ANN models for nano silica/ silica fume concrete strength prediction
Osama A. Hodhod, Mohamed S. Khalafalla and Mohamed S.M. Osman

Department of Civil Engineering, Faculty of Engineering, Cairo University, Giza, Egypt; Construction Research Institute, National Water Researches Center - NWRC, Cairo, Egypt

ABSTRACT
Artificial Neural Network (ANN) models were built to expect the compression strength of various types of concrete incorporating nano silica (NS) and silica fume (SF) as partial cement replacement. The mixtures data used in the networks model, which was collected from previous researchers, studied the effect of NS and SF on concrete. The previous researchers experimentally tested the specimens containing up to 10% of NS and up to 20% of SF as a partial cement replacement at age 7 and 28 days. A total of 488 experiments were used as data sets to train and test the network. The input parameters, like cement content, nano silica, water to cement ratio, and aggregate type and proportion, were varied for each experiment. Three sets of data were modeled using ANNs for both ages 7 and 28 days to predict the strength. The maximum average error of the three models did not exceed 10% of the exact result.

Review
For two centuries, concrete has been used as a construction material in the construction industry. Therefore, it has become important to find new technologies or alternative materials to decrease the amount of material used in concrete. In megastructures, which consume a large amount of cement, reducing or replacing Portland cement without reducing the properties of concrete is an important goal. Accordingly, this would decrease the amount of energy consumed and preserve the environment by limiting emitted carbon dioxide. For example, to produce one ton of clinker, which is the main content of cement, about 1.7 tons of raw material is needed. That production leads to the emission of greenhouse and other gases into the atmosphere (Meyer, 2009). The production of 1-ton of clinker emits about 1 ton of carbon dioxide (Gartner, 2004). This example illustrates the importance of pozzolan and cementitious materials in concrete production, which may lead to reduced production of cement without the negative effect of concrete performance (Heidari & Tavakoli, 2013). Many studies mention the effect of SF as a pozzolan in reducing cement in concrete and that it is significantly effective for concrete durability and performance (Bhajan & Sengupta, 2002; Mazloom, Ramezanianpour, & Brooks, 2004). In addition, other compositions of pozzolan and silica fume were used to manufacture high strength concrete (Shang, 2000). In recent years, NS has been more widely used. This is because of the positive effect of NS on cement mortar and concrete properties, according to various studies (Byung, Chang-Hyun, Ghi-ho, & Jong-Bin, 2007; Byung, Kim, & Lim, 2008; Li, Yao, & Ou, 2004; Qing, Zenan, Deyu, & Rongshen, 2007; Tao, 2005; Wan, Hyun, Tae, & Park, 2007). There was clearly an improvement in the mechanical properties of concrete which contained NS and pozzolan in the same mixing (Heidari & Tavakoli, 2013; Li, 2004). Min–Hong Zhang and Jahidul Islam reported that using NS accelerates the rate of cement hydration and has an effect on the high early strength of concrete (Zhang & Islam, 2012). Various studies mention that replacing cement with NS improves the microstructure of concrete while mixing, which leads to enhanced concrete performance (Beri, Benjain, Omran, Nik, & Nikbin, 2013; Gaitero, Campillo, & Guerrero, 2008; Jalal, Mansouri, Sharifipour, & Pouladkhan, 2012; Singh et al., 2013).

The human brain is the most complicated organ in the human body. It can be described as a complicated computer with super abilities. The human brain consists of small elements called neurons. There are complex connections between these neurons, which construct a network. (Chithra, Kumar, Chinnaraju, & AlfinAshmita, 2016). This is the main concept of ANNs. They are like the human brain in construction and in the way they work. They consist of neurons with connections between them. ANNs use the idea of the natural neural system, using software and industrial electronic subject to certain conditions. They also have the ability to apply nonlinearity to predict input-output relationships (Chithra et al., 2016). There are...
different engineering applications in more than one branch such as meteorology, hydrology that have applied ANNs. They have implemented ANNs for both qualitative and quantitative expectation of the variables implicated in water resource modeling (Kisi, 2007).

