Predicting the longitudinal dispersion coefficient by radial basis function neural network

Abbas Parsaie¹ · Amir Hamzeh Haghiabi¹

Received: 1 September 2015 / Accepted: 1 October 2015 / Published online: 29 October 2015
© Springer International Publishing Switzerland 2015

Abstract Study on the pollution transmission in rivers is the main part of environmental engineering studies. In this paper, the performance of empirical formulas, the radial basis function (RBF) and multilayer perceptron (MLP) neural network for predicting the longitudinal dispersion coefficient (D_L) in rivers were assessed. To this purpose, 150 data set related to the most influence parameters on the D_L was collected. The results of MLP model indicated that the performance of model with error indices such as correlation coefficient ($R^2 = 0.90$) and root mean square error (RMSE = 0.09) had suitable performance for predicting D_L . The result of RBF model with correlation R^2 (0.61) and RMSE (0.82) also has suitable ability to predict the D_L during the MLP model development, found that the MLP model included one hidden layer which ten neurons with log-sigmoid equation as transfer function and pure line equation as transfer function in the output hidden layer. The RBF model contained one hidden layer with 15 neurons.

Keywords River pollution · Soft computing · Multilayer neural network

Introduction

Study on the river water quality is the main part of environmental engineering. To this purpose, governments usually construct the sampling stations along the rivers. Investigation on the mechanism of the pollution transmission in rivers is the main subject in the field of water quality studies. Field study on the river water quality is the safest way to measure the pollution or water quality parameters such as total dissolved solids (TDS), electrical conductivity (EC), free mineral acidity, hardness, total suspended solids, turbidity, biological oxygen demand (BOD), chemical oxygen demand, total organic carbon, coliform group and etc. Research on the pollution transmission most time conducted by tracer routing along the river (Benedini 2013; Wang et al. 2015). To this purpose, a tracer with known concentration releases in river and after fully mixing numbers of samples is considered along the river. Researchers by routing the tracer concentration, the pollution mechanism and refinement of the river have studied. Several field study with regarding to routing the tracer concentration have been conducted on the several rivers. In this regard, study which was conducted by Atkinson and Davis (2000), Davis and Atkinson (2000), Davis et al. (2000) on the mechanism of the pollution transmission in Severn River. They considered 14 km of the Severn River and six sampling stations for their case study. The aim of their study was to investigate the effect of river geometry such as dead zone on the pollution transmission. Palancar (2003) studied on the water quality of two main river in Spanish. They studied the effect of nuclear power leakage on the river water quality. The aim of their study was focused on the defining the dispersion coefficient or in other to characterizing the capacity of self-refinement of these rivers. It can be said that the ultimate

---

✉ Abbas Parsaie
abbas_parsaie@yahoo.com

Amir Hamzeh Haghiabi
ahaghiabi@gmail.com

¹ Department of Water Engineering, Lorestan University,
Khorramabad, Iran

aim of all empirical studies have to determine the dispersion coefficient. Due to high cost of experiment and filed study equipment, researchers have tried to use mathematical approaches specifically soft computing methods for modeling and predicting the DL. In this regard, using the artificial neural network (MLP), genetic programming (GP), support vector machine (SVM), adaptive neuro fuzzy inference system (ANFIS), gene expression programming (GEP), group method of data handling (GMDH) has been reported. Based on the reports, the accuracy of soft computing methods are much more than the empirical formulas (Azamathulla and Ghani 2011; Azamathulla and Wu 2011; Noori et al. 2009, 2011, 2015; Parsaie and Haghiabi 2015a; Parsaie et al. 2015; Najafzadeh and Sattar 2015; Piotrowski et al. 2012; Riahi-Madvar et al. 2009; Toprak and Cigizoglu 2008; Toprak et al. 2004, 2014). In this research, the radial basis function (RBF) neural network which has high performance in pattern recognition and image processing is developed for predicting the DL and its performance is compared with empirical formula and multilayer perceptron neural network as common ANN model which uses by most of the researchers. In the other words, the aim of this study is assessing the performance of RBF model to predict the DL in rivers.

