

## DATA-DRIVEN ASSESSMENT OF ARTIFICIAL NEURAL NETWORK AND REGRESSION CURVE FITTING APPROACHES FOR DIMENSIONLESS TURBULENT FLOW HEAT TRANSFER PERFORMANCE OF A HEXAGONAL DUCT

ALI H. ABDULKARIM<sup>1</sup>, OGUZ TURGUT<sup>2</sup>, AHMED GHAREEB<sup>1</sup>,  
EYUB CANLI<sup>3,\*</sup>, MEHMET SARI<sup>2</sup>, KEVIN HALLINAN<sup>4</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Kirkuk, University Road, 36001, Kirkuk, Iraq

<sup>2</sup>Department of Mechanical Engineering, Gazi University, Eti Yukselis Sk. No. 5, Cankaya, 06570 Ankara, Turkey

<sup>3</sup>Department of Mechanical Engineering, Selcuk University, Alaaddin Keykubat Campus, Ardicli Mh. Hamdullah Suphi Tanriover Cd. Selcuklu 42250, Konya, Turkey

<sup>4</sup>Department of Mechanical and Aerospace Engineering, University of Dayton, 300 College Park, Dayton, 45469, OH, United States of America

\*Corresponding Author: ecanli@selcuk.edu.tr

### Abstract

Data reduction and mathematical analysis are always an important part of heat transfer related studies. Regression curve (RCF) and Artificial Neural Network (ANN) fitting methods are used extensively and regarded as reliable tools for this purpose. ANN and RCF approaches are used to predict the Nusselt number (Nu), and Darcy friction factor ( $f$ ) based upon a single input, namely, the Reynolds number (Re) varying from  $2.3 \times 10^3$  to  $52 \times 10^3$ . Experimental results from a previous comprehensive study on forced thermal convection through a hexagonal duct was used to develop models and test the efficacy of the methods. Several ANN architectures, hyperparameter values and RCF functions were tested. Normalized and non-normalized datasets were considered. Models were compared with each other by means of statistical indicators. Therefore, this work is distinguished from the literature by its experimental data driven assessment and comparison between RCFs and ANNs. Additionally, single input and double output design is very scarce in the literature. The results show that the regression schemes based upon RCF are sufficient and accurate for predicting Nu and  $f$ , and the trend associated with the variation to Re is captured. The best single output ANN yields better accuracy; but the best double output ANN is unable to capture the expected trend between Re and the targeted responses. In terms of correlation coefficients, 0.98-0.99 is possible for RCFs and 0.99 for ANNs when normalization is done. Normalization becomes prominent as neuron number increases. As a conclusion, regression is preferred over ANN for a single input relationship between Nu and  $f$  to Re for hexagonal duct flow and heat transfer. An ANN's worth will only potentially be seen if more inputs; e.g., geometrical factors such as the aspect ratio, surface roughness, ambient temperature, duct material and relative wall thickness, among others, are included.

Keywords: ANN, Forced thermal convection, Hexagonal duct, Normalization, Pressure drop.

## 1. Introduction

Duct flow is characterized by boundary layer formation. The boundary layer formation that is affected by several factors such as gravity, hydraulic diameter, thermo-physical properties and especially duct geometry. In terms of geometry, ducts and channels can be divided into two primary categories as with and without corners. The presence of corners indicates that the duct has planar surfaces. Boundary layers on these planar surfaces interact with other boundary layers at the intersecting edges of the planar surfaces. The interacting boundary layers impact both the pressure drops and heat transfer due to mixing of the merging boundary layers. Mass, momentum, and thermal diffusion increase. The mixing phenomenon, turbulence generation and shear due to momentum transfer to channel walls increase pressure drop and thus pumping power.

A hexagonal cross-sectional duct geometry offers compactness. Synthetic honeycombs are exemplars of this fact. Therefore, hexagonal ducts are preferable where compact bundles are needed. Aside their structural and production advantages, hexagonal cross-sectional ducts, channels, and passages have become quite prominent in heat exchangers. Such ducts and channels are especially relevant to nuclear reactors. Hexagonal duct flow is characterized by complex interactions between the boundary layers formed on the six planar surfaces and corners. Prior research has sought to better understand and characterize the impact of these interactions on heat and mass transfer and flow loss. The most prominent study to date is associated with Turgut and Sari [1]. Authors experimentally and numerically investigated the pressure drop and heat transfer for turbulent flow in the Reynolds number ( $Re$ ) range of  $2.3 \times 10^3 \leq Re \leq 52 \times 10^3$ . Tabulated coefficient values for correlating  $Re$  with Nusselt number ( $Nu$ ) and Darcy friction factor ( $f$ ) were given.

In another study, laminar flow in hexagonal duct for two typical thermal boundary conditions, i.e., constant temperature and constant heat flux, was investigated numerically by Turgut [2]. Hexagonal duct inner angles were also considered as a variable. Correlations for  $Nu$ ,  $f$ , and the entrance lengths were developed as a function of inner angles. The percentage deviation of the correlations relative to experimental results was reported to be about 1 percent. Iwaniszyn et al. [3] focused on laminar flow in short hexagonal ducts. Their study included both experimental and numerical simulations [3]. Results showed that the conjugate heat transfer which occurs as a result of having finite thickness steel channel walls causes deviation relative to results for zero wall thickness. Ismail et al. [4] used volume based finite elements method to investigate hexagonal channels in plate fins of a heat sink for electronic cooling. It was concluded that both thermal and hydrodynamic performances of circular and hexagonal cross-sections are close to each other.

Chen et al. [5] investigated twisted hexagonal tubes as a heat exchanging medium. It was concluded that heat transfer enhancement factor for the hexagonal twisted tube relative to square and elliptical twisted tubes with equal cross-sectional areas was documented to be about two times higher. Yadav et al. [6] experimentally investigated the effect of applying twisted tapes to the surfaces of hexagonal, circular and rectangular channels. The authors concluded that increasing the edge length of the tape improves thermal performance while reducing hydrodynamical resistance. The effect of twisting the duct on heat transfer performance was numerically analysed by Cheng et al. [7]. The results of the analysis showed that

twisted tubes yield better thermal performance in comparison to plain tubes. Mahato et al. [8] employed Computational Fluid Dynamics (CFD) to study twisted hexagonal duct flow. The authors concluded that hexagonal ducts are preferred over square ducts for laminar flow and square ducts are preferred for turbulent flow.

Hexagonal channels enable construction of very compact structures such as honeycombs. Such structures can be used for adsorbent beds as stated in the study of Zhang [9]. The author considered conjugate heat transfer due to thick walls of hexagonal tubes for simultaneously developing flow. Results showed a 10% increase in Nu and Sherwood numbers for thick wall channels when compared to channels with zero wall thickness. Tasnim et al. [10] investigated theoretically and experimentally wave propagation in hexagonal channels that are denominated porous. Hou et al. [11] conducted an analytical thermal and hydrodynamic investigation on hexagonal and re-entrant honeycomb structures. The authors concluded that re-entrant type honeycomb is better than other types of honeycombs in terms of thermal and structural performance. Another honeycomb heat transfer investigation was reported by Subasi et al. [12].

A number of recent studies have investigated hexagonal micro-channels. For example, Alfaryjat et al.'s study [13] compared hexagonal micro-channels to circular and rhombus cross-section micro-channels in terms of heat transfer and pressure drop. Hexagonal micro channels with a greater number of channels for a fixed width yielded both higher  $f$  and Nu.

Artificial Neural Networks (ANN) is now a common tool for scientific works. A specific case for ANN is in heat transfer studies, where the heat transfer performance is affected by numerous parameters. Mohanraj et al. [14] reviewed refrigeration, air conditioning and heat pump systems in respect of ANN. Energy and exergy results were mainly the outputs of the reviewed ANN systems. A specific work that used ANN for data interpolation in heat transfer through channels can be viewed from the work of Beigzadeh and Rahimi [15]. Nine hidden layers were used for predicting Nu and twelve were used for predicting  $f$ , which emphasizes the difference between the two parameters. Cong et al. [16] provided an extensive review of the use of ANN in nuclear engineering heat transfer problems. A wide error interval was observed by the authors. An interesting instance was reported by Chelang et al. [17] using an adaptive neuro-fuzzy interface with 5 input parameters and one output parameter. They compared their results with a multi-layered neural network approach. Azizi and Ahmadloo [18] collected 440 data points (observations) from literature for condensation heat transfer coefficient of a refrigerant in order to model the heat transfer coefficient with ANN for 4 inputs. They found a 0.995 correlation coefficient value after some trial-and-error process to improve the ANN hyperparameters to yield the best result.

After reviewing literature, it is seen that ANN modeling of heat transfer and pressure drop through hexagonal channels is not common. There are very few and limited studies. Therefore, starting from simplistic cases and growing towards complex schemes are necessary. Hexagonal channels lead to better trade-off between heat transfer and pressure drop compared to other non-circular ducts as suggested by the literature, where hexagonal ducts also provide compactness [8, 19, 20]. This is due to the fact that a positive change in heat transfer can be achieved by changing channel cross-section and channel configuration, as it is indicated in the related literature [21]. Edges emerging from the hexagon geometry modify flow

within the tubes and change thermal and hydrodynamic performance. Although ANNs have been used to tackle thermo-hydraulic problems since the 1990s and they have been used intensively to predict flow regimes and heat transfer [16], deciding which approach better is still a case particular issue. There are numerous reports stating that ANNs are better in performance compared to the common statistical models [17, 18, 22]. However, it is hard to say, if not impossible, that conventional regression is out of date, since the above literature survey reveals very close results between ANN and conventional regression, especially where the number of input parameters are relatively small.