Currently, ANNs are being used by various researchers to predict the compressive strength of concrete (Siddique, Aggarwal, & Aggarwal, 2011). ANNs were applied to relate the strength parameters of the experimental values to the data obtained from the literature for self-compacting concrete containing ash. A feedforward neural network with a gradient descent technique been used to create models for predicting the strength of High-Performance Concrete by Yeh (1999). Because of the complications of determining the compressive strength of lightweight concrete subject to high temperature using traditional methods, Bingol employed neural networks (Bingöl et al., 2013). Trocoli demonstrated the ability ANNs have to predict the compressive strength for a complex system of concrete having Construction and Demolition Wastes (CDW) in order to relate seventeen input variables to one output variable (Dantas, Leite, & Nagahama, 2012).

This comprehensive study focuses on applying ANNs to predict the compressive strength of concrete mixtures that contain nano silica and silica fume, based on previously studied data and experiments. The goal was to create a model and use it as the kernel of an interface application; the input data entered would be the required compressive strength (Fcu) required, cement, and aggregate type available. The application simulated the network to produce a mixture that could meet the requirements. The main goals to be achieved in this study included the following:

(A) Use the Artificial Neural Networks in creation probabilistic models for the expectation of the compressive strength of concrete with silica fume and nano silica.

(B) Verification of the performance of each model.

Data collection

A total of 488 concrete mixes were collected from 24 papers focused on studying the effect of NS and SF on concrete properties. The data collected had inputs that varied as follows. The amount of cement ranged from 280 Kg/m³ to 650 Kg/m³. The water to cement ratio (W/C) ranged from 0.2 to 0.63. The amount of coarse aggregate (CA) ranged from 372 kg/m³ to 2200 kg/m³; whereas the amount of fine aggregate (FA) ranged from 492 kg/m³ to 1263.2 kg/m³. The maximum percent of cement replaced with silica fume (SF) in the mixing proportion was 20%. The maximum percent of cement that was replaced with nano silica (NS) in the mixing proportion was 10%. The NS diameter ranged from 7 nm to 100 nm. The maximum percentage of superplasticizer (SP) added to the mix was 5.5%.

Material properties

Data was collected about the effect of NS on concrete properties or comparing properties of concrete containing NS and different admixtures. There was clearly variation and diversity in the material used in the studies. It has been considering that the accuracy of the models been acceptable if the difference between the predicted and the target result not exceed 10%. To create an accurate ANN model, it was recommended that the number of differences in input data be limited.

Therefore, the data collected were the common inputs, such as Ordinary Portland Cement of 42.5 grade with a specific gravity of 3.15. The fine aggregate used was river sand with a specific gravity of 2.66 and bulk density of 1780 kg/m³. The input had few variations in properties of coarse aggregate and NS. These varieties were taken into consideration for the model. Crushed basalt, crushed limestone, crushed granite, crushed dolomite, crushed dolerite, crushed scoria, and recycled aggregate were the types of coarse aggregate used. NS was used as cement partial replacement with size ranges from 7 to 100 nm and a specific gravity of 1.3 to 1.32.

Data input

The data collected differed in water to cement ratio. Some of the studies mentioned the total amount of water used in the study and others mentioned the water-binder ratio. It was necessary for the water data from all the samples to be in the same format; so the water-binder ratio was calculated for all samples. Also, the differences in types of coarse aggregate had a significant effect on the compressive strength of early and later age concrete.

It was necessary to create categories for the different types of aggregate. The inputs for the model were categorized as discussed below. Samples that used crushed basalt were categorized as category No. (01). The samples that used crushed limestone were categorized as category No. (02). Samples that used crushed granite were categorized as category No. (03). The samples containing crushed dolomite were categorized as category No. (04). Samples that used crushed dolerite were categorized as category No. (05). The samples containing crushed scoria were categorized as category No. (06). Lastly, samples containing recycled aggregate were categorized as category No. (07). It was difficult to justify an accurate size of the NS because of its amorphous property. X-ray diffraction (XRD) diagrams for NS, Figure 1, showed that the peak of the scan had more than one value specifying it, so the size of the Nanoparticles was calculated as the average of the range.

