Computers and Concrete, *Vol. 12, No. 2 (2013) 151-167*
DOI: http://dx.doi.org/10.12989/cac.2013.12.2.151

Prediction of lightweight concrete strength by categorized regression, MLR and ANN

S. Tavakkol[1,2], F. Alapour[2], A. Kazemian[*2], A. Hasaninejad[3], A. Ghanbari[4], A. A. Ramezanianpour[5]

[1]*Hydraulic Structures Division, Water Research Institute, Tehran, Iran*
[2]*Dept. of Civil and Environmental Engineering, Amirkabir University of Technology, Tehran, Iran*
[3]Dept. of Civil Engineering, Shahed University, Tehran, Iran
[4]*Dept. of Computer Engineering and IT, Amirkabir University of Technology, Tehran, Iran*
[5]*Concrete Technology and Durability Research Center, Amirkabir University of Technology, Tehran, Iran*

(Received November 11, 2012, Revised January 10, 2012, Accepted February 9, 2013)

Abstract. Prediction of concrete properties is an important issue for structural engineers and different methods are developed for this purpose. Most of these methods are based on experimental data and use measured data for parameter estimation. Three typical methods of output estimation are Categorized Linear Regression (CLR), Multiple Linear Regression (MLR) and Artificial Neural Networks (ANN). In this paper a statistical cleansing method based on CLR is introduced. Afterwards, MLR and ANN approaches are also employed to predict the compressive strength of structural lightweight aggregate concrete. The valid input domain is briefly discussed. Finally the results of three prediction methods are compared to determine the most efficient method. The results indicate that despite higher accuracy of ANN, there are some limitations for the method. These limitations include high sensitivity of method to its valid input domain and selection criteria for determining the most efficient network.

Keywords: compressive strength; lightweight aggregate concrete; LECA; multiple linear regression; neural networks; data cleansing

1. Introduction

Lightweight Aggregate Concrete (LWAC) is not a recent invention in concrete technology and it has been known since ancient times (Chandra and Berntsson 2002); however the mechanical properties of LWAC are still subject of research (Cui *et al*. 2012, Andiç-Çakır and Hızal 2012, Golias *et al*. 2012). There are two major classes of LWAC, non-structural lightweight concretes and structural lightweight concretes. The former class has a lower density because of higher air voids in the cement paste matrix and lighter aggregates and is commonly used for insulation properties.

Structural Lightweight Aggregate Concrete (SLWAC) is generally used to reduce the dead

*Corresponding author, MSc student, E-mail: ak_civil85@hotmail.com

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

weight of a structure which then allows the designer to reduce the size of columns, footings, beams, plates, etc (Topcu 1997, Mazaheripour 2011). Furthermore, the weight reduction consequently reduces the risk of earthquake damages to the structure, since the earthquake forces which should be resisted by the structures are proportional to mass of those structures (Yasar *et al*. 2003, Campione *et al*. 2005). The weight reduction is considerable in long-span bridges, because the live load is a minor part of the total load of bridges and a reduction in density leads to reductions in not only mass, but also in section size of those bridges (Popovics 1992, Chandra and Berntsson 2002).

Having large number of voids in the aggregates, SLWAC has a lower thermal conductivity and a lower coefficient of thermal expansion and therefore resists fire better than normal weight concrete. The large number of voids improves its sound insulation ability too (ACI-213, Mor 1992, Al-Jabri *et al*. 2008).

Lightweight concretes can be produced with a density range of approximately 300-2000 kg/m^3, corresponding cube strengths from approximately 1 to over 60 N/mm^2. These values can be compared with those for normal weight concrete with density range of 2100-2500 kg/m^3 and compressive strength range of 15-100 N/mm^2 (Clarke 1993). However, the definition of SLWC is not consistent in different specifications (Kılıç *et al*. 2009). The density and the uniaxial compressive strength of SLWC are defined by EN 206-1 to be 800 to 2000 kg/m^3 and 8 to 80 MPa, respectively (TS EN 206-1). SLWAC as defined in ASTM C 330 has a minimum 28-day compressive strength of 17 MPa, an equilibrium density between 1120 and 1920 kg/m^3. It consists entirely of lightweight aggregate or a combination of lightweight and normal-density aggregate (ASTM C 330).

Prediction of concrete properties has always been important for structural engineers and different methods are developed for this purpose. Most of these methods are based on experimental data and employ them for output estimation. Two typical methods of output estimation are Multiple Linear Regression (MLR) and Artificial Neural Networks (ANN). MLR has been widely used for estimation of concrete properties (Zain and Abd 2009, Huang *et al*. 2010, Dafico and Prudencio 2002), and ANN was a subject of great interest in the recent decades (Kasperkiewics 1995, Lai and Serra 1997, Yeh 1998, Lee 2003, Pala *et al*. 2005, Oztas *et al*. 2006, Topcu 2008, Alshihri 2009, Yuzer 2011, Oreta and Ongpeng 2011)

In this paper, firstly, a statistical method based on cross-sectional linear regression is introduced for data cleansing. Afterwards, both MLR and ANN are employed to predict the compressive strength of SLWAC. The valid input domain of the applied methods is briefly discussed. Considering the proper input domain is of high important; since out-range inputs can lead to wrong predictions. However, it is rarely discussed in the literature. Finally the results of two prediction methods are compared to determine the most efficient method for estimation of SLWAC compressive strength.

