PAPER • OPEN ACCESS

Characterization of Carbon nanotube reinforced Silica refractory nanocomposite using Artificial Intelligence Modelling: PART B

View the article online for updates and enhancements.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Characterization of Carbon nanotube reinforced Silica refractory nanocomposite using Artificial Intelligence Modelling: PART B

Y Tijjani[1*], F M Yasin[2], M H S Ismail[3], M A Azmah Hanim[4]

[1*] Department of Mechanical Engineering, Faculty of Engineering, Bayero University, Kano, Kano, Nigeria
[1,2,3] Department of Chemical and Environmental Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia
[4] Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia

*Corresponding author: ytijjani.mec@buk.edu.ng

Abstract: The present study has dwelled on the implementation and evaluation of an artificial intelligence model for the determination of predicted foundry physical properties; linear expansion, bulk density, apparent porosity, thermal shock resistance cycles and cold crushing strength of carbon nanotube (CNT) reinforced silica refractory nanocomposite. A multi input and multi output Artificial Neural Network (ANN) models were developed using the Levenberg Marquardt Back Propagation algorithm (LMBPA) in the neural network toolbox of MATLAB R2015a to train/predict the foundry physical properties of the CNT-silica refractory nanocomposite bricks obtained experimentally from the previous study. The predicted models were compared with the experimental test results in order to evaluate the power and the accuracy of the artificial intelligence model for the characterization of the entire series of CNT-silica refractory nanocomposite bricks. The developed (LMBPA ANN) model satisfactorily predicts the foundry physical properties of CNT reinforced silica nanocomposite with a coefficient of determination (R^2) in the range $0.75 \geq R^2 \leq 1$.

1. Introduction to Artificial Intelligence Modelling

Recently, artificial intelligence has been proven to be an important approach which has fascinated both the academic and industries in expanding the attainable study output via the enhancement of inputs from the experimental results [1, 2]. The needs for artificial intelligence has been rooted as mentioned in the previous study; [3] as; (i) Nonlinearity permits improved data fit. (ii) Noise-insensitivity gives proper prediction in the presence of unreliable data and measurement errors. (iii) High parallelism signifies fast processing and hardware failure-tolerance. (iv) Learning and flexibility enable the system to modify its internal structure in response to environmental change. (v) And generalization allows application of the model to un-learned data.

Example of such artificial intelligence tools, which has been employed extensively in addressing numerous scientific problems is the artificial neural network (ANN) [4]. The artificial neural network is a computer-based programs inspired by the ability of the human brain in processing information. An outstanding attribute of the ANN is offering high computational power to model sophisticated nonlinear processes by repeated training of a set of data until the output provides an agreeable value that correlates with the expected target rather than depending on mere computer programs [5]. Generally, the artificial

neural network technique is highly powerful and provides satisfactory level of accuracy. In addition, their characteristics are of pivotal importance in capturing the different material characteristics. Beside this, ANNs are very useful in preliminary prediction of characteristics of new materials prior to experimental work.

