



The use of GA–ANNs in the modelling of compressive strength of cement mortar

Sedat Akkurt^{a,*}, Serhan Ozdemir^{a,1}, Gokmen Tayfur^{b,2}, Burak Akyol^{c,3}

^aMechanical Engineering Department, Izmir Institute of Technology, 35437 Izmir, Turkey

^bCivil Engineering Department, Izmir Institute of Technology, 35437 Izmir, Turkey

^cQuality Department, Çimentaş İzmir Çimento Fabrikası Türk A.Ş. Işıkkent, 35070 Izmir, Turkey

Received 27 March 2002; accepted 6 December 2002

Abstract

In this paper, results of a project aimed at modelling the compressive strength of cement mortar under standard curing conditions are reported. Plant data were collected for 6 months for the chemical and physical properties of the cement that were used in model construction and testing. The training and testing data were separated from the complete original data set by the use of genetic algorithms (GAs). A GA–artificial neural network (ANN) model based on the training data of the cement strength was created. Testing of the model was also done within low average error levels (2.24%). The model was subjected to sensitivity analysis to predict the response of the system to different values of the factors affecting the strength. The plots obtained after sensitivity analysis indicated that increasing the amount of C_3S , SO_3 and surface area led to increased strength within the limits of the model. C_2S decreased the strength whereas C_3A decreased or increased the strength depending on the SO_3 level. Because of the limited data range used for training, the prediction results were good only within the same range. The utility of the model is in the potential ability to control processing parameters to yield the desired strength levels and in providing information regarding the most favourable experimental conditions to obtain maximum compressive strength.

© 2003 Elsevier Science Ltd. All rights reserved.

Keywords: Modelling; Compressive strength; Artificial neural networks; Genetic algorithms

1. Introduction

Compressive strength of Portland cement mortar is the major property that defines its quality and depends on several factors that need to be controlled during manufacture. These factors range from the C_3S content to the fineness of the milled product and have varying degrees of effect on strength. An analytical model to describe the effects of each of these factors on strength can be very complex. Artificial neural networks (ANNs) can be used for this purpose as a tool for prediction modelling of

strength. Its use for concrete strength prediction was previously studied [1], but the use of genetic algorithms (GAs)–ANNs for prediction of the strength of cement mortar has not yet been reported. Other modelling studies involved extrapolation method, regression analysis methods and fuzzy logic.

The extrapolation method proposed by de Siquera Tango [2] involves the prediction of 28-day strength based on the 2- and 7-day strengths via the use of the AMEBA method. Tsivilis and Parissakis [3], on the other hand, applied stepwise regression analysis techniques to develop a mathematical model. The predicted variables in that study were 2-, 7- and 28-day strengths as a function of chemical and physical parameters. In a recent paper of Fa-Liang [4], fuzzy logic was proposed as a modelling technique of compressive strength development of cement. Sensitivity analysis and the resulting response plots of the prediction models were not performed in these studies.

Data from a local cement plant for standard curing were employed in this research by feeding them to a GA-based

* Corresponding author. Tel.: +90-232-498-6523; fax: +90-232-498-6505.

E-mail addresses: sedatakkurt@iyte.edu.tr (S. Akkurt), serhanozdemir@iyte.edu.tr (S. Ozdemir), gokmentayfur@iyte.edu.tr (G. Tayfur), kalite@cimentas.com (B. Akyol).

¹ Tel.: +90-232-498-6585; fax: +90-232-498-6505.

² Tel.: +90-232-498-6280.

³ Tel.: +90-232-472-1050/254.

ANN to create a model to describe the cement strength. (Because testing the strength of cement is ordinarily performed by mixing cement with sand and water [5], the cement mortar is also commonly referred to as cement.) The data were collected during 6 months of plant operation. Because the plant operational parameters occasionally vary, the data had some variations that must be taken care of before modelling. In other words, the average strength of the shipped product varied as a function of time. Therefore, modelling the system for the data for the first 4 months and testing the quality of the model for the remaining 2 months could produce a biased model. To alleviate this issue, a GA-based neural network simulator algorithm for the prediction of cement strength is developed in this study. ANNs are extensively used to model complex systems in a wide range of fields [6–9].

