Using Steepness Coefficient to Improve Artificial Neural Network Performance for Environmental Modeling

Selami Demir*, Aykut Karadeniz, Neslihan Manav Demir

Yıldız Technical University, Faculty of Civil Engineering, Environmental Engineering Department,
34220, Esenler, Istanbul, Turkey

Abstract

This paper presents results from a research study in which the effects of steepness coefficient (S) for the activation function of a back propagation neural network (BPNN) were investigated, and optimum values of S for each activation function were suggested for environmental modeling purposes. A BPNN algorithm was implemented in Excel Visual Basic for Applications with built-in activation functions of sigmoid, hyperbolic tangent, and sinc. Various steepness coefficients were employed for modeling cyclone Euler numbers for pressure drop estimation with three different activation functions. Best results for sigmoid function were obtained for $S = 1.00$ with a median value of mean square errors (MSEs) of $4.33*10^{-4}$. For hyperbolic tangent function, the optimum value of S was found as 0.2 with a median MSE value of $2.02*10^{-4}$. The median value of MSEs obtained with BPNN sinc function was $1.20*10^{-3}$ for $S = 0.50$. Results showed, for environmental modeling problems, that any activation function can be used with satisfactory results provided that an optimized value of the steepness coefficient is used, which is considered problem-specific.

Keywords: artificial neural networks, environmental modeling, activation functions, steepness

Introduction

With the help of modern computational technologies, artificial neural networks (ANNs) have gained great attraction because of their reliable and robust characteristics in explaining the nonlinear relationships between inputs and outputs of complex systems [1, 2]. In recent decades ANNs have gained widespread application in all engineering fields, specifically environmental engineering, in which physical, chemical, and biochemical processes are employed together for water and wastewater treatment, air pollution and its control, and other related fields [3-8] where models capable of simulating complicated phenomena are needed. Representative research in which neural networks were applied for simulating processes highly complicated for solving with regression or any other means include Zhao and Su [9], Huang et al. [10], Aghav et al. [11], Bhatti et al. [12], Bayram and Kankal [13], Samli et al. [14], Hernandez-Ramirez et al. [15], Wu et al. [16], and Elsayed and Lacor [17].

An artificial neural network is a group of processing elements, called neurons, capable of replicating human

*e-mail: seldemir@yildiz.edu.tr

behavior for learning and simulating a given linear or nonlinear relationship. The neurons in an ANN are arranged in successive layers in a manner that each neuron in a given layer receives input signals from each of the neurons in a precedent layer and generates an output signal for the following layer of neurons. The connections between the neurons are provided by coefficients, called weights. In addition, each neuron utilizes an optional constant for shifted output signals, which is usually called a bias. Successive calculations of the neurons' output signals in each layer finally produce an estimate of dependent variables in the system.

Artificial neural networks are very similar to the human nervous system in which tremendous numbers of neurons perform a given task together [14]. The neurons in an ANN can communicate by sending signals to each other with the ultimate purpose of learning the complex relationship between inputs and outputs [13]. The learning process comprises a forward run and a backward run for each data point in each learning cycle, called epoch. In the forward run, a neuron employs its own bias as well as the weights and the output signals of the neurons in the preceeding layer to calculate an input signal, which is transmitted to the neurons in the successive layer via an activation fuction – finally producing an estimate of the system's output for the given inputs. In the backward run, on the other hand, the network updates all the biases and weights by backpropagating the the prediction error based on the discrepancy between the target and predicted values of the output signal, which is the main learning process. Couples of forward and backward runs for each sample in the training set complete an epoch and the learning continues for predefined times of epochs.

Artificial neural networks can be classified under a number of groups with respect to their structure, activation functions, and learning mechanisms, including back propagation neural networks (BPNN), radial basis function neural networks (RBFNN), and generalized regression neural networks (GRNN) [2].

Zhao and Su [9] reported that although both exhibit similar levels of prediction accuracy, RBF-type neural networks should be used instead of BPNN to reduce the computational efforts because of the fact that RBFNN learns faster. Although their conclusion is based on sufficient evidence, computational time and load is not very important in environmental modeling problems most of the time. Generally, being simpler in implementation is more of a value than being faster.

The key to success of a neural network in simulating a nonlinear relationship is based on the fundamental knowledge on how the neural network performs depending on a number of parameters related to network topology such as number of layers and number of neurons in each layer. Besides, the characteristics of the activation function and its defining parameters such as steepness coefficient (S) affect how fast and how good the neural network can learn the nonlinear pattern. Talebi et al. [18] reported that both the choice and the shape/slope of the activation function are the most important parameters affecting the network learning. Since steepness of the activation function defines its main characteristic, selection of the steepness coefficient (S) is equally important for successful application of the neural network. Talebi et al. [18] reported that selecting a large value for S may lead to various learning problems and yield results similar to those when a large learning rate is selected due to the fact that the weights are updated in proportion to S.

The main aim of this study is the improvement of BPNN for environmental modeling purposes by incorporating the steepness coefficient into most commonly used activation functions and optimizing the value of the steepness coefficient. The paper also presents results of the optimization work by the use of a test case in which the BPNN is optimized for simulating cyclone pressure drop as a function of cyclone geometrical parameters. Finally, a generalized protocol for selecting the optimum value of steepness coefficient is suggested.

Materials and Methods

Artificial Neural Network

An Excel Visual Basic for Applications (VBA) code was implemented for back propagation neural network. The program allows the user to select the number of neurons in the input and output layers. For the output layer, the user can also select an activation function for the neurons. The number of hidden layers and number of neurons in each layer as well as the activation function for each layer are user-defined inputs for the neural network program related to network topology. The program also allows the user to select the number of epochs and learning rate as the global options for the network. The percent of training data set is also an input for the program and the program selects the training data randomly in each run. Table 1 shows global options for the BPNN in this study and Fig. 1 shows the general structure of the implemented BPNN algorithm.

Activation Functions and Steepness Coefficients

Activation function is a property of each neuron in hidden layers and the output layer. It determines how the neuron behaves depending on total input to the neuron.

Table 1. Global options for BPNN.

