

Intelligent Prediction of Initial Setting Time for Cement Pastes by Using Artificial Neural Network

PINAR AKPINAR^{1,2}, MARIYA A. ABUBAKAR^{1,3}

¹Civil Engineering Department

Near East University

Near East Boulevard, Nicosia

N. CYPRUS

²pinar.akpinar@neu.edu.tr, ³aamariss02@gmail.com

Abstract: - Concrete has become a major construction material all around the world with over ten billion tons consumed annually. One of the major issues to be kept monitored during manufacture of concrete is its initial setting time; that is to say, the time needed for the initiation of fresh concrete's solidification. Information on the initial setting time of fresh concrete and/or cement paste are essential in construction scheduling, as well as in management of the haulage and placement of concrete. Conventional civil engineering laboratory tests for the determination of setting time are time and resources consuming in nature. This study aims to propose an intelligent model that will provide efficient prediction of setting time of cement pastes. An artificial Neural Network (ANN) model was proposed for the setting time predictions in this study; and its prediction performance was investigated systematically by using two training functions, under two different train:test data distributions together with five varying hidden neuron values. Setting time of cement pastes was predicted considering 12 input parameters. The results obtained indicates that the prediction accuracy of the employed ANN model is satisfactory; since it yielded remarkably high values of correlation coefficient and low mean square error such as 0.998 and 0.0003, respectively.

Key-Words: - Civil Engineering; Artificial Neural Network; Cement pastes; Initial setting time; Feedforward backpropagation; Train:test data distributions; Hidden neurons.

1 Introduction

Cement is the fundamental material used in concrete manufacture. Setting is described as the solidifying of the cement paste. Setting mainly occurs through the 'hydration' reaction of cement compounds. Hydration is described as the reaction that happens instantly after the cement powder comes in contact with water with the consequent development of heat [1,2]. Setting of fresh concrete is significant, since the fresh concrete should be maintained in the plastic stage for sufficient time period in order to ensure completion of mixing, placing and compaction processes of concrete in a feasible manner [1]. Therefore, setting time of a particular cement paste provides useful information for planning, construction and completion of the final concrete product. Conventional setting time determination tests are done with several trial and error batches of materials in the laboratory and each trial costs significant amount of materials, time and labour.

Artificial Neural Network perform a vital role in imitation of complex and indirect procedures [3]. Subsequently, for minimizing time loss and the plan

cost, help of Artificial Neural Networks (ANN) is adopted to create models, with the goal that the information extracted from these neural system models, can be used to anticipate any cement mix. Artificial Neural Networks (ANN) have been utilized as an effective device for several applications in different fields of civil engineering [3,4].

This study aims to propose an alternative intelligent method for the prediction of initial setting time of cement pastes in a non-destructive manner; without consuming materials, time and labour by using ANN. Even though ANN has been employed for several other aspects of civil engineering studies previously, it was observed that setting time predictions with ANN for cement pastes and concrete was not studied in detail in the related literature. Hence, another objective covered within this study is to provide insight on the effects of varying train:test data distributions, as well as varying hidden neuron values on the performance of the ANN model employed for the case of cement paste setting time predictions.

The methodology used for defining the proposed model parameters as well as the used ANN

architecture are described in detail in the below sections.

2 Methodology

2.1 Data Selection

In this study, the dataset used includes 174 cases having comparable parameters, that were extracted from 6 experimental studies published in the related literature. Table 1 presents the information about these experimental studies.

Table 1: References used for data acquisition

Data Source	Number of test cases
Brooks et al., 2000 [5]	13
Gulbandilar & Kocak, 2013 [6]	18
Khan & Ullah, 2004 [7]	39
Tamas, 1960 [8]	23
Yurdakul, 2010 [9]	16
Güneyisi et al., 2009 [10]	65
Total	174

Twelve (12) input parameters having progressive effects on the setting time cement pastes were identified, by taking the discussions presented in the related experimental literature as a basis. The details of the parameters are presented in Table 2.