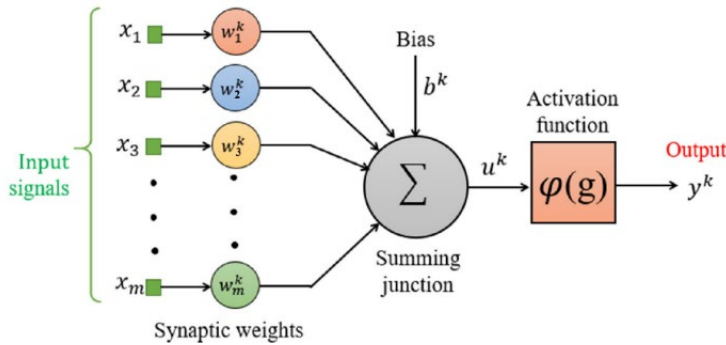
In this work, several ANN network setups and Regression Curve Fitting (RCF) methods are presented in order to model heat transfer performance of a hexagonal duct using  $Re$  as the input, and the  $Nu$  and  $f$  as outputs. This type of single input and double output curve fitting is a unique feature of the work. Experimental data of a previous work [1] was used to train, test, and validate the ANNs and RCFs. Statistical indicators for checking preferable representation are given and compared to ones that emerge from conventional regression schemes. A comprehensive layout and method indicators distinguish the work from literature. This effort is due to the search for optimal designs in heat transfer mediums via various methods and approaches as stated in the literature. Prior literature reveals that there is an effort towards better heat transfer with lower pumping costs [14, 23-25]. This work mainly focuses on comparison of conventional regression with ANN when a single dimensionless input is used for two dimensionless outputs. Also, effect of normalization for ANNs is presented.

## 2. Method

One of the major problems associated with heat transfer literature is vast amount of data relating to various design parameters. RCF had been the leading method for data reduction after tabulated data presentation of the past. Today, machine learning and advance statistics are tried to be implemented in the field for fast decision making. Most of machine learning approaches or data-driven approaches focus on the collected data in order to create links between input parameters, their values and outputs [26]. One may prefer simple tools such as basic linear regression or decision trees, and others may prefer selecting more advanced tools such as ANNs for establishing links between input and output. Although it is always possible to improve the predictive capability for the simplified models, this can not necessarily guarantee capture of the complex relationship between the variables. James et al. [27] provided a distinction between model flexibility and interpretability and showed the reasons for selecting each method. A reliable tool can be introduced as ANNs. This work focuses on comparison of ANN with conventional RCFs since it is desired to show its applicability on heat transfer by a specific case that can be encountered in industry.

The ANNs are enlivened by the human cerebrum's simultaneous structure and depend on the principles of organic cellular structures [28]. They involve the development of numerical processes which are similar to organic processes such as learning and memory. Normal neurons transmit information through neurotransmitters located on the dendrites; the equivalent of this is artificial neurons [29]. In an ANN, there can be multiple layers of interconnecting neurons (nodes). Each neuron has a matrix of weights ( $w$ ) and bias ( $b$ ) for the inputs and a vector of outputs [18]. In a biological system, when the receiving indications are adequate (perform

a particular quantity), the neuron is implemented and discharges a prompt through the axon, and the signal can be sent to another neuro-transmitter and can energize distinct neurons [30]. When neurons receive signal information from the inputs, they transmit output information to the next layer. There are three types of layers; namely input layer, hidden layer, and output layer, and typically the signals transferred from input to output as shown in Fig. 1. The number of nodes within the input and output layers is dictated by the nature of the problem to be solved and the number of input and output variables needed to define the problem [31, 32]. For example, if there is only one variable predicted, the output layer has only one neuron. The optimal number of hidden layers and nodes in each hidden layer needed to realize the lowest error between the predicted and actual outputs is generally decided through a trial-and-error process.



**Fig. 1. Schematic model of a neuron.**

A damped least-squares method has been used to solve the problem of fitting the nonlinear least-squares curve to predict the measured or simulated average Nu and average  $f$  estimation objectives. This technique is also regarded as the Levenberg–Marquardt algorithm, which was initially suggested by Levenberg in 1944 [33] and subsequently reclaimed by Marquardt [34]. This optimization method incorporates Gauss-Newton and steep descent methods to converge to an ideal alternative. It is undoubtedly one of the efficient learning algorithms for feed-forward neural networks [35]. It also operates progressively as a gradient-descent approach when the parameters are a long way from their desired location and works increasingly like the Gauss-Newton method when the parameters are close to their optimal solution avoiding the downsides of both methods [36]. Levenberg–Marquardt equation is given as:

$$[J^T W J + \lambda I] h_{lm} = J^T W (y - \hat{y}) \tag{1}$$

where small amounts of the damping parameter  $\lambda$  lead to overview of Gauss-Newton and consequently higher values of  $\lambda$  lead to overview of the gradient curve.

The damping parameter  $\lambda$  is initialized to be large so that the first changes are small measures in the steepest descent path. In Marquardt’s updated relationship is [37]:

$$[J^T W J + \lambda \text{diag}(J^T W J)] h_{lm} = J^T W (y - \hat{y}) \tag{2}$$

The  $\lambda$  values are normalized with the  $J^T W J$  values [38]. The update of the ANN weights can occur in Levenberg-Marquardt as shown below:

$$\Delta w = \left[ \mu I + \sum_{p=1}^P J^p(w)^T J^p(w) \right]^{-1} \nabla E(w) \tag{3}$$

where  $J^p(w)$  is the Jacobian matrix of the error vector  $e^p(w)$  evaluated in  $w$ ,  $I$  is the identity matrix, and the vector error  $e^p(w)$  is the error of the network for pattern  $p$ , i.e.  $e^p(w) = t^p - o^p(w)$  [39]. The algorithm for Levenberg-Marquardt calculates the network output, the error vectors, and the Jacobian matrix. It then calculates  $\Delta w$  and recalculates the error as network weights using  $w + \Delta w$ . Once the error has decreased,  $\mu$  is degraded by  $\beta$ , new weights are retained, and the process starts once again; else,  $\mu$  is multiplied by  $\beta$  and  $\Delta w$  is newly calculated, and it is again repeated [40]. MATLAB software was used for ANN fit of the data. The dataset for RCF and ANN fitting tasks is given in Table 1. The data set consists of 51 observations and those observations consist of dimensionless numbers. Inputs are Re values while outputs are Nu and  $f$ . However, in order to see the effect of parameter normalization, 0-1 normalization of all values are also given, which are done by their respected intervals. The mathematical expression of the normalization is shown in Eq. (4). In Eq. (4),  $\phi$  represents any parameter in the dataset. Test numbers will be used in the following part of the paper to introduce training, test, and validation steps. In Fig. 2, three-dimensional line plot of dataset according to Re, Nu and  $f$  is given in order to provide more insight of the dataset.

$$\phi_{i,N} = \frac{(\phi_i - \phi_{min})}{(\phi_{max} - \phi_{min})} \tag{4}$$

**Table 1. Dimensionless numbers in the dataset and their normalized values.**

Exp. No.	Re	Nu	$f$	Re <sub>N</sub>	Nu <sub>N</sub>	$f_N$
1	2322	8.9	0.0556	0	0	1
2	2634	9.42	0.0521	0.006273	0.005508	0.90085
3	2670	9.75	0.0519	0.006997	0.009004	0.895184
4	2761	10.86	0.0516	0.008827	0.020763	0.886686
5	3609	12.04	0.047	0.025877	0.033263	0.756374
6	3759	12.4	0.044	0.028893	0.037076	0.671388
7	4112	12.88	0.0428	0.03599	0.042161	0.637394
8	4114	13.1	0.0422	0.03603	0.044492	0.620397
9	4337	13.58	0.0417	0.040514	0.049576	0.606232
10	4482	13.87	0.04	0.043429	0.052648	0.558074
11	4533	14.16	0.0408	0.044455	0.05572	0.580737
12	4751	14.51	0.0404	0.048838	0.059428	0.569405
13	4940	15.1	0.0394	0.052638	0.065678	0.541076
14	5567	16.34	0.0393	0.065244	0.078814	0.538244
15	5648	17.93	0.038	0.066873	0.095657	0.501416
16	6347	18.96	0.0365	0.080927	0.106568	0.458924
17	7737	20.42	0.0342	0.108875	0.122034	0.393768
18	7794	20.66	0.0334	0.110021	0.124576	0.371105
19	8463	23.46	0.033	0.123472	0.154237	0.359773
20	8980	24	0.0324	0.133867	0.159958	0.342776
21	11000	30.4	0.0314	0.174481	0.227754	0.314448
22	11047	31.8	0.0308	0.175426	0.242585	0.29745
23	12591	36.9	0.029	0.20647	0.29661	0.246459
24	14644	40.4	0.0294	0.247748	0.333686	0.25779
25	14731	40.2	0.0291	0.249497	0.331568	0.249292
26	16597	43.3	0.029	0.287015	0.364407	0.246459
27	16624	44.2	0.0283	0.287558	0.373941	0.226629
28	16879	45.6	0.0281	0.292685	0.388771	0.220963