1.1. ANN Model Structure and Training Algorithm

Figure 1, is the flow chart depicting a sequence and procedural steps in the development of the ANN models for the characterization of CNTs-reinforced silica refractory nanocomposites. As demonstrated in the figure, the predictive modeling in the present research can be categorised via three fundamental steps consisting of data acquisition, model training, and predictions.
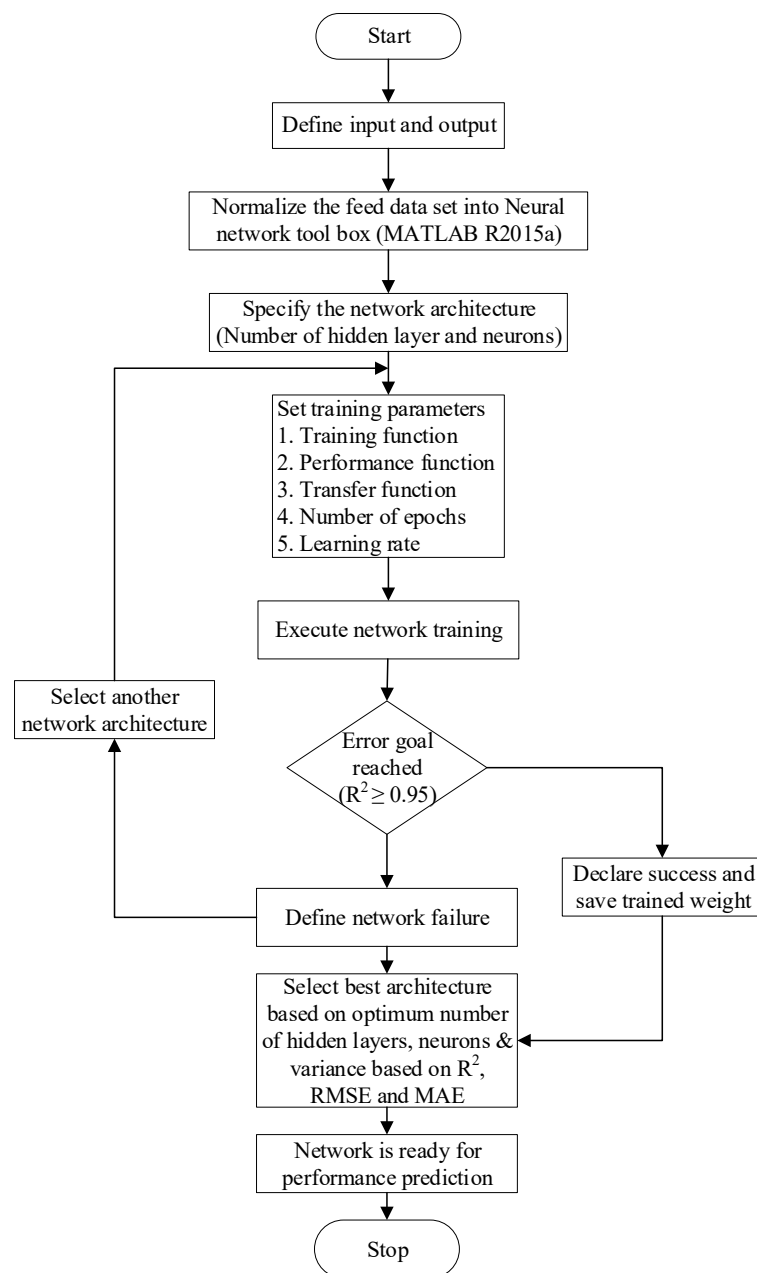

Figure 1. The flow chart for developing the neural network models

1.2. The ANN Modeling

A total of 11 data patterns; input and output, sourced from all sample formulations were employed in training the ANN. The overall data in the present model were divided into three classes, consisting of 80% training, 10% validation and 10% testing. Furthermore, the input and output data have been normalized between 0 and 1 with the aids of the equation (1), to establish the proper implementation of the sigmoid transfer function.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where X_{norm} is the normalized value, X is the true value of the variable, X_{min} & X_{max} are the minimum & maximum values of the dataset respectively.

The network model ANN, was programmed by a supervised multilayer perceptron (MLP) trained with the back propagation (BP) algorithm (section 1.3; Levenberg Marquardt (LM) algorithm) using the neural network toolbox of MATLAB R2015a. For the ANN; Figure 2, MLP ANN has two input parameters (Quartzite content and CNTs content) and five output parameters (linear expansion, bulk density, apparent porosity, thermal shock resistance cycles and cold crushing strength) was developed and evaluated.
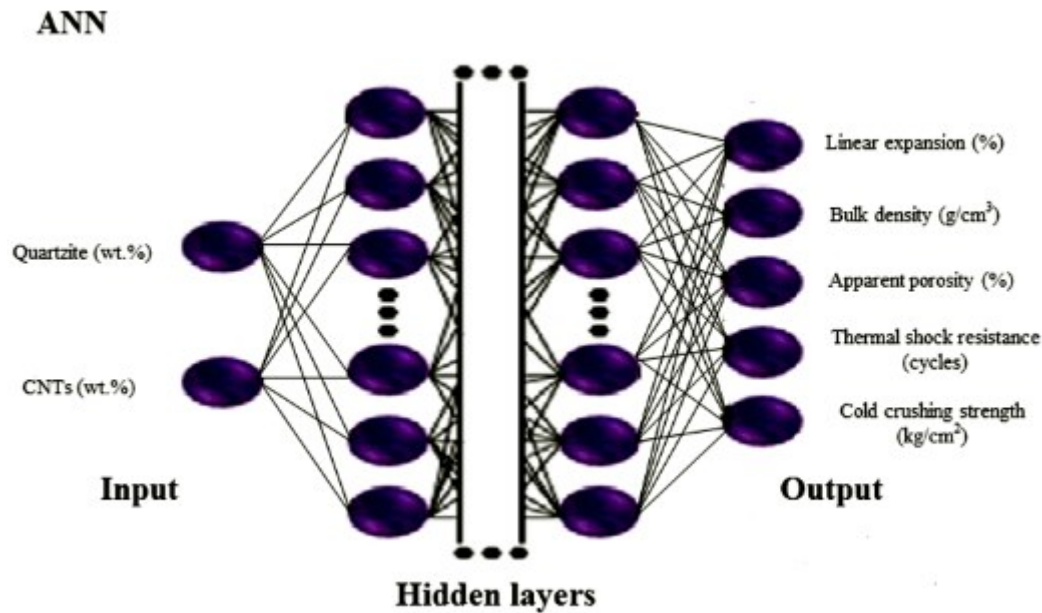