Table 2: Input and Output Parameters

Description	Parameters	Units
Inputs	1 Cement's CaO content	%
	2 Cement's Al ₂ O ₃ content	%
	3 Cement's Fe ₂ O ₃ content	%
	4 Cement's SiO ₂ content	%
	5 water/cement ratio	%
	6 Cement content	kg/m ³
	7 Cement fineness	cm ² /g
	8 Temperature	°C
	9 Slag	%
	10 Fly ash content	%
	11 Silica fume content	%
	12 Cementitious materials' fineness	cm ² /g
Output	1 Initial setting time	mins

Within these twelve inputs, the contents of cement's chemical compounds were considered as individual parameters for the very first time in this study, as they are known to hydrate and set at different rates [2]. The novel inclusion of each cement compound's content individually, enables the use of proposed ANN model for any kind of cement type of interest, as the cement type would be defined by considering the compositional information. Other input parameters included were the water to cement ratio, temperature and admixture used in the manufacturing of the cement paste.

The target output considered in this study was the cement paste's initial setting time.

2.2 Data normalization

Normalization of input and output dataset is considered as an important data pre-processing in training of an ANN model, since reduces the percentage of saturation of neurons which might lead to slower convergence where Sigmoid function is used on a network training [11-12]. According to the Sigmoid function principles, the dataset was normalized between 0 and 1. Also, to ensure that each parameters retains its importance within the network, normalisation was considered within each parameters. Each parameter was normalized within its minimum and maximum values [3]. The equation below was adopted for the normalization [13]:

$$S_T = \frac{S_i - S_{min}}{S_{max} - S_{min}} \quad (1)$$

Where; S_{min} = minimum, S_{max} = maximum,

S_t = normalized, S_i = actual values used.

2.3 ANN Training

2.3.1 Network architecture

Within the hidden layer, a Feedforward Multilayer Perceptron Networks was created having only one layer. In this study, the input layer contains 12 nodes corresponding to the twelve input parameters considered (See Fig. 1). The number of hidden layer nodes varied and its effect on the progress of the network was also studied and discussed in the following sections.

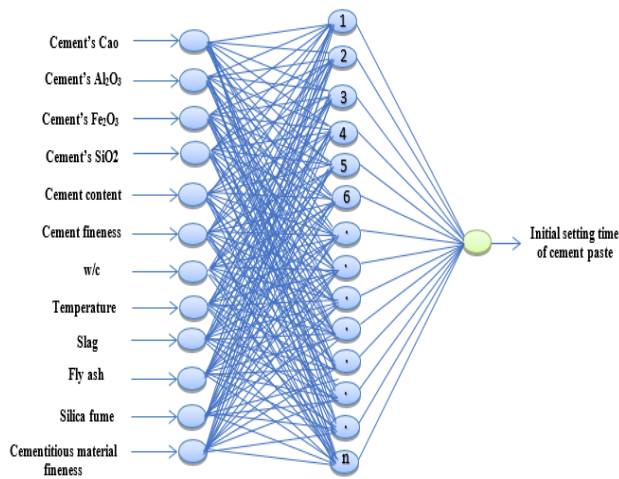


Fig. 1: ANN architecture used in this study

2.3.2 Network learning structure

The term learning structure (LS) considered in this study refers to the two different data partitions that were used with varying hidden neuron values. Two considered data partitioning ratios were; 70:30 and 50:50 for train:test. The details of the combinations of data partitioning ratios and varying values of hidden neurons are presented in Table 3 below.

Table 3: ANN learning structure

	Data distribution	
	L.S 1	L.S 2
	70/30	50/50
Hidden neurons	5	5
	12	12
	18	18
	23	23
	30	30

In this way, it is aimed to study the influence of the training dataset ratio, as well as varying hidden values on the accuracy of the setting time predictions.

2.3.3 Network training parameters

In training a feedforward neural networks, certain parameters are adjusted with respect to the training function used in the network. In this study two training functions were considered; Levenberg-Marquardt and Scaled Conjugate Gradient. Levenberg-Marquardt has been a commonly used training algorithm in modelling in neural nets [14]. Also the algorithm is a default algorithm in MATLAB[®] neural networks tool box. The algorithm is favored due to its speed of convergence. Similarly,

Scaled Conjugate Gradient was also selected based on the results and outcomes of the research with similar cementitious input parameters [15]. The algorithm also has similar converging speed to that of the Levenberg-Marquardt. The two referenced training function trained respective networks with Sigmoid function.

2.3.4 Network training

ANN models were developed using MATLAB[®] 2017a command line. For each training, the networks were trained until an optimal performance obtainable could be reached. It is noteworthy to mention here that, in training neural nets, optimisation could decrease or increase with each successive training, this is attributed to the different path a network might traced in its gradient descend. This is particularly associated with gradient descent based algorithm. In order not to lose the results of the previous optimal performance, multiple windows were opened and trained. MATLAB[®] software window with optimal performance was held in place while other windows were retrained until there is a higher result otherwise the results of the window are selected.