2. Artificial neural networks

In this study, the common three-layer feed-forward type of ANNs, as shown in Fig. 1, is considered. In a feed-forward network, the input quantities are first normalized to a range of 0.1 to 0.9 via Eq. (1), and then fed into input layer neurons, which in turn pass them on to the hidden layer neurons after multiplying by a weight. A hidden layer neuron adds up the weighted input received from each input neuron, associates it with a bias, if any, and then passes the result on through a nonlinear transfer function. The output neurons do the same operation as that of a hidden neuron.

$$X_i = 0.1 + 0.8(X_i - X_{min_i}) / (X_{max_i} - X_{min_i}) \tag{1}$$

where X_{max_i} and X_{min_i} are the maximum and minimum values of the i th node in the input layer for all the feed data vectors, respectively. The weights were assigned a random value between -1 and 1 .

Before its application to any problem, the network is first trained, whereby the difference between the target output and the calculated model output at each output neuron is minimized by adjusting the weights and biases through some training algorithm. During training, a neuron receives inputs from a previous layer, weights each input with a

prearranged value, and combines these weighted inputs. The combination of the weighted inputs is represented as:

$$net_j = \sum x_i v_{ij} \tag{2}$$

where net_j is the summation of the weighted input for the j th neuron, x_i is the input from the i th neuron to the j th neuron, and v_{ij} is the weight from the i th neuron in the previous layer to the j th neuron in the current layer.

The net_j is passed through a transfer function to determine the level of activation. If the activation of a neuron is strong enough, it produces an output that is sent as an input to other neurons in the successive layer. In this study, sigmoid function is employed as an activation function in the training of the network:

$$f(net_j) = \frac{1}{1 + e^{-net_j}} \tag{3}$$

The learning of ANNs is accomplished by a back-propagation algorithm where information is processed in the forward direction from the input layer to the hidden layer and then to the output layer (Fig. 1). The objective of a back-propagation network is, by minimizing a predetermined error function, to find the optimal weights that would generate an output vector $Y=(y_1, y_2, \dots, y_p)$ as close as possible to target values of output vector $T=(t_1, t_2, \dots, t_p)$ with a selected accuracy.

A predetermined error function has the following form [9]:

$$E = \sum_P \sum_p (y_i - t_i)^2 \tag{4}$$

where y_i is the component of an ANN output vector Y , t_i is the component of a target output vector T , p is the number of output neurons; and P is the number of training patterns.

The least square error method, along with a generalized delta rule, is used to optimize the network weights. The gradient descent method, along with the chain rule of derivatives, is employed to modify network weights as:

$$v_{ij}^{new} = v_{ij}^{old} - \delta \frac{\partial E}{\partial v_{ij}} \tag{5}$$

where δ is the learning rate that is used to increase the chance of avoiding the training process being trapped in a local minima instead of a global minima.

The details of ANNs can be obtained from the literature [6–7].

3. GA as an average equalizer

3.1. Description of the problem

In ANN programming, the data is unequally divided in two; the first batch is used for training and the rest is for

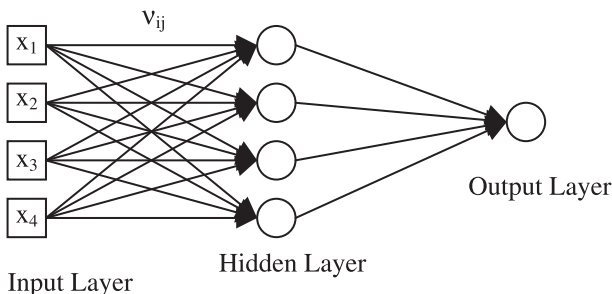


Fig. 1. A typical back-propagation ANN model with four neurons.

testing of the model. For robust modelling, the network needs to be fed with a statistically balanced data. That is, the training and testing data sets should have approximately the same minimum to maximum range and average strength values as in the parent data set. In this study, there were a total of 150 data sets with minimum, average and maximum strength values of 47.6, 53.14 and 58.4 MPa, respectively. We initially used the first 100 of these data sets, in chronological order, for training of the model and the remaining 50 for testing. We found out that the resulting model was biased due to the different average strength values between the training (52.47 MPa) and testing (54.49 MPa) data sets. This difference was about 4%. We eliminated this bias problem by the use of GAs, which sorted the training and testing data set in a balanced manner. The resulting data sets had the same average strength of 53.14 MPa.