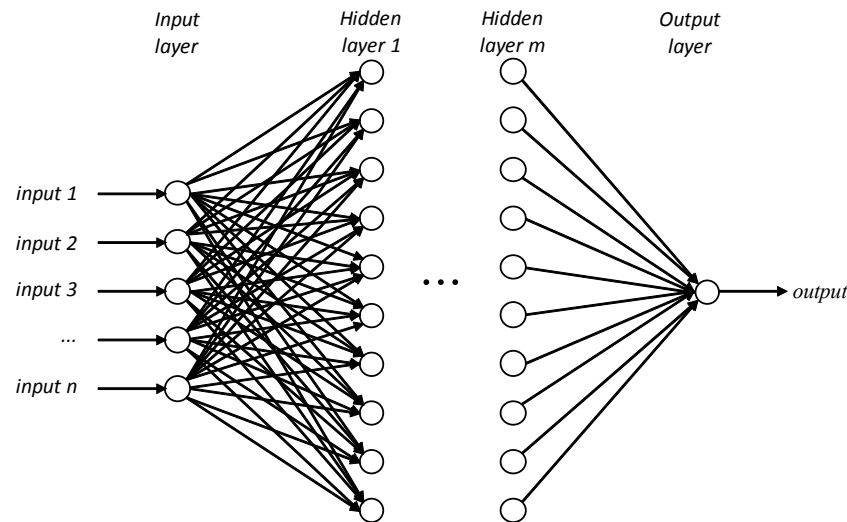| Parameter	Value
Number of hidden layers	1
Number of hidden neurons	10
Percentage of training data set (%)	70
Learning Rate	0.75
Number of epochs	5000

Fig. 1. Artificial neural network structure.

Results of a previous work showed that the sigmoid and the sinc functions were capable of representing highly complicated, nonlinear relationships. On the other hand, the hyperbolic tangent produced very weak correlations between measured and predicted results, contrary to the expected results, and the reason for this unexpected performance was considered \the probable effect of the steepness coefficient.

In this study, sigmoid, sinc, and hyperbolic tangent activation functions with various steepness coefficients were separately employed to determine the effects of steepness coefficient on the BPNN performance. These activation functions are shown in Table 2 and their behaviors with various steepness coefficients are shown in Fig. 2.

Test Case

Cyclone pressure drop data was obtained from results of a research project and part of measurement results that previously appeared in Demir [19] and Karadeniz [20]. The data include 972 pressure drop measurements in 162 cyclones of various geometrical dimensions at six different inlet velocities between 10 and 24 m/s. Measured pressure drops ranged between 84 and 2,045 Pa.

Cyclone pressure drop is usually described by Euler number (also called number of inlet velocity heads) as in Eq. (1):

$$\Delta P = \left(\frac{1}{2}\rho V_{in}^2\right)E_U$$

(1)

...where ΔP is the pressure drop in cyclone (Pa), ρ is density of gas (kg/m^3), V_{in} is inlet velocity (m/s), and E_U is Euler number. The Euler number is a function of solid loading to the cyclone, gas density, viscosity, and cyclone geometry. Cyclone geometry includes inlet height (a), inlet width (b), cylinder height (h_b), cone height (h_c), vortex finder length (S_L), vortex finder diameter (D_e), and cone-tip diameter. For describing the shape of a cyclone, it is customary to use these parameters as their ratios to the cyclone's body diameter (K_a, K_b, K_{hb}, K_{hc}, K_{SL}, K_{De}, and K_B).

Table 2. Behaviors of activation functions used.

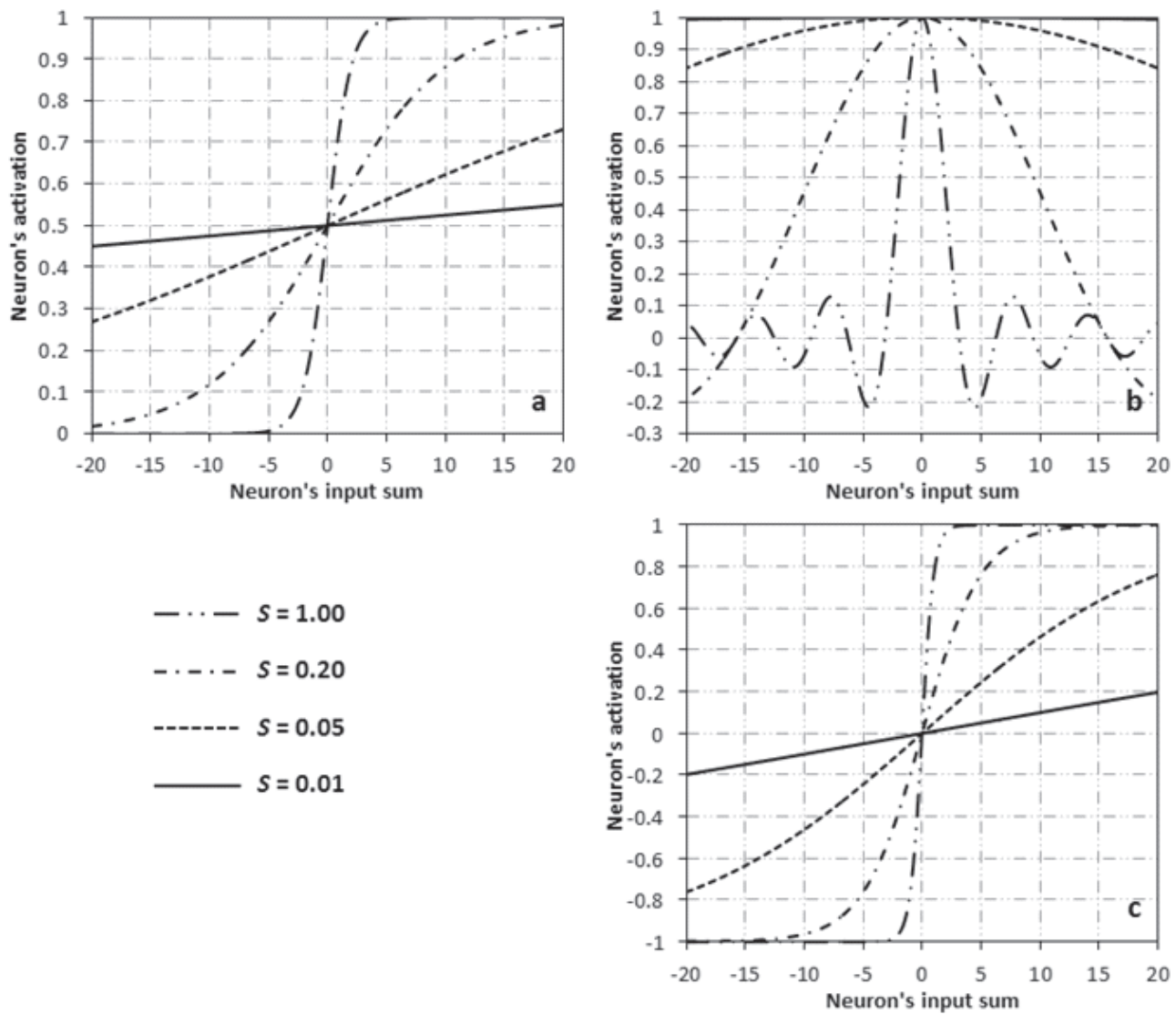| Function Name	Mathematical Expression	Derivative
Sigmoid	$y = \dfrac{1}{1+e^{-Sx}}$ *Span:* $0 < y < 1$	$\dfrac{dy}{dx} = Sy(1-y)$
Hyperbolic tangent	$y = \dfrac{e^{Sx}-e^{-Sx}}{e^{Sx}-e^{-Sx}}$ *Span:* $-1 < y < 1$	$\dfrac{dy}{dx} = S(1-y^2)$
Sinc	$y = \begin{cases} 1 & \Leftrightarrow \quad x=0 \\ \dfrac{\sin(Sx)}{Sx} & \Leftrightarrow \quad x \neq 0 \end{cases}$ *Span:* $-0.2172 \leq y \leq 1$	$y = \begin{cases} 0 & \Leftrightarrow \quad x=0 \\ \dfrac{Cos(Sx)}{x} - \dfrac{Sin(Sx)}{Sx^2} & \Leftrightarrow \quad x \neq 0 \end{cases}$