2. Experimental program

2.1 Materials

Cement used in the mixtures was ASTM C150 type I portland cement. The used lightweight aggregate was an artificial Light Expanded Clay Aggregate (LECA). Fig. 1 compares the grading of

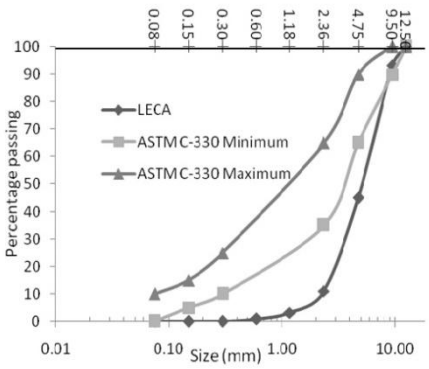


Fig. 1 Grading of LECA used in this study

Table 1 Physical properties of aggregates

Aggregate	Specific density	Water absorption (%)
Coarse LECA	1.31	5.1 (30 - min)
Fine LECA	0.9	8.5 (30 - min)
Natural sand	2.65	2.4 (24 - hour)

Table 2 Characteristics of portland cement

Property	Percentage
Chemical composition	
CaO	61.5
SiO_2	21.5
MgO	4.8
Al_2O_3	3.7
Fe_2O_3	2.8
SO_3	2.5
Na_2O	0.1
K_2O	0.95
MgO	4.8
Physical properties	
Blaine (m^2/kg)	320
Specific gravity	3.12

the used LECA with grading requirements specified in ASTM C 330 for SLWAC. It is observed that there is low amount of fine aggregates in LECA used in this study. Furthermore, trial batches indicated that by using only LECA as the aggregate, the compressive strength of the mixture would not satisfy the ASTM C 330 minimum of 17 MPa. Therefore, the use of ordinary sand was inevitable in order to satisfy both grading and strength criteria of ASTM C330. The physical properties of aggregates are shown in Table 1, while characteristics of portland cement are presented in Table 2.

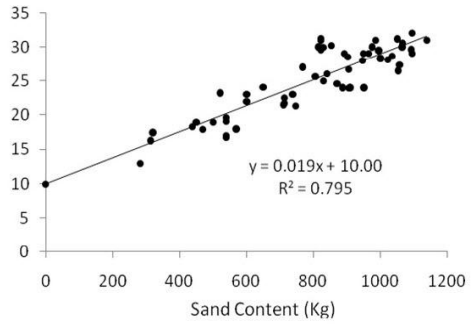


Fig. 2 Relationship between strength and sand content

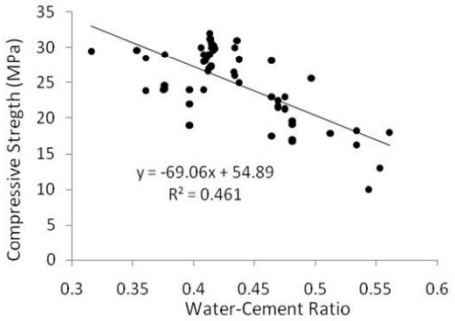


Fig. 3 Relationship between strength and w/c

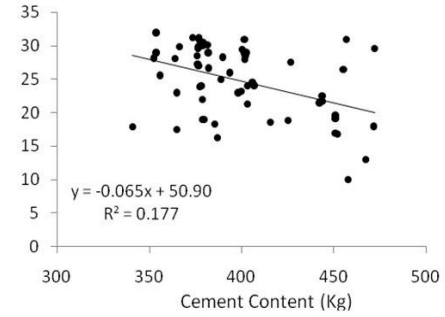


Fig. 4 Relationship between strength and cement content

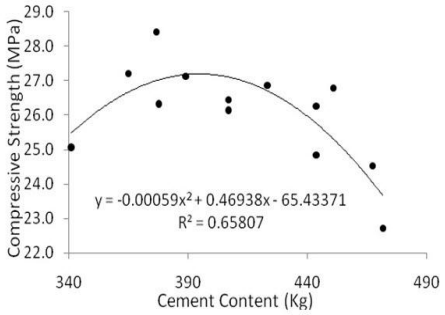


Fig. 5 Cross sectional plot for strength and cement content (w/c = 0.41 and 800 kg < S < 950 kg)

2.2 Mix designs and specimens

A total number of 59 lightweight concrete mixtures were prepared. The proportions were determined according to the absolute volume method. In this method, the volume of fresh concrete is considered equal to the sum of the absolute volumes of cementitious materials, aggregate, free water, and entrained air (ACI 213R-03). Table 3 shows the range of modeling parameters and the mechanical properties of the concrete mix designs in this study. Furthermore, Table 4 shows the parameters for all mid designs.

2.3 Curing and testing of LWAC specimens

The specimens were moved into a lime-saturated water tank after 24 hours. At the age of 7 days, they were cured for 48 hours in 105℃. Finally, prior to measuring the compressive strength, they were cooled down to the room temperature during 24 hours. Three 100-mm cube specimens of each mixture were tested for compressive strength.

In order to compare the compressive strength of the 28-day specimens with the oven cured ones, 21 specimens of 7 different mix designs were cured in the lime-saturated water tank until the age of 28 days. Results showed that there is only a maximum difference of 5 percent between the compressive strength of specimens cured in the two mentioned conditions.

Table 3 Range of parameters inputs

Parameter		Min.	Max.
Cement content (C) *	kg	340.9	472.2
Water-cement ratio (w/c)	-	0.316	0.561
Sand content (S)*	kg	0	1138.6
Oven dried density (D)	kg/m^3	1119.4	1802.4
Compressive strength (σ)	MPa	10.0	32.0

