CHAPTER 6

Modeling using artificial neural network (ANN)

Artificial neural network architecture.

6-1 Introduction.

The architecture of ANN mimics that of biological neurons and their operation essentially simulatesthe internal operation of the human brain**[Liu, Y. A., Baughman, 1995].** In recent years, ANNhas shown exceptional performanceas a regression tool, especially when used for pattern recognition and function estimation.Neural networks have been trained to performcomplex functions in various fields that are difficult forconventional computers such as; aerospace,automotive, banking, credit card activity checking, defense,electronics, entertainment, financial, industrial, insurance,manufacturing, medical, oil and gas, robotics, speech, securities,telecommunications, transportation, and civil engineering**[E. Guneyisi, et al, 2009]**.

They are highly nonlinear, and can capture complex interactions among input/output variables in the systemwithout any prior knowledge about the nature of these interactions. A neural network is an empiricalmodeling tool, and it does operate by "curve-fitting". However, some notable differences exist betweenneural networks and typical, traditional empirical models. In comparison to traditional methods, ANNstolerate relatively imprecise, noisy or incomplete data, approximate results, are less vulnerable tooutliers, have better filtering capacity, and are more adaptive. Moreover, ANNs are also massively parallel, that is, their numerous independent operations can be executed simultaneously. Some of thelimitations of the neural networks are possible long training times, the need for large amount of reliabletraining data, and no guarantee of optimal results.

Also the mostimportant property of ANN in civil engineering problems isits capability of learningdirectly from examples. The other important properties of ANN are their correct or nearlycorrect response to incomplete tasks, their extraction of information from noisy or poor data,and their production of generalized results from the novel cases.The above-mentionedcapabilities make ANN a very powerful tool to solve many civil engineering problems,particularly problems, where data may be complex or in an insufficient amount **[A.B. Goktepe, et al. 2006].** Thebasic strategy for developing an ANN system based models for material behavior is to trainan ANN system on the results of a series of experiments using that material **[W.P.S. Dias, et al 2001].** If theexperimental results include the relevant information about the material behavior, then thetrained ANN system will contain enough information about material's behavior to qualify asa material model. Such a trained ANN system not only would be able to reproducethe experimental results, but also they would be able to approximate the results in otherexperiments through their generalization capability [**W.P.S. Dias, et al 2001**].

On the other hand neural network is adjusted, based on a comparison of theoutput and the target, when the network output matches thetarget. Typically, many such input/target pairs are needed totrain a network.

A neural networkconsists basically of a number of simple processing units called neurons. Typically,the neurons are organized logically into groupings called layers. The network ishierarchical, consisting of three or more layers: the input layer, one or more hiddenlayers, and the output layer as presented in **Fig. 6.1**. Neural networks might be single-or multi layered. The singlelayerneural networks present processing units of the neural networks, which takeinput from the outside of the networks and transmit their output to the outside ofthe networks; otherwise, the neural networks are considered multi layered.Each neuron in a given layer is connected to allthe neurons in the next layer. The neurons in each layer interact with neurons inother layers through weighted connections.Each unit in the network receives an input from the lowest-level units andcomputes a weighted sum. The input data are propagated through the network bymultiplying the values of the neurons in the input layer by the connecting weights.The weights should be randomized initially in order to avoid the system to belocked at the starting point **(Baxter, 2001).** These products are summed at thetarget neurons in the first hidden layer. The summed products are operated on bya transfer function to determine the output level of each neuron. A linear sigmoidcould be used as the transfer function producing an analog output with valuesranging from 0 to 1 **(Deboeck, 1994)**. The signals from the first hidden layer arepropagated to the following hidden layers, if they exist, and to the output layer ina similar fashion.Although in theory a single hidden layer is sufficient to solve any function approximationproblem, some problems might be easier to solve using more than a singlehidden layer, especially two hidden layers **(Partovi and Anandrajan, 2002)**. The number of neurons in a hidden layer was usually determinedvia a trial and error procedure.

In the most general sense, the neural network is created for two different phases. The first phase is the training phase and the second phase is the testing (simulation) phase **(Tapkin et al. 2006)**.ANNs have the ability of performing with a good amount of generalization from the patterns onwhich they are trained. A) Training consists of exposing the neural network to a set of known input-output patterns **(Rafiq et al. 2001, Ashour and Alqedra, 2005)**. Several methods do exist to train a network. One of the most successful and widely usedtraining algorithms for multi-layered perceptron (MLP) is the feedbackpropagation **(Kartam et al. 1997)**. The neural network is operated using feedbackpropagation training algorithmin this study. Feedbackpropagation neural networks generally have a layered structure with an input, anoutput, and one or more hidden layers.The modification process is continued in the output layer, where the error between the networkoutputs and desired targets is calculated, and then propagated back to the network through alearning mechanism.

B) The computer program code for the ANN simulation, including neural networks toolbox, waswritten in MATLAB software. Different ANN architectures are tried and then the appropriate modelstructure is determined for the data sets. Numerous trials are carried out in the neural networkenvironment to determine neuron number of the hidden layers. Optimum hidden neuron numbersare obtained for different cases. The ANN model is then tested and the results are compared bymeans of root mean squared error(RMSE), and coefficient of determination (R^2) statistics.

Any difference between the output values and expected from the input pattern is interpreted as anerror in the system.In the previous researches the feasibility of using ANN to create an intelligentmodel for predicting the temperature of the reinforcement inside concrete under elevated temperature or the residual compressive strength has been demonstrated.

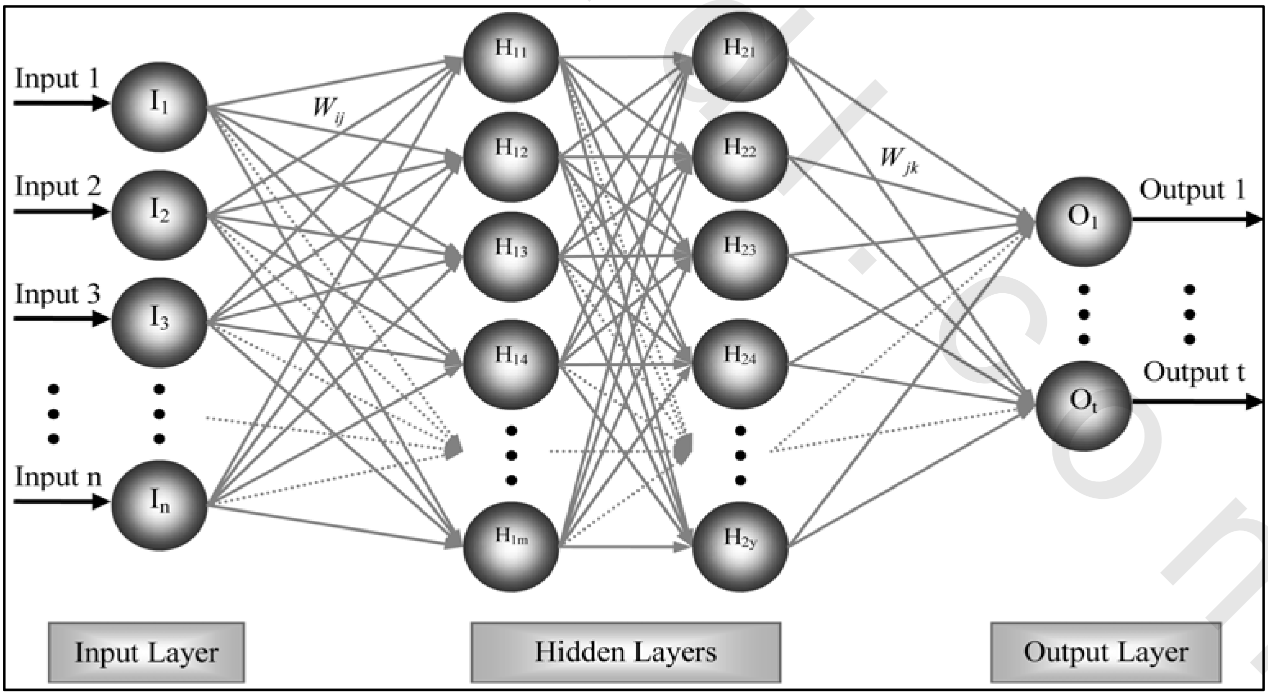

Fig. 6.1- A typical ANN topology with (n) input nodes, m and y hidden nodes, and t output nodes.

This study is aimed at finding the prediction temperature of the steel inside the small scale concrete slabs with different concrete cover thickness ranging from 1 in. to 3 in. and the prediction residual compressive strength of concrete cubes after exposed to 20, 200, 400 and 600°C for 2 hours through a comparative analysis by ANN. All of those results are mainly compared using the coefficients of determination calculatedfor the models with highest R^2 values.

6.2 ANN based Model 1: Predicting the temperature of the reinforcement inside small scale concrete slabs with different concrete cover thickness using ANN.