Figure 2. The ANN model for characterizing the CNTs-silica refractory nanocomposite

Due to the lack of standards for obtaining the required hidden layers and neurons, preliminary study revealed that two hidden layers provided the optimum performance value; minimum absolute error values for the training and testing datasets compared to one hidden layer. Thus, while evaluating the best number of neurons needed for building the neural network models, the number of hidden layers was set at two for the ANN model. Furthermore, after many iterations the optimum ANN topologies were chosen based on: (i) The hyperbolic tan-sigmoid transfer function; equation (2) in both hidden layers. (ii) A linear transfer function; equation (3) in the output layer. (iii) And best performance criteria.

$$Tansig\ (n) = \frac{2}{(1 + \exp(-2n)) - 1} \tag{2}$$

$$Purelin\ (n) = n \tag{3}$$

While, the network performance and the goodness of fit of the various ANN topologies tested in the present study were assessed via three statistical parameters: R^2 (coefficient of determination); equation (4). RMSE (root mean square error); equation (5). And MAE (mean absolute error); equation (6) [6].

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{pred}^i - y_{targ}^i\right)^2}{\sum_{i=1}^{N}\left(y_{pred}^i - \bar{y}\right)^2} \qquad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_{pred}^i - y_{targ}^i\right)^2}{N}} \qquad (5)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_{pred}^i - y_{targ}^i| \qquad (6)$$

Where y_{pred}^i is the predicted value, y_{targ}^i is the real value retrieved from the experimental data, $\bar{y}$ is the average of the actual observation and N is the number of observations.

1.3. Levenberg-Marquardt Back Propagation Algorithm (LMBPA)

Levenberg Marquardt (LM) algorithm; a trust-region type of Gauss-Newton technique and a second order back propagation algorithm has been recommended as the preferred supervised training algorithm [7-9]. In order to minimize the error function until a particular minimum or stopping gradient is attend, the algorithm trains a neural network via repeated modification and update of the weight and bias values [10]. Thus, despite consuming a comparatively greater memory and processing resources than other algorithms, the Levenberg-Marquardt Back Propagation Algorithm has been selected in the present research as the training algorithm because of its high efficiency and accuracy. Figure 3 summarizes the standard LMBPA.

Step 1: The weights and parameter are initialize μ (μ = 0.01 is suitable).
Step 2: The sum of the squared errors over all inputs F(w) applying equation (7) is compute:
F(w) = e^{T}e   (7).
Where w = [w$_1$, w$_2$… w$_n$] includes all weights of the network, e is the error vector containing the error for all the training examples.
Step 3: The weight increment Δw is computed via equation (8):
Δw = [J^{T}J + μI]$^{-1}$J^{T}e   (8).
Where J is the Jacobian matrix, μ is the learning rate that is to be updated employing the β based on the outcome. Specifically, μ is multiplied by decay rate, where β (0<β<1).
Step 4: Applying w+Δw as the trial w, and judge:
IF performance index F(w) < F(w) in step 2 THEN
W = w + Δw
μ = μ. β
go back to step 2
ELSE
μ = μ/ β
go back to step 4
END IF

Figure 3. Pseudo-code for LMBPA [11, 12].