Data test and distribution

It was necessary to test the distribution of the samples that were used to train the model in order to
justify the range of data and to ensure that all samples were distributed throughout the range between the maximum and minimum of each of the inputs. As shown in Figure 2(a), the distribution of cement content of the sample was uniform from the minimum to the maximum. Figure 2(b) shows the distribution of the water-binder ratio from the minimum to the maximum values of the samples.

**Data scale**

It was important to scale the data by fixing the minimum and maximum values of the variables in the collected data to \(-1.00\) and \(+1.00\) and to confine the collected data to values between the minimum and maximum. This step was very important and allowed the training program to work within a fixed range, which allowed for more perfect results. The data limits were as shown in Table 1.

| Inputs          | Minimum | Maximum |
|-----------------|---------|---------|
| (Cement) Kg/m³  | 280     | 650     |
| (W/C) %         | 0.2     | 0.6364  |
| (CA/FA) %       | 0.4898  | 2.33    |
| (CA Type)       | 1       | 7       |
| (SF) %          | 0       | 20      |
| (NS) %          | 0       | 10      |
| (NS Size) nm    | 0       | 100     |
| (SP) %          | 0       | 5.5     |
| Age (days)      | 7       | 28      |

**Methodology**

According to (Kisi, 2007), there are several algorithms that can be implemented in ANN modeling, illustrated as follows:

- Bayesian regularization.
- Levenberg–Marquardt.
- BFGS Quasi-Newton.
- Powell–Beale conjugate gradient.
- gradient descent.
- gradient descent with either momentum or adaptive learning rate or both.
- Fletcher–Powell conjugate gradient.
- random order incremental training with learning.
- Polak–Ribiere conjugate gradient.
- batch training with weight and bias learning rules.

Sobhani checked all various algorithms available and concluded that the Levenberg–Marquardt algorithm is the most commonly used training algorithm because of its robustness and speed (Sobhani, Najimi, Pourkhorsheid, & Parhizkar, 2009). Therefore, in this paper research, the Levenberg–Marquardt (LM) algorithm has been implemented to create ANN models. The working flow of this algorithm is to use layered feed-forward networks in which the neurons are
arranged in layers, signals are moved forward, and errors are spread backward. (Figure 3).

There are several ways to generate the ANN networks, in this study the neural network models were developed using the Neural Network Toolbox in the MATLAB software. three ANN models were created, namely Net1, Net2, and Net3. The models were generated with different hidden layers and various neurons in the hidden layers. From the total data, approximately 70% was considered for training. Out of the remaining 30%, 15% was considered for testing and 15% was considered for validation. A total of 464 experimental concrete mixing proportions were used to train the network. Also, there was another test for the network which used 24 samples that were not simulated to the network before. In Artificial Neural Networks are based on training the inputs data with actual values in outputs data and adjustments of weights of each parameter according to the differences between the predicted and actual values. the transfer function has been applied in this process is a nonlinear sigmoidal function. this process should be limited by the number of iterations, the numbers of iterations is termed an epoch. The epoch number is set to stop at a number of times that the weights between layers neurons were reinitialized until a satisfactory model with the most accurate possible correlation was acquired.

The structure of network Net2, as shown in Figure 3, consisted of an input layer, hidden layers, and an output layer. All networks that were modeled have the same input and output layers and neurons and differ only in the number of hidden layers and neurons on each hidden layer. The input layers consisted of 10 neurons and the output layers consisted of one neuron. The number of hidden layers varied between the networks created.

The design of networks with different hidden layers and different neurons in each layer come from try and error and the final design of the three networks are the most accurate design which obtains an accurate result within a small time., as shown in Table 2.