* The amounts are given in 1 m^3 of fresh concrete mixture

Table 4 Parameters of mix designs

No	C (Kg)	S (Kg)	W/C	σ (MPa)	No	C (Kg)	S (Kg)	W/C	σ (MPa)	No	C (Kg)	S (Kg)	W/C	σ (MPa)
1	352.6	1001.8	0.41	28.2	21	377.0	822.8	0.41	31.0	41	388.9	830.2	0.44	25.0
2	376.0	903.4	0.36	28.5	22	400.2	994.6	0.32	29.5	42	389.9	1008.3	0.39	28.3
3	377.5	976.0	0.41	30.0	23	401.5	964.3	0.38	29.0	43	393.5	840.1	0.41	26.0
4	377.4	907.0	0.36	23.9	24	401.5	986.3	0.34	31.0	44	385.6	438.6	0.53	18.3
5	377.8	887.4	0.40	24.0	25	401.6	947.2	0.41	28.0	45	386.9	312.9	0.53	16.3
6	378.8	501.1	0.41	19.0	26	402.4	949.2	0.41	29.0	46	398.2	737.6	0.47	23.0
7	378.0	608.4	0.40	22.0	27	405.8	870.1	0.38	24.6	47	403.3	747.1	0.47	21.3
8	379.8	449.8	0.45	19.0	28	406.8	650.0	0.38	24.1	48	455.1	1053.1	0.43	26.5
9	378.4	909.3	0.38	24.0	29	400.0	521.4	0.47	23.2	49	456.9	1138.6	0.44	31.0
10	353.7	1094.2	0.41	32.0	30	403.2	951.0	0.41	24.0	50	442.2	711.5	0.47	21.5
11	353.7	1094.2	0.41	29.0	31	381.4	852.9	0.42	30.2	51	443.6	713.8	0.47	22.5
12	373.6	1050.3	0.41	31.2	32	381.9	892.8	0.41	29.0	52	443.6	713.8	0.47	21.7
13	376.6	821.8	0.42	29.6	33	382.2	905.9	0.41	26.5	53	450.9	538.7	0.48	19.1
14	376.6	821.9	0.42	29.9	34	402.1	1034.5	0.41	28.0	54	450.9	538.7	0.48	19.6
15	376.7	768.4	0.41	27.0	35	377.1	813.9	0.38	30.0	55	450.9	538.7	0.48	17.0
16	378.7	1064.7	0.41	30.0	36	340.9	469.0	0.51	17.9	56	452.0	540.1	0.48	16.8
17	379.0	1065.5	0.41	30.5	37	355.8	805.4	0.50	25.6	57	457.7	0.0	0.54	10.0
18	366.3	829.3	0.42	29.9	38	363.9	1023.0	0.46	28.1	58	467.2	281.7	0.55	13.0
19	376.1	1057.5	0.41	27.3	39	364.9	321.3	0.44	17.5	59	471.6	568.9	0.56	18.0
20	376.9	822.4	0.41	31.3	40	366.1	598.7	0.41	23					

Table 5 Classification of parameters

Parameter	Group	Range
C	Low	$C < 380$ kg/m^3
	Medium	380 $kg/m^3 \leq C < 430$ kg/m^3
	high	430 $kg/m^3 \leq C$
w/c	Low	$w/c < 0.42$
	high	$w/c \geq 0.42$

Table 6 Defined groups and number of experiments

C	w/c	Notation	No. of experiments
High	High	HH	12
Medium	High	MH	7
Low	High	LH	6
High	Low	HL	1
Medium	Low	ML	13
Low	Low	LL	20

3. Prediction model set up

The compressive strength of SLWAC depends on the aggregates characteristics (grading, size and strength), cement content, water-cement ratio, curing conditions, etc. (kilic *et al*. 2009). In the present paper the compressive strength (σ) relationship with sand content (S), water-cement ratio (w/c), and cement content (C) is studied. A basic method for this investigation is plotting the σ values against S, w/c, and C separately (Figs. 2-4). It can be seen that there is significant linear relation between σ and S and a recognizable linear relation between σ and w/c, while the plot of σ and C is almost randomly scattered. It can be postulated that σ is depended linearly on S and w/c, but the relationship between σ and C needs further investigation, because absence of a linear regression does not mean a relationship is not present. Fig. 4 shows the cross-sectional plot of σ and C, where S varies in the range of 800-950 Kg and w/c is 0.41. It can be seen that for constant values of S and w/c, the relationship between σ and C is not monotonic and it can be well estimated with a parabola.

4. Categorized linear regression and data cleansing

Simple linear regression is a popular statistical method for predicting the value of a dependent parameter, Y, based on an independent parameter (predictor), X. In problems with higher dimensions, it is possible to choose the most effective parameter as the main predictor and categorize the rest into groups; then, fit a linear regression between the dependent variable and the main predictor in each category. Consider Y depended on X_p, as the main predictor, and X_1, X_2, …, X_n. Categorizing this problem will lead into $m_1 \times m_2 \times \ldots \times m_n$ categories and a corresponding linear regression, where m_i is the number of stratums for independent variable X_i.

The main predictor is supposed to be the sand content (S) of mixture. The cement content (C) is stratified into three groups, while water-cement ratio (w/c) is divided in two groups as presented in Table 4.

This categorization divides the data into 6 groups. Table 5 shows these groups and the number of experiments in each one. It can be seen that there is only one experiment with high amount of cement and low water-cement ratio, therefore modeling is abortive in this group.

Categorizing continuous predictors is not recommended to be used in prediction models (Royston *et al*. 2006, Fedorov *et al*. 2009); In fact, it is used in this paper for data cleansing. The proposed method is based on the strength and significance of the correlation in each group.

The strength of a linear regression is indicated by the correlation coefficient, R, but is actually measured by the coefficient of determination, R^2. R^2 is always less than 1, and the greater it is, the

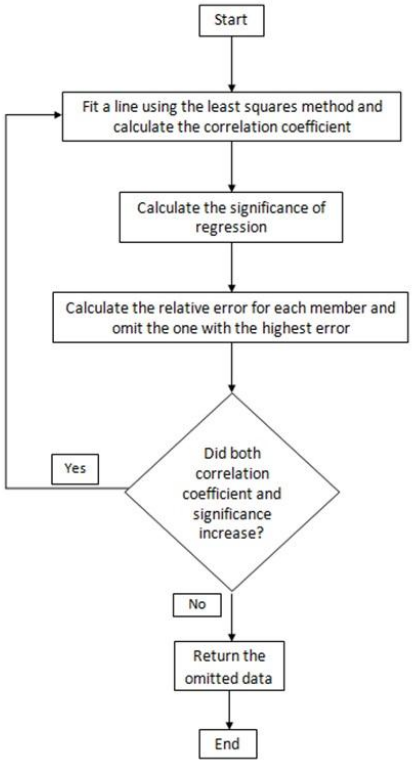


Fig. 6 Flow chart for proposed cleansing method

stronger the correlation is. There is always a possibility that even a high value of R^2 is merely by chance. This possibility increases by decreasing the count of samples for a constant value of R^2. The significance of a relationship is expressed in probability levels, p (for example significant at $p = 0.05$). This tells how unlikely a given correlation coefficient, R, will occur given no relationship in the population. It is assumed that the sampling distribution of i is t-students. The simplest formula for computing the appropriate t value to test significance of a correlation coefficient is (Rawlings *et al*. 1998, Chattefuee and Hadi 2006)

$$t = R\sqrt{\frac{n-2}{1-R^2}} \tag{1}$$

where n is the size of sample. According to the t value of a regression, the probability level in which the regression is significant can be found from the table of t-student distribution. Note that the degree of freedom of the distribution is n-2.