The data sorting for equalizing the averages can be done manually only if the number of data is limited. When more data are involved, a separate GA program may be employed before the application of NN. This code may also be thought of as an average filter.

3.2. Genetic algorithms

In nature, the species are evolved to survive under the harshest conditions. This evolution (survival of the fittest) process has attracted the attention of computing society for solving problems in various disciplines. As a result, GAs that mimic the evolution process were developed in the last decade of the 20th century.

The GAs employ Darwinian selection and Mendelian crossover principles. GAs contain steps that range from simulating fertilization of any diploid organism to the extinction of a certain species during a computing session. Because GAs are robust and guided random search methods, they have found a niche in the nonlinear programming field. An in-depth analysis is given in Ref. [7].

4. Data collection

The data used in GA-ANNs modelling were collected from a local cement plant that uses strength testing for process control. The data belonged to the period between the months of January and July 2001. Cement strength testing is carried out according to European standard EN 196-1[5]. Premanufactured sand with controlled particle size distribution and chemical composition is mixed with known amounts of cement and water. The mixture is molded into rectangular shapes ($4 \times 4 \times 16$ cm) and stored in a humidity cabinet at >90% relative humidity and 20 ± 1 °C for 24 h. Standard curing samples are stored in a water bath for 1, 6 and 27 more days for compressive strength testing. The type of cement used in

this research was Cem I 42.5R European standard EN 197-1 [10]. In all types of strength tests, six identical sample bars were tested for better statistics.

5. Model construction

In this study, the ANN architecture was of feed-forward type composed of three layers (Fig. 1). There were 20 neurons in the input layer for the 20 input variables. The middle layer had 20 neurons although different numbers of neurons were also tested. In the output layer, one neuron was used for the output variable of cement strength. The input variables, their means and ranges are listed in Table 1.

Bias term was not used during modelling but a momentum term was used to help obtain faster convergence during iterations. This helped the iteration process not to get stuck in local minima, but rapidly reach the desired global minima. There were a total of 150 data sets each with 21 components ($x_1, x_2, \dots, x_{20}; y$) 20 of which are the input variables whereas the 21st one is the output variable (Table 1). The program was instructed to run for 40000 iterations and the optimal weights were calculated with an average percentage training error of 1.88% for the

Table 1
The input variables used in model construction

Code	Input variable	Data used in model building		
		Minimum	Average	Maximum
x_1	SiO ₂ (%)	18.60	19.54	20.40
x_2	Al ₂ O ₃ (%)	4.60	5.07	5.70
x_3	Fe ₂ O ₃ (%)	3.50	3.64	4.00
x_4	CaO (%)	62.7	64.1	65.3
x_5	SO ₃ (%)	2.2	2.7	3.1
x_6	Loss on ignition (%)	1.30	1.87	2.70
x_7	Free lime (%)	0.60	1.13	1.70
x_8	C ₃ S (%)	51.70	60.96	68.30
x_9	C ₂ S (%)	3.60	10.01	18.30
x_{10}	C ₃ A (%)	6.30	7.29	8.90
x_{11}	C ₄ AF (%)	10.60	11.05	11.80
x_{12}	Aluminate modulus (Al ₂ O ₃ /Fe ₂ O ₃)	1.30	1.39	1.60
x_{13}	Silicate modulus (SiO ₂ /(Al ₂ O ₃ + Fe ₂ O ₃))	2.00	2.24	2.50
x_{14}	Na ₂ O (%)	0.10	0.21	0.30
x_{15}	K ₂ O (%)	0.70	0.78	0.80
x_{16}	Initial setting time (min)	95.00	156.77	225.00
x_{17}	Final setting time (min)	150.0	248.0	365.0
x_{18}	Specific surface (cm ² /g)	3120.0	3657.8	4100.0
x_{19}	Sieve residue on 90 μm (%)	0.10	0.72	2.40
x_{20}	Sieve residue on 32 μm (%)	8.20	15.52	25.50
y	Compressive strength (MPa (or N/mm ²))	47.60	53.14	58.40