2. Performance of the LMBPA-ANN Models

To validate the performance of the LMBPA ANN in determining the properties of various grades of CNT-silica refractory nanocomposites developed in the present work, different ANN topologies were deployed in training the network by altering the number of hidden neurons. Table 1, highlight the network performance for each of the ANN topology; hidden neuron configuration used in developing the ANN model. Over 15 different ANN topologies were tested by systematically permuting the number of neurons in the hidden layers (between 4 & 20) [13].

Table 1. Variation of hidden neurons employed in developing the ANN model

Hidden neurons	Train			Validation			Test		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
4-6	0.9764	0.1990	0.1649	0.9757	0.0412	0.0412	0.8448	0.0124	0.0124
4-8	0.9744	0.1925	0.1521	0.8645	0.0113	0.0113	0.9989	0.1602	0.1602
4-18	0.9596	0.1792	0.1418	0.9848	0.0066	0.0066	0.9726	0.2840	0.2840
5-5	0.9753	0.2115	0.1724	0.9780	0.0022	0.0022	0.8370	0.0199	0.0199
6-6	0.9023	0.2068	0.1634	0.9845	0.0091	0.0091	0.9407	0.0505	0.0505
5-10	0.7930	0.1999	0.1628	1.0000	0.0620	0.0620	0.8680	0.0017	0.0017
5-15	0.9209	0.1865	0.1392	0.9418	0.5071	0.5071	0.9900	0.0503	0.0503
10-5	0.9798	0.2113	0.1665	0.9692	0.0505	0.0505	0.7887	0.0204	0.0204
10-6	0.9209	0.2209	0.1600	0.9994	0.0781	0.0781	0.8980	0.0044	0.0044
10-12	0.9656	0.2000	0.1659	0.9779	0.0027	0.0027	0.9596	0.0405	0.0405
10-15	0.9264	0.3181	0.2016	1.0000	0.0323	0.0323	1.0000	0.0096	0.0096
10-20	0.9644	0.1843	0.1442	0.9749	0.0028	0.0028	0.9860	0.2428	0.2428
12-10[*]	0.9690	0.1762	0.1392	0.9994	0.1808	0.1808	0.9838	0.0407	0.0407
15-20	0.9803	0.2108	0.1691	0.9773	0.0562	0.0562	0.8162	0.0277	0.0277
20-20	0.9593	0.1995	0.1678	0.9618	0.0247	0.0247	0.9234	0.0005	0.0005

[*] Best model based on training, validation and test data.