Method and materials

The longitudinal dispersion coefficient is proportional of the properties of fluid, hydraulic condition and the river geometry (cross-sections and path line). All the effective parameters can be written as below function:

$$D_L = f_1(\rho, \mu, u, u_*, h, w, s_f, s_n) \quad (1)$$

Where ρ is fluid density; μ is dynamic viscosity; w is the width of cross-section; h is flow depth; u_* is shear velocity, s_f is longitudinal bed shape and s_n is sinuosity. To derive the dimensionless parameter affective on D_L , the Buckingham theory as dimensional analysis approach was considered and dimensionless parameter will be derived as shown below (Seo and Cheong 1998),

$$\frac{D_L}{hu_*} = f_2\left(\rho \frac{uh}{\mu}, \frac{u}{u_*}, \frac{w}{h}, s_f, s_n\right) \quad (2)$$

The flow in the nature especially in the river is always turbulent. Therefore, the Reynolds number $\rho \frac{uh}{\mu}$ can be ignored and the bed form and sinuosity path parameters cannot be measured clearly, as well. Therefore, the effect of them can be considered as flow resistant, which is seen in the flow depth. The dimensionless parameters that can be clearly measured, are given as below (Seo and Cheong 1998; Seo and Baek 2004).

$$\frac{D_L}{hu_*} = f_2\left(\frac{u}{u_*}, \frac{w}{h}\right) \quad (3)$$

Developing the empirical formulas and AI models are based on these dimensionless parameters. Table 1 presents the most famous empirical formula which have presented for D_L calculation.

Preparing the soft computing models is based on the data set, so to this purpose about 100 data set related to the Eq. (3) was collected and range of them is given in the Table 2 (Etemad-Shahidi and Taghipour 2012).

Table 1 Empirical equations for estimating the longitudinal dispersion coefficient

Row	Author	Equation
1	Elder (1959)	$D_L = 5.93hu_*$
2	McQuivey and Keefer (1974)	$D_L = 0.58\left(\frac{h}{u_*}\right)^2 uw$
3	Fisher (1967)	$D_L = 0.011 \frac{u^2 w^2}{hu_*}$
4	Li et al. (1998)	$D_L = 0.55 \frac{wu_*}{h^2}$
5	Liu (1977)	$D_L = 0.18\left(\frac{u}{u_*}\right)^{0.5} \left(\frac{w}{h}\right)^2 hu_*$
6	Iwasa and Aya (1991)	$D_L = 2\left(\frac{w}{h}\right)^{1.5} hu_*$
7	Seo and Cheong (1998)	$D_L = 5.92\left(\frac{u}{u_*}\right)^{1.43} \left(\frac{w}{h}\right)^{0.62} hu_*$
8	Koussis and Rodriguez-Mirasol (1998)	$D_L = 0.6\left(\frac{w}{h}\right)^2 hu_*$
9	Li et al. (1998)	$D_L = 5.92\left(\frac{u}{u_*}\right)^{1.2} \left(\frac{w}{h}\right)^{1.3} hu_*$
10	Kashefipour and Falconer (2002)	$D_L = 2\left(\frac{u}{u_*}\right)^{0.96} \left(\frac{w}{h}\right)^{1.25} hu_*$
11	Tavakollizadeh and Kashefipour (2007)	$D_L = 7.428 + 1.775\left(\frac{u}{u_*}\right)^{1.752} \left(\frac{w}{h}\right)^{0.62} hu$
12	Rajeev and Dutta (2009)	$D_L = 10.612\left(\frac{u}{u_*}\right) hu$

Table 2 Range of collected data related to the D_L

Data range	W (m)	H (m)	U (m/s)	U* (m/s)	$D_L(m^2/s)$
Min	11.9	0.2	0.0	0.0	1.9
Max	711.2	19.9	1.7	0.6	1486.5
Avg	73.2	1.5	0.5	0.1	115.3
SD	106.9	2.3	0.4	0.1	218.7

Artificial neural networks (ANNs)

ANN is a nonlinear mathematical model that is able to simulate many complexes mathematical that relate the inputs and outputs. Multilayer Perceptron (MLP) networks are common types of ANN that are widely used in the researches. To use MLP model, definition of appropriate functions, weights and bias should be considered. Due to the nature of the problem, different activity functions in neurons can be used. An ANN maybe has one or more hidden layers. Figure 1 demonstrates a three-layer neural network consisting of inputs layer, hidden layer (layers) and outputs layer. As shown in Fig. 1. w_i is the weight and b_i is the bias for each neuron. Weight and biases' values will be assigned progressively and corrected during training process comparing the predicted outputs with known outputs. Such networks are often trained using back propagation algorithm. In the present study, ANN was trained by Levenberg–Marquardt technique because this technique is more powerful and faster than the conventional gradient descent technique (Parsaie and Haghiabi 2015b; Zhang 2010).