Artificial neural networks mimic the structure and operation of biological neurons and have theunique ability of self-learning, mapping, and functional approximation. They are highly nonlinear,massively parallel and can capture complex interactions among input/output variables in thesystem without any prior knowledge about the nature of these interactions.

Before proceeding with model development, some modelparameters were selected from the experimental program test results of this thesis. The total database size in the present study was 39 cases, considering 3 inputs and one output foreach model. These data are divided into 70% for training, 15%for testing (also called verification) and 15% for validation. The preliminary architecture of the neural networkaccording to MATLAB manual, see **Fig. 6.2** was conceived asfollows:

(a) Type of neural network: Multilayer perceptron feed-forwardwas trained through the error back-propagationalgorithm (this is the most commonly used type ofANN and its application to function approximationhas already been proven in several studies).

(b) Neurons in the first layer: Three neurons were specified using MATLAB manual according to model size.

(c) Hidden layers: It has been found that two hiddenlayer presents satisfactory results for many problems**[F. Ozcan, C.D. Atis, et al. 2009].**

(d) Neurons in hidden layer: Ten neurons were specifiedfrom empirical criteria.

(e) Number of outputs: Single output for the model (Temperature of the steel inside the conventional concrete slabs with different concrete cover).

Thus, 3 different ANN models have been used for the system model.For all of those ANN models, the following network parameters are taken the same:

Network type: Feed-forward backprop

Training function: TRAINLM

Performance function: MSE

Transfer function layer 1: LOGSIG as shown in **Fig. 6.3**.

Transfer function layer 2: PURELIN as shown in **Fig. 6.3**.

The activation function of the layer-1 used is the log-sigmoidal function. A sigmoid curve is produced by a mathematical function having an "S" shape. Often, sigmoid function refersto the special case of the logistic function**[Ren LQ,2002]** shown in **Fig. 6.4** and definedby the formula:

S (t) = 1 / (1+e-t)

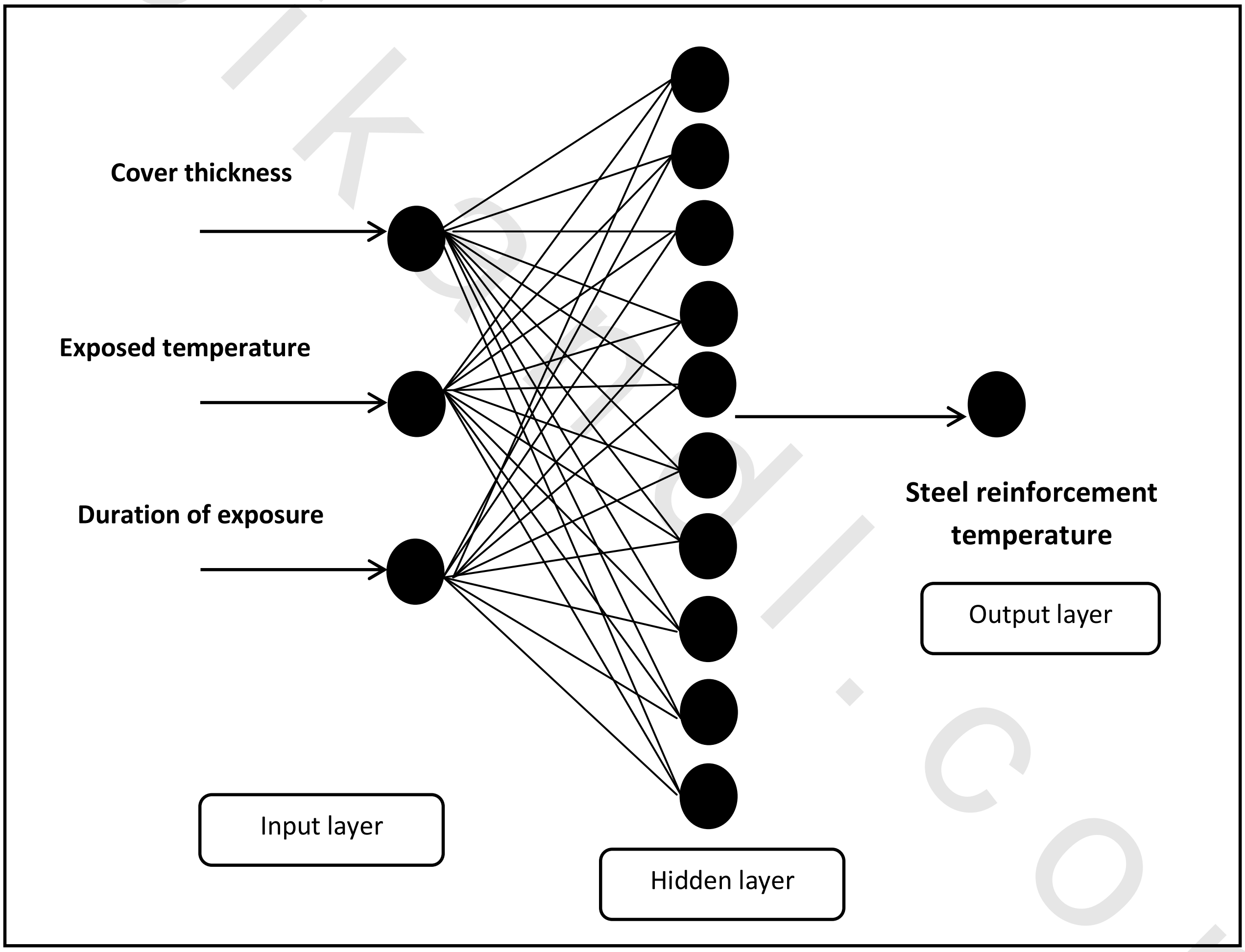

Fig. 6.2- Topology of the prediction model

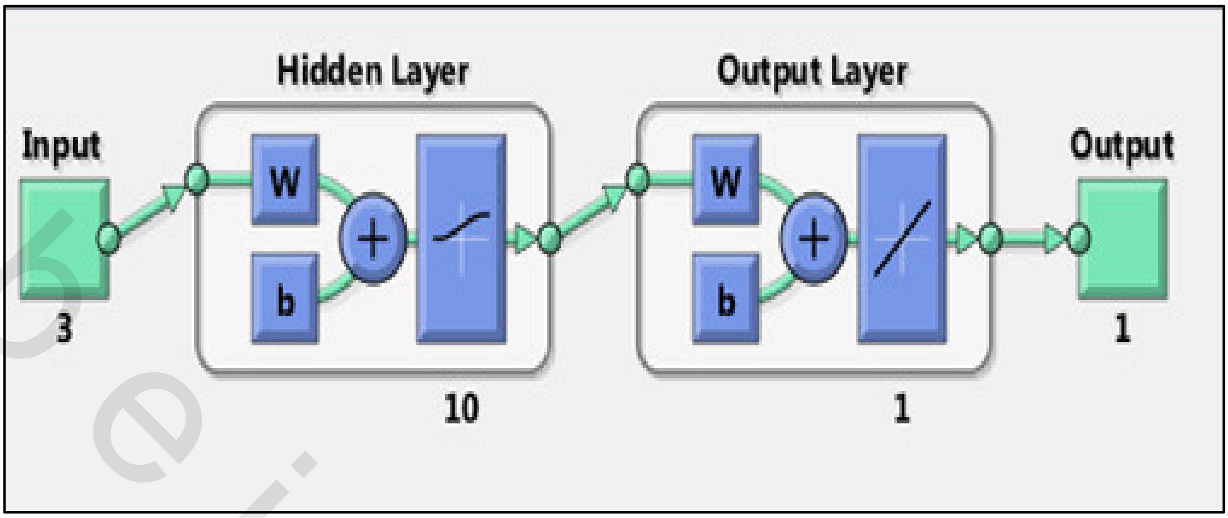

Fig. 6.3- Schematic of the input layer, hidden layer and output layer of the model as presented in ANN.

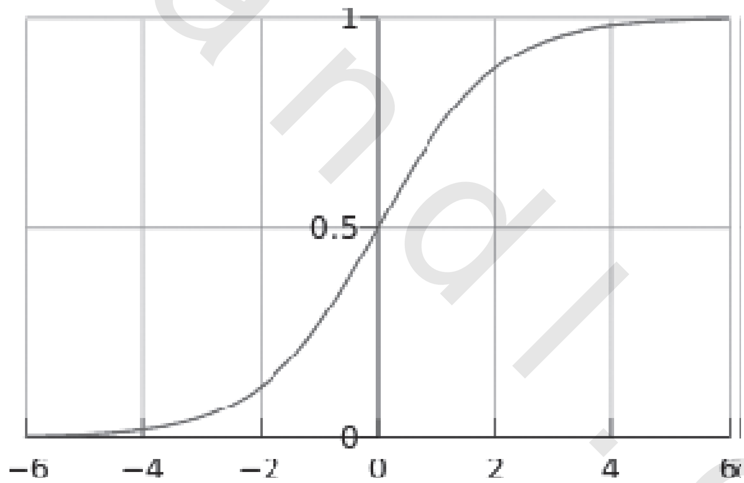

Fig.6.4- Logistic Curve

This thesis arranged the data used inthe neural network model in a format of three input parametersthat cover the concrete cover thickness, exposed surface temperature of the reinforced concrete slab and duration of fire exposure, the output parameter is steel reinforcement temperature inside the conventional concrete slab as seen in **Fig.6.2**.The ranges of the used variables in the database are presented in **Table 6-1.**

Table 6.1: Ranges of used variables in database.