A relationship can be strong and yet not significant, conversely, a relationship can be weak but significant. The key factor is the size of the sample. For small samples, it is easy to produce a strong correlation by chance and "significance" must be considered. For large samples, it is easy to achieve significance, and one must pay attention to the "strength" of the correlation. Therefore, removing a suspicious experiment from the samples may increase the correlation coefficient, but simultaneously, it may decrease the significance of the relationship.

In the present paper a cleansing method based on the significance and strength of regression is employed which can be seen in Fig. 6 and is described in below steps:

1) Calculate the correlation coefficient between the compressive strength and sand content in each group.

2) Calculate the significance of the correlation coefficient, which is defined as 1-*p*.

3) Calculate the relative error of predicted compressive strength for each member of the group, and omit the member with the highest error.

4) Repeat the previous steps for the remained members. If both the significance and correlation coefficient improved, the omission would be confirmed, and the process would continue. Otherwise, the omitted member would be restored, and the process finishes.

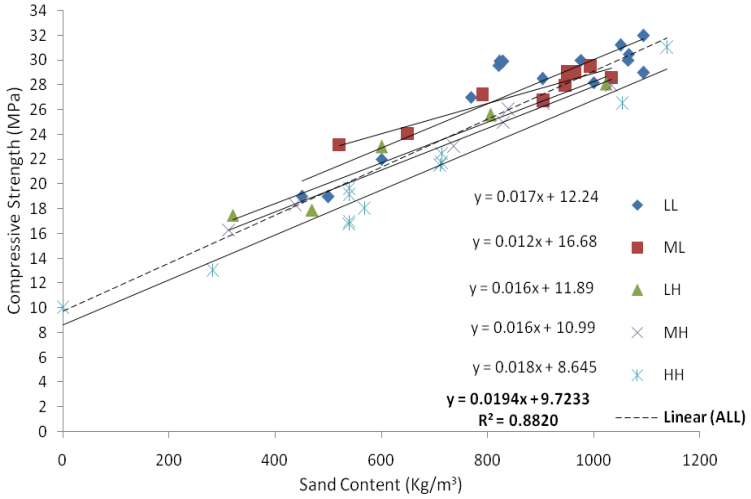


Fig. 7 Cleansed experimental data and corresponding regression lines

Table 7 Key parameters of the linear regression in each group before and after cleansing

Group	Before cleansing			After cleansing		
	Count	R^2	Significance	Count	R^2	Significance
HH	12	0.9559	0.9999999	12	Not changed	
MH	7	0.9456	0.9999402	6	0.9902	0.9999947
LH	6	0.8308	0.9971475	5	0.9411	0.9984575
ML	13	0.4258	0.9960943	8	0.9013	0.9999601
LL	20	0.5343	0.9999582	15	0.8146	0.9999990

Table 8 Capabilities of neural networks with different number of hidden layers (after Heaton 2008)

No. of hidden layers	Capability
none	Representing linear separable functions
1	Approximating any function that contains a continuous mapping from one finite space to another
2	Representing an arbitrary decision boundary to arbitrary accuracy with rational activation functions and can approximate any smooth mapping to any accuracy

The cleansing method was applied on the experiments. Table 6 shows the strength and significance of the linear regression in each group before and after cleansing. The cleansed experimental data and corresponding regression lines are shown in Fig. 7. In the following sections the cleansed data are used.

5. Multiple linear regression

MLR is a method used to model the linear relationship between a dependent variable, *Y*, and one or more independent variables, X_i. In fact, the simple linear regression is a special case of MLR. The relationship between *Y* and $X_1, X_2, \dots, X_p$, is formulated as a linear model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_P X_P \tag{2}$$

where $\beta_0, \beta_1, \beta_2, \ldots,$ and β_P are constants referred to as the model partial regression coefficients or simply as the regression coefficients. The regression coefficients are found based on least squares method. Eq. (2) is a linear relation between the response and predictors; however, the relationship between the response and a certain predictor is not always linear. In this case the data should be transformed to make the relationships linear.

According to Figs. (2)-(4), the relationship between σ and S as well as the relationship between σ and w/c are assumed to be linear, and considering, Eq. (3) describes the relationship between σ and *C* is supposed to be parabolic. Eq. 3 describes the relationship between σ and its predictors

$$\sigma = \beta_1 C^2 + \beta_2 C + \beta_3 S + \beta_3 (WC) + \beta_0 \tag{3}$$

To transform the data and make the relationship linear, C^2 is introduced as a new independent variable, say C_2. Eq. (4) is the transformed equation

$$\sigma = \beta_1 C_2 + \beta_2 C + \beta_3 S + \beta_3 (WC) + \beta_0 \tag{4}$$

The regression coefficients can be found using any spreadsheet software such as MS-Excel (LINEST function). Eq. (5) is the final resultant equation for predicting the compressive strength of SLWAC using MLR. The coefficient of determination, R^2, is 0.9352. In this equation *S* and *C* are expressed in tons.

$$\sigma = -261.049C^2 + 189.057C + 16.037S - 15.909(WC) - 14.263 \tag{5}$$

6. Artificial neural network

An Artificial Neural Network is similar to a function with a set of inputs and a set of outputs. A neural network consists of many simple elements called neurons. A neuron may have many inputs but a single output. The value of each input multiplies by its corresponding coefficient, referred to as its weight. The products is then summed up and feed to a function referred as transfer function (or activation function), to produce the single output. The neurons are arranged in layers. Eq. (6) expresses the process in a single neuron

$$O = f(\sum w_i x_i + b) \tag{6}$$

where O is the output, f is the activation function, x_i is an input value, w_i is its corresponding weight, and b is the bias value which can be considered as weight of a constant input equal to 1. The appropriate values for weights are calculated in a process called training. A set of inputs with known targets (resulted from experiment) is fed into the network, and in each epoch the weight changes in a manner which decreases the error between the output of the network and the target values. Many alternative training processes are available such as back propagation (BP) and cascade correlation (CC) schemes (Alshihri *et al*. 2009), but the BP is the most popular training method (Topcu 2008, Alshihri *et al*. 2009).