Fig. 2. Results of training of 100 data sets.

output variable whereas the maximum percentage training error was 6.30% (Fig. 2). As can be seen in Fig. 2, the training of the model was successfully accomplished. The model was very close to the actual data and was able to follow the trend. The extreme values of the targeted strengths, on the other hand, could not be obtained from the model. This was to be expected because the model was conservative and needed more extreme training data to learn the extremes.

The trained model was tested by comparing it to actual measured data that forms a group of 50 data sets sorted after the application of GAs, and the testing results are given in Fig. 3. As can be seen in Fig. 3, the performance of the model was very good with an average error of 2.24% and a maximum error of 8.67%.

model was able to closely follow the trend of the actual data.

6. Results of sensitivity analysis and discussion

Sensitivity analysis is performed by feeding input parameters at varying levels into the developed model and producing prediction outputs of cement strength. The whole range of each input factor is divided into 10 equal parts to have a continuous plot for factor effects. These ranges are listed in Table 2 and are the same as for training data because the model cannot be used to predict strength values for input parameter ranges for which it is not trained. For example, the SO₃ range of 2.25–3.12 is divided into 10

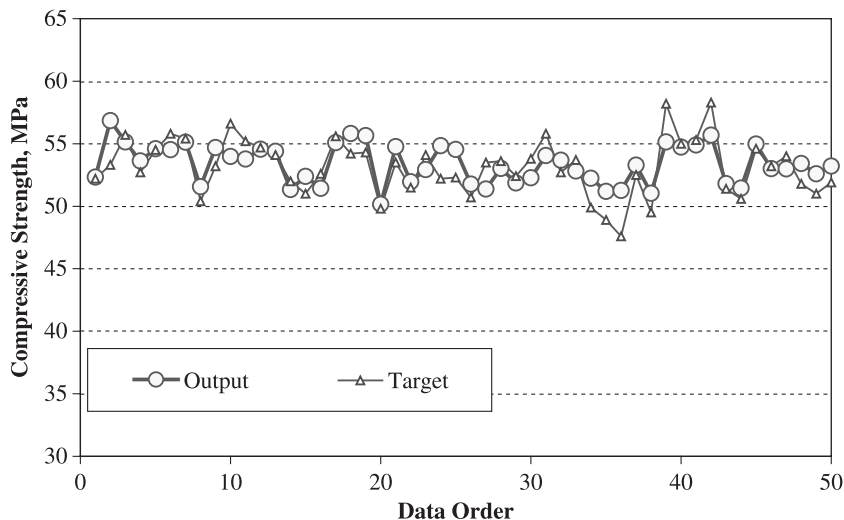


Fig. 3. Results of testing of the model.

Table 2

Factor ranges that are used for sensitivity analysis

Factor effect	Range studied in sensitivity analysis
SO ₃	2.25–3.12
Free CaO	0.58–1.64
Na ₂ O	0.1–0.3
K ₂ O	0.7–0.8
C ₃ S	51.66–68.35
C ₂ S	3.63–18.28
C ₃ A	6.26–8.86
Surface area (Blaine)	3120–4080
CPCT 32 μm	8.2–25.5

subdivisions and the surface area (Blaine=3150–4080) divided into 10 subdivisions, and a total of 100 predicted percent cement strength values are obtained (Fig. 4). The utility of sensitivity analysis in this research is that it enables researcher or plant operators to easily identify the experimental conditions for higher cement strength.