29	17122	45.9	0.0279	0.297571	0.391949	0.215297
30	17302	45.4	0.0276	0.301119	0.386653	0.206799
31	17375	48.8	0.0274	0.302658	0.422669	0.201133
32	20424	55.4	0.0271	0.363962	0.492585	0.192635
33	20471	55.8	0.0269	0.364907	0.496822	0.186969
34	24677	58.1	0.0267	0.449473	0.521186	0.181303
35	27025	62.4	0.0252	0.496682	0.566737	0.13881
36	27159	63.2	0.0249	0.499377	0.575212	0.130312
37	30232	70.3	0.0248	0.561163	0.650424	0.127479
38	30363	70.96	0.0247	0.563797	0.657415	0.124646
39	30645	72.9	0.0246	0.569467	0.677966	0.121813
40	31220	73.2	0.0239	0.581028	0.681144	0.101983
41	32061	74.6	0.0233	0.597937	0.695975	0.084986
42	32388	75.32	0.023	0.604512	0.703602	0.076487
43	34376	75.2	0.0225	0.644483	0.702331	0.062323
44	35824	78.8	0.0224	0.673597	0.740466	0.05949
45	38386	81.4	0.025	0.725109	0.768008	0.133144
46	42336	88.1	0.0214	0.804528	0.838983	0.031161
47	42951	89.1	0.0213	0.816893	0.849576	0.028329
48	47394	96.9	0.021	0.906225	0.932203	0.01983
49	51203	100.2	0.0203	0.982809	0.967161	0
50	51483	101.4	0.0205	0.988439	0.979873	0.005666
51	52058	103.3	0.0203	1	1	0

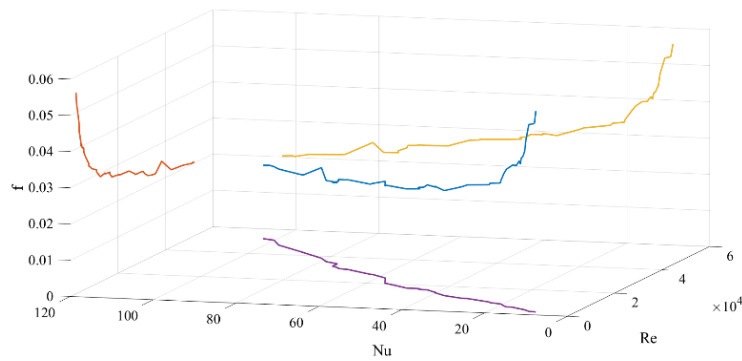


Fig. 2. Three-dimensional line plot of the parameters.

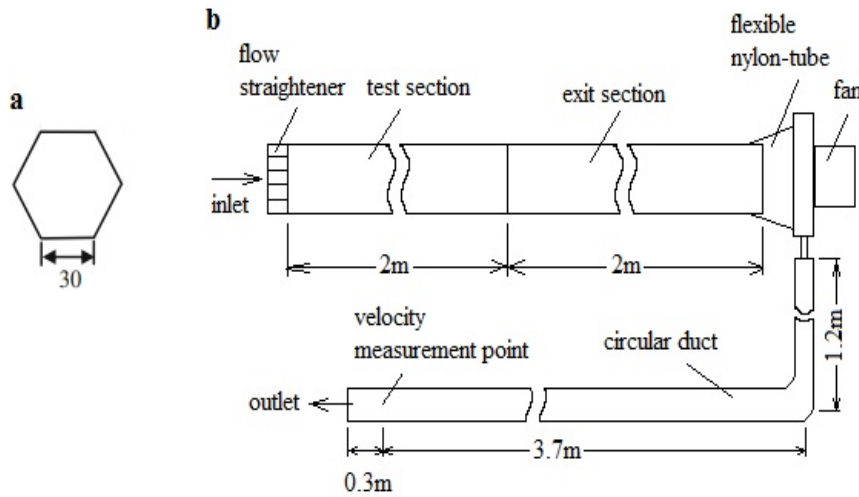
The dataset given in Table 1 and shown in Fig. 2 belongs to the experimental work of Turgut and Sarı [1]. A schematic diagram of the experimental setup and the cross-section of the hexagonal channel in that work is shown in Fig. 3. In Fig. 3(a), dimensions are in millimetre. Measurements were carried out at steady state conditions under constant temperature boundary condition. Experimental uncertainties are ignored in the present work and data are assumed absolute for comparing conventional RCF and ANN fitting. Definitions of  $Re$ ,  $Nu$  and  $f$  are given with below sequence. The characteristic length in dimensionless numbers was chosen as hydraulic diameter ( $D_h$ ). The cross-sectional area is denoted with  $A_c$  and perimeter of the hexagonal cross-section is denoted with  $P_c$ . The average inlet velocity,  $u_{avg}$ , is used in Eqs. (6) and (8). All thermo-physical properties were evaluated at the bulk temperature of the flow. Viscosity is indicated by  $\mu$ ; density is shown by  $\rho$ , and thermal conductivity is symbolized by  $k$ . The convection heat transfer coefficient,  $h$ , was calculated from measurements, and pressure difference  $\Delta P$  was calculated from the measured inlet and outlet static pressures. The length of the test section,  $L$ , is predetermined.

$$D_h = 4A_c / P_c \tag{5}$$

$$Re = \frac{D_h u_{avg} \rho}{\mu} \tag{6}$$

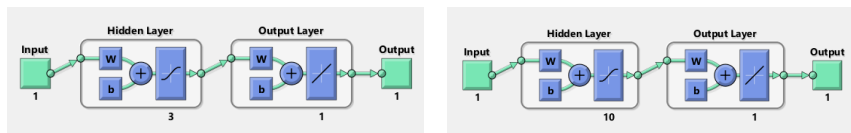
$$Nu = \frac{h D_h}{k} \tag{7}$$

$$f = \frac{\Delta P}{\left(\frac{L}{D_h} \frac{\rho u_{avg}^2}{2}\right)} \tag{8}$$



**Fig. 3. Schematic diagram of the experimental set-up.**  
**(a) Cross-section of the channel, (b) Layout and dimensions.**

In order to realize the ANN for the dataset of experimental flow and heat transfer characteristics, the available data set from the experimental work was divided randomly into training, testing and validation sets by 70%, 15% and 15%, respectively. Selected data for training, testing, and validating are shown in Table 2 by test numbers in Table 1. An ANN's efficiency is influenced by the network's characteristics such as the number of hidden layers and the number of nodes in each hidden layer. Two different hidden layer neuron numbers were tried for single output ANNs, and three different hidden layer neuron numbers were tried for double outputs ANNs. Figure 4 shows schematic illustrations of the ANN structures in this work. The ANN labels in this figure are used in Table 2 and rest of the paper including results section graphics and tables. Similar ANN visual definitions are common in literature [41, 42]. Both normalized and non-normalized data were tried with the ANNs. The indicators in Table 2, i.e., “Tr”, “Tt” and “V” are “Training”, “Testing” and “Validation” respectively.







Both RCF and ANN fittings are examined with statistical measures. The evaluation metrics used to compare the ANN structures in addition to the goodness of the fit and to evaluate regression approaches are: The Mean Absolute Errors (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination ( $R^2$ ) described as follows, respectively:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

$$\text{MAPE} = \left[ \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right] \times 100 \quad (12)$$

$$R^2 \equiv 1 - \frac{\left[ \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]}{\left[ \sum_{i=1}^n (y_i - \bar{y})^2 \right]} \quad (13)$$

where  $y_i$  is the actual values,  $\hat{y}_i$  is the predicted or interpolated output values,  $\bar{y}$  is the average and  $n$  are the total number of data points.

The ANN fitting process with its indicators for 14 ANNs is presented below. As it is known, ANN fitting in MATLAB uses limits to end the ANN training, validation, and test process. The limits are epoch number (1000), time (is not needed for this work since elapsed time is too small), performance (0), gradient ( $1 \times 10^{-7}$ ), mu value ( $1 \times 10^{10}$ ), and validation checks (6). Among them, performance, gradient, and validation checks are more prone to reach the limit value. After one of the limit values is reached, the epoch that gives the limit value first is used for later predictions. Gradient shows the variability in the error rate, mu is the threshold value of each iteration that is updated for each iteration, and the validation control states whether the iteration currently completed has minimized the error compared with the preceding iterations.

The epoch number versus MSE plots during training, validation and testing show whether the ANNs fit is truly converged. Figure 5 shows those plots for both different Re and Nu and ANN architectures. In Fig. 5, bottom right figure denotes the curve denominations and axes. From the figure, single input and single output Re-Nu ANNs show that normalization gives much smaller MSE values with increasing epoch number. Since the process time is undetectably small, normalization seems favourable in terms of the MSE of the training phase. Also, three hidden layers seems more proper. However, these figures are not enough to make the ultimate evaluation. Surprisingly, normalization does not seem to have a significant impact on single input single output Re-f ANNs by viewing Fig. 5(f) to 5(h). Nevertheless, both 10 neurons in the hidden layer and normalization cured relatively high MSE of three neuron single input single output Re-f ANN. Use of 10 neurons in the hidden layer of non-normalized ANN seems favourable in terms of MSE. In case of single input and double output ANNs, increasing neuron numbers of hidden layer clearly and significantly increases the lowest MSE of the validation results. It is also apparent

that normalization decreases the best MSE values greatly. Figure 5 shows that 3 neurons for hidden layer is favourable in general for the present case. On the other hand, if the best epoch is smaller than the last epoch with 6 units, than it means that validation limit stopped the process; if not, it means that gradient limit stopped the process. Except two ANNs, all processes were stopped by validation limits. Similar ANN evaluations are given in the literature based on below visual types [41, 42]. The evidenced negative slope of the validation error with increased epoch number shows that no overfitting exists. It should be noted that when the number of neurons in the hidden layer is increased to 100, the validation error slope at the neuron limit is flat, indicating overfitting.

