit line.

6. Conclusions

In this study, prediction models for compressive strength of cement blended with natural pozzolan were developed using regression models (linear and quadratic) and ANNs. The ANNs were developed using two different approaches, namely traditional single-phase learning algorithm and 2-phase learning algorithm. In the 2-phase learning algorithm, different learning algorithms, such as BP, CGD and LM, were used in combination as a 2-step process for training of ANNs. All the models were developed using 5-fold cross validation approach to avoid any sampling errors. The performance of all the models developed in this study was compared using RMSE, MAPE and R^2 parameters.

• Analysis of the data indicated that the linear regression model gave better results in terms of its accuracy compared with the quadratic model. Furthermore, it was found that both the percentage of supplementary cementing material and the curing period have statistically significant effect on the compressive strength of concrete prepared with the natural pozzolan. The number of hidden neurons and iterations for each ANN model were optimized by iterating different values of these parameters and comparing their RMSE, MAPE and R2 values. It was found that LM is the best single learning algorithm on the basis of these parameters for ANN model development for the dataset used in this study. This ANN model gave better prediction performance than the regression model developed in this study, with an average difference of 1 MPa in terms of RMSE and 4% in terms of MAPE. Furthermore, the 2-phase learning algorithm employing LM-CGD gave the least error values (highest R^2) among all the models studied.

• The superiority of the proposed 2-phase algorithm approach was further reinforced by the fact that the mean values of all performance parameters for this approach were better than those for the single-phase algorithm. Moreover, the range of these parameters was also narrow for 2-phase models compared to the single-phase model in most of the cases. This observation shows that the 2-phase approach proved to be more consistent in terms of accuracy, or less sensitive to sampling errors, compared to the single-phase approach.

• A comparison with earlier studies shows that the proposed 2-phase learning algorithm can give equivalent performance to that achieved by the ensemble models. Hence, the 2-phase learning approach can be useful for ANN model development for forecasting purposes. It is also expected that this approach will also give better prediction accuracy for other datasets compared to the traditional approach.

Acknowledgments

This work was supported by King Fahd University of Petroleum and Minerals (KFUPM) and King Abdulaziz City for Science and Technology (KACST) under National Science, Technology and Innovation Plan (Project No. 10-ADV1365-04).

References

Aticin, U. (2011), "Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network", *Exp. Syst. Appl.*, **38**(8), 9609-9618.

Awwad, M.T. (2004), "Developing a forecasting model of concrete compressive strength using relevance vector machines", *J. Eng. Technol.*, **3**(2), 224-229.

Bingöl, A.F., Tortum, A. and Gül, R. (2013), "Neural networks analysis of compressive strength of lightweight concrete after high temperatures", *Mater. Des.*, **52**(1), 258-264.

Cheng, M.Y., Chou, J.S., Roy, A.F.V. and Wu, Y.W. (2012), "High-performance concrete compressive strength prediction using time-weighted evolutionary fuzzy support vector machines inference model", *Automat. Constr.*, **28**(1), 106-115.

Chou, J. and Pham, A. (2013), "Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength", *Constr. Build. Mater.*, **49**(1), 554-563.

Chou, J., Chiu, C., Farfoura, M. and Al-Taharwa, I. (2011), "Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques", *J. Comput. Civil Eng.*, **25**(June), 242-253.

Davoodi, E. and Khanteymoori, A.R. (2010), "Horse racing prediction using artificial neural networks", *Proceedings of the Recent Advances in Neural Networks, Fuzzy Systems & Evolutionary Computing*, 155-160.

De Larrard, F. and Sedran, T. (1994), "Optimization of ultra-high-performance concrete by the use of a packing model", *Cement Concrete Res.*, **24**(6), 997-1009.

Devore, J. (2012), *Probability and Statistics for Engineering and Sciences*, 8th Edition, Cengage Learning, California, U.S.A.

Engelbrecht, A.P. (2007), *Computational Intelligence: An Introduction*, John Wiley & Sons, West Sussex, U.K.

Erdal, H.I. (2013), "Two-level and hybrid ensembles of decision trees for high performance concrete compressive strength prediction", *Eng. Appl. Artif. Intell.*, **26**(7), 1689-1697.