Fig. 2. Behaviors of a) sigmoid, b) sinc, and c) hyperbolic tangent activation functions with various steepness coefficients (S).

The effects of solid loading and gas density were examined previously by Cortes and Gil [21], and Chen and Shi [22]. In this study, geometrical parameters and Euler numbers for 162 cyclones were used as a test case for determining the performance of BPNN with respect to steepness coefficients. The Euler number for each cyclone was calculated by Eq. (1) as an average of measurement results at six inlet velocities and used as target output for the neural network. Since no data was present on how the inlet width and cone-tip diameter affect Euler number, K_a, K_{hb}, K_{hc}, K_{SL}, and K_{De} for 162 cyclones were used as inputs to the neural network.

Calculated Euler numbers ranged between 1.09 and 9.17, and during the simulations the target outputs were normalized in accordance with the span of the respective activation function (Table 2). The neural network outputs were denormalized for comparing and assessing the performance of BPNN with various steepness coefficients.

Results and Discussion

A total of 162 cyclones with varying geometries was used for testing the back propagation neural network using three different activation functions, namely sigmoid, hyperbolic tangent, and sinc functions with various steepness coefficients between 0.01 and 2.00. 70% of data points were used as a training set with a learning rate of 0.75. Since the network randomly selects the training data set, it was run 25 times for each activation function and each steepness coefficient. The mean square errors from each run were averaged and compared.

Fig. 3 shows mean square errors (MSEs) for various steepness coefficients (S) ranging from 0.01 to 2.00. Fig. 3a is a box-whisker plot. In Fig. 3a, the line in the boxes represent the median value from 25 runs, the upper and lower boundaries of boxes represent the 75th and 25th percentiles, respectively, while the tip of the whiskers represent 90th and 10th percentiles, respectively. It is clear that MSEs showed similar tendencies for $S = 0.01$,
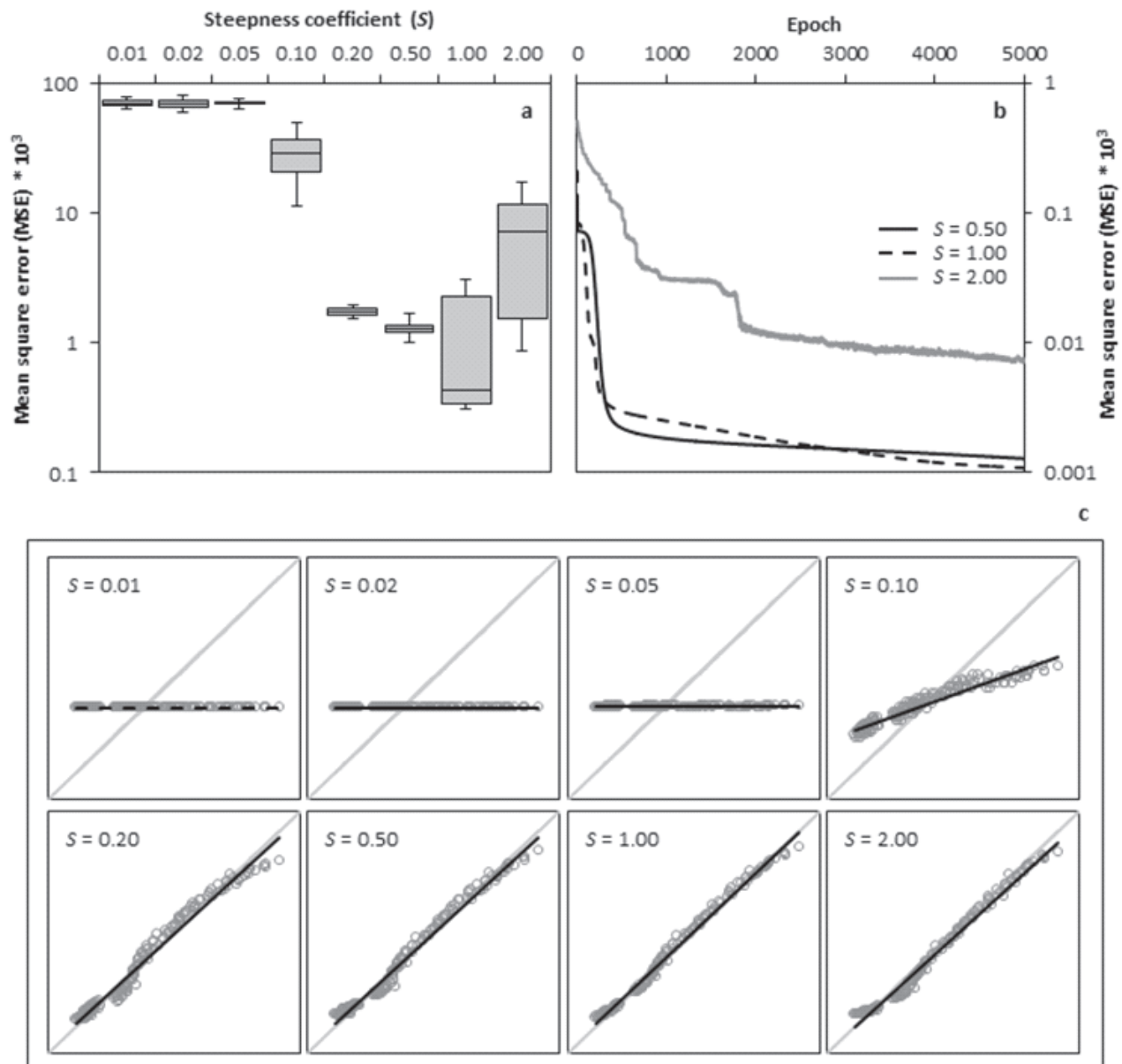
Fig. 3. BPNN results with sigmoid activation function: a) mean square errors with various steepness coefficients, b) learning trends, c) measured vs. predicted Euler numbers for varying steepness coefficients.