There are three types of layers, namely input layer, hidden layer, output layer. There is one neuron for each input value in the input layer. No calculation is performed in this layer. The computations are carried out in hidden layer(s). The output layer, which consists of neurons with linear activation function, produces the output of hidden layer and transforms it to the target scale.

6.1 Topology of neural network

The first step to determine the topology of an ANN is to decide on the number of hidden layers. Currently, there is no theoretical reason to employ more than two hidden layers. In fact, for many practical problems, there is no reason to use an ANN even with more than one hidden layer (Heaton 2008). Table 7 summarizes the capabilities of neural network architectures with various hidden layers.

According to Table 7 , only one hidden layer is employed in the present research. The training method is the back-propagation with the Levenberg–Marquardt algorithm. In this method the experimental data are divided into three groups, namely training, validation, and test. The training data are used to train the network and find the appropriate weight values. The validation set, stops the training process when overfitting is observed. Overfitting is modeling the noises and errors of data which may cause diversion in interpolations. The test set, is an unseen data set which is employed to evaluate the performance of the network.

There is no general rule for selecting the number of neurons in hidden layer (Kewalramani and Gupta 2006, Lin *et al*. 2003). Some authors suggested some rules of thumbs which relate the number of neurons to the number of input and output variables and the number of training patterns

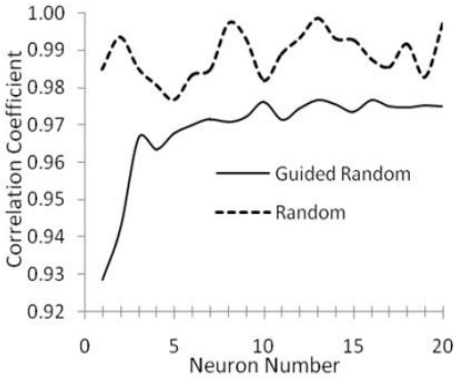


Fig. 8 Comparison of maximum R for test data for two different methods of trial and error procedure

(Rogers 1994, Swingler 1996, Heaton 2008), but these rules do not guarantee generalization of the network (Alshihri *et al*. 2009). In the present paper, trial and error is employed to develop a stable network with high accuracy prediction.

The trial and error process is performed using a MATLAB code and its ANN toolbox. The number of neurons and their activation functions were subject of the trial and error process. Furthermore, each network is initialized and trained 100 times to increase the possibility of occurrence of the best initial weights. Note that the training, validation, and test sets should not change too many times randomly. Random selection of sets increases the probability of selecting a misleading set for testing the network. The misleading set is a set which produces a high coefficient of determination with inappropriately located data, for example very close to the training data. In the present research the sets are chosen once randomly but in a guided and inspected manner. The neuron number of the hidden layer was in the range of 1-20 and the layer activation function is chosen among three MATLAB built-in functions: Logsig, Tansig, and Purelin. Logsig and Tansig are non-linear sigmoid functions, while Purelin is a simple linear transfer.

Although the value of correlation coefficient shows the strength of the network, but it should be noted that there are always several sources of uncertainty in input and output values. Furthermore, not all of the effective parameters can be considered in modeling. Therefore, there must be a difference between the correlation coefficient of any modeling with the perfect value, 1. The difference is inversely proportional to the accuracy of experimental data and modeling complexity. Thus, there should be something wrong with a network with an extremely high R. For example, Fig. 8 compares two different method of trial and error procedure. In the first method, the subsets (train, validation, test) are once chosen guided randomly and are not a subject of numerous trial and error steps. In the second method, the data set is partitioned in the trial and error loop, and therefore so many different kind of partitioning is tested. As mentioned before, this method increases the possibility of choosing an inappropriate set for testing data, for example data so close to the training data or inappropriately jammed in a limited range of values. It can be seen in Fig. 8 that the *R* values of the second method are very close to 1 with the maximum value of 0.9985. This high value of *R* is meaningless due to the presence of error and uncertainty in the experimental data. Furthermore, the *R* value is not growing with increasing neuron numbers in the second method. Therefore the second method of trial and error should be avoided.

6.2 Finding the most efficient topology

Fig. 9 shows the correlation coefficient for training, testing and all data, which were obtained after 100 times of weight initialization, for each totpology. Both maximum and average values of all 100 attempts are included. According to the values of maximum R (Fig. 9 (a), (b), and (c)), It can be seen that increasing the neuron numbers while their activation function is purelin, does not increase capability of the network. This is due to the fact that the combination of several linear functions is a linear function itself; even increasing the number of hidden layers would not increase the network performance. The capability of networks with either tansig or logsig activation function is similar and it is growing by increasing the number of neurons, but the capability is approximately constant for networks with more than 10 neurons. Studying Fig. 9 (d), (e), and (f) indicates that the average *R* value for networks with tansig and logsig activation function moves apart after reaching a certain number of neuron numbers, which hints logsig activation function to be a better choice. This might be because of the output range of these

functions. The output range of tansig is (-1,1), however, the target values are mapped in the range of [0,1] which almost coincides with the output range of logsig, (0,1). This fact would not influence the maximum strength of the network, but it may decrease occurrence possibility of the maximum strength, which will lead to a lower average value of *R*.