Also, Figure 4 depicted the coefficient of determination (R^2) values for the selected best topology models based on the training, validation and test data. 12-10 is the best topology for the developed ANN.
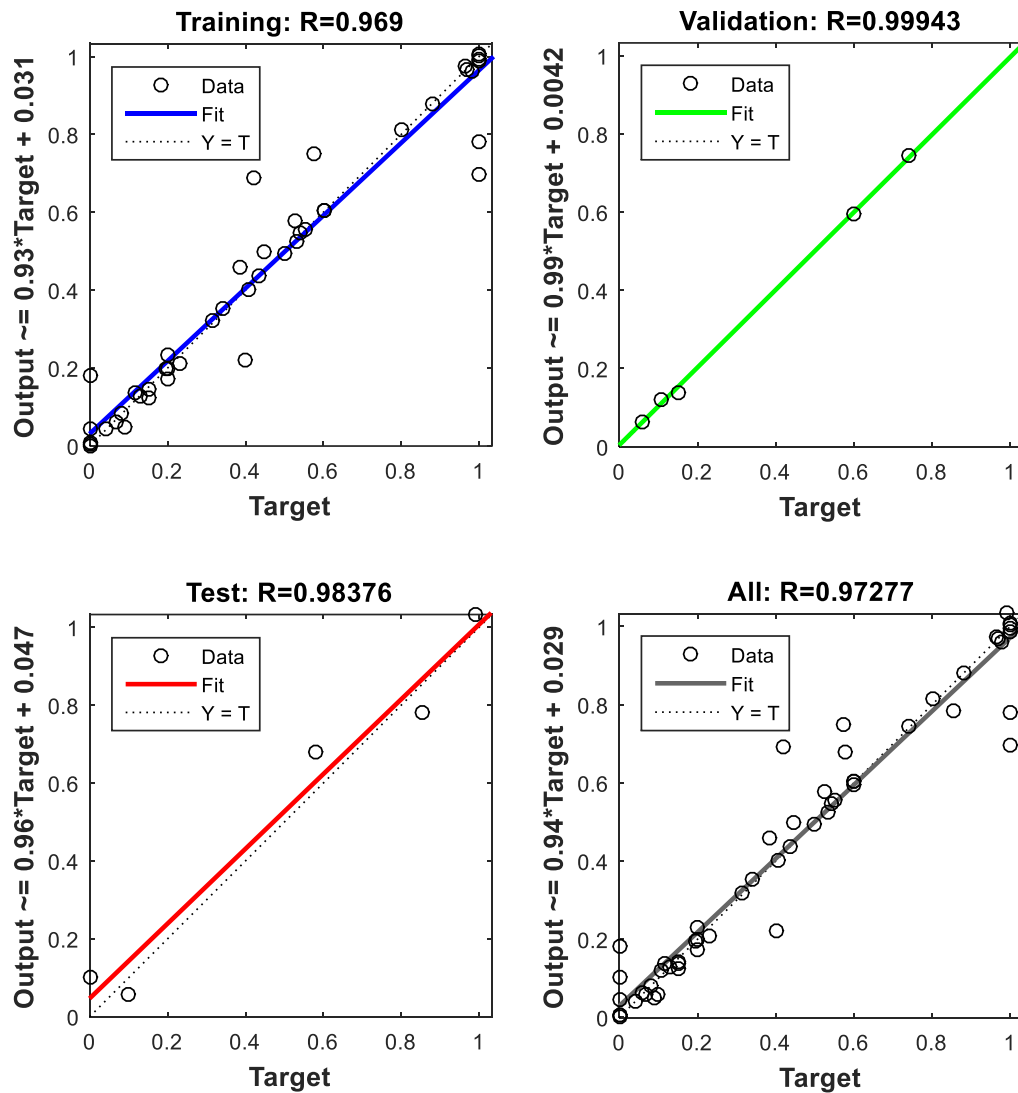
Figure 4. Coefficient of determination (R^2) of the selected topology. 12-10 for the ANN model, processed with the aid of MATLAB R2015a.

Based on the cumulative RMSE values of the various ANN topologies tested, the performance of the ANN model in predicting the output parameters is depicted in Figure 5. From both the Table 1 and Figures 4 and 5, it is clear that the best hidden neuron configurations on the basis of highest coefficient of determination (R^2) and the lowest training error (RMSE and MAE) is 12-10 for the dataset trained in the model ANN. Hence, the 2-12-10-5 is selected as the preferred and the most accurate ANN architectures deployed in training and predicting the output parameters of the ANN.
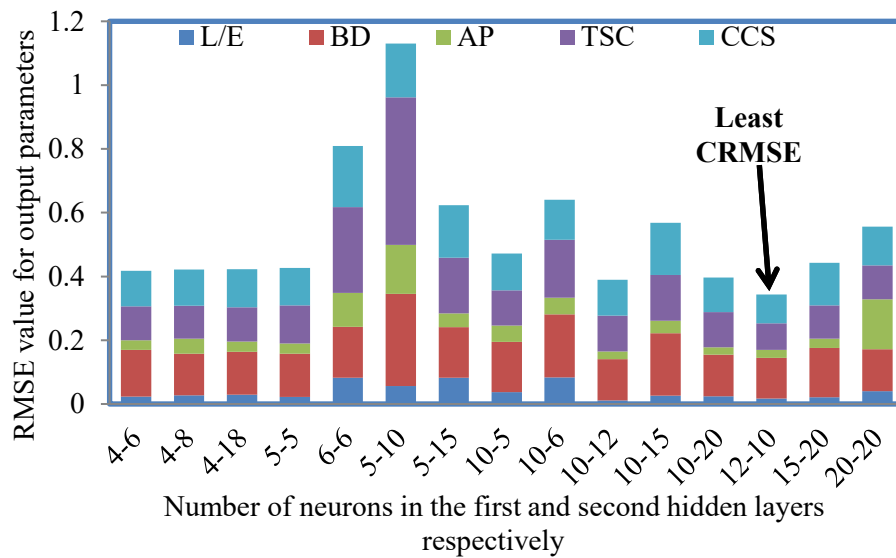
Figure 5. The performance of the ANN in characterizing CNT-silica refractory nanocomposites in terms of RMSE values. Where: L/E stands for linear expansion. BD; bulk density. AP; apparent porosity. TSC; thermal shock resistance cycles. CCS; cold crushing strength. And CRMSE; cumulative root mean square error.