Figs. 4–9, based on the results of prediction runs of the model, show the effects of two factors at a time on each surface plot of the cement strength. We preferred paired-variable plots to single-variable plots for the sake of brevity. In addition, one can see the effects of two variables at once on a single surface plot. We plotted a fraction of all possible variable pair combinations to save space in the paper. However, the selections adequately represent the critical parameters. The effects of SO₃ and surface area (Blaine) on cement strength are shown in Fig. 4. As can be seen in Fig. 4, increasing Blaine leads to a gradual increase of cement strength at all levels of SO₃. The effect of increasing SO₃, likewise, results in increasing strength at all levels of Blaine. There is no interaction between the two factors, that is, the rate of increase of strength with Blaine was the same at all levels of SO₃. Because of the limited range of factors covered during training, the surface plot of Fig. 4 was a slightly inclined surface without much topographical vari-

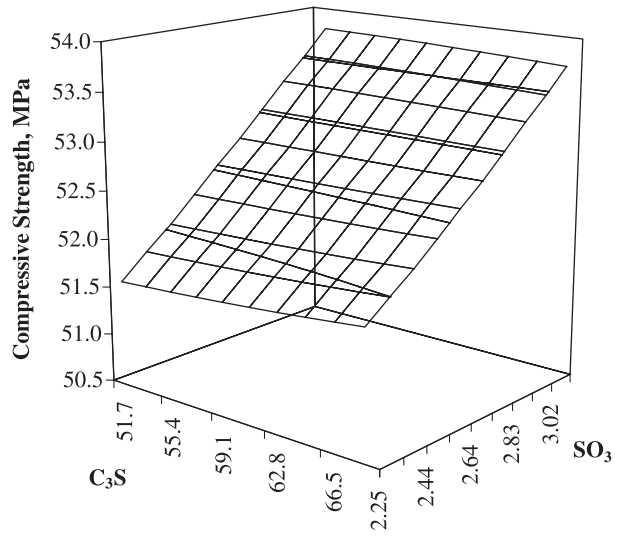


Fig. 5. Combined effects of SO₃ and C₃S on cement strength.

ation. In short, the effects of Blaine and SO₃ were both linear. According to Czernin [11], the effect of increasing Blaine on strength is countered by the slowdown of hydration due to decelerated water diffusion through the gels. He also suggests that the fineness of grinding is less important than the chemical composition of the cement when final strength is considered. A wider test range for the effect of Blaine could have produced the slowdown predicted by Czernin. A further study by Tsvivilis et al. [12] stated that the Blaine should be considered along with the particle size distribution data (uniformity factor). Lea [13] suggested that addition of gypsum has a positive effect on mechanical properties of cement up to a certain optimum content.

The effects of C₃S and SO₃ are plotted in Fig. 5. As can be seen in Fig. 5, slightly higher strengths could be obtained when C₃S was increased at low levels of SO₃. The effect of

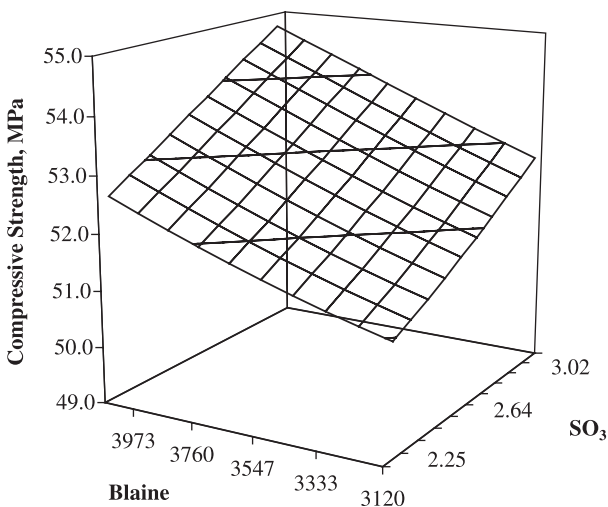


Fig. 4. Combined effects of SO₃ and Blaine on cement strength.