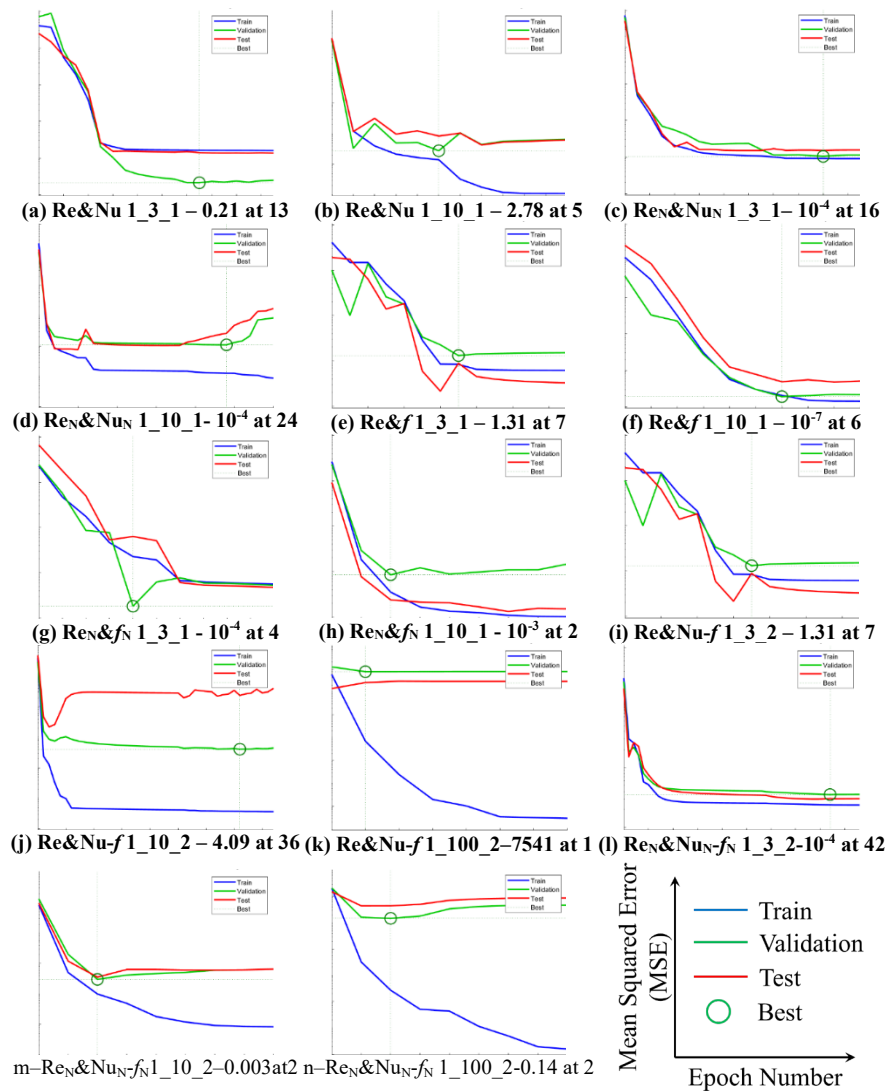
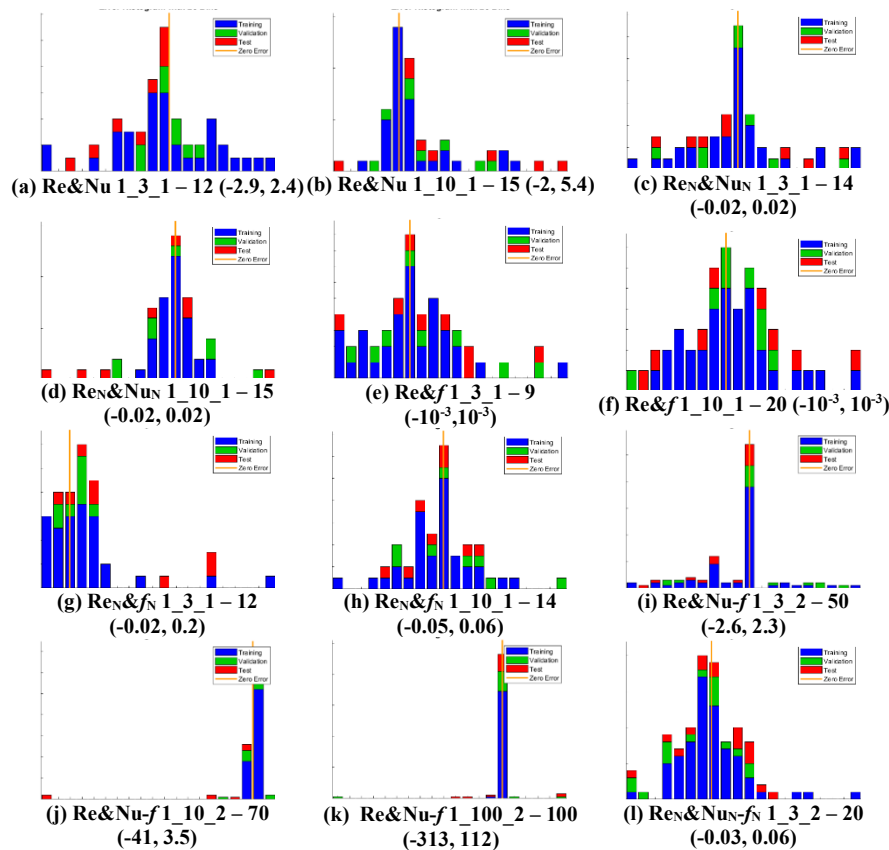


Fig. 5. MSE versus epoch numbers for different ANNs of the present work.

Error bins are shown in Fig. 6 and ANN fitting curves are given in Fig. 7, in a similar manner to Fig. 5. The zero error is displayed with an orange line. Increasing the hidden layer neuron number increases the minimum error and increases the interval of error. Normalization reduces the error associated with prediction of Nu, but does not significantly change the prediction of  $f$ . For double output ANNs, increasing the hidden layer neuron number increases the interval of error values; therefore, a big portion of the errors accumulate in the central bin.

It is seen from Fig. 7 that increasing neuron number in hidden layer makes the curve to approach exact shape of the experimental curve. However, this is not a desired behavior in case of hexagonal duct heat transfer since the main desire is to have a trend of data, filtering out the exceptions arising from uncertainties or experimental events. Normalization, on the other hand, somehow contributes to the mentioned phenomenon. In case of double output ANNs, fit curve is only given for Nu. Later the best correlation for  $f$  will be provided. For now, the Nu fit curve is enough. Similar to the results shown in Fig. 2, it is seen that double output fit curve exhibits occasional deviations at some test observations. Examination of the results shown in Fig. 7(j) and 7(k) and Fig. 7(m) and 7(n) supports a conclusion that three neurons in hidden layer is the best for ANN fitting with and without normalization. Also, 100 neurons in the hidden layer deteriorates the ability of ANN.



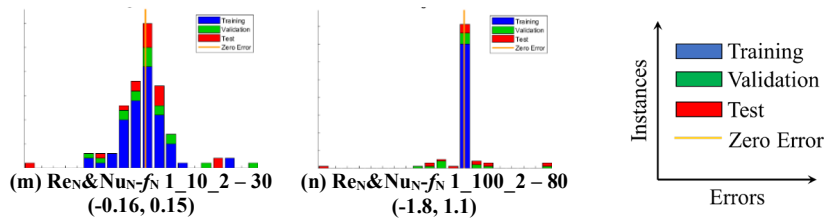


Fig. 6. Error histogram with 20 bins of ANNs training, validation, and testing.

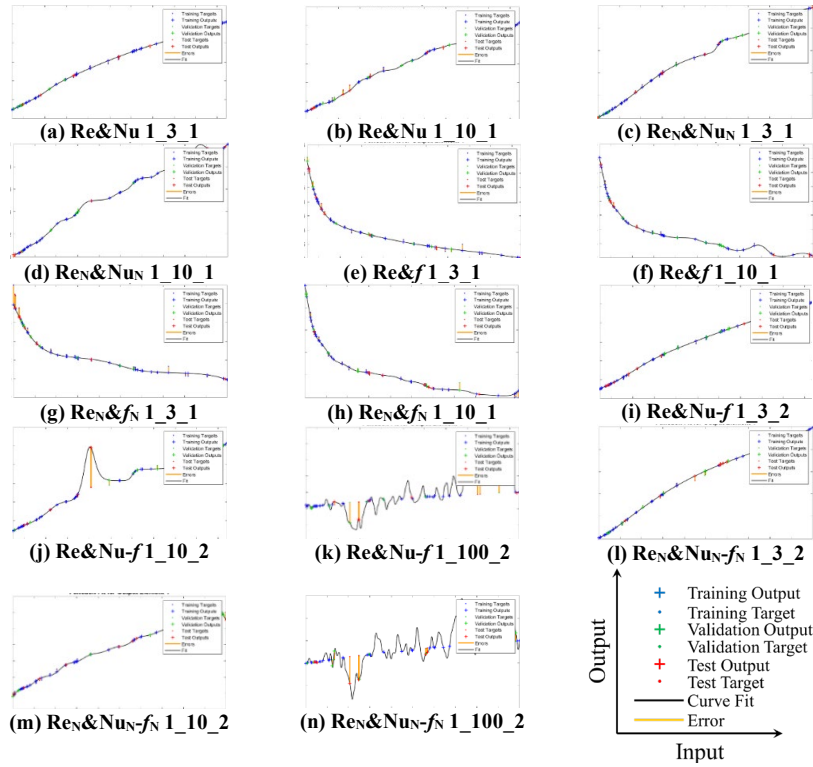


Fig. 7. Plot of ANNs fit curves.