As can be seen from Figures 6, the predicted values determined by using the optimum ANN topologies in the developed ANN model are close to the experimental results. Eventhougth the deviation of some points is obvious, especially for bulk density, thermal shock resistance cycles and cold crushing strength, which could be attributed to limited experimental dataset employed for training the ANN model. From the ANN model, R^2 values ≥ 0.92 was obtained for the output parameters, with the exception of the bulk density where R^2 value of approximately 0.75 was registered.
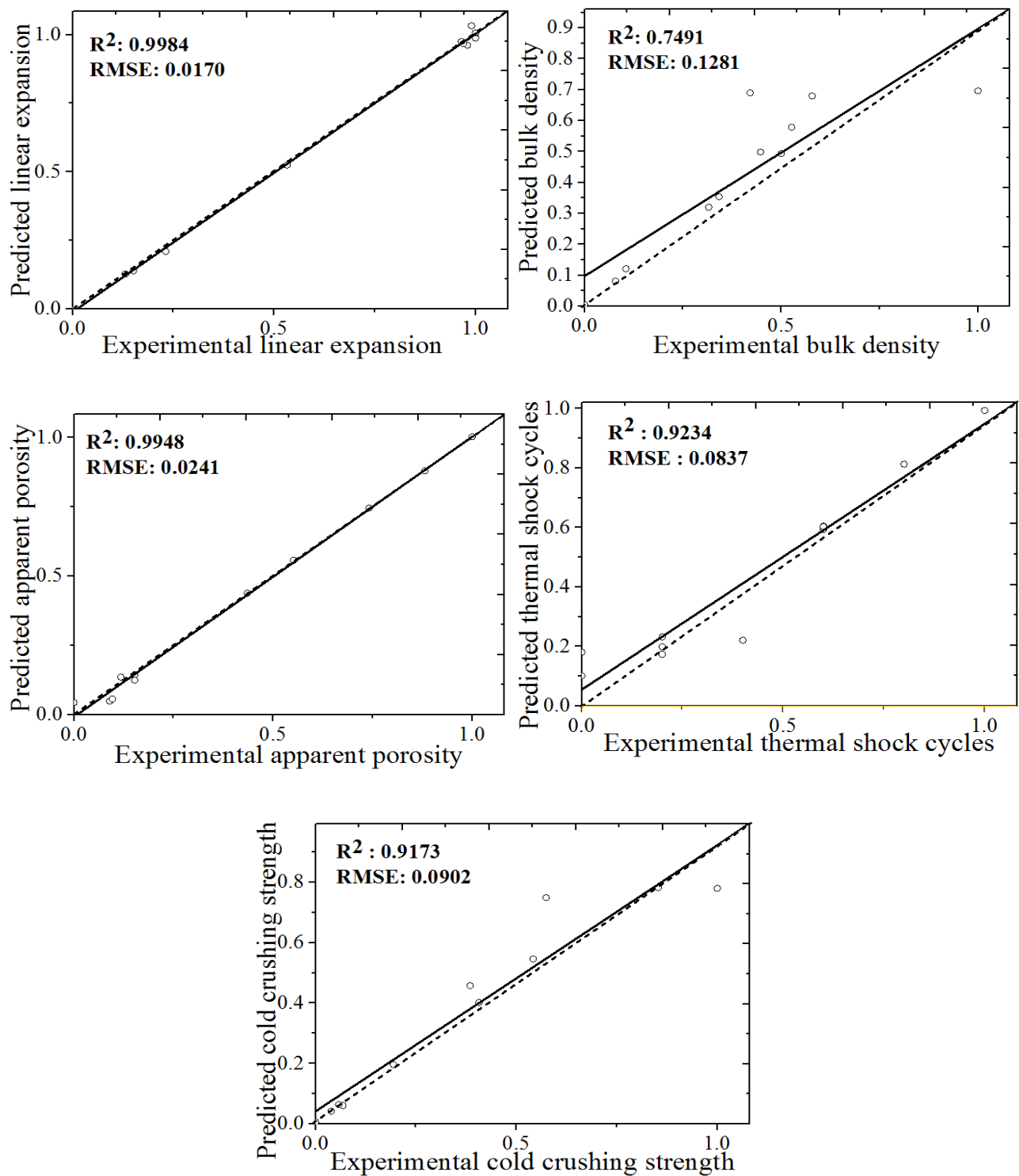
Figure 6. Comparison between predicted and experimental results for the CNT-silica refractory nanocomposites