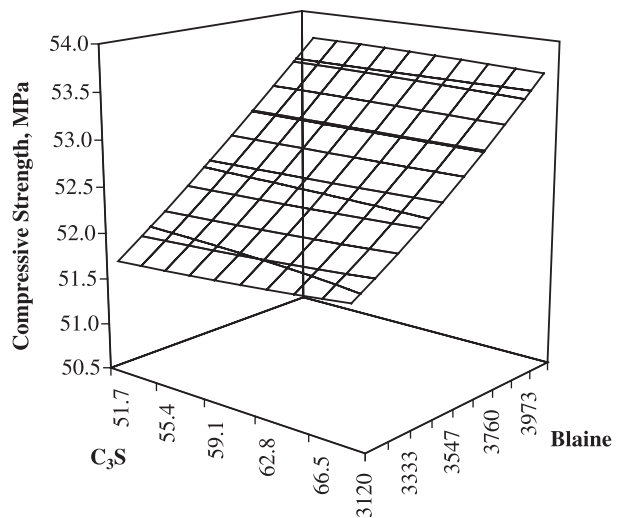


Fig. 6. Combined effects of C₃S and Blaine on cement strength.

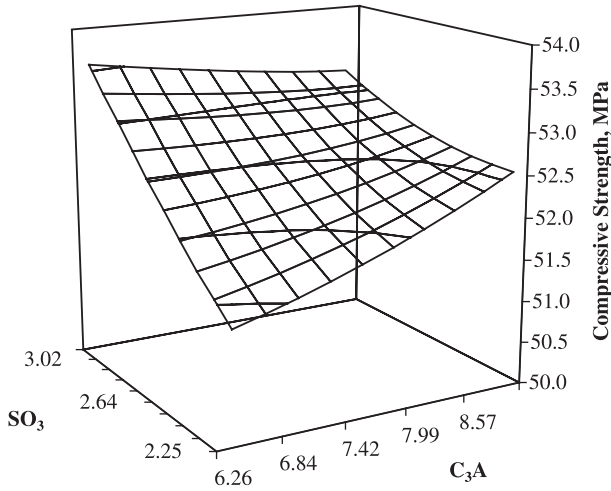


Fig. 7. Combined effects of SO₃ and C₃A on cement strength.

C₃S was insignificant when SO₃ levels were high. Increasing C₃S is normally expected to provide higher strength [11,13]. However, the narrow margin of C₃S levels used in modelling might explain the reason for this insignificant contribution from C₃S.

The combined effects of C₃S and Blaine on strength are shown in Fig. 6. Again C₃S has little effect on strength whereas Blaine significantly increases the strength. The same effect was observed for Blaine in Fig. 4, but here it was observed to be much more effective in increasing strength.

Fig. 7 shows the effects of SO₃ and C₃A on strength. Increasing SO₃ increases strength at all levels of C₃A but at different rates, suggesting the presence of interaction between the two factors. The effect of C₃A, however, was to increase strength at low levels of SO₃ and to decrease strength at high levels of SO₃. This is negative interaction between the two factors. Lea [13] suggested that increasing

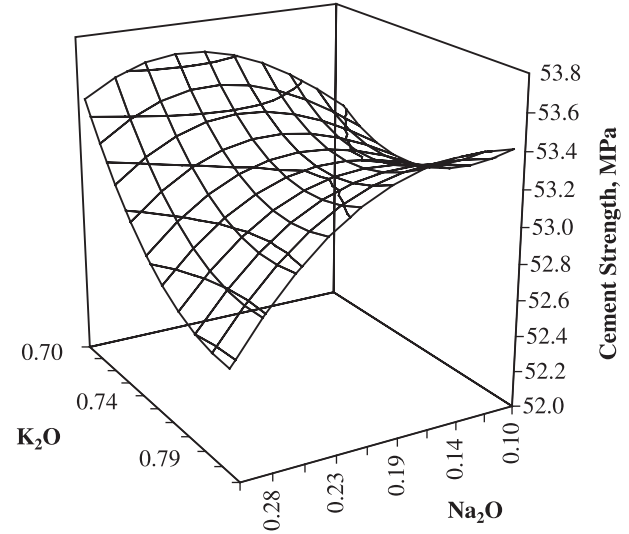


Fig. 9. Combined effects of Na₂O and K₂O on cement strength.

content of C₃A increases the strength to a certain degree. Decreasing strength at high levels of SO₃ could be explained by the excessive amount of gypsum in the cement that may have led to expansion of the cement mortar [14].