$S = 0.02$, and $S = 0.05$ (Fig. 3a). For $S = 0.10$, a decreasing trend started and the results were suddenly improved for $S = 0.20$. The best MSEs were obtained with $S = 1.00$, with median and mean values of $4.33*10^{-4}$ and $1.089*10^{-3}$, respectively, over 25 runs. Increasing the steepness coefficient further leads to increased MSEs, and a steepness coefficient of $S = 1.00$ was observed to be the optimum value.

Fig. 3 also shows the change of MSEs through epochs (Fig. 3b). The mean square errors for $S = 2.00$ were very high compared to those for $S = 1.00$ and $S = 0.50$, and they also fluctuated during the epochs. On the other hand, MSEs for $S = 1.00$ and $S = 0.50$ steadily decreased during the learning process with better results for $S = 1.00$.

Fig. 3c shows measured versus predicted values of E_U for 162 cyclones with varying steepness coefficients. The

diagonal represents one-to-one line, while the the black line is the regression line between measured and predicted E_U. As obviously seen, the predicted values of E_U were very weakly correlated with the measured values for $S < 0.20$. On the other hand, the correlation improves with $S = 0.20$ and the best correlation was obtained for $S = 1.00$. Although a strong correlation was also obtained for $S = 2.00$, the slope of the regression line was not as close to unity as the one obtained for $S = 1.00$.

Fig. 4 shows results from 25 runs of BPNN with hyperbolic tangent activation function for various steepness coefficients. Mean square errors showed a gradual decreasing trend starting from $S = 0.01$ to $S = 0.20$ (Fig. 4a). The least MSEs were obtained for $S = 0.20$ with median and mean values of $2.02 * 10^{-3}$ and $2.69 * 10^{-3}$, respectively, over 25 runs. From this point

forward, increasing the value of steepness coefficient led to substantial increases in MSEs. Therefore, the optimum value of steepness coefficient was determined to be $S = 0.20$ for hyperbolic tangent function.

Fig. 4b shows the learning pattern of the BPNN with hyperbolic tangent activation function for steepness coefficients of $S = 0.50$, $S = 0.20$, and $S = 0.10$. The learning pattern showed similar trends as those observed for the sigmoid function. For $S = 0.50$, the network's ability of learning the complex relationship between cyclone geometry and Euler number was greeatly reduced and the MSEs showed great fluctuations. On the other hand, MSEs gradually decreased as the network was learning with the best MSEs for the optimum value of steepness coefficient.

Fig. 4 also shows measured versus predicted values of Euler numbers for various steepness coefficients (Fig. 4c).

Increasing the steepness coefficient from 0.01 to 0.02 leads to a sudden improvement of correlation between measured and predicted values. From this point on, the correlation coefficients gradually increased with increasing steepness coefficient and the slope of the regression line gradually approached unity. Although the strongest correlation was obtained for $S = 0.20$, increasing the steepness coefficient further did not result in improved correlations. For very small and very large steepness coefficients, the network did not learn the relationship between cyclone geometry and Euler numbers at all.

Fig. 5 shows MSEs calculated for steepness coefficients between $S = 0.01$ and $S = 2.00$. For $S < 0.10$, the MSEs were larger than 0.1 and the network's learning ability was very weak (Fig. 5a). The network performance suddenly improved starting from $S = 0.10$ and MSEs
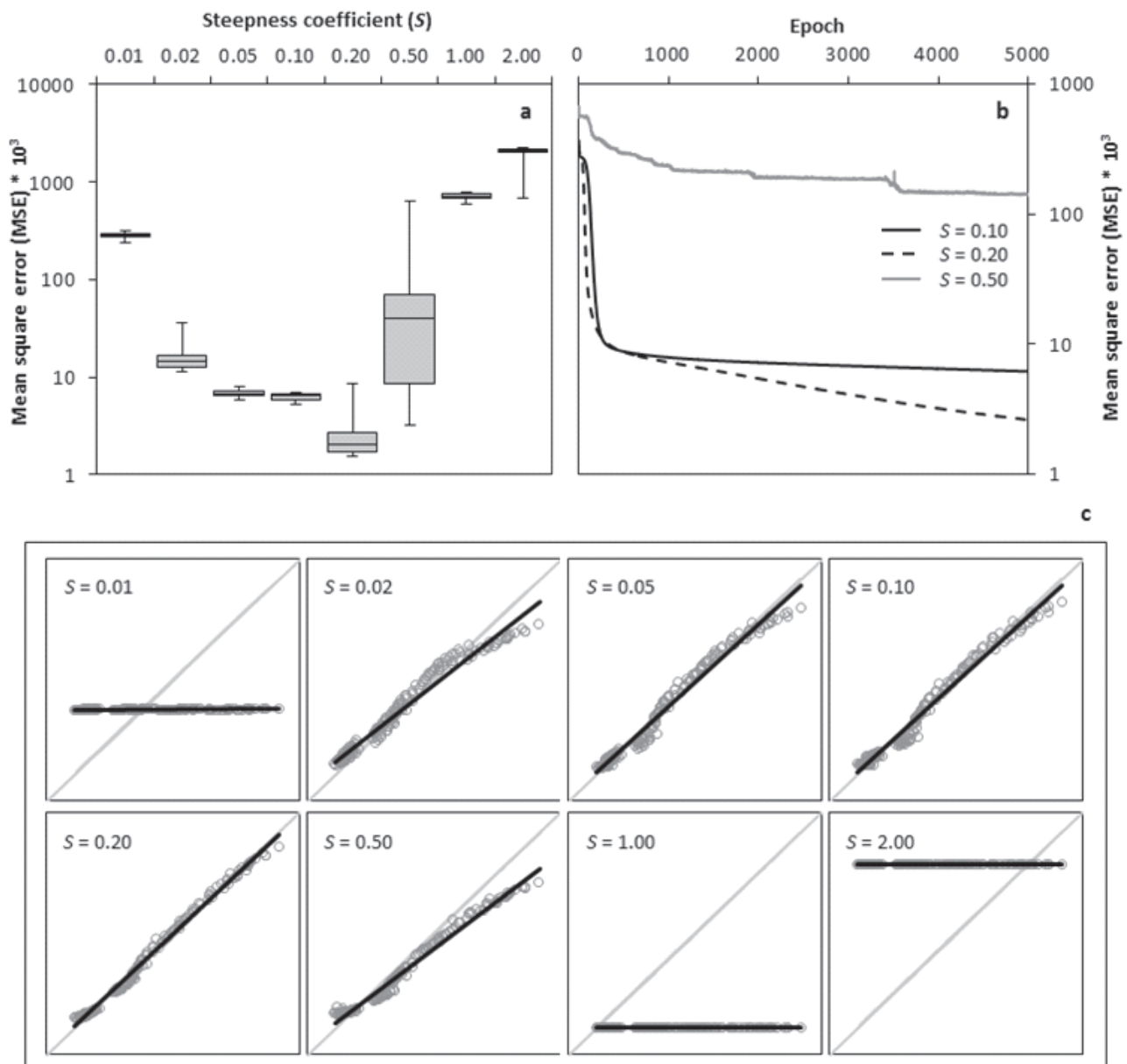