3 Results and Discussions

This study evaluated the effect of two data distribution ratios on the employed ANN model's performance for the prediction of cement setting time. The results of 50:50 and 70:30 train:test data distributions use together with varying hidden neurons are presented in Tables 4 and 5, respectively. Levenberg-Marquardt and Scaled Conjugate Gradient Learning functions were used to optimize the network.

The prediction performance obtained with different train:test distributions was evaluated based on the values of correlation coefficient (R) and the mean square error (MSE).

The high R and low MSE values presented for all cases in Tables 4 and 5, indicate that ANN can be used as an efficient prediction method for the determination of setting time of cement pastes. In this aspect, ANN is observed to have a high potential of serving satisfactorily as an alternative to conventional civil engineering laboratory tests.

Table 4: Prediction performance results with 50:50 data distribution

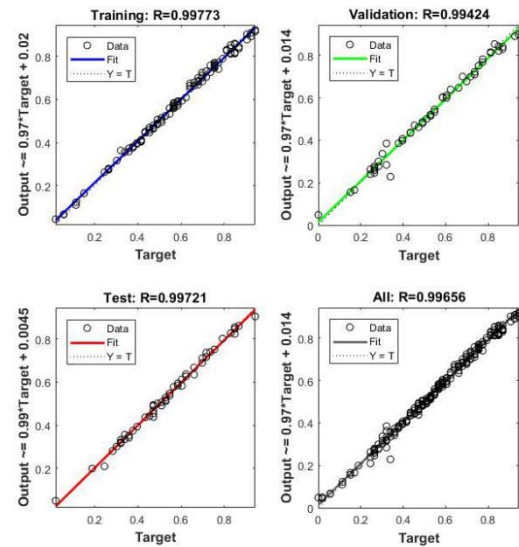
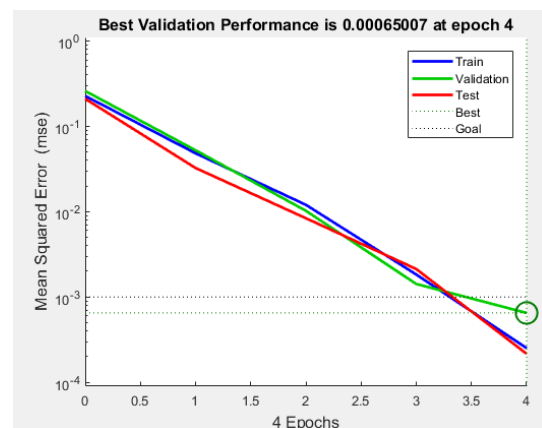
Training Func.	H-N	R-Train	R-Test	R-Overall	MSE
LM	5	0.9977	0.9972	<i>0.9966</i>	<i>0.0007</i>
SCG		0.9909	0.9819	<i>0.9883</i>	<i>0.0128</i>
LM	12	0.9974	0.9716	0.9721	0.0073
SCG		0.9907	0.9287	0.9732	0.0020
LM	18	0.9967	0.9877	0.9933	0.0010
SCG		0.9905	0.8688	0.9633	0.0024
LM	23	0.9964	0.9841	0.9917	0.0010
SCG		0.9903	0.9644	0.9837	0.0036
LM	30	0.9974	0.9928	0.9945	0.0008
SCG		0.9866	0.8990	0.9498	0.0046

Table 5: Prediction performance results with 70:30 data distribution

Training Func.	H-N	R-Train	R-Test	R-Overall	MSE
LM	5	0.9986	0.9989	<i>0.9986</i>	<i>0.0003</i>
SCG		0.9898	0.9928	<i>0.9901</i>	<i>0.0009</i>
LM	12	0.9984	0.9757	0.9934	0.0007
SCG		0.9903	0.9850	0.9893	0.0008
LM	18	0.9987	0.9973	0.9985	0.0003
SCG		0.9893	0.9115	0.9682	0.0021
LM	23	0.9985	0.9975	0.9982	0.0005
SCG		0.9875	0.9850	0.9871	0.0011
LM	30	0.9972	0.9812	0.9935	0.0004
SCG		0.9892	0.9684	0.9806	0.0056

Figures 2 and 3 presented below, illustrate the correlation coefficient (part-a) and mean square error (part-b) for the best performance cases among all combinations having 50:50 data distribution presented in Table 4 and among all combinations having 70:30 data distribution presented in Table 5, respectively.