Input variables	Minimum	Maximum
Concrete cover thickness (in.)	1	3
Exposed surface temperature (°C)	20	800
Duration of exposure (min)	0	120
Output variables		
Steel reinforcement temperature (°C)	20	275

The commercial software MATLAB was used for thedevelopment of the model. Thebasic methodology of neural networks consists of three processes: Networktraining, testing, and validation.As an important point the connection weights of the neuralnetwork are adjusted through the training process without any change in this study,neural networkslearn from examples and exhibit some capability for generalization beyond thetraining data. Then, other testing data are used to check the generalization.A script was developed andadjusted several times until the error criteria were met. In order to avoid overtraining of the network,the training was stopped when the testing error increased. Thisfeature is automatically set up in the software. Thetraining method used in the model development was theLevenberg–Marquardt algorithm which exhibits the fastestconvergence in similar problems.

In the present study, three forms were used to comparativeevaluation of the performance of the multilayer feed-forwardneural network model. These forms are root-mean-squared error(RMSE) error, coefficient of determination(R^2) and meanabsolute percentage error (MAPE) as given in **Eqs. (1)–(3).**These forms were calculated between model's results andexperimental results**:**

$$\mathbf{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(ti-oi)^2} \quad \ldots\ldots \mathbf{(1)}$$

$$\mathbf{R^2} = \frac{(n\sum_i^t oi-\sum_i^t\sum_i^o\ )^2}{(n\ \sum ti^2-(\sum ti\ )^2)\ (n\sum oi^2-(\sum oi)^2\ )} \quad \ldots\ldots \mathbf{(2)}$$

$$\textbf{MAPE} = \frac{1}{n}\left[\frac{\sum_{i=1}^{n}|ti-oi|}{\sum_{i=1}^{n} ti}\ X\ 100\right] \ldots\ldots\ldots\ldots\ldots\ldots \textbf{(3)}$$

Where, (n) is the number of observations,(t) is the target value, (o) is the output value.

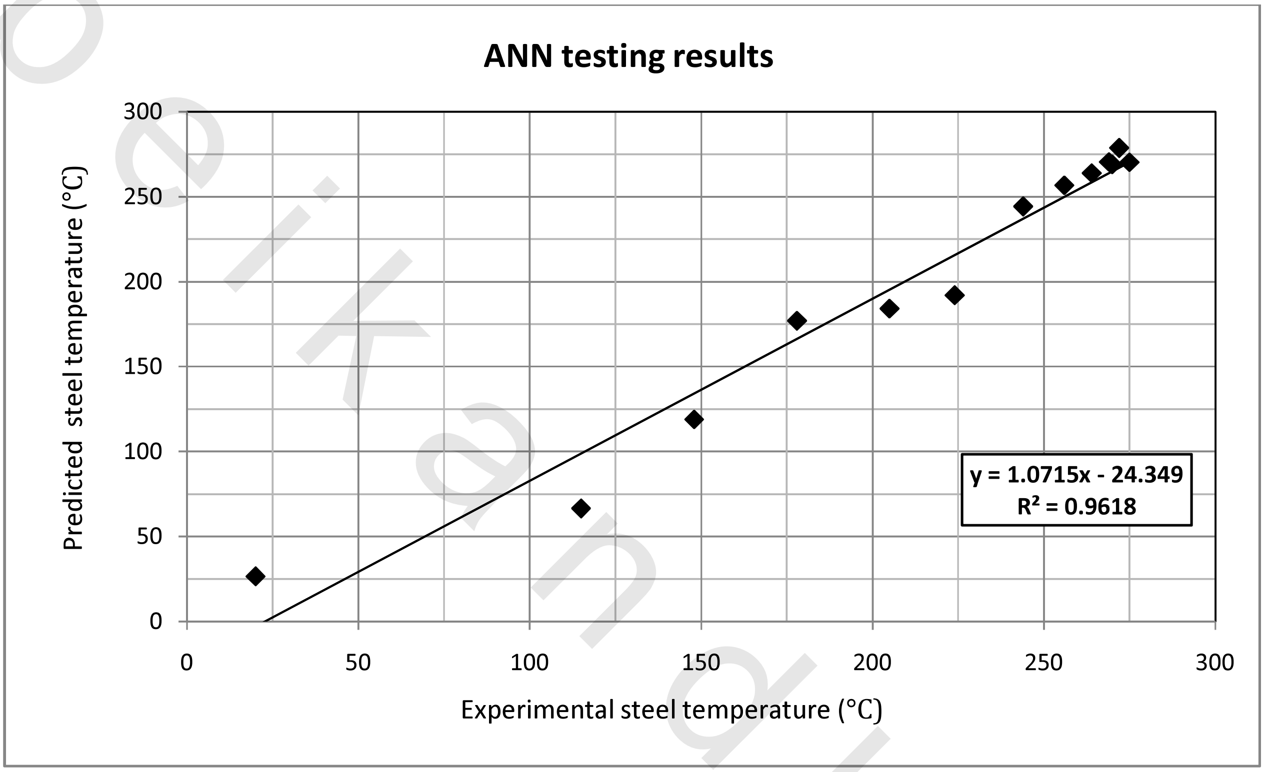

Fig. 6.5-Model-1: predicted vs. experimental steel temperature inside small scale concrete slab with 1 inch concrete cover thickness exposed to 800°C for 2 hrs.

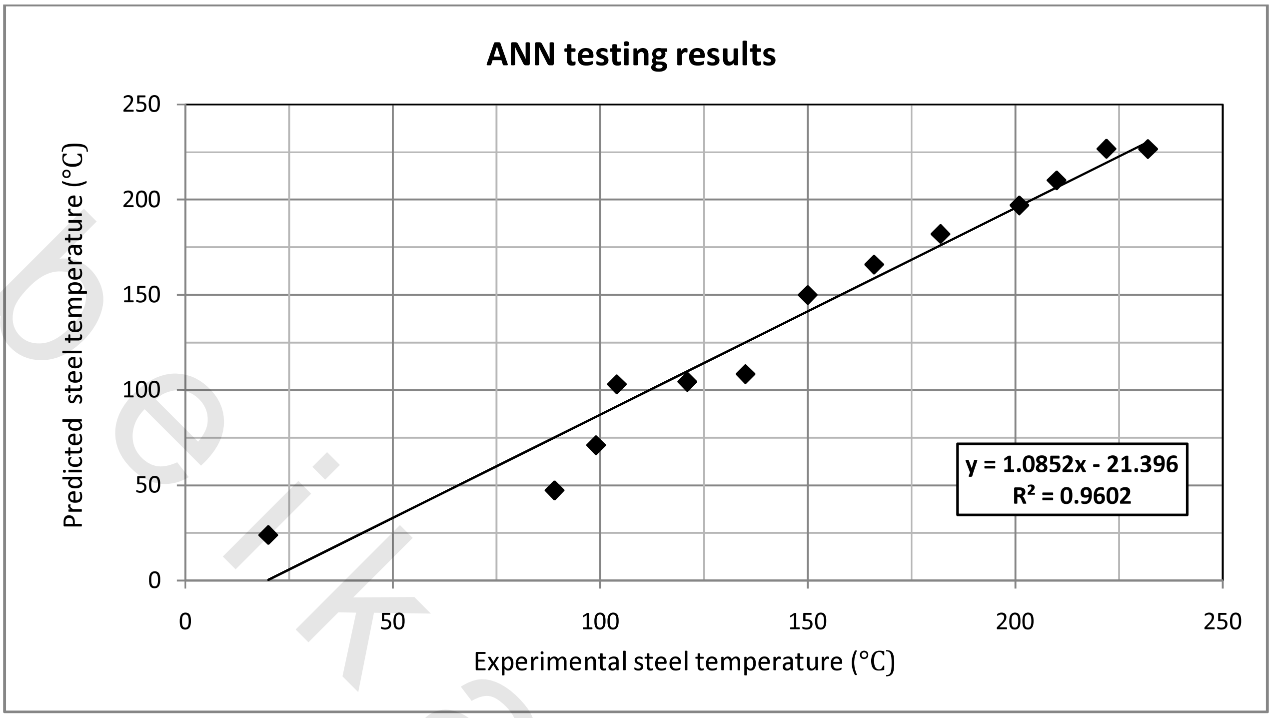

Fig. 6.6-Model-2: predicted vs. experimental steel temperature inside small scale concrete slab with 2 inch concrete cover thickness exposed to 800°C for 2 hrs.