Note that the maximum values plotted in Fig. 9 for different data subsets are not corresponding to each other, in other words, a network giving the highest *R* value for testing data does not guarantee to produce the highest value of *R* for the whole data. Therefore, one should choose the best network according to the correlation coefficient of all three subsets (train, validation, and test) and the whole data. Fig. 10 shows a cross sectional plot, i.e. where two parameters out of three are kept

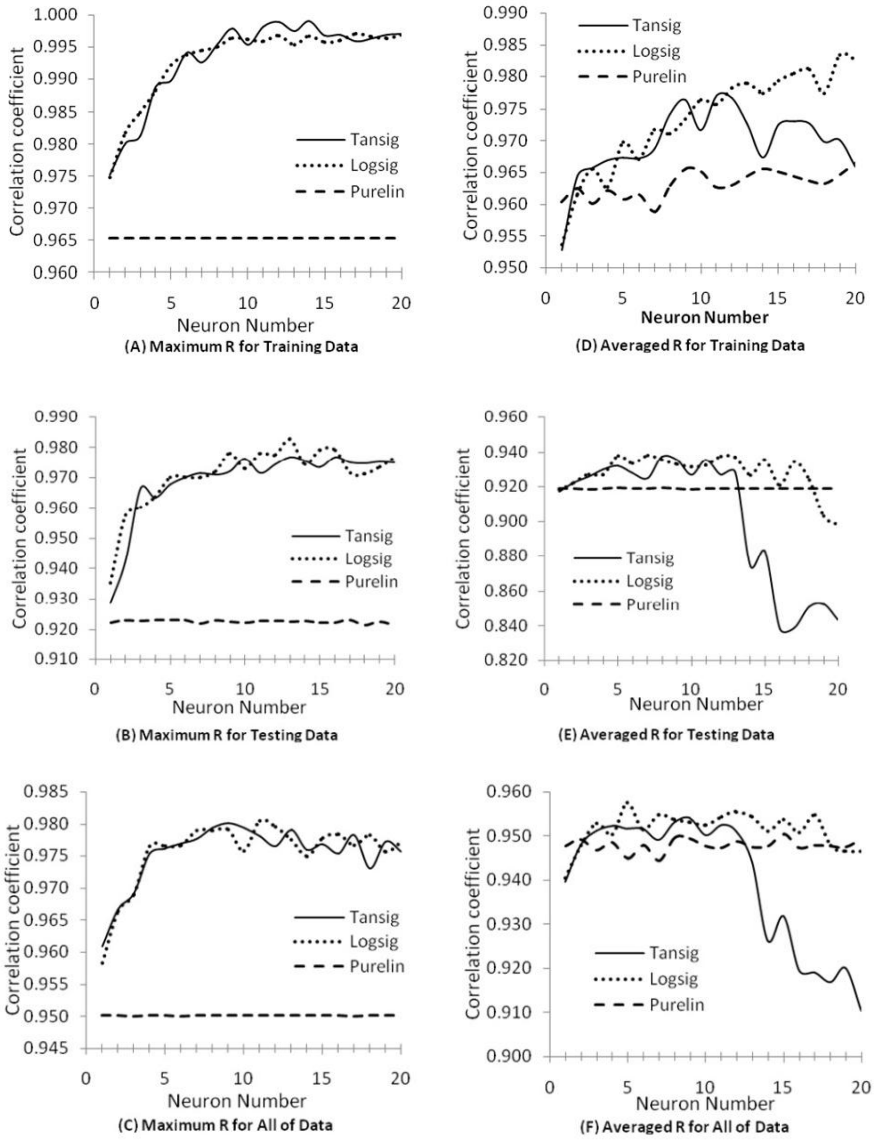


Fig. 9 Relationship between correlation coefficient and neuron numbers

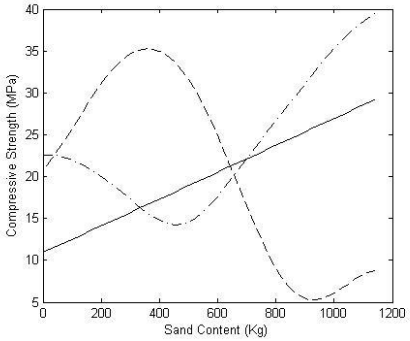


Fig. 10 Cross sectional plots (C = 432 kg ; w/c = 0.478)

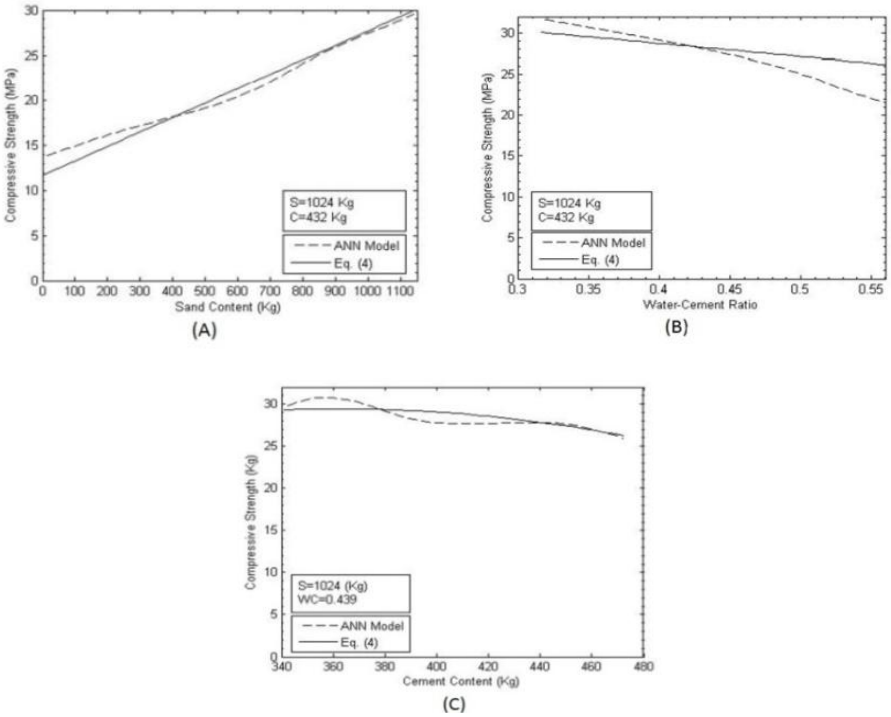


Fig. 11 Cross sectional plots of the selected ANN model-Relationship between compressive strength and (a) sand content (b) w/c (c) Cement content

constant, of two ANN models and their *R* values. Eq. (4) is also plotted as a reference value. It can be seen that deciding on the correlation coefficient of only one subset can be misleading.