Conventional RCFs are compared with ANNs as a basis for comparison to the best ANN results. The utilized RCF equations are given below while the coefficients are presented in Table 3.

$$(Nu \text{ or } f) = a Re + b \tag{14}$$

$$(Nu \text{ or } f) = a Re^2 + b Re + c \tag{15}$$

$$(Nu \text{ or } f) = a Re^b + c \tag{16}$$

$$(Nu \text{ or } f) = ae^{bRe} \tag{17}$$

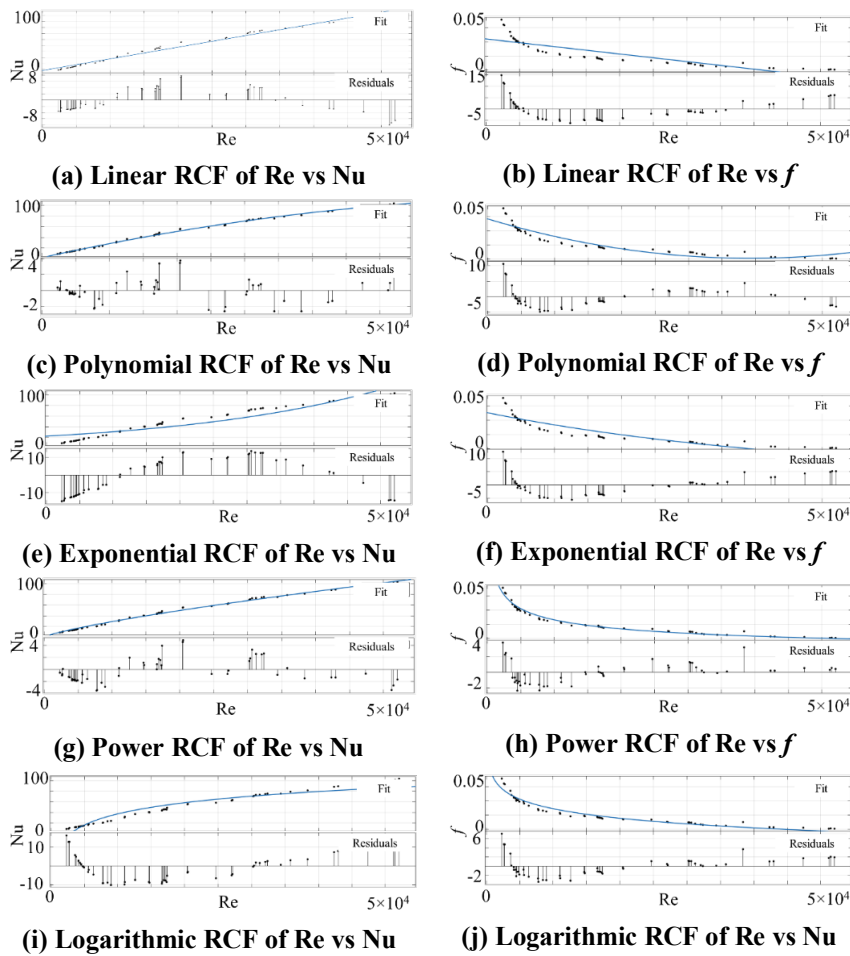
$$(Nu \text{ or } f) = a \ln(Re()) \tag{18}$$

**Table 3. Coefficients of RCF equations between Eqs. (14)-(18).**

RCF Function	Re-Nu			Re-f			Re-Nu <sub>N</sub>			Re-f <sub>N</sub>		
	a	b	c	a	b	c	a	b	c	a	b	c
Linear	0.001956	8	-	-5.331×10 <sup>-7</sup>	0.04197	-	1.031	0.03858	-	-0.7511	0.5787	-
Polynomial	-1.702×10 <sup>-8</sup>	0.002786	2.1521.972×10 <sup>-11</sup>	-1.407×10 <sup>-6</sup>	0.04812	-0.446	1.426	-0.003925	1.255	-1.864	0.6983	-
Power	0.02315	0.7753	-	0.5646	-0.3081	-	1.018*	0.8195*	-	NA	NA	NA
Exponential	22.18	3.207×10 <sup>-5</sup>	-	0.04519	-2.1×10 <sup>-5</sup>	-	0.1817	1.852	-	0.8027	-4.96	-
Logarithmic	30.26	-241.5	-	-0.009707	0.1238	-	NA	NA	NA	-0.1262*	0.08859*	-

\* Normally not available due to normalization that caused zero value related singularity. A very small number such as 1×10<sup>-7</sup> was added to have the coefficients. NA – Not available.

Statistical indicators and one-to-one plots between predicted and measured values are given in results section. However, RCF plots and residual plots are shown in Fig. 8 for non-normalized parameters in order to give an idea about the RCFs. Trends are same for normalized and non-normalized parameters since normalization only changes coefficients as seen in Table 3. Some RCFs in Fig. 8 indicate that Re-Nu change is mostly linear and can also be represented by second order polynomial. Normalization only increases values of constants which is a benefit for users for future users of the expressions. In terms of *f*, power and logarithmic RCFs seem proper.



**Fig. 8. RCF and residual plots for non-normalized parameters.**

### 3. Results and Discussion

Statistical indicators defined in Eqs. (9)-(13) are used to compare and evaluate the developed ANN and RCF models. Table 4 shows the statistical evaluations for all cases considered. Also, two-dimensional line plots are given for comparison of experimental results and predicted results in Figs. 9-12 for  $Nu$  and  $f$ . Figures 9 and 11 are for RCFs and Figs. 10 and 12 are for ANNs. RCFs and ANNs with and without normalization are also compared.

RCFs results for  $Re$ - $Nu$  couple show that  $Nu$  dependence on  $Re$  does not give a complex curve. The change curve is not completely linear, but it can be represented with a linear curve within acceptable error. Therefore, linear, second order polynomial and power curves are conveniently fitted to the data. The exponential and logarithmic functions are not proper for accounting the change in  $Nu$  relative to  $Re$  in the case of turbulent heat transfer in the hexagonal channel. Normalization of  $Re$  and  $Nu$  makes expression coefficient of RCFs to approach about unity as can be seen from Table 3 and, therefore, provide an application ease. However, exponential and logarithmic expressions have problem with the zero-value emerging from the normalization. When a very small number, such as  $1 \times 10^{-7}$  is added to normalized  $Re$  values in order to have a result from the exponential and logarithmic expressions, exponential expression gives result with relatively high error. The logarithmic expression gives enormous error and interestingly this error increases as the added small number magnitude decreases.

In case of the friction factor, RCFs show that the change depending on change of  $Re$  is more complex than the change of  $Nu$ . Table 4 gives the error values and values of determination coefficient for different RCFs in case of  $f$ - $Re$  change. The best function seems to be the power curve. The logarithmic curve, on the other hand, gives accurate results. However, it is very clear that the linear regression is not sufficient. Also, second order polynomial expression is not sufficient for curve fitting of  $Re$ - $f$  change. The exponential function is again not proper. Further, it is clear that normalization has more impact on  $f$  data than  $Nu$  data. The most accurate regression expression, i.e., power curve, cannot be used with normalized  $f$  data. Also, logarithmic regression performance becomes worse for the normalized data.

In contrast, the exponential function statistical performance improves with normalization. Nevertheless, in terms of accuracy, ease, and applicability of RCF for non-normalized values seems favourable. On the other hand, as seen from Table 3, normalization again bring the coefficients in the RCF expressions about unity. By comparison, ANNs perform well for single outputs when predicting  $Nu$  from  $Re$ , as seen in Table 4. Error values are very close and small for the 1\_3\_1 and 1\_10\_1 ANNs. On the other hand, the 3 neurons in the hidden layer of double output ANN led to relatively close values to single output ANNs. However, increasing number of neurons in the hidden layer of double output ANNs increases error values and reduces determination coefficient values. The problem here is associated with the small fluctuations of the experimental data, which are then amplified with the increasing number of neurons in the hidden layer. This amplifying effect is similar to overtraining phenomenon of ANNs. Nevertheless, with up to ten neurons in the hidden layer of double output ANN, the determination coefficient is above 0.9 value.

Normalization, on the other hand, greatly improves the performance of double output ANNs. Normalization reduces the error values and increases the value of determination coefficient. Still, it is seen that increasing hidden layer neuron numbers after a certain number

deteriorates results due to aforementioned amplification phenomenon. Table 4 also shows the error values and values of determination coefficient for ANNs of  $Re-f$  changes. Single output ANNs are very successful in prediction and interpolation of the friction factor data. Absolute error values are very small even for small number values of the inputs. Also, ten neurons in the hidden layer of single output ANN seems to perform better comparing to the three-neuron single output ANN. However, double output ANNs are very problematic in terms of predicting  $f$ . Since normalization gives a range between one and zero for the friction factor, which is same for  $Nu$  during ANN training, validation and testing of the normalized parameters, normalization greatly increases double output ANNs' performance. However, increasing hidden layer neuron numbers to a hundred increases error values both for non-normalized and normalized parameters. Although error magnitudes and values of determination coefficient in Table 4 give an idea about RCFs and ANNs in the study, the distribution of one-to-one data of predicted values versus experimental ones is important to view. By viewing those graphics, trends and biases can be observed. As mentioned earlier, those graphics are grouped in Figs. 9-12 for  $Nu$  and  $f$  predictions.