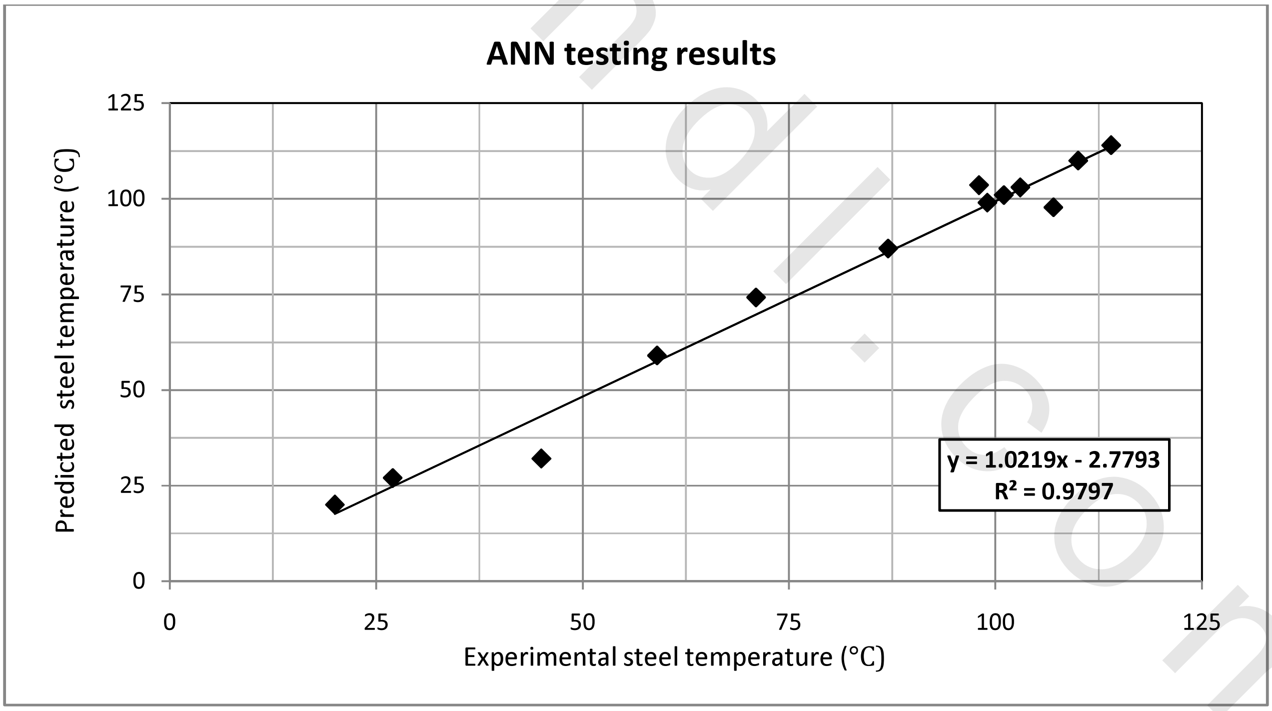

Fig. 6.7-Model-3: predicted vs. experimental steel temperature inside small scale concrete slab with 3 inch concrete coverthickness exposed to 800°C for 2 hrs.

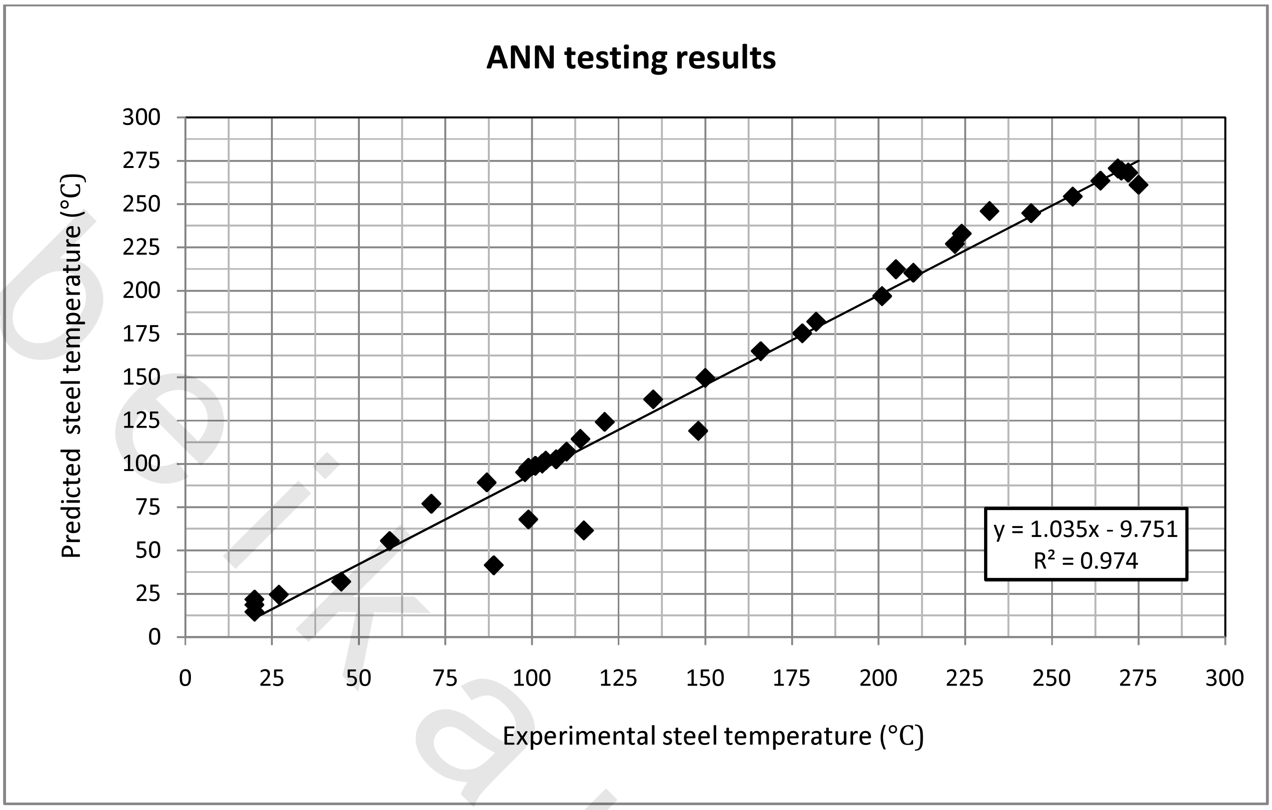

Fig. 6.8-Model-4: Over all predicted vs. experimental steel temperature inside small scale concrete slab with 1, 2, 3 inch concrete coverthickness exposed to 800°C for 2 hrs.

R^2 (coefficient of determination) is usually used to test the accuracy of thetrained network. The coefficient of determination is a measure of how well theindependent variables considered account for the measured dependent variable.A higher R^2 value indicates a better prediction relationship. Predictive capabilityof the method on the other hand was within the range of data employed for modelfitting **(Yeh, 1999).**

Profile plot is the best way to visualize thefitted model. By varying only one predictor (parameter) between two values and keeping all theothers fixed at some pre-specified values we get the profile plot – which is a one dimensional crosssection of the high dimensional fitted surfaces.

All results,obtained from experimental studies and the predicted values for steel temperature inside small scale concrete slab with different concrete cover thickness vs. timeare shown from**Figs. 6.5**to **Fig. 6.7.**From these figures, the values obtained from the trainingand testing using the ANN model were very close to the experimentalresults. The results of ANN model demonstrate that the(ANN) system can be successfully applied to establish accurateand reliable prediction models. The statistical parametervalues of RMS, R^2 and MAPE showed obviously this behavior.The statistical values of RMS, R^2 and MAPE includingall the ANN models, is given in **Table 6.2**. The best R^2valueobtained is 0.9797 for training set ANN at 3 inch concrete cover thickness, while, the minimumvalue of R^2 is 0.9602 for testing set ANN at 2 inch concrete cover thickness.

Table 6.2: The statistical values of proposed models.

Model no.	**ANN**		
	RMS	**R^2**	**MAPE**
1	19.10483549	09618	9.60989709
2	16.55034081	0.9602	10.42419948
3	4.755115537	0.9797	3.660961265

Over all predicted vs. experimental steel temperatures inside small scale concrete slab with 1, 2, 3 inch concrete coverthickness exposed to 800°C for 2 hrs were also studied to make a broader vision on using ANN as plotted in **Fig. 6.8**. To train the ANN models, first, the entire training data file is randomly divided into training and testing data sets. 27 sets are used to train the different network architectures. The remaining 12 patterns are used for testing and validation to verify the prediction ability of each trained ANN model. Linear correlation can be observed and the correlation coefficient is found to be 0.9747. Thus it can be concluded that the model successfully predicted the temperature of the steel inside the concrete slabs in good manner.