Radial basis function (RBF) neural network

Radial basis function (RBF) neural network is a type of artificial neural networks which widely uses in image processing, pattern recognition and nonlinear system modeling. The RBF model as shown in the Fig. 2 consist of two layers, the first layer considered as hidden layer and second layer as output layer. The radial function is considered as transfer function for the neurons which are in the hidden layer and linear function as output layer transfer

Fig. 1 Sketch of three-layer ANN architecture

function. Designing the RBF neural network is based on the defining the center of these functions in other to the aim of RBF model training is mapping the input space to output space as $f: R^n \rightarrow R$. The transfer function of the RBF model is defined as Eq. (4).

$$f(v) = \sum_{i=1}^n w_i \varphi(\|v - c_i\|) \quad (4)$$

Where v is the inputs variable, w_i is the weight coefficients, φ is Gaussian function which is the basic function that uses as kernel function in RBF model development and defines as Eq. (5).

$$\varphi(v) = e^{\left(\frac{-v^2}{2\sigma^2}\right)} \quad (5)$$

The RBF model training usually is carried out by gradient descent approach. The aim of the RBF model is defining the value of the kernel function parameters and weights. Initial value of the weights is defined randomly. The error for each sample of the data set is calculated as Eq. (6).

$$e_i = t_i - y_i = t_i - \sum_{j=1}^N w_j \varphi(\|v_i - c_j\|) \quad (6)$$

The error for the total input data set is calculated as Eq. (7).

$$E = \frac{1}{2} \sum_{i=1}^p |e_i|^2 \quad (7)$$

The RBF model is finished when the error of the RBF model for all the data set is lower than the threshold error which is defined by designer.

Results and discussions

The performance of empirical formulas was assessed for predicting the D_L . To this purpose, three standard error indices such as coefficient of correlation (R^2 -Eq. 8), root mean square error (RMSE-Eq. 9) were used. The results of the error indices calculation for each empirical formula are given in the Table 3.

Fig. 2 A RBF model structure

Table 3 Performance of empirical formulas to calculate the D_L

Equation	R^2	RMSE
Elder (1959)	0.076	230.05
McQuivey and Keefer (1974)	0.004	4,612.990
Fisher (1967)	0.326	1301.03
Li et al. (1998)	0.056	230.38
Liu (1977)	0.326	24,932.62
Iwasa and Aya (1991)	0.238	183.83
Seo and Cheong (1998)	0.383	394.92
Koussis and Rodriguez-Mirasol (1998)	0.200	312.85
Li et al. (1998)	0.351	4404.65
Kashefipour and Falconer (2002)	0.343	525.30
Tavakollizadeh and Kashefipour (2007)	0.451	5861.86
Rajeev and Dutta (2009)	0.395	262.20

$$R^2 = \left(\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right)^2 \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (9)$$

Table 4 The performance and summary of the MLP model during the development stage

Row	N-H-L	F-HL&TF	S-HL&TF	R^{2*}	MSE*	RMSE*	R^{2**}	MSE**	RMSE**
1	1	5-tansig	–	0.76	0.086	0.02	0.77	0.04	0.015
2	1	8-tansig	–	0.83	0.084	0.03	0.75	0.03	0.018
3	1	5-logsig	–	0.86	0.04	0.016	0.85	0.05	0.012
4	1	8-logsig	–	0.91	0.046	0.024	0.92	0.052	0.016
5	1	10-logsig	–	0.98	0.022	0.015	0.95	0.053	0.01
6	2	9-logsig	5-tansig	0.93	0.053	0.028	0.86	0.060	0.012
7	2	12-tansig	5-tansig	0.95	0.063	0.035	0.95	0.068	0.015

N-H-L number of hidden layer(s), *F-HL&TF* first hidden layer and transfer function, *S-HL&TF* second hidden layer and transfer function

* Error indices of MLP during the training stage

** Error indices of MLP during the testing stage

As seems from the Table 3, the Tavakollizadeh and Kashefipour (2007) is accurate among the empirical formulas. Developing the MLP model is included choosing the number of hidden layer, choosing the data set for training and defining the portfolio of training and testing groups from the total data set, number of the neuron in the each hidden layer(s), choosing the transfer function for the neurons and at the end choosing the learning algorithm. To provide the MLP model for predicting the D_L , the Matlab software utilities was used. Table 4 gives a summary of the process of the MLP model development. Table 4 presents a summary of the try and error process for preparing the MLP model. To develop the MLP model 70 % of the total data set was considered for training and others (30 %) for validation and testing. To avoid the overlearning, during the model development, the 15 % of the data set was considered for validation. As seems from the Table 4, the MLP model which has ten neurons with Log-sigmoid transfer function(logsig) as transfer function in the first layer is accurate and adding the number hidden layer and number of the neurons in the layer has not significant effect on increasing the model performance. The architect of developed MLP model is shown in the Fig. 3.