The effects of SO₃ and C₂S are also plotted based on the model as shown in Fig. 8. Increasing amount of C₂S leads to a decrease in strength at low levels of SO₃. This was to be expected because the C₂S level is not measured but calculated from C₃S percentages. The effect of SO₃ was similar to the previous plots. Again, the two factors were found to be interacting. The effect of SO₃ was again to produce increased strength.

Fig. 9 shows the effects of alkali content of cement on its strength. The graph is saddle shaped but the trend is in increasing strength with decreasing K₂O at high levels of Na₂O. The effect of varying K₂O was always concave whereas that of Na₂O was convex. Maximum strength appeared to be obtained at low K₂O and moderately high Na₂O (0.28–0.20) levels. Osbaeck [15] reported that the effect of alkalis on strength was significant and that the effect of soluble alkalis correlated better with strength than

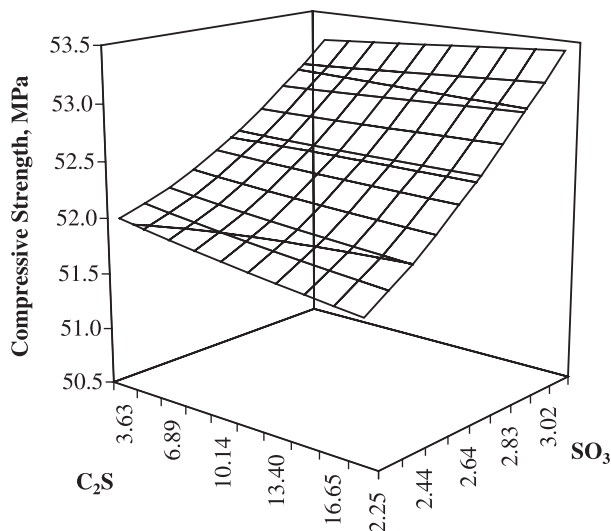


Fig. 8. Combined effects of SO₃ and C₂S on cement strength.

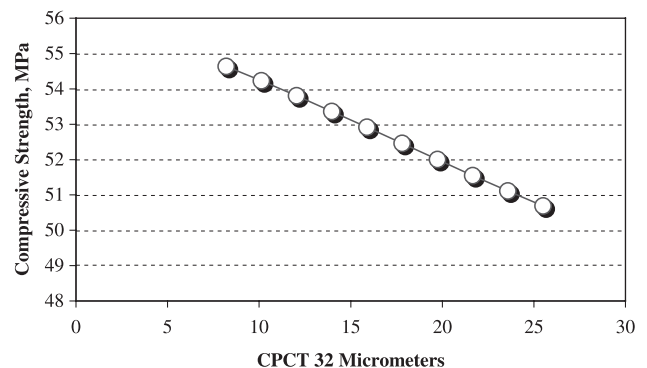


Fig. 10. Effect of CPCT 32 μm on strength.

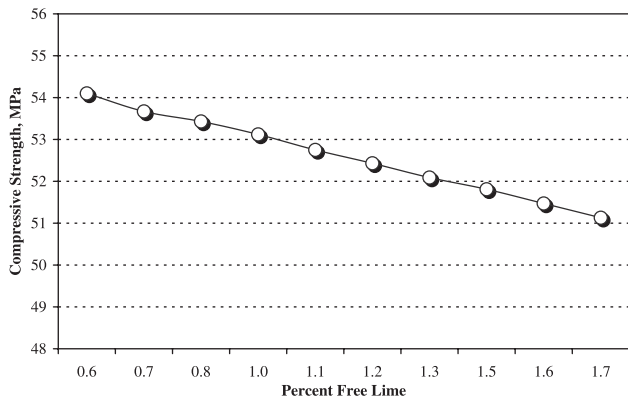


Fig. 11. Effect of percent free lime on strength.

the total alkalis. The saddle shape of the graph may perhaps be explained by the effects of soluble and insoluble parts of alkalis.