<u>6.3 ANN based Model 2: The use of neural networks in residual concrete compressive strength estimation.</u>

Experiments intended for determining the resistance of structures to mechanical effects under heat conditions, Thus two types of tests were carried out: "I" – determination of compressive resistance of the cube concrete specimen under normalambient conditions, "II"– determination of compressive resistance of the cube concrete specimen under elevatedtemperature conditions.The data were gathered at seven different concrete mixtures that exposed to different elevated temperature 20, 200, 400 and 600°C for 2 hrs as presented in **Table 6.3**. The basic parameters considered in this study were water to cement ratio, cement content, silica fume content, limestone powder content, bentonite content and exposed temperature.The ranges of various input and the output parameters used in data mining techniques are given in **Table 6.4**. The successful model to predict the 28 days compressive strength after exposed to elevated temperature depends upon the magnitude of the training data. Every model was analyzed separately and after that overall model was developed. The predicted results were compared with the values obtained experimentally,**Table 6.5** give the actual, predicted values along with error as has been shown.

Table 6.3: Serial number that used in modeling.

S.No	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Mix.	M-Control	M-Silica	M-(L.P.10R)	M-(L.P.15R)	M-(L.P.10A)	M-(L.P.15A)	M-(Ben.10R)

Table 6.4: Input and Output variables.

Variables	**Parameter**	**Abbreviation**	**Database Range**	
			Minimum	**Maximum**
Input	Water to cement ratio	W/C	0.40	0.40
	Cement (kg/m³)	Cement	340	400
	Silicafume(kg/m³)	SF	0	40
	Limestone powder (kg/m³)	LP	0	60
	Bentonite(kg/m³)	Ben	0	40
	Exposed temperature (°C)	T	20	600
Output	Compressive strength (MPa)	28-day CS	13	47.1

Table 6.5: Actual, predicted and error values of 28 days cube compressive strength of concrete at different elevated temperature 20, 200, 400 and 600°C using ANN technique.

S.No.	A (20)	P (20)	Error	A (200)	P (200)	Error	A (400)	P (400)	Error	A (600)	P (600)	Error
1	36.20	35.85	-0.35	30.10	32.20	2.10	28.00	27.00	-1.00	20.00	20.12	0.12
2	39.00	38.99	-0.01	35.00	38.70	3.70	29.00	26.93	-2.07	16.00	16.25	0.25
3	35.50	35.11	-0.39	28.50	30.95	2.45	24.70	23.15	-1.55	16.70	16.84	0.14
4	32.00	31.20	-0.80	23.00	25.34	2.34	19.00	15.19	-3.81	13.00	13.37	0.37
5	45.00	44.21	-0.79	36.00	38.28	2.28	33.00	30.17	-2.83	25.00	25.15	0.15
6	47.10	45.74	-1.36	36.46	38.33	1.86	34.48	32.26	-2.21	26.37	26.52	0.15
7	32.59	32.36	-0.23	27.90	29.65	1.75	26.41	25.10	-1.30	19.24	19.34	0.10

A: Actual compressive strength (MPa).

P: Predicted compressive strength (MPa).

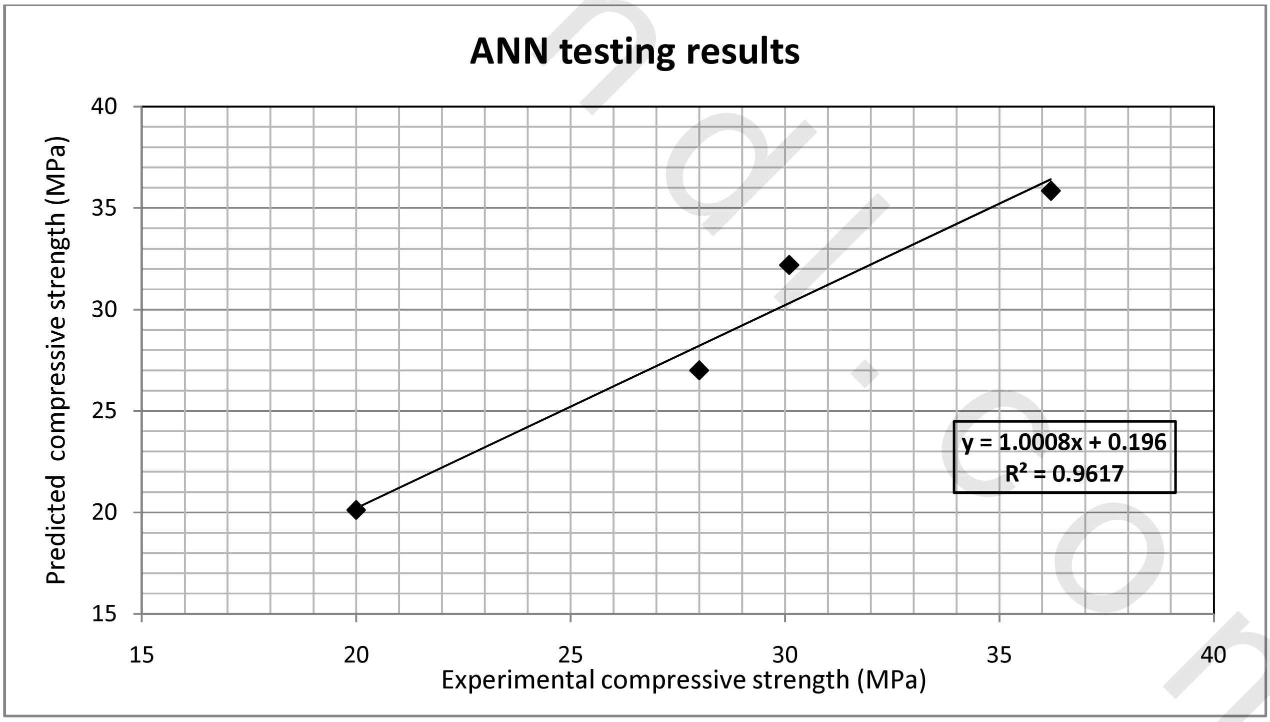

Fig. 6.9-predicted vs. experimental cube compressive strength of the M-Control after exposed to 20, 200, 400 and 600°C for 2 hrs.

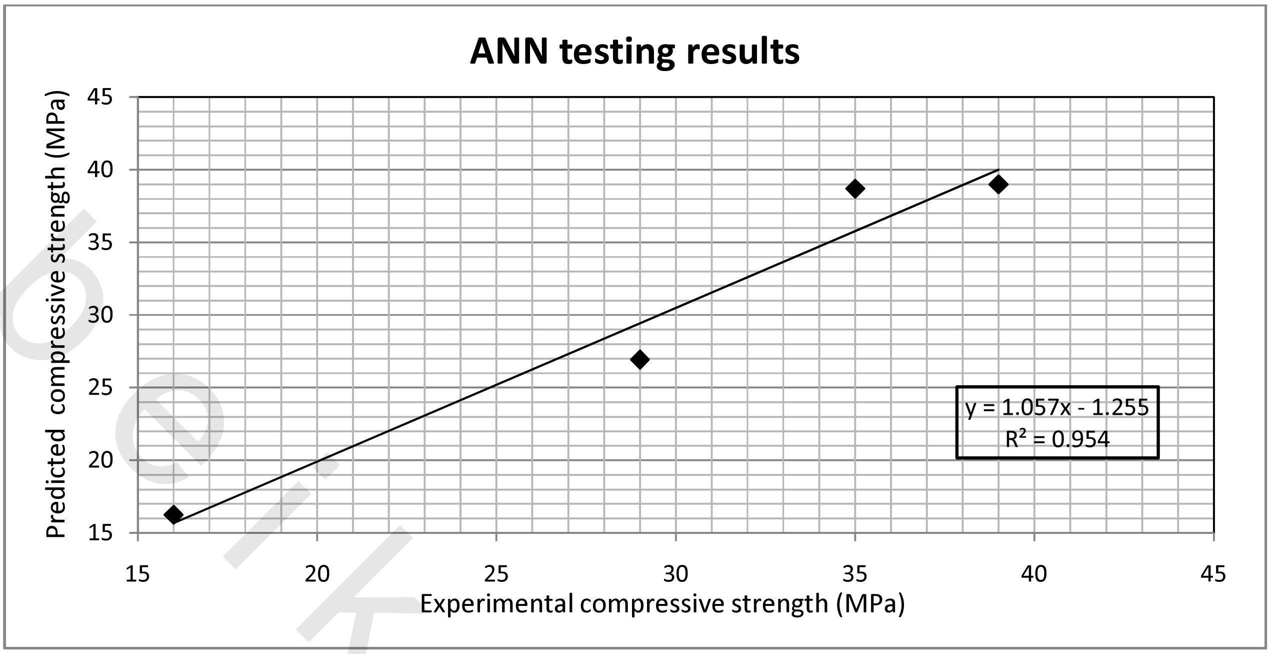

Fig. 6.10-predicted vs. experimental cube compressive strength of the M-Silica after exposed to 20, 200, 400 and 600°C for 2 hrs.