In both data distributions, it was observed that the best performance cases (i.e. highest R and lowest MSE) for both Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) functions occurred at 5 hidden neuron values.

**Figure 2-a:** Coefficient of correlation (R) values for train-validation-test and overall for the best case of 50:50 data distribution; with LM at 5 Hidden neurons**Figure 2-b:** Mean square error (MSE) values for train-validation-test and overall for the best case of 50:50 data distribution; with LM at 5 Hidden neurons

In both cases, LM was observed to yield R and MSE values indicating higher accuracies than the values yielded with SCG; even though the difference between the correlation functions yielded with LM and SCG become almost insignificant when the train data set percentage was increased to 70% (see the values in italics in Tables 4 and 5).

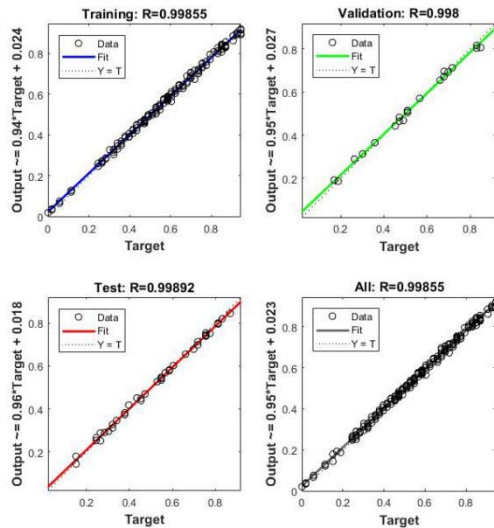


Figure 3-a: Coefficient of correlation (R) values for train-validation-test and overall for the best case of 70:30 data distribution; with LM at 5 Hidden neurons

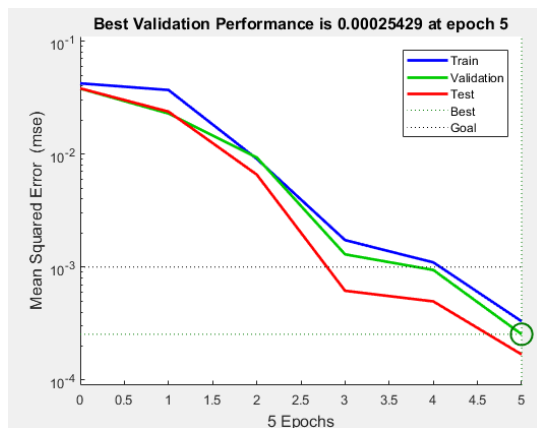


Figure 3-b: Mean square error (MSE) values for train-validation-test and overall for the best case of 70:30 data distribution; with LM at 5 Hidden neurons

More detailed evaluation of the effect of train:test data distribution and varying hidden neuron values on the evolution of R and MSE are presented in the below subsections.

3.1 Evaluation of the Effect of Varying Train:Test Data Distribution on the Performance of Employed ANN Model for the Setting Time Predictions

Figure 4 shows the evolution of correlation coefficient for each hidden neuron value *at both* 50:50 and 70:30: train:test ratios in a comparative manner.

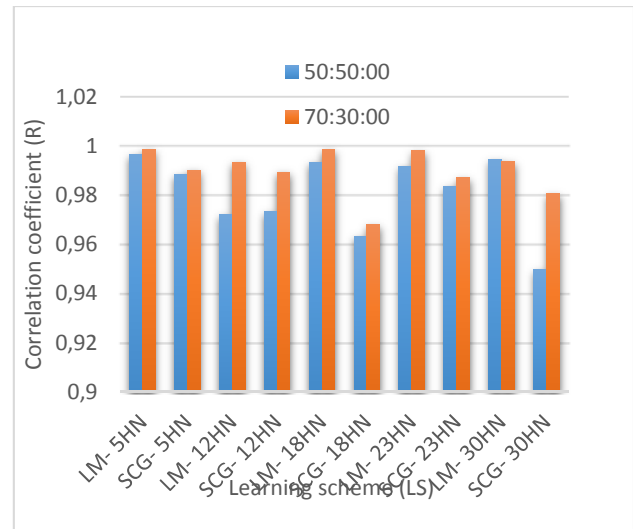


Figure 4: Evolution of R-value with changing hidden neurons together with data distribution

It is clearly seen that 70:30 train:test distribution yielded a higher R value, indicating more accurate prediction of setting time, for all employed hidden neurons with both LM and SCG training functions. The increase in the R value when the train data set was increased from 50% to 70% is observed to be more significant at certain hidden neuron values (e.g. at HN:12), and much less significant at other hidden neuron values, as in the case of employing 5 hidden neuron values.