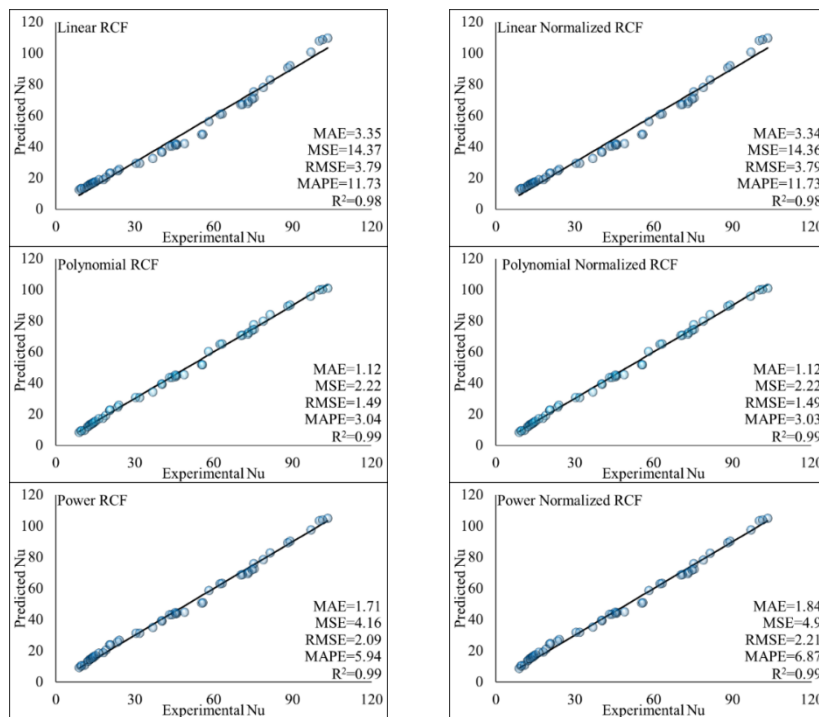
**Table 4. Statistical measures of RCFs and ANNs.**

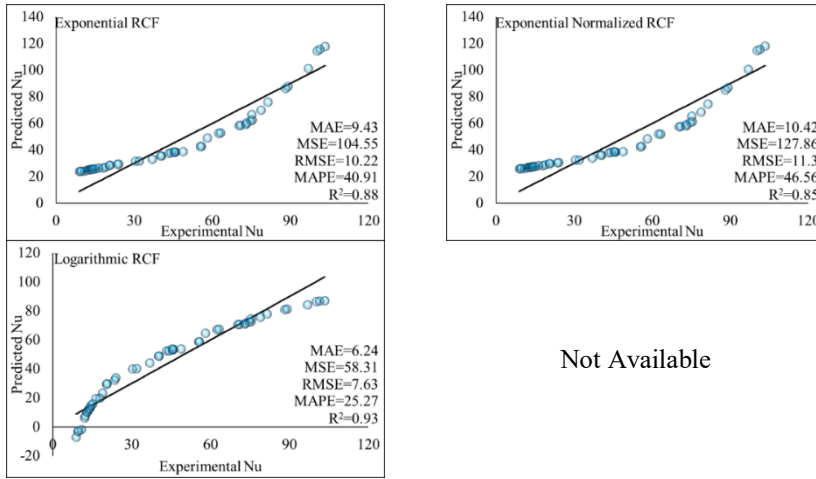
		MAE	MSE	RMSE	MAPE	R <sup>2</sup>	
Re-Nu	RCFs	Linear	3.35	14.37	3.79	11.73	0.98
		Polynomial	1.12	2.22	1.49	3.04	0.99
		Power	1.71	4.16	2.09	5.94	0.99
		Exponential	9.43	104.55	10.22	40.91	0.88
		Logarithmic	6.24	58.31	7.63	25.27	0.93
Re-Nu Nu <sub>N</sub>	RCFs	Linear	3.34	14.36	3.79	11.73	0.98
		Polynomial	1.12	2.22	1.49	3.03	0.99
		Power	1.84	4.9	2.21	6.87	0.99
		Exponential	10.42	127.86	11.3	46.56	0.85
		NA	NA	NA	NA	NA	
Re-Nu	ANNs	1_3_1	0.94	1.49	1.22	2.4	0.99
		1_10_1	0.62	0.73	0.85	2.73	0.99
		1_3_2	1.17	1.84	1.35	4.16	0.99
		1_10_2	2.59	74.96	8.65	5	0.91
		1_100_2	19.2	3168	56.28	42.63	0.42
Re-Nu Nu <sub>N</sub>	ANNs	1_3_1	1.45	2.07	1.44	4.78	0.99
		1_10_1	0.48	0.72	0.85	1.46	0.99
		1_3_2	0.96	1.67	1.29	2.68	0.99
		1_10_2	1.56	4.77	2.18	5.64	0.99
		1_100_2	13.48	1397	37.38	30.05	0.47
Re-f	RCFs	Linear	0.0039	2.47×10 <sup>-5</sup>	0.0049	12.39	0.71
		Polynomial	0.0028	1.42×10 <sup>-5</sup>	0.0037	9.22	0.83
		Power	0.001	1.84×10 <sup>-6</sup>	0.0013	3.09	0.98
		Exponential	0.0032	1.75×10 <sup>-5</sup>	0.0041	10.17	0.8
		Logarithmic	0.0017	4.96×10 <sup>-6</sup>	0.0022	5.24	0.94
Re-Nu f <sub>N</sub>	RCFs	Linear	0.0039	2.47×10 <sup>-5</sup>	0.0049	12.39	0.71
		Polynomial	0.0026	1.12×10 <sup>-5</sup>	0.0033	8.16	0.86
		Power	NA	NA	NA	NA	NA
		Exponential	0.0019	5.55×10 <sup>-6</sup>	0.0023	6.02	0.93
		0.003	3.92×10 <sup>-5</sup>	0.0062	8.37	0.6	
Re-f	ANNs	1_3_1	0.0004	4.67×10 <sup>-7</sup>	0.0006	1.55	0.99
		1_10_1	0.0002	2.16×10 <sup>-7</sup>	0.0004	0.82	0.99
		1_3_2	0.015	0.0002	0.016	48.24	0.87
		1_10_2	0.003	2.15×10 <sup>-5</sup>	0.0046	11.3	0.76
		1_100_2	0.023	0.0009	0.031	83.31	0.47
Re-Nu f <sub>N</sub>	ANNs	1_3_1	0.001	3.23×10 <sup>-6</sup>	0.0017	2.78	0.97
		1_10_1	0.0005	5.18×10 <sup>-7</sup>	0.0007	1.81	0.99
		1_3_2	0.0004	4.33×10 <sup>-7</sup>	0.0006	1.57	0.99
		1_10_2	0.0013	3.6×10 <sup>-6</sup>	0.0018	4.08	0.96
		1_100_2	0.0028	3.71×10 <sup>-5</sup>	0.006	10.34	0.63



In Fig. 9, predicted Nu values versus experimental Nu values are given. The predictions in which normalization is utilized are given on the right column. Of course, Fig. 9 is for RCFs of Nu according to Re. It is seen that normalization has almost no effect in case of Nu prediction with RCFs. Also, as stated according to Table 4, polynomial, power, and linear expressions have high accuracies as seen in Fig. 9. Their extrapolation potential also seems good looking to the change of distribution. However, both the interpolation and extrapolation abilities of exponential and logarithmic expressions are evaluated as inconvenient in terms of Nu. In Fig. 10 for  $f$ , it should be stressed that same performance of RCFs cannot be mentioned in terms of  $f$ .

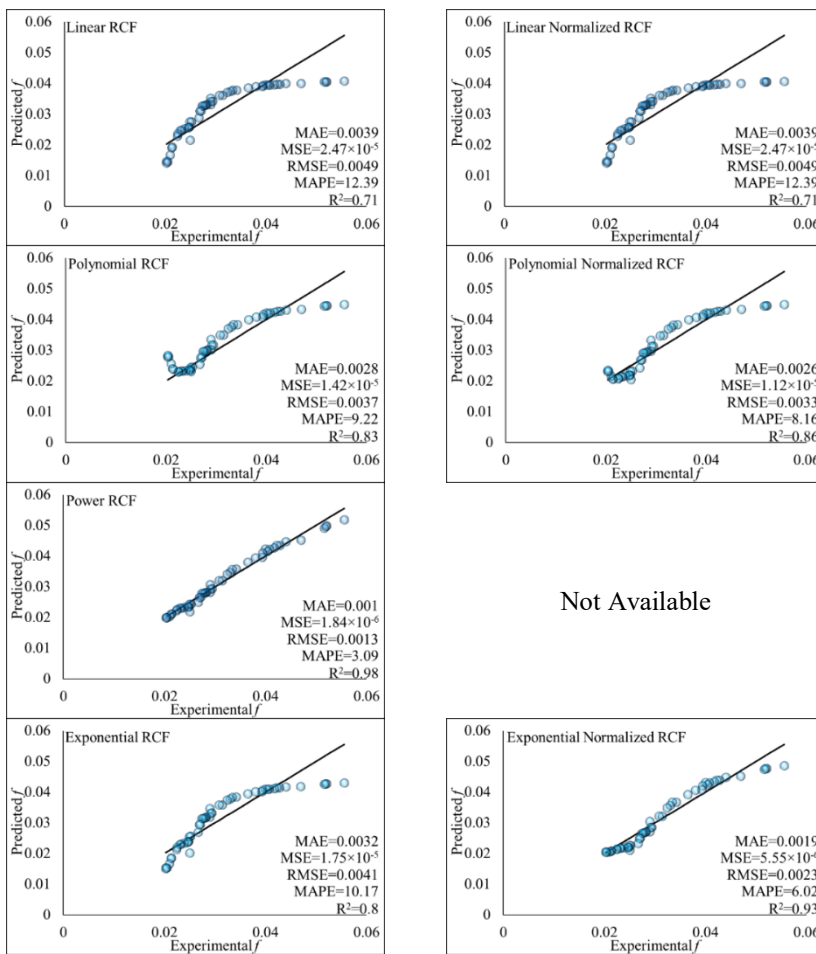
By viewing Fig. 10, only interpolation of data is recommended. The most convenient expression is power formula for non-normalized  $f$  values. However, when normalization is applied, power formulation cannot be used. Therefore, only option left seems to be exponential expression in terms of RCFs according to Fig. 10. Figure 11 presents ANNs for prediction, interpolation, and extrapolation of Nu. Comparatively, the distribution of data is favourable than RCFs. Only 100 neurons double output ANNs have problems by only looking to the Fig. 11. Therefore, it can be said that interpolation with ANNs is superior to RCFs in terms of accuracy. However, effect of normalization cannot be observed in Fig. 11. On the other hand, Fig. 12 shows ANNs for prediction of  $f$ . Especially, effects of normalization are clearly seen in Fig. 12. Normalization cures result of ANNs for  $f$ . Also, ten neurons in hidden layer for prediction of  $f$  both for single and double output ANNs give a more favourable data point distribution as seen in Fig. 12.