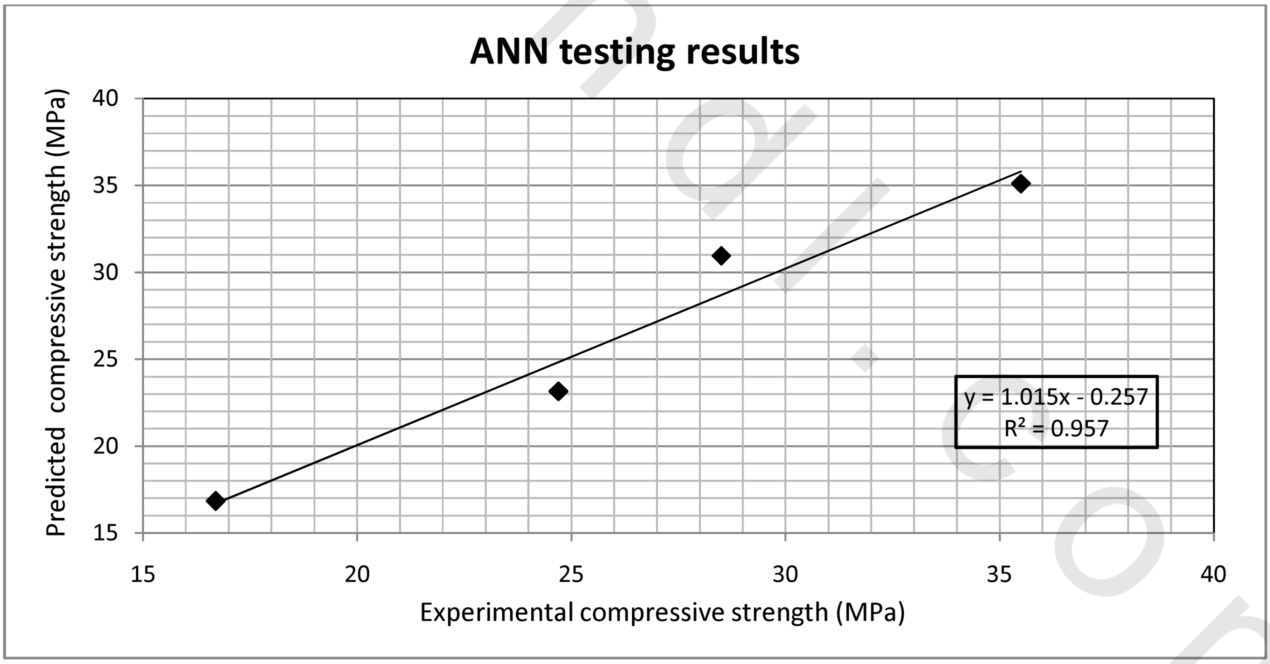

Fig. 6.11-predicted vs. experimental cube compressive strength of the M-(L.P.10R) after exposed to 20, 200, 400 and 600°C for 2 hrs.

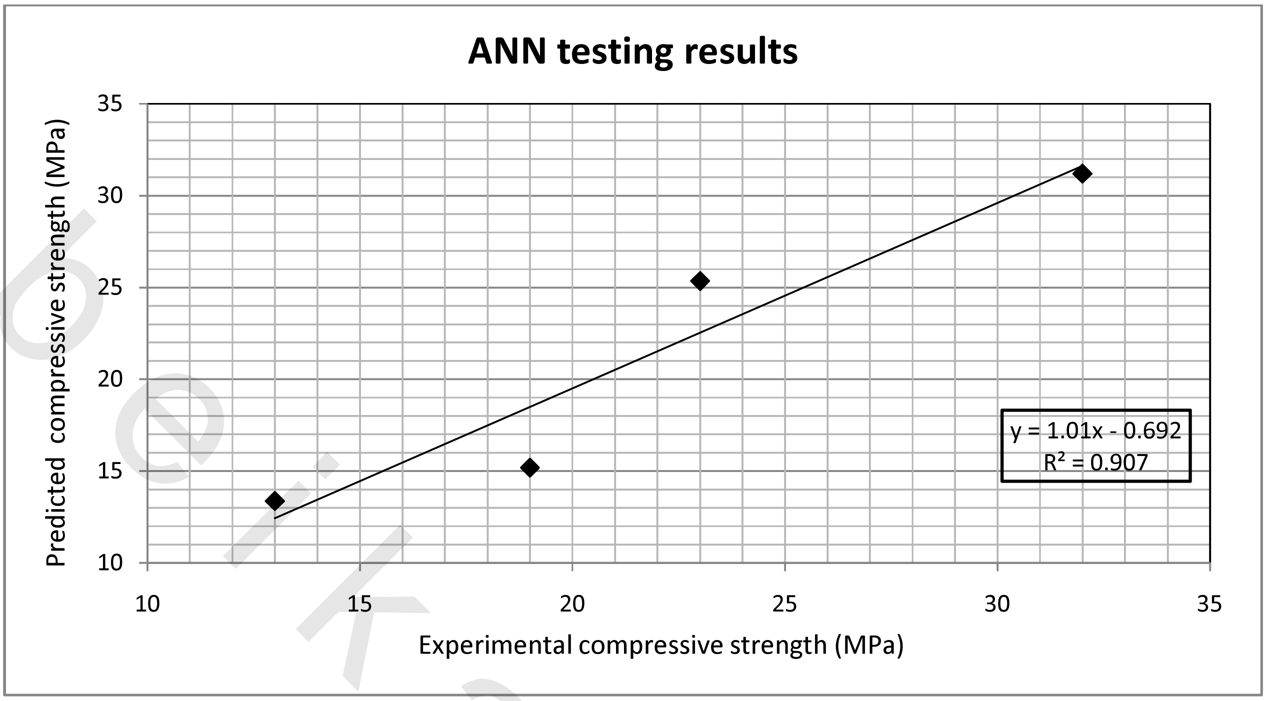

Fig. 6.12-predicted vs. experimental cube compressive strength of the M-(L.P.15R) after exposed to 20, 200, 400 and 600°C for 2 hrs.

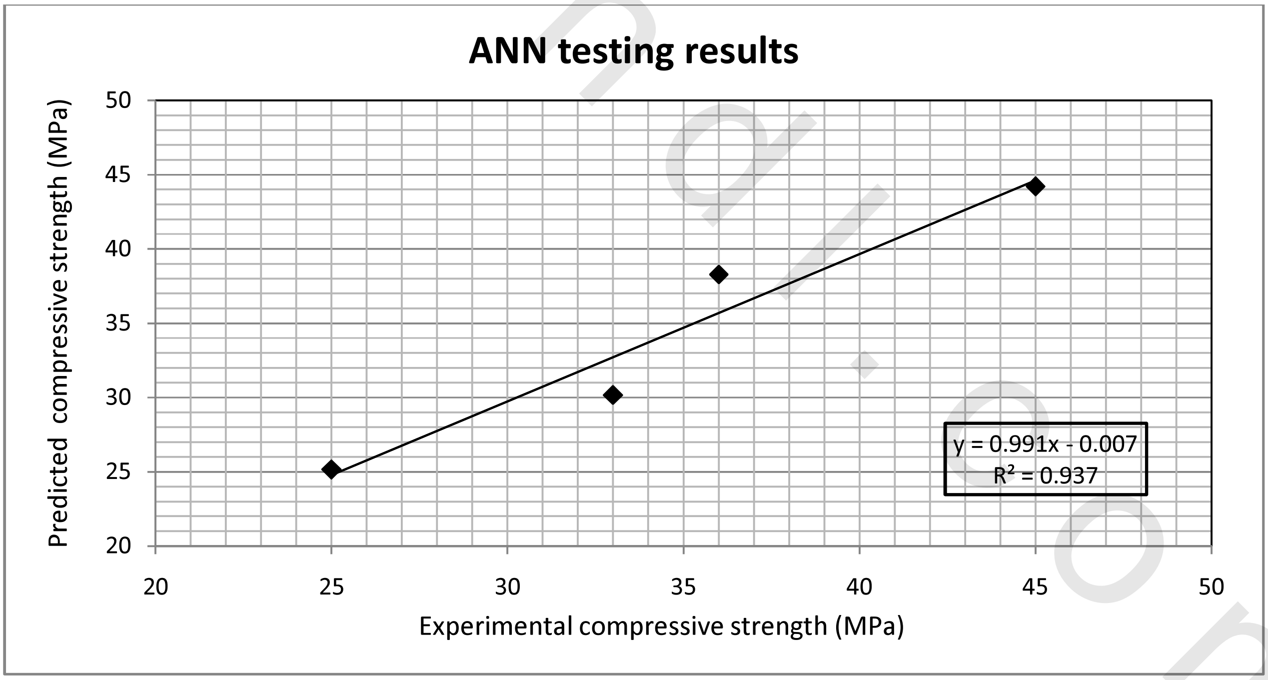

Fig. 6.13-predicted vs. experimental cube compressive strength of the M-(L.P.10A) after exposed to 20, 200, 400 and 600°C for 2 hrs.