Fig. 4. BPNN results with hyperbolic tangent activation function: a) mean square errors with various steepness coefficients, b) learning trends, c) measured vs. predicted Euler numbers for varying steepness coefficients.

started increasing after $S > 1.00$. The minimum value for MSEs calculated as the meadian value over 25 run was for $S = 1.00$; however, very large MSEs were also calculated for this steepness coefficient. The results were not stable for $S = 1.00$. On the other hand, best MSEs were obtained for $S = 0.50$ with mean and meadian values of $1.21*10^{-3}$ and $1.20*10^{-3}$, respectively, over 25 runs. Therefore, the optimum value of steepness coefficient for sinc function was selected as $S = 0.50$.

Fig. 5b shows the learning patterns of the BPNN network with sinc activation function for $S = 0.20$, $S = 0.50$, and $S = 1.00$. The patterns were very similar to those obtained for hyperbolic tangent and sigmoid functions. Again, the steepness coefficient of $S = 0.50$ provided minimum average MSEs.

The correlation plots between measured and predicted values of Euler numbers for varying steepness coefficients were shown in Fig. 5c. The sinc function did not produce satisfying correlations for steepness coefficients of $S = 0.01$, $S = 0.02$, $S = 0.05$, and $S = 2.00$. Fig. 5c also supports the selection of the optimum steepness coefficient as $S = 0.50$ for sinc activation function since the slope of regression line for $S = 0.50$ was very close to unity with very small values of mean square error.

In Fig. 3 through 5, average values of results from 25 runs of BPNN for each steepness coefficients were given. Thus, the results reflect the mean tendencies of BPNN instead of actual performance. Besides, the figures also includes the results from the training set, which was randomly selected for each run and no differentiation
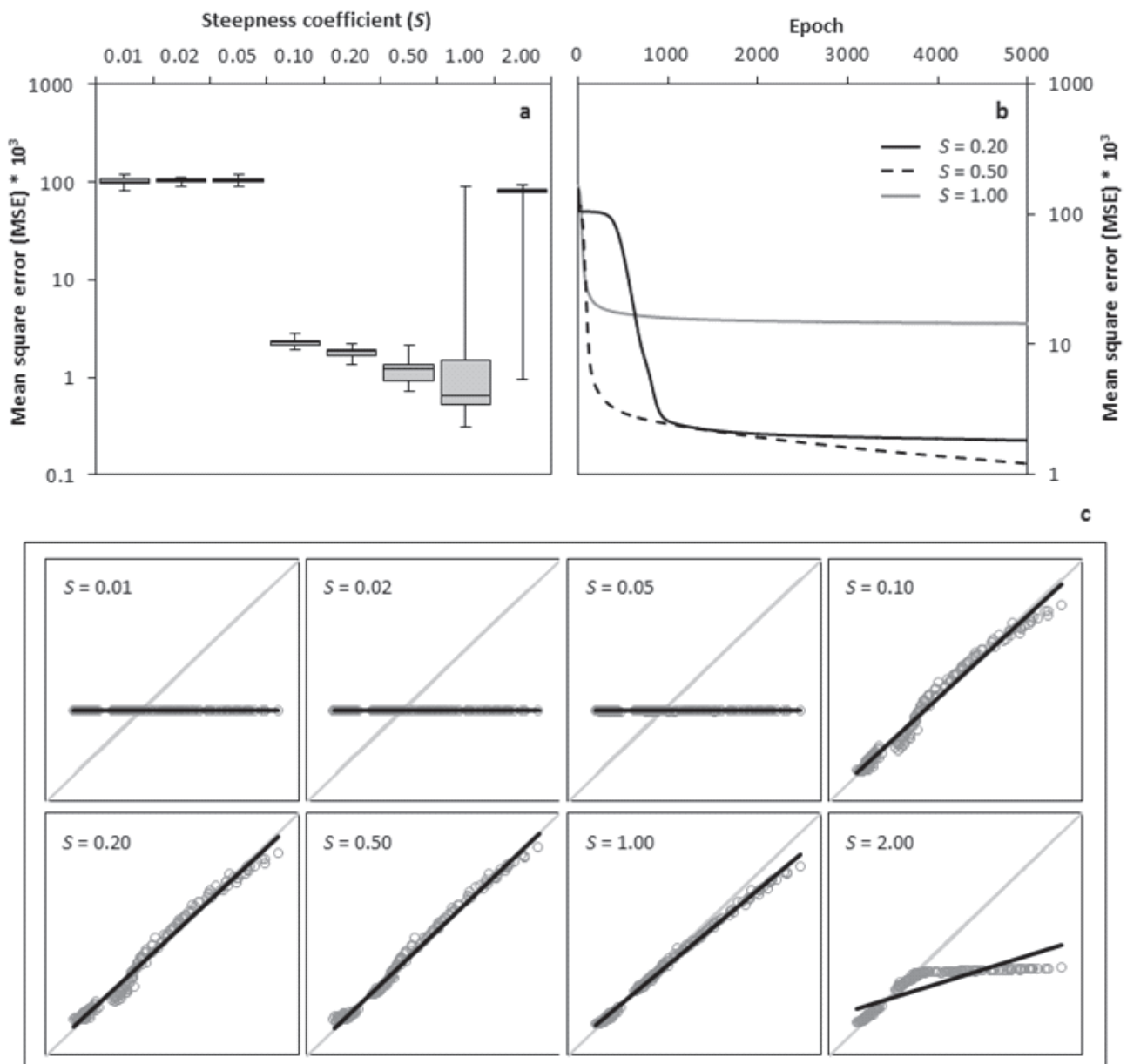