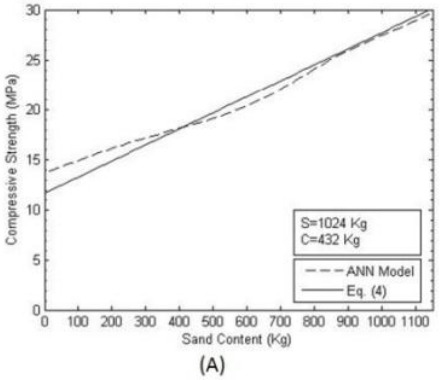
According to aforementioned criteria, a network with 8 neurons and logsig activation function is chosen as the best network. The *R* value of the testing subset is 0.978 while it is 0.980 for the whole data set. The corresponding R^2 values are 0.956 and 0.961, respectively. Fig. 11 shows a few cross sectional plots of the selected ANN model in which the predictions of MLR and ANN are compared.

7. Input domain of models

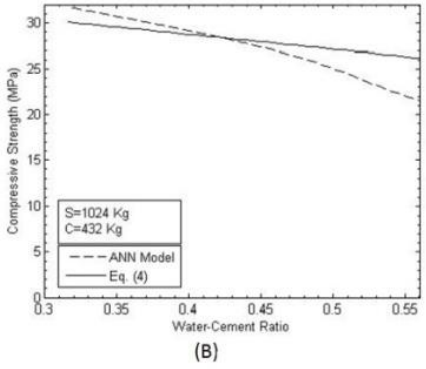
The input domain of a modeling can be defined as the range of experimental data fed to it. Although it is a very important property of the modeling, it is often neglected. A model predicts the response value by interpolation in its input domain, and it may be able to extrapolate to some extent out of its input domain. A first guess is to limit the domain to the minimum and maximum values of input parameters, which is called min-max domain in this paper. For example Table 3 presents the min-max domain in the present research. But in fact, the real range of experimental data is a subset of its min-max domain.

Fig. 12 shows the pair-wise plot of input parameters of experimental data. It can be seen that considering the pair-wise distribution of parameters, the input domain shrinks to a smaller area and any model fed by these data will extrapolate out of the domain limited to the dashed line.

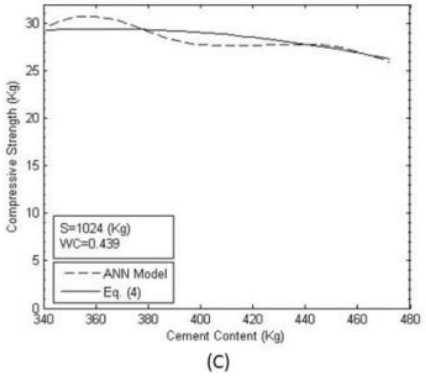
ANN is not suitable for extrapolation (Capecchi *et al.* 2010, Nazari and Mozafari 2011) because of its flexibility. Therefore the ANN model trained in this paper is not suitable to be used out of the range specified in Fig. 12.



(a) Cement content vs. normalized w/c



(b) Cement content vs. normalized sand content



(c) Normalized w/c ratio vs. noormalized sand content

Fig. 12 Pair- wise plots of input parameters for experimental date

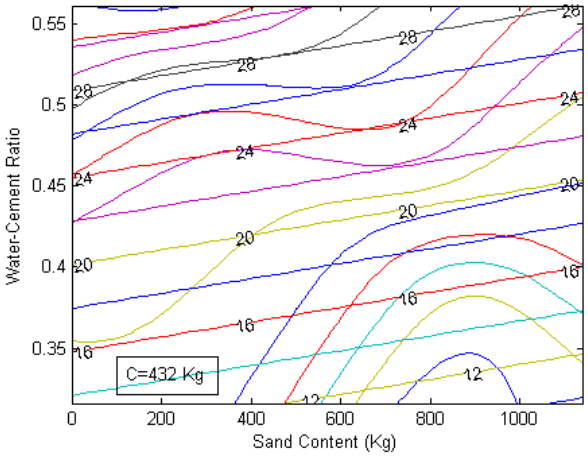


Fig. 13 Contour plot of Eq. (4) (straight lines) and ANN model (curved lines)

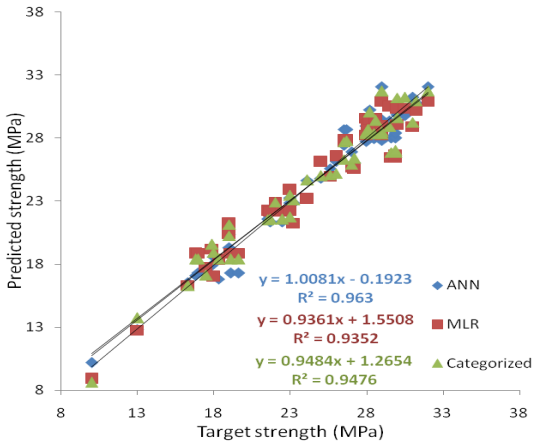


Fig. 14 Target values vs. predicted values for ANN, MLR and categorized linear regression

MLR can extrapolate to some extent because it fits a relation to the data which can be true even out of experimental output range. Fig. 13 shows the contour plot of Eq. (4) (straight lines) as the reference and the contour plot of ANN model (curved lines). It can be seen that the contour plots are only close to each other near the main diagonal of the figure. Superposing Fig. 13 on Fig. 12 (C) can explain why these two models are too much far in some area. Obviously the two series of contour plots fall apart in areas which are out of the input domain. Although none of the contours are experimental data and one cannot easily decide which one is correct, the pattern of ANN contours seems to be wrong in the area out of the input domain (Fig. 12(C)).