Figs. 10 and 11 show the effects of CPCT 32 μm (cumulative percent coarser than) and free lime on strength, respectively. The effect of free lime was only a minor effect that led to a slight linear decrease in strength with increasing free lime. Similarly, CPCT 32 μm was also slightly effective in decreasing strength at higher fractions of coarse particles. These are the expected results and there is a widely documented literature on these effects [11–13].

7. Conclusions

In this study, a three-layer GA-ANN model is developed for the prediction of 28-day cement strength. Input parameters used in the model creation process included the chemical composition of cement, surface area, particle size distribution, and C_3S and silicate moduli. The model is created for a local cement plant process control data. GAs provided a balanced means of separating training and testing data at the beginning of the modelling task. The learning rate, momentum term and the number of hidden layer neurons were adjusted to yield a model with the least error.

Sensitivity analysis was performed on the model to predict strength values for selected combinations of factor effects. These predictions were then shown on surface plots. The effects of varying SO_3 , C_3S , C_2S , C_3A , K_2O , Na_2O and surface area (Blaine) were plotted on surface graphs based on the developed model. The effects of free lime and CPCT 32 μm were plotted on linear plots, which showed that their effects are limited to very low levels. The effect of increasing the C_3A level was found to produce higher strength at

low levels of SO_3 , and lower strength at higher levels of SO_3 . A similar interaction was observed for the effects of SO_3 and C_2S . Other factors showed the expected results.

The satisfactory predictions of the observed cement strength by the model indicates that ANNs could be a useful tool for understanding such systems. Consequently, the model could be utilized by plant operators to optimally choose strength as a function of measured cement properties.

Acknowledgements

The authors are grateful for the support of Izmir Institute of Technology and Çimentas Cement Company.

References

- [1] H.-G. Ni, J.-Z. Wang, Prediction of compressive strength of concrete by neural networks, *Cem. Concr. Res.* 30 (2000) 1245–1250.
- [2] C.E. de Siquera Tango, An extrapolation method for compressive strength prediction of hydraulic cement products, *Cem. Concr. Res.* 28 (1998) 969–983.
- [3] S. Tsivilis, G. Parissakis, A mathematical-model for the prediction of cement strength, *Cem. Concr. Res.* 25 (1995) 9–14.
- [4] G. Fa-Liang, A new way of predicting cement strength-fuzzy logic, *Cem. Concr. Res.* 27 (1997) 883–888.
- [5] European Committee for Standardization (CEN), Methods of Testing Cement Part 1: Determination of Strength, European Standard EN 196-1.
- [6] S. Haykin, *Neural Networks a Comprehensive Foundation*, MacMillan, New York, 1994.
- [7] T. Munakata, *Fundamentals of the New Artificial Intelligence: Beyond Traditional Paradigms*, Springer-Verlag, New York, 1998.
- [8] S. Akkurt, S. Ozdemir, G. Tayfur, The use of GA-ANNs in the prediction of germanium recovery from zinc plant residues, *Trans. Inst. Min. Metall., C, Miner. Process. Extr. Metall.* 111 (12) (2003) (in press).
- [9] ASCE Task Committee, Artificial neural networks in hydrology: I. Preliminary concepts, *J. Hydrol. Eng.*, ASCE 5 (2) (2000) 115–123.
- [10] European Committee for Standardization (CEN), Cement Part 1: Composition, specifications and conformity criteria for common cements, European Standard EN 197-1, June 2000.
- [11] W. Czernin, *Cement Chemistry and Physics for Civil Engineers*, Bauverlag, Berlin, 1980, pp. 58–72.
- [12] S. Tsivilis, S. Tsimas, A. Benetatou, E. Haniotakis, Study on the contribution of fineness on cement strength, *Zem.-Kalk-Gips* 43 (1990) 26–29.
- [13] F.M. Lea, *Chemistry of Cement and Concrete*, Chemical Publishing, New York, 1971.
- [14] A.M. Neville, *Properties of Concrete*, Halsted Press, New York, 1963, pp. 10–20.
- [15] B. Osbaeck, The influence of alkalis on the strength properties of Portland cement, *Zem.-Kalk-Gips* 32 (1979) 72–77.