Not Available

Fig. 9. Consistency of the predicted Nu with experimental Nu by RCFs.



Not Available

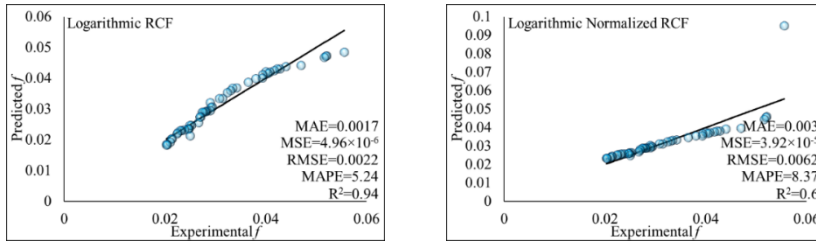


Fig. 10. Consistency of the predicted  $f$  with experimental  $f$  by RCFs.

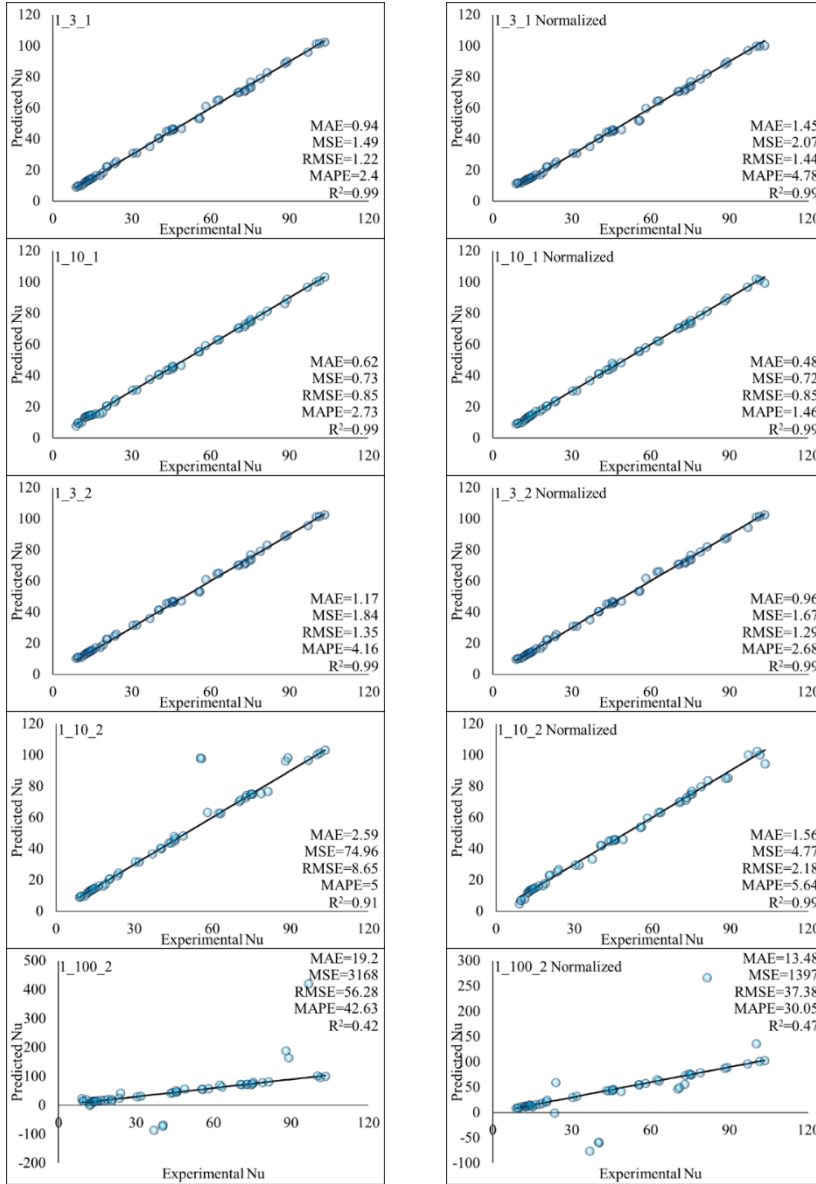


Fig. 11. Consistency of the predicted Nu with experimental Nu by ANNs.

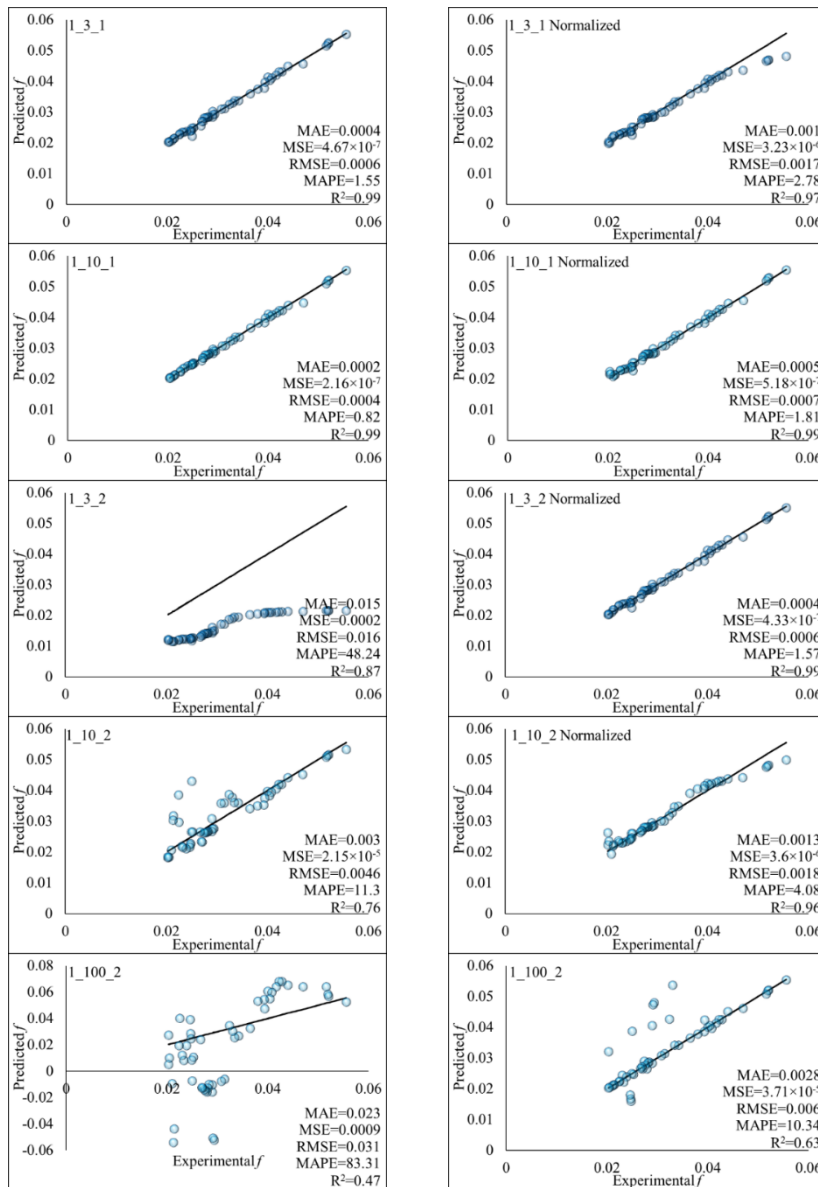


Fig. 12. Consistency of the predicted  $f$  with experimental  $f$  by ANNs.

#### 4. Conclusion

Experimental data of heat transfer and fluid flow in a hexagonal channel for turbulent flow are used for comparison of several RCFs and ANNs. The experimental data consist of  $Re$  of the flow as the independent input parameter, and two other dependent parameters of  $Nu$  and  $f$ .  $Re$  was fed to RCFs and ANNs in order to have the  $Nu$  and  $f$  values. Those predicted values are compared with the dataset values in the paper. The evaluation and assessment on how the predictions fit with dataset inform about RCFs and ANNs. Also, normalized and non-normalized data were used during utilizations of RCFs and ANNs for showing

effects of normalization on the data fitting process. A comprehensible literature survey was conducted on heat transfer and pressure drop in hexagonal ducts and channels, heat transfer and pressure drop in non-circular conduits, ANNs and some other machine learning and statistical approaches in heat transfer and pressure drop research. Details of the methods as well as some metrics for method evaluation are provided. Dataset is given in tabular and graphical forms. The work is distinguished from the literature with its unique single input and double output design. Comprehensive method indicators and comparison of RCF versus ANN stress the offered value. Nonnormalized dataset is used in order to show effectiveness of the normalized data in ANNs.