Fig. 5. BPNN results with sinc activation function: a) mean square errors with various steepness coefficients, b) learning trends, c) measured vs. predicted Euler numbers for varying steepness coefficients.

was made between training and validation data sets for the purpose of averaging. In order to see the actual performance, another figure is provided in which best results from each activation function are shown.

Fig. 6 shows the correlation plots between measured and predicted values of Euler numbers by BPNN. For each activation function a randomly selected 70% amount of the data set (113 data points) was used for training the BPNN, and the remaining 30% (49 data points) was used for validating the network. The mean square errors were calculated as $3.32*10^{-4}$ for sigmoid, $1.55*10^{-3}$ for hyperbolic tangent, and $7.20*10^{-4}$ for sinc activation function.

Sigmoid activation function with $S = 1.00$ produced the best correlation between measured and predicted values of Euler numbers with a determination coefficient

of $R^2 = 0.9950$ (Fig. 6a). The slope of the regression line was equal to 0.9900 and the intercept was very close to zero, meaning that most of the data points concentrated around a one-to-one line and that there was a strong linear relationship between measured and predicted values of Euler numbers.

The coefficient of determination calculated for the hyperbolic tangent activation function with $S = 0.20$ was 0.9922 (Fig. 6b). The slope of the regression line for this function was close to unity as well. Various steepness coefficients produced unstable results and learning was very limited for small and large values of steepness coefficients.

Fig. 6c shows the best results from sinc activation function BPNN with $S = 0.50$. The coefficient of determination for sinc function was calculated as
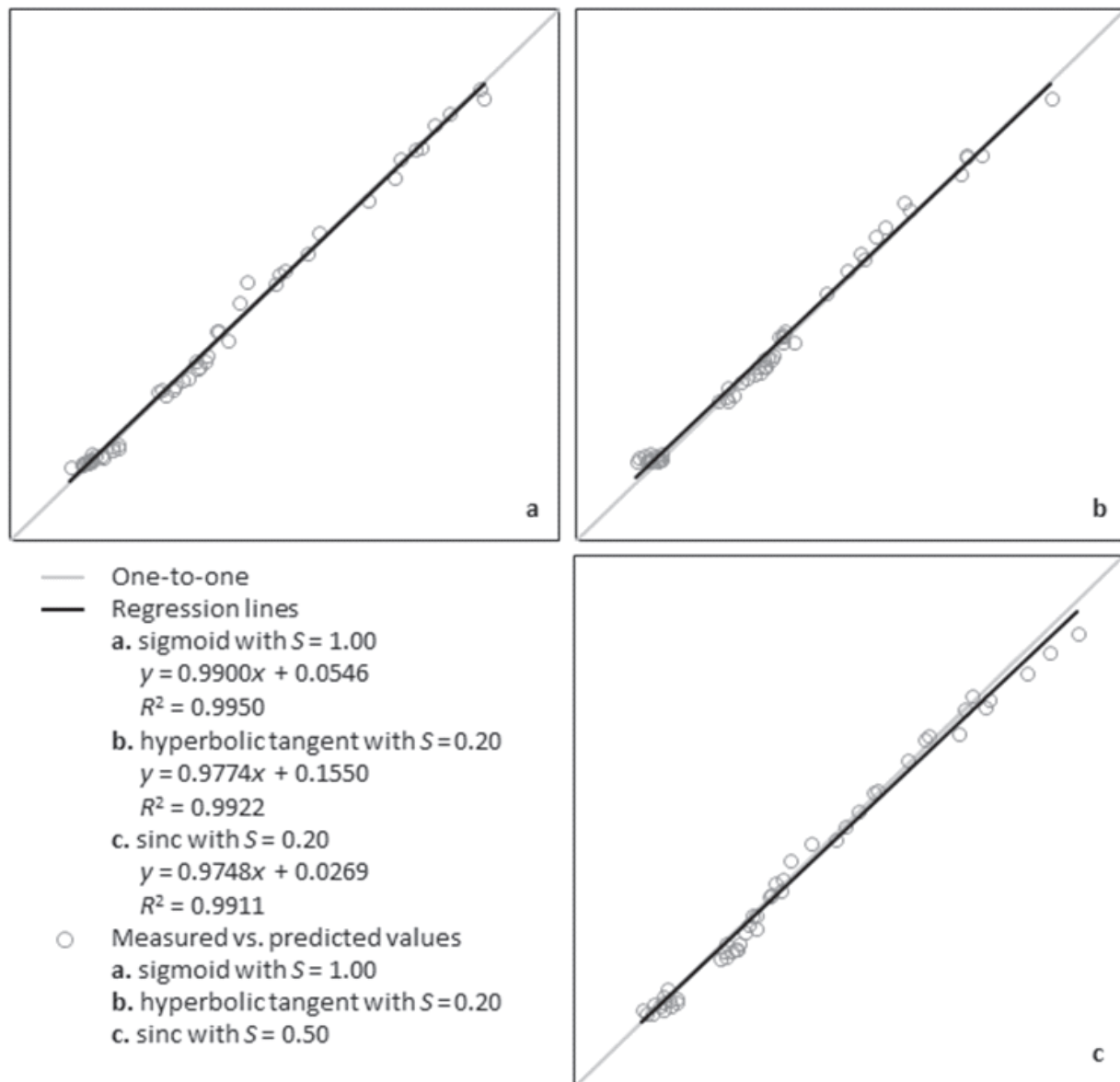