Another noteworthy remark that can be seen with Figure 4 is that; the effect of using increased train data distribution can be more or less significant depending on the training function employed for certain hidden neuron values. This feature can be observed with increased R difference with SCG at 30 hidden neuron values, rather than the much less significant R difference observed with LM when train data percentage is increased (Figure 4).

A similar observation can be made with Figure 5, which shows the evolution of MSE for each hidden neuron value *at both* 50:50 and 70:30: train:test ratios in a comparative manner. In this figure as well, the effect of yielding significantly higher accuracies for the case of increased train data set, seems to be training function-dependent also, at certain hidden neuron values. This can be observed with 5 hidden neurons in Figure 5; where for SCG, increasing the train data set from 50% to 70% made a clear difference in MSE, but this was not the case for LM at the same hidden neuron value.

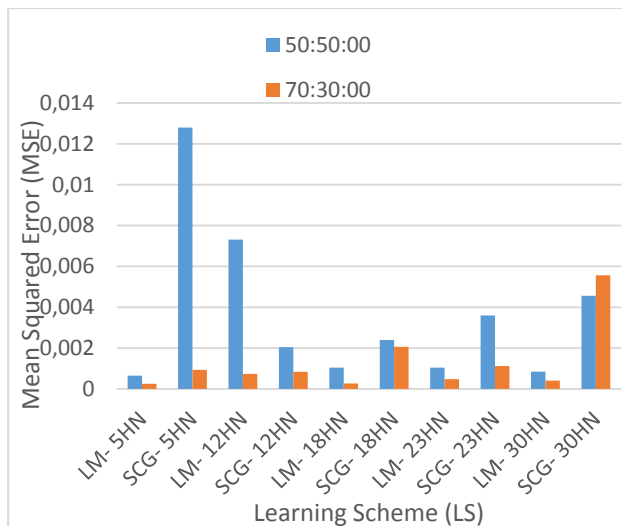


Figure 5: Evolution of MSE with changing hidden neurons together with changing data distribution

Similarly, in the case of using 12 hidden neurons; this time the difference in MSE for increasing train data set was more significant for LM but not for SCG. This finding implies that the selection of efficient train:test data distribution should be made carefully, considering the combinations of hidden neurons and training functions to be employed in each case. As in the case of Figure 4, Figure 5 also indicates that an increased train data set percentage implies an increased accuracy in setting time predictions, since the use of 70% train data set yielded lower MSE values, in all cases.

In case of using 50:50 data distribution ratio (see Table 4 and Fig 4), it can be seen that the highest overall R-values for both LM and SCG were obtained at hidden neuron 5; as 0.996 and 0.9883 respectively. For this data distribution, in certain selected hidden neurons (i.e. 18, 23 & 30), LM was found to yield higher R-values, except for the case of 12 hidden neuron where the percentage difference between the two algorithms was quite negligible (i.e. 0.11%).

In the case of the 70:30 train:test ratio (see the Table 5 and Fig 4), it can be seen that the highest overall R-values for both LM and SCG were obtained at hidden neuron 5; as 0.9989 and 0.9928 respectively, with a negligible difference of 0.0061.

Figure 6 illustrates the best performance cases with highest R values for both train:test data distributions with both LM and SCG, which occurred with five hidden neuron values.