Main conclusions from the work are; RCFs clearly show the main trends and functional expressions of the dependent parameters. Applying and using RCF are easier comparing to ANNs. Normalization has no significant effect on RCF considering statistical error measures and some RCFs cannot deal with zero edge of the data interval. Normalization greatly increases effectiveness and accuracy of the ANNs for increased neuron numbers. Power RCF gives good approximation and interpolation for both  $Nu$  and  $f$ . In case of  $Nu$ , polynomial regression expression, and in case of  $f$ , logarithmic regression expression are the other RCFs that give accurate results. ANNs have higher determination coefficient values and lower error indicators comparing to RCFs. There is an overtraining risk of ANNs. For cases similar to the present one, RCFs are recommended by the authors. ANNs may be considered for more input parameter numbers. It is inferred from the results that an interference is needed to the ANN code in order to group hidden layer numbers into the number of outputs. After a number of hidden layer neurons, increasing neuron numbers amplifies the fluctuations in the dataset.

## References

1. Turgut, O. and Sari, M. (2013). Experimental and numerical study of turbulent flow and heat transfer inside hexagonal duct. *Heat and Mass Transfer*, 49(4), 543-554.
2. Turgut, O. (2014). Numerical investigation of laminar flow and heat transfer in hexagonal ducts under isothermal and constant heat flux boundary conditions. *Transactions of Mechanical Engineering (IJST)*, 38(M1), 45-56.
3. Iwaniszyn, M.; Jodłowski, P.J.; Sintera, K.; Gancarczyk, A.; Korpyś, M.; Jędrzejczyk, R.J. and Kołodziej, A. (2021). Entrance effects on forced convective heat transfer in laminar flow through short hexagonal channels: Experimental and CFD study. *Chemical Engineering Journal*, 405, 126635.
4. Ismail, M.F.; Hasan, M.N. and Ali, M. (2014). Numerical simulation of turbulent heat transfer from perforated plate-fin heat sinks. *Heat and Mass Transfer*, 50(4), 509-519.
5. Chen, J.; Ji, X.; Lu, X. and Wang, C. (2017). Mechanism study of heat transfer enhancement using twisted hexagonal tube with slurry from biogas plant. *Energy Procedia*, 142, 880-885.
6. Yadav, R.J.; Kore, S.S. and Joshi, P.S. (2018). Correlations for heat transfer coefficient and friction factor for turbulent flow of air through square and hexagonal ducts with twisted tape insert. *Heat and Mass Transfer*, 54(5), 1467-1475.

7. Cheng, J.; Qian, Z.; Wang, Q.; Fei, C. and Huang, W. (2019). Numerical study of heat transfer and flow characteristic of twisted tube with different cross section shapes. *Heat and Mass Transfer*, 55(3), 823-844.
8. Mahato, S.K.; Rana, S.C.; Barman, R.N. and Goswami, S. (2018). Numerical analysis of heat transfer and fluid flow through twisted hexagonal and square duct and their comparisons. *Chemical Engineering*, 71
9. Zhang, L.-Z. (2015). Transient and conjugate heat and mass transfer in hexagonal ducts with adsorbent walls. *International Journal of Heat and Mass Transfer*, 84, 271-281.
10. Tasnim, S.H.; Mahmud, S. and Fraser, R.A. (2012). Compressible pulsating convection through regular and random porous media: The thermoacoustic case. *Heat and Mass Transfer*, 48(2), 329-342.
11. Hou, X.; Deng, Z. and Yin, G. (2014). Application of transfer matrix method in heat transfer performance analysis of multi-re-entrant honeycomb structures. *Heat and Mass Transfer*, 50(12), 1765-1782.
12. Subasi, A.; Ozsipahi, M.; Sahin, B. and Gunes, H. (2017). Performance evaluation of RANS-based turbulence models in simulating a honeycomb heat sink. *Heat and Mass Transfer*, 53(7), 2435-2443.
13. Alfaryjat, A.; Mohammed, H.; Adam, N.M.; Ariffin, M. and Najafabadi, M.I. (2014). Influence of geometrical parameters of hexagonal, circular, and rhombus microchannel heat sinks on the thermohydraulic characteristics. *International Communications in Heat and Mass Transfer*, 52, 121-131.
14. Mohanraj, M.; Jayaraj, S. and Muraleedharan, C. (2012). Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems - A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1340-1358.
15. Beigzadeh, R. and Rahimi, M. (2012). Prediction of heat transfer and flow characteristics in helically coiled tubes using artificial neural networks. *International Communications in Heat and Mass Transfer*, 39(8), 1279-1285.
16. Cong, T.; Su, G.; Qiu, S. and Tian, W. (2013). Applications of ANNs in flow and heat transfer problems in nuclear engineering: a review work. *Progress in Nuclear Energy*, 62, 54-71.
17. Chelang, A.; Abulkarim, A. and Adnan, I. (2018). Prediction of discharge coefficient for cylindrical weirs using adaptive neuro fuzzy inference system ANFIS and multilayer neural networks MLP. *International Journal of Applied Engineering Research*, 13(9), 7042-7051.
18. Azizi, S. and Ahmadloo, E. (2016). Prediction of heat transfer coefficient during condensation of R134a in inclined tubes using artificial neural network. *Applied Thermal Engineering*, 106, 203-210.
19. Taimoor, I.; Rahman, M.L.; Aankhy, N.S. and Khalid, M.B. (2019). Thermal performance analysis of plate fin arrays with hexagonal perforations under turbulent flow regime. *AIP Publishing LLC*. 030002.
20. Yang, D.; Wang, Y.; Ding, G.; Jin, Z.; Zhao, J. and Wang, G. (2017). Numerical and experimental analysis of cooling performance of single-phase array microchannel heat sinks with different pin-fin configurations. *Applied Thermal Engineering*, 112, 1547-1556.

21. Wung, T.; Niethammer, J. and Chen, C.-J. (1986). Measurements of heat-mass transfer and pressure drop for some non-standard arrays of tubes in crossflow. *Begel House Inc.*
22. Dou, M.; Qin, C.; Li, G. and Wang, C. (2020). Research on calculation method of free flow discharge based on artificial neural network and regression analysis. *Flow Measurement and Instrumentation*, 72, 101707.
23. Kalogirou, S.A. (2001). Artificial neural networks in renewable energy systems applications: a review. *Renewable and sustainable energy reviews*, 5(4), 373-401.
24. Huminic, G. and Huminic, A. (2016). Heat transfer and flow characteristics of conventional fluids and nanofluids in curved tubes: a review. *Renewable and Sustainable Energy Reviews*, 58, 1327-1347.
25. Kumar, P. and Pandey, K. (2017). Effect on heat transfer characteristics of nanofluids flowing under laminar and turbulent flow regime—a review. *IOP Publishing*. 012168.
26. Ghareeb, A.; Wang, W. and Hallinan, K. (2019). Data-driven modelling for building energy prediction using regression-based analysis. 1-5.
27. James, G.; Witten, D.; Hastie, T. and Tibshirani, R. (2013). *An introduction to statistical learning* Springer
28. Jain, A.K.; Mao, J. and Mohiuddin, K.M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44.
29. Sengupta, R.N.; Gupta, A. and Dutta, J. (2016). *Decision sciences: Theory and practice* Crc Press
30. Braspenning, P.J.; Thuijsman, F. and Weijters, A.J.M.M. (1995). *Artificial neural networks: an introduction to ANN theory and practice* Springer Science & Business Media
31. Karsoliya, S. (2012). Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International Journal of Engineering Trends and Technology*, 3(6), 714-717.
32. Panchal, F.S. and Panchal, M. (2014). Review on methods of selecting number of hidden nodes in artificial neural network. *International Journal of Computer Science and Mobile Computing*, 3(11), 455-464.
33. Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2), 164-168.
34. Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2), 431-441.
35. Hagan, M.T. and Menhaj, M.B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6), 989-993.
36. Gavin, H. (2011). The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering, Duke University*, 28, 1-5.
37. Akdag, U.; Komur, M.A. and Ozguc, A.F. (2009). Estimation of heat transfer in oscillating annular flow using artificial neural networks. *Advances in engineering software*, 40(9), 864-870.

38. Gavin, H.P. (2019). The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering, Duke University* <http://people.duke.edu/~hpgavin/ce281/lm.pdf>, 1-19.
39. Khan, K. and Sahai, A. (2012). A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context. *International Journal of Intelligent Systems and Applications*, 4(7), 23.
40. Touthmalani, R. (2013). Comparison result of inversion of gravity data of a fault by particle swarm optimization and Levenberg-Marquardt methods. *SpringerPlus*, 2(1), 1-6.
41. Aghayari, R.; Rohani, S.; Ghasemi, N.; Heiran, E.N.K. and Mazaheri, H. (2020). Numerical investigation of heat transfer in a helically coiled tube using copper/water nano-fluid under constant heat flux and prediction of the results using perceptron and radial basis function networks. *Heat and Mass Transfer*, 56(4), 1051-1075.
42. Ghasemi, N.; Maddah, H.; Mohebbi, M.; Aghayari, R. and Rohani, S. (2019). Proposing a method for combining monitored multilayered perceptron (MLP) and self-organizing map (SOM) neural networks in prediction of heat transfer parameters in a double pipe heat exchanger with nanofluid. *Heat and Mass Transfer*, 55(8), 2261-2276.