Fig. 6. Correlation plots between measured and predicted values of Euler numbers by a) sigmoid function with S = 1.00, b) hyperbolic tangent function with S = 0.20, c) sinc function with S = 0.50.
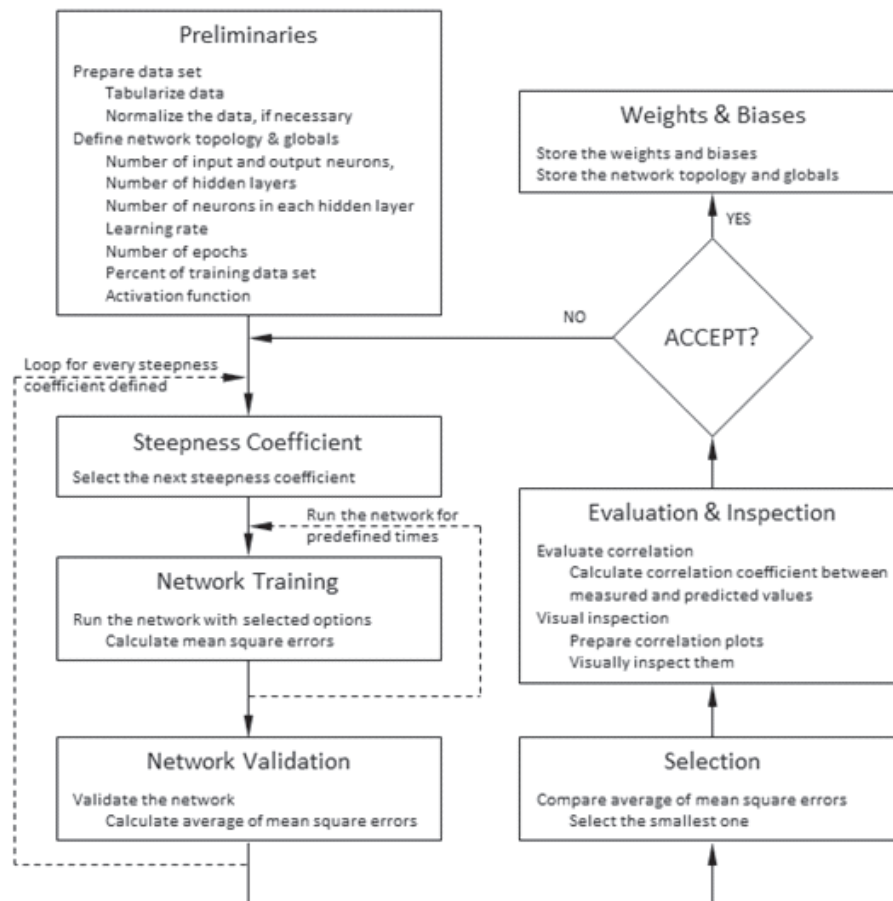
Fig. 7. Protocol for selecting optimum value of steepness coefficient.

R^2 = 0.9911. However, the slope of regression line was somewhat smaller than the sigmoid function and calculated as 0.9748. The results showed that the discrepancies between measured and predicted values of Euler numbers by sinc activation function BPNN increases with increasing measured values. The higher the expected value of Euler number, the higher the predicted error is.

A final issue related to the methodology followed for selecting the optimum value of the steepness coefficient is the development of a generalized protocol on how to implement an algorithm to select an optimum value of steepness coefficient that is considered to be problem-specific. Without an implementation protocol it could be very difficult to repeat the ANN modeling study. On the other hand, the random selection of training data by VBA code implemented for this study complicates the development of a protocol for an automated algorithm. Still, an outer loop may be incorporated into the procedure employed in this study in order to provide an easy-to-implement protocol for selecting the optimum value of steepness coefficient. A flowchart is provided in Fig. 7. Although some of the steps in this protocol may be accomplished automatically by a computer program, visual inspection of the measured versus predicted results are required to ensure that the selected value of steepness coefficient is suitable.

Conclusions

This study focused on the effective use of back propagation neural network (BPNN) for environmental modeling problems. The main aim of the study was to investigate the effects of steepness coefficients (S) on the performance of BPNN with various activation functions, including sigmoid, hyperbolic tangent, and sinc functions. For this purpose a BPNN algorithm was implemented in Excel VBA with built-in activation functions and the option for selecting the steepness coefficient. Cyclone pressure drop data was obtained for 162 cyclones of various geometrical shapes and used as a test case for the implemented software. Inlet height, cyclinder height, cone height, vortex finder diameter, and vortex finder length were used as inputs, while the Euler numbers calculated for each cyclone were used as target outputs. BPNN algorithm was employed for simulation separately with three activation functions and eight steepness coefficients for each. Since the software randomly selects the training data set for each run, the BPNN software was run 25 times for each variable (a total of three functions *8 steepness coefficient *25 = 600 times). Averages of the results from each batch (25 runs) were calculated and compared. For sigmoid function, best results were obtained with $S = 1.00$ and the median of mean square errors (MSEs) were calculated as $4.33*10^{-4}$. The median of MSEs for

hyperbolic tangent function with $S = 0.20$ was calculated as $2.02*10^{-3}$. Finally, for sinc function best results were obtained with $S = 0.50$, resulting in a median value of MSEs equal to $1.20*10^{-3}$.

The following conclusions can be withdrawn from the results of this study:

- Back propagation neural network can handle activation functions of any kind and can be used for environmental modeling problems in which usually highly complicated phenomena are involved.
- The selection of the activation function does not have a strong impact on network performance, provided that the target outputs are normalized in accordance with the activation function's span.
- Steepness coefficient incorporated in the activation function is the main parameter that affects network performance. Thus, optimized values of steepness coefficient must be used. For this study, optimized values of steepness coefficients were 1.00, 0.20, and 0.50 for sigmoid, hyperbolic tangent, and sinc activation functions, respectively.
- The optimum value of steepness coefficient might be problem-specific. Therefore, a preliminary step must be taken to determine the optimum value of steepness coefficient for the respective simulation problem. The point of start for optimization process is intuitive and one should select it based on personal experience.
- Back propagation neural network can be confidently used for estimating the Euler number for a given cyclone geometry.