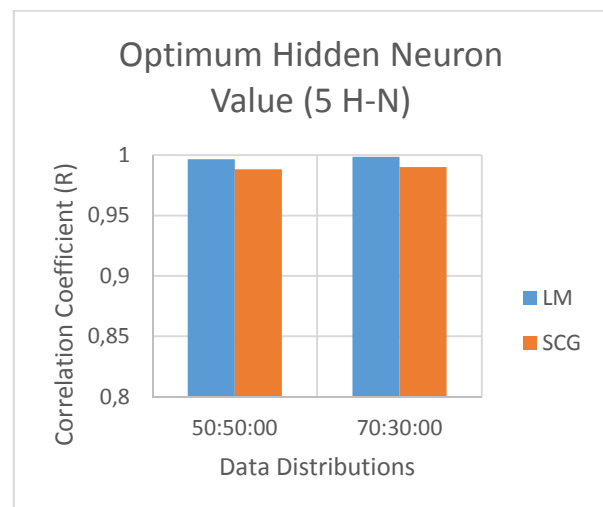


Figure 6: Coefficient of correlation values for optimum hidden neuron (5 H-N) for both train:test data distributions used.

3.2 Evaluation of the Effect of Varying Hidden Neurons on the Performance of Employed ANN Model for the Setting Time Predictions

ANN's "blackbox" nature is expected to be also affected, to some extent, by the choice of hidden neuron values to be employed.

Figures 7-a and 7-b illustrate the effect of increasing hidden neuron values for R and MSE values respectively, both in the case of 50:50 data distribution ratio.

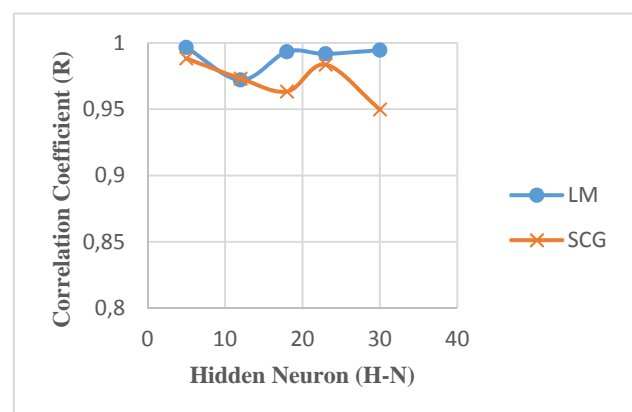


Figure 7-a: Evolution of R-value with increasing hidden neuron values for 50:50 data distribution ratio

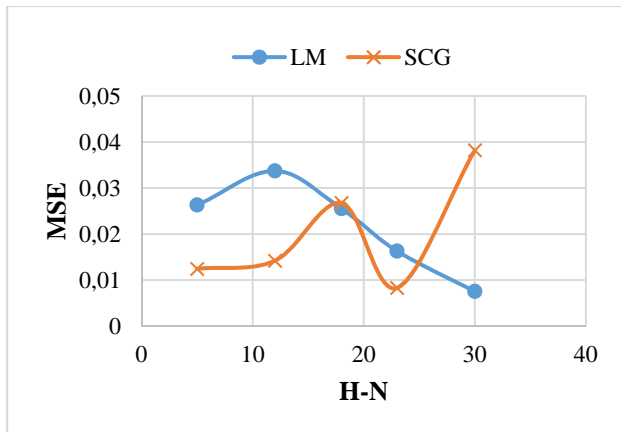


Figure 7-b: Evolution of MSE-value with increasing hidden neuron values for 50:50 data distribution ratio

In both Figures 7-a and b, SCG is observed to be relatively more sharply affected by the changes in hidden neuron values. This finding is more evident especially with the evolution of MSE presented in Fig 7-b, which may imply that the selection of an efficient hidden neuron value at a certain data distribution ratio is highly dependent of the training function employed, and changes in the hidden neurons yield in different prediction accuracy evolution for each training function to be considered.

Figures 8-a and 8-b illustrate the effect of increasing hidden neuron values for R and MSE values respectively, both in the case of 70:30 data distribution ratio.

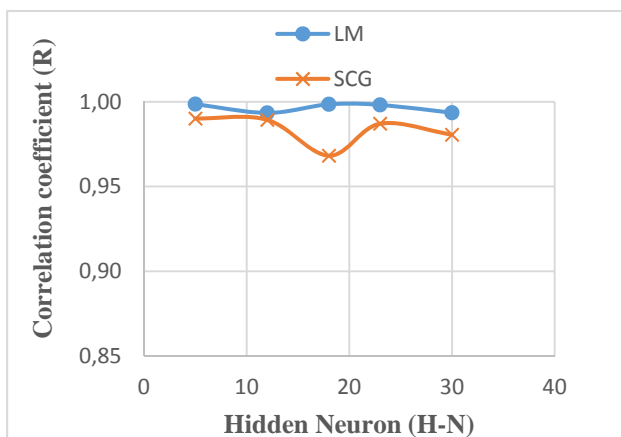


Figure 8-a: Evolution of R-value with increasing hidden neuron values for 70:30 data distribution ratio

It can be observed in Fig. 8-a that, when R is taken as a basis for accuracy, SCG when employed with the at 70:30 data distribution ratio, seem to be more sharply affected by the changes in hidden neuron

values proposed to be used in the model. However, when MSE is considered (Fig 8-b), LM training function is observed to be more prone to the changes in the hidden neuron values proposed to be used in the prediction model.