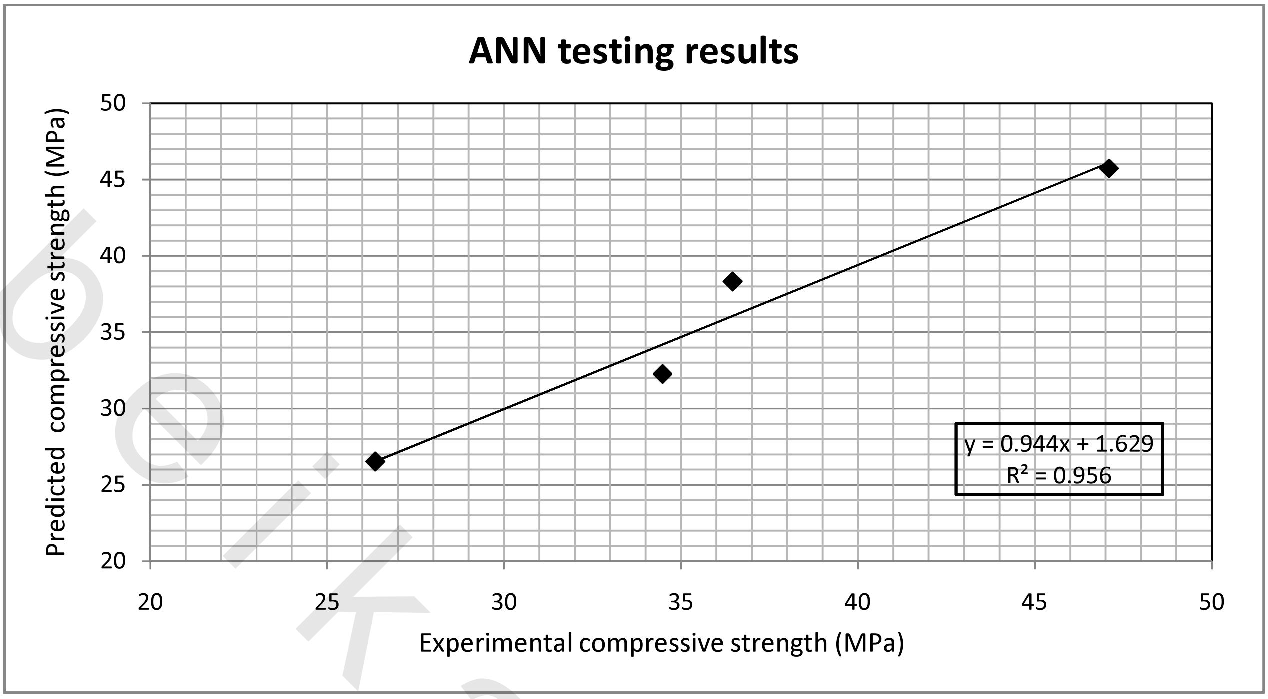

Fig. 6.14-predicted vs. experimental cube compressive strength of the M-(L.P.15A) after exposed to 20, 200, 400 and 600°C for 2 hrs.

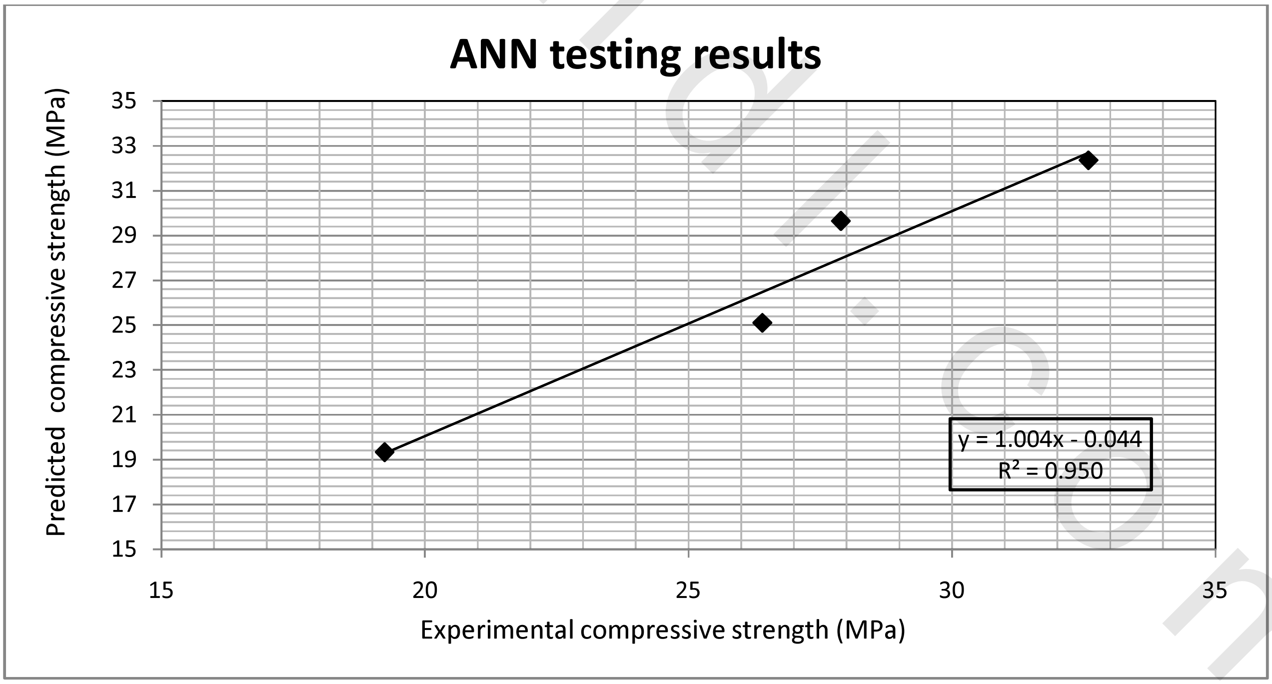

Fig. 6.15-predicted vs. experimental cube compressive strength of the M-(Ben.10R) after exposed to 20, 200, 400 and 600°C for 2 hrs.

The acceptance or rejection of the model developed is determined by its ability to predict the 28 days compressive strength of the mix used. Since the neural networks are trained on actual test data, they are trained to deal with inherent noisy or imprecise data. As new data become available, the neural network model can be readily updated by retraining with patterns which include these new data.

The obtained results are graphically plotted showing comparison of predictions through ANNanalysis method. From **Fig.6.9** to **Fig.6.15**were shownpredicted residual compressive strengths of the all studied concrete cubes at ambient temperature and after exposed to200, 400 and 600°C for 2 hrsthrough ANN. Thepredictions on figures are based on data from the testing set implemented to samples that are not inthe training set. These figures clearly show that experimentally evaluated values of cube concretecompressive strength are in strong consistency with the values predicted through ANN for most ofthe specimens.The correlation coefficient (R^2), root mean square error (RMS), and mean absolute error (MAE) is used to judge the performance of the neural network approach in predicting the results.**Table 6.6** gives the summary of correlationcoefficient, mean absolute error and root mean square error obtained topredict the 28 days compressive strength of the all concrete mixtures after exposed to 20, 200, 400 and 600°C for 2 hrs**.** From all these graphs it is observed that the modeling results are exceptionally close to the real compressivestrength test results; where the R^2 values for all the modeling not less than o.90 which considered excellent results therefore there is no doubt regarding the accuracy of the RMS values.

Table 6.6: The statistical values of proposed models.

Model no.	**ANN**		
	RMS	**R^2**	**MAPE**
1	3.212561	0.9617	10.18677
2	1.72201	0.9545	4.320918
3	1.494145	0.9576	5.148777
4	1.359371	0.9074	6.837533
5	2.781508	0.9371	6.576491
6	3.220589	0.956	7.777054
7	1.224084	0.9507	3.757757

On the other hand, overall model was used all concrete mixture at the same model. There are six nodes in the input layer corresponding to the 6 variables: water to cement ratio, cement content, silicafume content, limestone powder content, bentonite content and exposed temperature as sketched in **Fig 6.16**. The input parameter ranges used in this study are presented in **Table 6.4**. As usual in this this research a logsig transfer function is used within the network. This model was considered more complex than the former models. The total database size in the present model was 28 cases, considering 6 inputs and one output foreach model.It should be noted that, this model has been developed to change the number of neurons from 10 to 12as presented in **Fig. 6.17** because of from previous readings, it was preferred to make the number of neurons is equal to twice the input variables including at least 10 neurons to gain the best results.