Acknowledgements

Abbreviations

a = Inlet height of cyclone, mm
b = Inlet width of cyclone, mm
B = Cone-tip diameter of cyclone, mm
BPNN = Back propagation neural network
D = Cyclone's body diameter, mm
D_e = Vortex finder diameter of cyclone, mm
E_U = Euler number (dimensionless)
GRNN = Generalized regression neural network
h_b = Cylinder (body) height of cyclone, mm
h_c = Cone height of cyclone, mm
K_a = Ratio of inlet height to cyclone body diameter (dimensionless)
K_b = Ratio of inlet width to cyclone body diameter(dimensionless)
K_B = Ratio of cone-tip diameter to cyclone body diameter (dimensionless)
K_{De} = Ratio of vortex finder diameter to cyclone body diameter (dimensionless)

K_{hb} = Ratio of cylinder height to cyclone body diameter (dimensionless)
K_{hc} = Ratio of cone height to cyclone body diameter (dimensionless)
K_{SL} = Ratio of vortex finder diameter to cyclone body diameter (dimensionless)
MSE = Mean square error
RBFNN = Radial basis function neural network
S = Steepness coefficient for activation function
S_L = Vortex finder length of cyclone, mm
V_{in} = Inlet velocity to the cyclone, m/s
ΔP = Pressure drop in cyclone, Pa
ρ = Gas density, kg/m.s.

References

1. SIBI P., JONES A.A., SIDDARTH P. Analysis of different activation functions using back propagation neural networks. Journal of Theoretical and Applied Information Technology **47** (3), 1264, **2013**.
2. ELSAYED K., LACOR C. Modeling and pareto optimization of gas cyclone separator performance using RBF type artificial neural networks and genetic algorithms. Powder Technol. **217**, 84, **2012**.
3. AZADI S., KARIMI-JASHNI A. Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars province, Iran. Waste Manage. **48**, 14, **2016**.
4. JAHANDIDEH S., JAHANDIDEH S., BARZEGARI E., ASKARIAN M., MOVAHEDI M.M., HOSSEINI S., JAHANDIDEH M. The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation. Waste Manage. **29**, 2874, **2009**.
5. DAVID C., ARIVAZHAGAN M., IBRAHIM M. Spent wash decolourization using nano-Al2O3/kaolin photocatalyst: Taguchi and ANN approach. Journal of Saudi Chemical Society **19**, 537, **2015**.
6. YE J., CONG X., ZHANG P., ZENG G., HOFFMANN E., WU Y., ZHANG H., FANG W. Operational parameter impact and back propagation artificial neural network modeling for phosphate adsorption onto acid-activated neutralized red mud. J. Mol. Liq. **216**, 35, **2016**.
7. BANERJEE P., SAU S., DAS P., MUKHOPADHAYAY A. Optimization and modelling of synthetic azo dye wastewater treatment using graphene oxide nanplatelets: Characterization toxicity evaluation and optimization using artificial neural network. Ecotox. Environ. Safe. **119**, 47, **2015**.
8. SIMIC V.M., RAJKOVIC K.M., STOJICEVIC S.S., VELICKOVIC D.T., NIKOLIC NC., LAZIC M.L., KARABEGOVIC I.T. Optimization of microwave-assisted extraction of total phenolic compounds from chokeberries by response surface methodology and artificial neural network. Sep. Purif. Technol. **160**, 89, **2016**.
9. ZHAO B., SU Y. Artificial neural network-based modeling of pressure drop cefficient for cyclone separators. Chem. Eng. Res. Des. **88**, 606, **2010**.
10. HUANG M., WAN J., MA Y., LI W., SUN X., WAN Y. A fast predicting neural fuzzy model for on-line estimation of nutrient dynamics in an anoxic/oxic process. Bioresour. Technol. **101**, 1642, **2010**.
11. AGHAV R.M., KUMAR S., MUKHERJEE S.N. Artificial

neural network modeling in competitive adsorption of phenol and resorcinol from water environment using some carbonaceous adsorbents. J Hazard. Mater. **188**, 67, **2011**.

12. BHATTI M.S., KAPOOR D., KALIA R.K., REDDY A.S., THUKRAL A.K. RSM and ANN modeling for electrocoagulation of copper from simulated wastewater: Multi objective optimization using genetic algorithm approach. Desalination **274**, 74, **2011**.

13. BAYRAM A., KANKAL, M. Artificial neural network modeling of dissolved oxygen concentrations in a Turkish Watershed. Pol. J. Environ. Stud. **24** (4), 1507, **2015**.

14. SAMLI R., SIVRI N., SEVGEN S., KIREMITCI V.Z. Applying artificial neural networks for the estimation of chlorophyll-a concentrations along the Istanbul Coast. Pol. J. Environ. Stud. **23** (4), 1281, **2014**.

15. HERNANDEZ-RAMIREZ D.A., HERRERA-LOPEZ E.J., RIVERA A.L., DEL REAL-OLVERA J. Artificial neural network modeling of slaughterhouse wastewater removal of COD and TSS by electrocoagulation. Stud. Fuzziness Soft Comput. **312**, 273, **2014**.

16. WU W., DANDY G.C., MAIER H.R. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development in drinking water quality modelling. Environ. Modell. Softw. **54**, 108, **2014**.

17. ELSAYED K., LACOR C. CFD modeling and multi-objective optimization of cyclone geometry using desirability function, artificial neural networks and genetic algorithms. Appl. Math. Model. **37**, 5680, **2013**.

18. TALEBI H.A., PATEL R.V., KHORASANI K. Control of flexible-link manipulators using neural networks. Lecture Notes in Control and Information Sciences **261**, 79, **2001**.

19. DEMIR S. A practical model for estimating pressure drop in cyclone separators: An experimental study. Powder Technol. **268**, 329, **2014**.

20. KARADENIZ A. Effect of modifications on stairmand high efficiency type cyclone geometry on particle collection efficiency and pressure drop. MSc Thesis, Graduate School of Natural and Applied Sciences, Yıldız Technical University, Istanbul (in Turkish), **2015**.

21. CORTES C., GIL A. Modeling the gas and particle flow inside cyclone separators. Prog. Energ. Combust. **33**, 409, **2007**.

22. CHEN J., SHI M. A universal model to calculate cyclone pressure drop. Powder Technol. **171**, 184, **2007**.