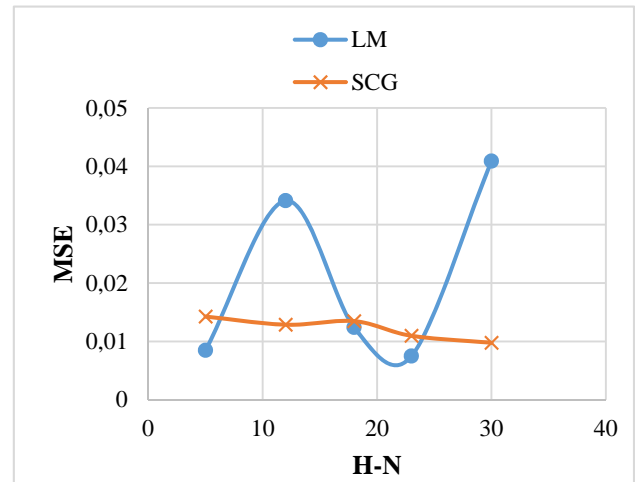


Figure 8-b: Evolution of MSE-value with increasing hidden neuron values for 70:30 data distribution ratio

Hence, this observation may imply that the effect of variations in the hidden neuron values employed, on the evolution of prediction accuracy is observed to be changing depending on the training function used in each case at a certain data distribution ratio.

As an overview, the results obtained within this study validates the sensitivity of the network performance depending on the hidden neuron size at each employed case. Similar effects are discussed in detail in the related literature [16–22]. The related literature further suggests that increase in hidden layer size could lead to network complexity due increased interconnectivity which may result to early saturation within the network [16–22].

4 Conclusive Remarks

The study was carried out to propose an intelligent method for cement paste setting time predictions that could serve as an alternative to the ‘time and resources consuming’ conventional civil engineering laboratory test methods. The ANN model’s performance was investigated *systematically* considering the use two optimization functions under 70:30 and 50:50 train:test data distributions together with five different hidden layer nodes number of 5, 10, 12, 23, and 30. With the interpretation of the

obtained results presented in the previous sections, the following conclusions were drawn:

- I. The employed ANN model yielded satisfactorily high R-values, such as 0.949 and above in all cases. Hence, ANN has the potential to serve efficiently as an alternative method for the determination of cement setting time in related civil engineering applications.
- II. It was observed that greater training data distributions such as 70% used in this study, yielded a higher accuracy in setting time predictions for both training functions employed.
- III. The best prediction performance for both train:test data distributions were obtained with five hidden neurons, with both training functions. However, increasing hidden neuron values were observed to yield different performance tendencies in the prediction of setting time for each type of training function used within ANN.
- IV. In the study, the higher prediction performances were observed to be obtained when Levenberg-Marquardt (LM) training function was employed, for all combinations of data distributions and hidden neurons used within this study. The highest coefficient of correlation was obtained as 0.9986 when LM was used with five hidden neurons and 70:30 train:test data distribution.

5 Recommendations for Future Studies

Authors of this study recommends further studies with data obtained preferably from one single laboratory, which would minimize the incongruity of the data and hence may yield in further increased accuracy. It is also recommended that different learning methods in ANN should also be additionally employed in future studies for the intelligent prediction of initial setting time of cement pastes.