It was noted that when a network was retrained, the weights of the links between the layers changed, which caused an accurate result sometimes and caused an unacceptable error in another retraining. Therefore, the model was retrained multiple times and the results were considered the most accurate for the three network model and were saved, see Figure 4. The results are presented in the following sections.

**Network performance**

Figures 5, 6, and 7 show the network performance and the best validation performance for the three networks. The best validation performance mean squared error (MSE) for Net1 was 0.0111265 at epoch 67 through 73, for Net2 was 0.0067809 at epoch 38 through 44, and for Net3 was 0.00843 at epoch 21 through 27.

The MSE as shown in Figures 5, 6, and 7 is close to zero. The smaller the MSE, the closer to finding the line of best fit. Depending on the data, it may be impossible to get a very small value for the MSE. For example, the network trained the data through 73 epochs for Net1 and the best MSE was at the epoch 67.

| Table 2. The number of hidden layers and neuron on each hidden layer for the networks. |
| No | Name of network | No. of hidden layers | No. of neurons in 1st hidden layer | No. of neurons in 2nd hidden layer | No. of neurons in 3rd hidden layer | No. of neurons in 4th hidden layer |
|----|----------------|---------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| 01 | Net1           | 4                   | 5                               | 4                               | 3                               | 2                               |
| 02 | Net2           | 4                   | 10                              | 5                               | 3                               | 2                               |
| 03 | Net3           | 3                   | 15                              | 10                              | 5                               | 0                               |

**Figure 3.** Artificial Neural Network (ANN) topology for Net2.
Regression and correlation factor

The correlation coefficients (R) obtained for training, testing, validation, and all three phases together as average for each model are presented in network regression charts in Figures 8, 9, and 10 for Net1, Net2, and Net3, respectively. The correlation coefficient (R) was almost equal to one. This indicated that the neural network models had a high degree of fitness to the actual values.

Figures 8, 9, and 10 show the regression of the network through the four phases of training, validation,
and testing, as well as the average of the three phases. In the figures, the target result is presented on the horizontal axis and the output from the trained data according to the weights between neurons is presented on the vertical axis. The dotted line across the figure is the exact result, which if the fitting line of all experiments matched or aligned with it that would mean the correlation coefficient (R) was equal to one and all data was trained perfectly. The figures presented are for the training data, validation data, test data, and all data together on one chart. It could be considered that the regression figures for all data present the correlation coefficient (R) and the equation for the network. All data that was trained is summarized in Table 3.
Figure 8. Correlation coefficient (R) for Net1 trial 01, best results.

Figure 9. Correlation coefficient (R) for Net2 trial 01, best results.
From Figures 7, 8, and 9 and from Table 3 it can be noted that network Net3 is close to Net2 in the simulation and both are not far from Net1. That means that all data which was yet to be simulated with the three networks would be close together, despite Net3 having the best accuracy in prediction.

Test the accuracy of the networks

The second test used 24 samples that had not already been used to train the network to check the accuracy of the machine and to get the result of the predicted values and the error, which are presented in Table 4. The acceptable limit of error was ±10%. Some samples had an unacceptable percent error, like sample No 11, 14, 15 and 24 for Net1; 05, 06, 14, 18, and 24 for Net2; and 2, 16, and 18 for Net3. It can be noted that not all samples that had a large error percentage result from Net2 had a large error percentage in Net1. The reasons behind that difference in the errors from samples to others in the different networks are the learning paths and weights between neurons and hidden layers been changed between the different networks and consequently lead to the networks became accurate in such samples than others and different through the different networks, so a third Network model, Net3, was created.

It is presumed that the reasons behind the various errors throughout the experiments from one network to another are the differences in weights between the hidden layers and the number of neurons. In addition, the differences between the network structures are the number of hidden layers and the neurons in each layer.