Fig. 14 presents target values against predicted compressive strength by ANN, MLR and categorized linear regression methods. As mentioned earlier, degree of freedom is higher for categorized linear regression method comparing to MLR approach; therefore it leads to a higher coefficient of determination. Despite this high R^2 value, application of categorized linear regression method is not recommended for prediction of SLWAC compressive strength due to its high degree of freedom, which is 10 in the present study. According to Fig. 14 and considering R^2

values, ANN is the most accurate method for prediction of SLWAC compressive strength. The main limitation for the method is its high sensitivity to the proper input data, which must be in its valid input domain. ANN should not be used as a black box tool, since it may lead to misleading results and predictions.

8. Conclusions

In this study, a statistical method based on cross-sectional linear regression was introduced for data cleansing. Then, MLR and ANN approaches were employed to predict the compressive strength of SLWAC and the results were compared. Results indicated that the compressive strength value of SLWAC is successfully predicted using cement content, sand content and w/c as inputs. Although the accuracy of ANN approach was higher than MLR, there are some limitations to be considered. These limitations for application of ANN include high sensitivity to validity of input data and selecting the most efficient network.

Furthermore, multi-linear regression method can be used to extrapolate to some extent because it fits a relation to the data which can be true even out of experimental output range; however, ANN is not suitable for extrapolation because of its high flexibility. For sure, extrapolation even by employing MLR should be always done bearing in mind the fact that it may digress from the reality.

References

Alshihri, M., Azmy, A. and El-Bisy, M. (2009), "Neural networks for predicting compressive strength of structural light weight concrete", *Constr. Build.mater.*, **23**(6), 2214-2219.

Andiç-Çakır, Ö. and Hızal, S. (2012), "Influence of elevated temperatures on the mechanical properties and microstructure of self consolidating lightweight aggregate concrete", *Constr. Build.mater.*, **34**, 575-583.

Capecchi, V., Buscema, M., Contucci, P. and D'Amore, B. (2010), Applications of Mathematics in Models, Artificial Neural Networks and Arts, Springer, Heidelberg, Dordrecht, London, New York.

Chandra, S. and Berntsson, L. (2002), Lightweight Aggregate Concrete Science: Technology and Application, Noyes Publications.

Chattefuee, S. and Hadi, A. (2006), Regression Analysis by Example, 4th Edition, JOHN WILEY & SONS.

Clarke, J.L. (1993), Structural Lightweight Aggregate Concrete, Taylor & Francis.

Cui, H.Z., Lo, T., Memon, S.H., Xing, F. and Shi, X. (2012), "Analytical model for compressive strength, elastic modulus and peak strain of structural lightweight aggregate concrete", *Constr. Build.mater.*, **36**, 1036-1043.

Dafico, D.A., and Prudencio, L.R. (2002), "A simple model to explain compressive strength of high-performance concrete", *High performance concrete: performance and quality of concrete structures proceedings*, Third International Conference, American Concrete Institute.

Fedorov, V., Mannino, F. and Zhang, R. (2009), Consequences of dichotomization. Pharm Stat, 8:50-61.

Heaton, J.T. (2008), Introduction to Neural Networks With Java, 2nd edition, Heaton Research.

Huang, J.L., Huang, P.Y., Li, Z.W. and Zheng, X. (2010), "Prediction of bond strength of FRP-concrete based on multiple linear regression method", *Adv. Mater. Res.*, 163-167.

Golias, M., Castro, J. and Weiss, J. (2012), "The influence of the initial moisture content of lightweight aggregate on internal curing", *Constr. Build.mater.*, **35**, 52-62.

Topçu, İ. and Saridemir, M. (2008), "Prediction of rubberized concrete properties using artificial neural network and fuzzy logic", *Constr. Build.mater.*, **22**, 532-540.

Oreta, A. and Ongpeng, J. (2011), "Modeling the confined compressive strength of hybrid circular concrete columns using neural networks", *Comput. Concr.*: *Int. J.*, **8**(5).
Nazari, A. and Mozafari, M. (2012), "Simulation of structural features on mechanochemical synthesis of Al2O3–TiB2 nanocomposite by optimized artificial neural network", *Adv. Powder Tech.*, **23**(2), 220-227.
Mazaheripour, H., Ghanbarpour, S., Mirmoradi, S.H., Hosseinpour, I. (2011), "The effect of polypropylene fibers on the properties of fresh and hardened lightweight self-compacting concrete", *Constr. Build. Mater.*, **25**(1), 351-358.
Rogers, J.L. (1994), "Simulating structural analysis with neural networks", *ASCE J. Comput. Civil Eng.*, **8**(2), 252-2265.
Swingler, K. (1996), Applying Neural Networks a Practical Guide, Academic Press, New York, London.
Kewalramani, M.A. and Gupta, R. (2006), "Concrete compressive strength prediction using ultrasonic pulse velocity through artificial neural networks", *Automat. Constr.*, **15**, 374-379.
Lin, Y., Lai, C.P. and Yen, T. (2003), "Prediction of ultrasonic pulse velocity (UPV) in concrete", *ACI Mater. J.*, **100**(1), 21-28.
Rawlings, J., Pantula, S. and Dickey, D., (1998), *Applied Regression Analysis: A Research Tool*, 2nd edition, Springer.
Royston, P., Altman, D. and Sauerbrei. W. (2006), "Dichotomizing continuous predictors in multiple regression: a bad idea", *Stat. Med.*, **25**, 127-141.
Yuzer, N., Akbas, B. and Kizilkanat, A.B. (2011), "Predicting the high temperature effect on mortar compressive strength by neural network", *Comput. Concr., Int. J.*, **8**(5), 491-511.
Zain, M. and Abd, S.M. (2009), "Multiple regression model for compressive strength prediction of high performance concrete", *J. Appl. Sci.*, **9**, 155-160.

CC