3. Conclusion

As seen from the Figure 6, a good fit between the experimental values (the straight line) and the predicted values (the data points) signifies that the proposed ANN model has the ability to adequately predict with a minimal error, the foundry physical properties; specifically, linear expansion, apparent porosity, thermal shock resistance cycles and cold crushing strength of CNT-silica refractory nanocomposites having formulations within the range of those values employed in the experimental process as mentioned elsewhere (PART A study).

Acknowledgement

References

[1]   Duan Y, Ma L, Qi H, Li R and Li P 2017 Developed constitutive models, processing maps and microstructural evolution of Pb-Mg-10Al-0.5 B alloy *Mater. Charact.* **129** 353-66

[2]   Gokceoglu C, Zorlu K, Ceryan S and Nefeslioglu H 2009 A comparative study on indirect determination of degree of weathering of granites from some physical and strength parameters by two soft computing techniques *Mater. Charact.* **60** 1317-27

[3]   Basheer I A and Hajmeer M 2000 Artificial neural networks: Fundamentals, computing, design, and application *J. Microbiol. Methods* **43** 3-31

[4]   Gardner M W and Dorling S 1998 Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences *Atmos. Environ.* **32** 2627-36

[5]   Jain A K, Mao J and Mohiuddin K M 1996 Artificial neural networks: A tutorial *Computer* **29** 31-44

[6]   Nadian M H, Rafiee S, Aghbashlo M, Hosseinpour S and Mohtasebi S S 2015 Continuous real-time monitoring and neural network modeling of apple slices color changes during hot air drying *Food and Bioproducts Processing* **94** 263-74

[7]   Kaminski M and Orlowska-Kowalska T 2015 An on-line trained neural controller with a fuzzy learning rate of the Levenberg–Marquardt algorithm for speed control of an electrical drive with an elastic joint *Applied Soft Computing* **32** 509-17

[8]   Mukherjee I and Routroy S 2012 Comparing the performance of neural networks developed by using Levenberg–Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process *Expert Systems with Applications* **39** 2397-407

[9]   Akbari M K, Shirvanimoghaddam K, Hai Z, Zhuiykov S and Khayyam H 2017 Al-TiB2 micro/nanocomposites: Particle capture investigations, strengthening mechanisms and mathematical modelling of mechanical properties *Mater. Sci. Eng., A* **682** 98-106

[10]   Wilamowski B M and Yu H 2010 Improved computation for Levenberg–Marquardt training *IEEE Transactions on Neural Networks* **21** 930-37

[11]   Nazari A and Riahi S 2013 Artificial neural networks to prediction total specific pore volume of geopolymers produced from waste ashes *Neural Computing and Applications* **22** 719-29

[12]   Sarıdemir M, Topçu İ B, Özcan F and Severcan M H 2009 Prediction of long-term effects of GGBFS on compressive strength of concrete by artificial neural networks and fuzzy logic *Construction and Building Materials* **23** 1279-86

[13]   Dele-Afolabi T, Hanim M A, Norkhairunnisa M, Sobri S, Calin R and Ismarrubie Z 2018 Agro-waste shaped porous Al_2O_3/Ni composites: Corrosion resistance performance and artificial neural network modelling *Mater. Charact.* **142** 77-85