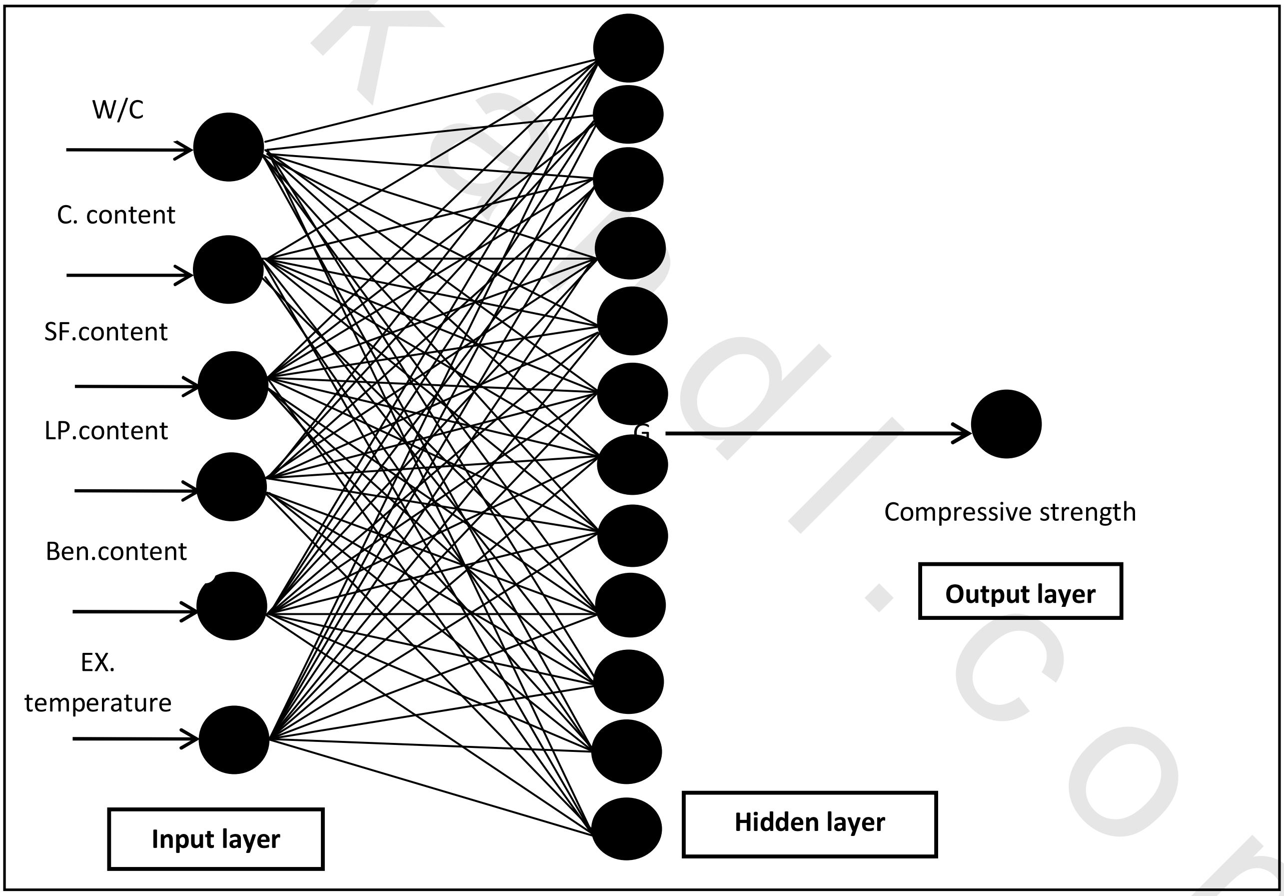

Fig. 6.16-The chosen model architecture

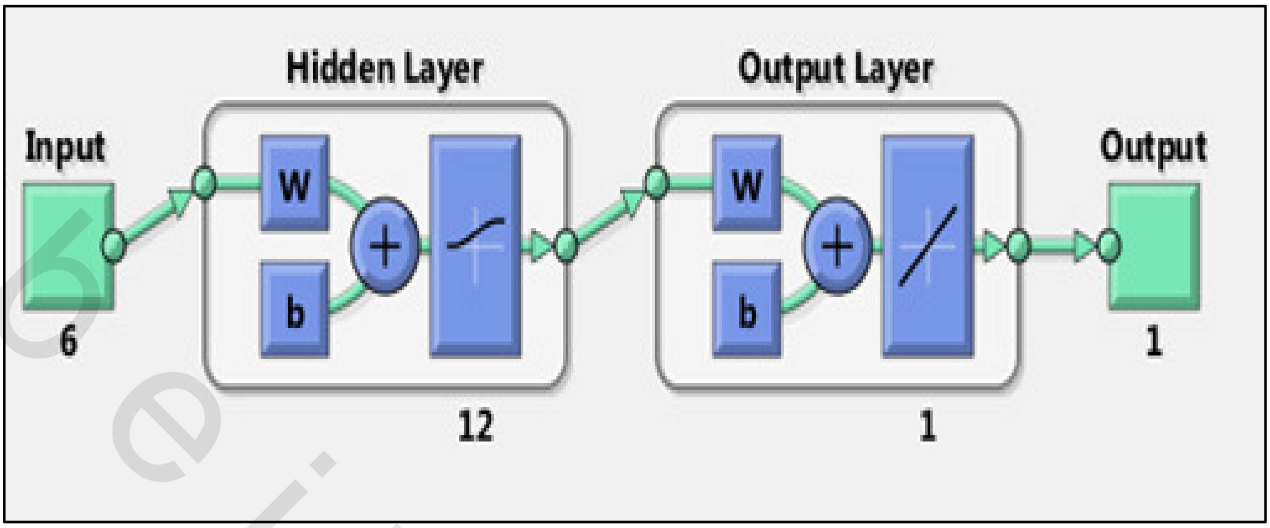

Fig. 6.17- Schematic of the input layer, hidden layer and output layer of the overall model as presented in ANN.

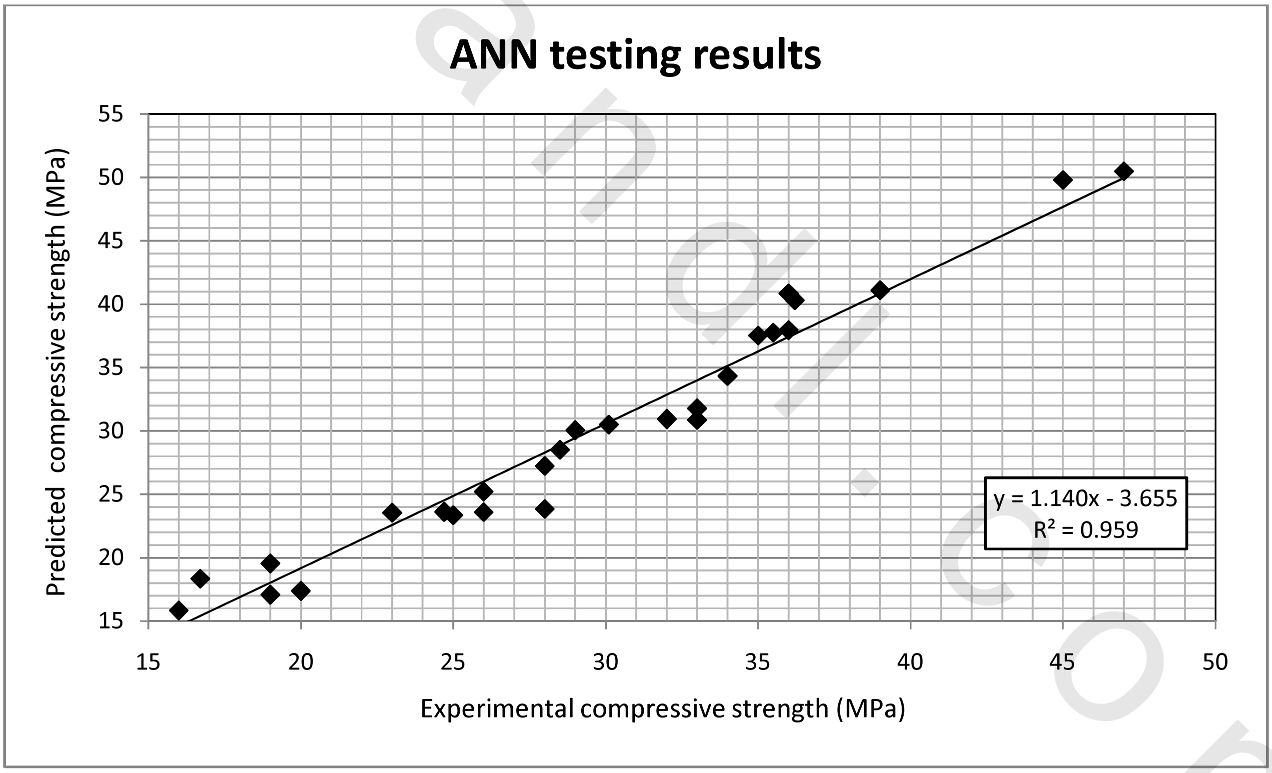

Fig. 6.18-Overallpredicted vs. experimental cube compressive strength of the all studied mixturesafter exposed to 20, 200, 400 and 600°C for 2 hrs.

The coefficient of determination (R^2) for this model was 0.9598 that was considered an excellent result. A more visual insight to the wholedata set's performance can be obtained and analyzed by this way. Lastly, the performance of the overall system with this amount of input data for fire resistance of concrete cubesstrength can be more meaningful and easier to analyze by this method of analysis.

.