Feng, M.Q. and Bahng, E.Y. (1999), "Damage assessment of jacketed RC columns using vibration tests", *J. Struct. Eng.*,

125(3), 265-271.

Feng, M.Q. and Kim, J.M. (1998), "Identification of a dynamic system using ambient vibration measurements", *J. Appl. Mech.*, **65**(2), 1010-1023.

Gazder, U. and Hussain, S.A. (2013), "Traffic forecasting for King Fahd Causeway using artificial neural networks", *Proceedings of the 15th International Conference on Computer Modelling and Simulation (UKSim)*.

Guang, N.H. and Zong, W.J. (2000), "Prediction of compressive strength of concrete by neural networks", *Cement Concrete Res.*, **30**(1), 1245-1250.

Hacene, S.M.A.B., Ghomari, F., Schoefs, F. and Khelidj, A. (2014), "Probabilistic modelling of compressive strength of concrete using response surface methodology and neural networks", *Arab. J. Sci. Eng.*, **39**(6), 4451-4460.

Hong-Guang, N. and Ji-Zong, W. (2000), "Prediction of compressive strength of concrete by neural networks", *Cement Concrete Res.*, **30**(1), 1245-1250.

Jain, A.K., Mao, J. and Mohiuddin, K.M. (1996), "Artificial neural networks: A tutorial", *IEEE Comput.*, **29**(3), 31-44.

Kasperkiewicz, J., Racz, J. and Dubrawski, A. (1995), "HPC strength prediction using artificial neural network", *J. Comput. Civil Eng.*, **9**(4), 279-284.

Khan, U.S., Ayub, T. and Rafeeqi, S.F.A. (2013), "Prediction of compressive strength of plain concrete confined with ferrocement using artificial neural network (ANN) and comparison with existing mathematical models", *Am. J. Civil Eng. Archit.*, **1**(1), 7-14.

Khatibinia, M., Feizbakhsh, A., Mohseni, E. and Ranjbar, M.M. (2016), "Modeling mechanical strength of self-compacting mortar containing nanoparticles using wavelet-based support vector machine", *Comput. Concrete*, **18**(6), 1065-1082.

Kim, J.M., Kim, D.K., Feng, M.Q. and Yazdani, F. (2004), "Application of neural networks for estimation of concrete strength", *J. Mater. Civil Eng.*, **16**(June), 257-264.

Martins, F.F. and Camões, A. (2013), "Prediction of compressive strength of concrete containing fly ash using data mining techniques", *Cement Wapno Beton*, **(1)**.

Oh, J.W., Lee, I.W., Kim, J.T. and Lee, G.W. (1999), "Application of neural networks for proportioning of concrete mixes", *ACI Mater. J.*, **96**(1), 61-67.

Russell, H.G. (1999), "ACI defines high-performance concrete", *Concrete Int.*, **21**(2), 56-57.

Sadrmomtazi, A., Sobhani, J. and Mirgozar, M.A. (2013), "Modeling compressive strength of EPS lightweight concrete using regression, neural network and ANFIS", *Constr. Build. Mater.*, **42**(1), 205-216.

Santamaría, A., Orbe, A., Losañez, M.M., Skaf, M., Ortega-Lopez, V. and González, J.J. (2016), "Self-compacting concrete incorporating electric arc-furnace steelmaking slag as aggregate", *Mater. Des.*, In Press.

Vlahogianni, E.I., Karlaftis, M.G. and Golias, J.C. (2005), "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach", *Transp. Res. Part C: Emerg. Technol.*, **13**(3), 211-234.

Wu, C., Pan, Z. and Meng, S. (2016), "Cyclic constitutive model for strain-hardening cementitious composites", *Mag. Concrete Res.*, **68**(22), 1-10.

Yeh, I.C. (1998), "Modelling of strength of high-performance concrete using artificial neural networks", *Cement Concrete Res.*, **28**(12), 1797-1808.

Zurada, J.M. (1992), *Introduction to Artificial Neural Systems*, St. Paul: West, Kentucky, U.S.A.

CC