References:

- [1] A. M. Neville and J. J. Brooks, "Concrete Technology," *Building and Environment*, vol. 11, pp. 1–9, 2010.
- [2] A. M. Neville, *Properties of Concrete*, 5th ed. Pearson Education Publisher, 2011.
- [3] B. K. Khademi F, "Evaluation of concrete compressive strength using artificial neural network and multiple linear regression models.," *Iran Univ. Sci. Technol. (Elmsford, N.Y.)*, vol. 6(3), pp. 423–432, 2016.
- [4] H. Baseri, S. M. Rabiee, F. Moztarzadeh, and M. Solati-Hashjin, "Mechanical strength and setting times estimation of hydroxyapatite cement by using neural network," *Mater. Des.*, vol. 31, no. 5, pp. 2585–2591, 2010.
- [5] J. J. Brooks, M. A. Megat Johari, and M. Mazloom, "Effect of admixtures on the setting times of high-strength concrete," *Cem. Concr. Compos.*, vol. 22, no. 4, pp. 293–301, 2000.
- [6] E. Gulbandilar and Y. Kocak, "Prediction of the effects of fly ash and silica fume on the setting time of Portland cement with fuzzy logic," *Neural Comput. Appl.*, vol. 22, no. 7–8, pp. 1485–1491, 2013.
- [7] B. Khan and M. Ullah, "Effect of a Retarding Admixture on the Setting Time of Cement Pastes in Hot Weather," *JKAU Eng.Sci.*, vol. 15, no. 1, pp. 63–79, 2004.
- [8] F. D. Tamas, "Acceleration and Retardation of Portland Cement Hydration by Additives," no. 8, pp. 392–397, 1960.
- [9] E. Yurdakul, "Optimizing concrete mixtures with minimum cement content for performance and sustainability," *Optim. Concr. Mix. with Minim. Cem. content Perform. Sustain.*, p. 112, 2010.
- [10] E. Güneyisi, M. Gesoglu, and E. Özbay, "Evaluating and forecasting the initial and final setting times of self-compacting concretes containing mineral admixtures by neural network," *Mater. Struct. Constr.*, vol. 42, no. 4, pp. 469–484, 2009.
- [11] F. Khademi, M. Akbari, S. M. Jamal, and M. Nikoo, "Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete," *Front. Struct. Civ. Eng.*, vol. 11, no. 1, pp. 90–99, 2017.
- [12] A. Sharifi, S. S., Delirhasannia, R., Nourani, V., Sadraddini, A. A., and Ghorbani, "Using Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS) for Modeling and Sensitivity Analysis of Effective Rainfall," p. 133–139., 2009.
- [13] F. D. Nourani, V., Alami, M. T., & Vousoughi, "Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling," *J. Hydrol.*, vol.

- 524, pp. 255–269, 2015.
- [14] F. Khademi and S. M. Jamal, “Predicting the 28 days compressive strength of concrete using artificial neural network,” *i-Manager’s J. Civ. Eng.*, vol. 6, no. 2, p. 1, 2016.
 - [15] P. Akpinar and I. D. Uwanuakwa, “Intelligent Prediction of Concrete Carbonation Depth Using Neural Networks,” *Bulletin of the Transilvania University of Brasov*, Vol. 9 (58), No. 2-2016, Series III:Mathematics, Informatics, Physics, pp. 99–108, 2016.
 - [16] C. Fahlman, S. E., & Lebiere, “The cascade-correlation learning architecture,” *Adv. Neural Inf. Process. Syst.*, pp. 524–532, 1990.
 - [17] M. W. Lee, Y., Oh, S. H., & Kim, “An analysis of premature saturation in back propagation learning,” *Neural networks*, vol. 6(5), pp. 719–728, 1993.
 - [18] C. S. G. Lin, C. T., & Lee, “Neural-network-based fuzzy logic control and decision system,” *IEEE Trans. Comput.*, vol. 40(12), pp. 1320–1336, 1991.
 - [19] R. G. Parekh, A. K., & Gallager, “A generalized processor sharing approach to flow control in integrated services networks: the single-node case,” *IEEE/ACM Trans. Netw.*, vol. 1(3), pp. 344–357, 1993.
 - [20] V. Parekh, R., Balakrishnan, K., & Honavar, “An empirical comparison of flat-spot elimination techniques in back-propagation networks,” *PROCEEDINGS-SPIE Int. Soc. Opt. Eng. SPIE Int. Soc. Opt.*, p. 55, 1992.
 - [21] V. Spartz, R., & Honavar, “Empirical analysis of the expected source values rule,” *Proc. 3rd Work. Neural Networks Acad. Publ by Soc. Photo-Optical Instrum. Eng.*, 1993.
 - [22] J. Vitela, J., & Reifman, “Enhanced backpropagation training algorithm for transient event identification,” *Trans. Am. Nucl. Soc. States*, vol. 69(CONF-93, 1993.