It was clear that the percent error for sample number 24 was unacceptable for Net1, Net2, and Net3. The reason behind this error may be that the limit of data is in the same range as these samples, so the machine didn’t learn the weight of these samples. Nevertheless, it would not be accurate to say that the exact reasons behind the errors of some values are known.

| No | Name of network | Correlation Coefficient (R) | Fitting line equation |
|----|-----------------|-----------------------------|-----------------------|
| 01 | Net1            | 0.9818                      | 0.69 Target – 0.014   |
| 02 | Net2            | 0.9663                      | 0.96 Target – 0.0014  |
| 03 | Net3            | 0.9867                      | 0.97 Target – 0.0086  |

Figure 10. Correlation coefficient (R) for Net3 trial 01, best results.

Table 3. The correlation coefficient (R) and the equation of the fitting line.
It was not clear which element had the most significant effect in predicting the best result; therefore, the best way for assessing all three networks was to determine the average of the outputs from the three ANN models that decreased the error percentage to below ±10%.

### Conclusion

In this study, 488 different concrete mixes incorporating Nano silica and Silica fume were collected and the compressive strength was tested for each mix at curing ages of 7 and 28 days. The compressive strength of different types of concrete mixes with various types of aggregate were found to increase up to the replacement of 20% cement with Silica fume and 10% cement with Nano silica. Models were generated to predict the compression strength of these concrete mixes by implementing Artificial Neural Networks. Three models were created using neural network techniques. The three networks varied in the number of hidden layers and number of neurons on each hidden layer. The performance, correlation coefficient, and simulation of each model were measured as pointed below.

- The Mean Squared Error (MSE) for all models was very low and close to be zero. The correlation coefficient (R) obtained for all the neural network models was greater than 0.95. This shows the acceptability of this ANN models.
- The three networks used to bean application to predict the accurate compressive strength, so it was recommended to simulate the three networks and consider the average result. That justified the diversity of the errors from the different networks and the average error not exceeds 9.14% and that was the best can be done.
- The number of hidden layers and neurons in each hidden layer had an effect on the result of the prediction. Because of the weights between hidden layers, it is not reasonable to say the main element has the most significant effect in predicting the best result.
- In applying the ANN or machine learning it is important to collect all data available and that data set is one of the important factors in machine learning to obtain an accurate result and due to the small numbers of the dataset, it is hard to obtain the best result.

### ORCID

Mohamed S. Khalafalla [http://orcid.org/0000-0001-8631-1675](http://orcid.org/0000-0001-8631-1675)

### References

Beigi, M. H., Berenjian, J., Omran, O. L., Nik, A. S., & Nikbin, I. M. (2013). An experimental survey on compressive strength of silica fume concrete using statistical methods. *Cement and Concrete Research*, 32, 1391–1394.
Boga, A. R., Ozturk, M., & Topcu, I. B. (2012). Using ANN and ANFIS to predict thermomechanical and chloride-depermeability properties of concrete containing GGBFS and CNI. Composite Part B, 45, 688–696.

Boukhatem, B., Ghrici, M., Kenai, S., & Hamou, A. T. (2011). Prediction of efficiency factor of ground – Granulated blast – Furnace slag of concrete using artificial neural network. ACI Material Journal, 108, 55–63.

Byung, W. J., Chang-Hyun, K., Ghi-ho, T., & Jong-Bin, P. (2007). Characteristics of cement mortar with nano-SiO2 particles. Construction and Building Materials, 21, 1351–1355.

Byung, W. J., Kim, C. H., & Lim, J. H. (2008). Investigations on the development of powder concrete with nano-SiO2 particles. KSCE Journal of Civil Engineering, 11(1), 37–42.

Chithra, S., Kumar, S. R. R. S., Chinnaraju, K., & AlfinAshmita, F. (2016). A comparative study on the compressive strength prediction models for high performance concrete containing nano silica and copper slag using regression analysis and artificial neural networks. Construction and Building Materials, 114, 528–535.

Dantas, A., Leite, M., & Nagahama, K. (2012). Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. Construction and Building Materials, 38, 717–722.

Gaitero, J. J., Campillo, I., & Guerrero A. (2011). Prediction of the compressive strength of concrete incorporating nano-SiO2. Cement and Concrete Research, 38, 1489–1498.

Gartner, E. (2004). Industrially interesting approaches to low-CO2 cements. Cement and Concrete Research, 34(9), 389–400.

Heidari, A., & Tavakoli, D. (2013). A study of the mechanical properties of ground ceramic powder concrete incorporating nano-SiO2 particles. Construction and Building Materials, 38, 255–264.

Jalal, M., Mansouri, E., Sharifipour, M., & Pouladkhan, A. R. (2012). Mechanical rheological, durability and microstructural properties of high-performance self-Compacting concrete containing SiO2 micro and nanoparticles. Material Design, 34, 389–400.

Khanzadi, M., & Behnood, A. (2009). Mechanical properties of high-strength concrete incorporating copper slag–gascoarse aggregate. Construction and Building Materials, 23, 2183–2188.

Kisi, O. (2007). Streamflow forecasting using different artificial neural network algorithms. ASCE, 12:5(532), 1084–1099.

Li, G. (2004). Properties of high-volume fly ash concrete incorporating nano-SiO2. Cement and Concrete Research, 34, 1043–1049.

Li, H., Xiao, H.-G., Yuan, J., & Ou, J. (2004). Microstructure of cement mortar with nano-particles. Composites, 35, 185–189.

Maitha, H. (2011). Using MATLAB to develop a partial artificial neural network model for predicting global solar radiation. In Ali Assi Ed., Engineering Education and Research Using MATLAB (pp. 219–238). Intech.

Mazloom, M., Ramezanianpour A. A., & Brooks, J. J. (2004). Effect of silica fume on mechanical properties of high-strength concrete. Cement & Concrete Composites, 26, 347–357.

Meyer, C. (2009). The greening of the concrete industry. Cement and Concrete Composition, 31, 601–605.

Qing, Y. (2001). Research on the comparison of pozzolanic activity between nano-SiO2 and silica fume. Concrete, 3, 19–22.

Qing, Y., Zenan, Z., Deyu, K., & Rongshen, C. (2007). Influence of nano-SiO2 addition on properties of hardened cement paste as compared with silica fume. Construction and Building Materials, 21, 539–545.

Ramana, K. V. S., Anita, T., Mandal, S., Kaliappan, S., Shaikh, H., Sivaprasad, P. V., . . . Khatak, H. S. (2009). Effect of different environmental parameters on pitting behavior of AISI type316L stainless steel: Experimental studies and neural network modeling. Material Design, 30, 3770–3775.

Shang, M. J. (2000). High strength concrete containing natural pozzolan and silica fume. Cement and Concrete Composites, 22, 399–406.

Siddique, R., Aggarwal, P., & Aggarwal, Y. (2011). Prediction of compressive strength of self-Compacting concrete containing bottom ash using artificial neural networks. Advancing Engineering Software, 42, 780–786.

Slonski, M. (2010). A comparison of model selection methods for compressive strength prediction of high–Performance concrete using neural networks. Computer Structure, 88, 1248–1253.

Sobhani, J., Najimi, M., Pourkhorsheid, A. R., & Parhizkar, T. (2009). Prediction of the compressive strength of no–Slump concrete: A comparative study of regression, neural network and ANFIS models. Construction and Building Materials, 24, 709–718.

Tao, J. (2005). Preliminary study on the water permeability and microstructure of concrete incorporating nano-SiO2. Cement and Concrete Research, 35, 1943–1947.

Wan, J. B., Hyun, K. C., Tae, C. H., & Park, J. B. (2007). Characteristics of cement mortar with nano-SiO2 particles. Construction and Building Materials, 21, 1351–1355.

Yeh, I. C. (1999). Design of high–Performance concrete mixture using neural networks and nonlinear programming. Computer Civil Engineering, 13(1), 36–42.

Zhang, M., & Islam, J. (2012). Use of nanosilicat in construction of high-volume fly ash-slag. Construction and Building Materials, 29, 573–580.