The performance of MLP model during the stages of model preparation (training, validation and testing) is shown in the Figs. 4, 5 and 6). In Figs. 4, 5 and 6 the results of the MLP model was plotted together and versus the measured data. In addition, to understand performance of the MLP during the model preparation stages, the error distribution also was plotted. Moreover, to define the error concentration, the histogram of error was plotted. As seems the histogram of error, density of error aggregated around the zero. In overall the performance of MLP model is so suitable especially in compare to empirical formulas.

To assess the performance of the RBF model for prediction of the D_c was developed 80 % of the data set was assigned for the RBF model training and others was

Fig. 3 The structure of developed MLP model

Fig. 4 Performance of MLP model during the training stage

Fig. 5 Performance of MLP model during the validation stage

Fig. 6 Performance of MLP model during the testing stage

Table 5 The performance and summary of the RBF model during the development stage

Row	N-H-L	R^2	MSE	RMSE	R^2	MSE	RMSE
1	5	0.43	0.005	0.07	0.50	0.008	0.09
2	8	0.54	0.007	0.08	0.79	0.003	0.05
3	12	0.56	0.006	0.07	0.50	0.002	0.04
4	15	0.61	0.005	0.07	0.81	0.015	0.12

considered as testing. Preparation of RBF model is a try and error process. Table 5 shows a summary of RBF model development process. As seems in the Table 5,

the RBF model number 4 is accurate in compare to the others. During the RBF model development found that increasing the number of neurons hasn't significant effect on the model precision. The architect of the RBF model is shown in the Fig. 7. The performance of the RBF model during the training, validation and testing model is shown in the Figs. 8 and 9. As shown from these figures, the performance of the RBF model to predict the D_c is not so suitable. Comparison the performance of the RBF model with MLP model in training, validation and testing stages shows that the MLP model is more accurate.

Fig. 7 The structure of developed RBF model

Fig. 8 Performance of RBF model during the training stage

Fig. 9 Performance of RBF model during the testing stage

Conclusion

In this study the longitudinal dispersion coefficient (D_c) was calculated and predicted by empirical formulas, multilayer perceptron (MLP) and radial basis function (RBF)

neural network. The results of this study indicated that the Tavakollizadeh and Kashefipour (2007) was accurate among the empirical formulas. To achieve more accuracy in D_c prediction, the MLP model and RBF model was developed, to prepare the MLP and RBF models the about 150 data set

related the D_c was collected. The result of the assessing the performance of the MLP shows that the MLP model has suitable performance to predict the D_c . The results of the RBF model development indicated the accuracy of this model is a little better than the empirical formals. In overall the performance of the MLP model is much more than the RBF and empirical formulas.

References

- Atkinson T, Davis P (2000) Longitudinal dispersion in natural channels: 1. Experimental results from the River Severn, UK. *Hydrol Earth Syst Sci Discuss* 4(3):345–353
- Azamathulla H, Ghani A (2011) Genetic programming for predicting longitudinal dispersion coefficients in streams. *Water Resour Manage* 25(6):1537–1544
- Azamathulla HM, Wu F-C (2011) Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Appl Soft Comput* 11(2):2902–2905
- Benedini M, Tsakiris G (2013) *Water Quality Modelling for Rivers and Streams*, Springer
- Davis P, Atkinson T (2000) Longitudinal dispersion in natural channels: 3. An aggregated dead zone model applied to the River Severn, UK. *Hydrol Earth Syst Sci* 4(3):373–381
- Davis P, Atkinson T, Wigley T (2000) Longitudinal dispersion in natural channels: 2. The roles of shear flow dispersion and dead zones in the River Severn, UK. *Hydrol Earth Syst Sci Discuss* 4(3):355–371
- Elder JW (1959) The dispersion of a marked fluid in turbulent shear flow. *J Fluid Mech* 5(4):544–560
- Etemad-Shahidi A, Taghipour M (2012) Predicting longitudinal dispersion coefficient in natural streams using M5' model tree. *J Hydraul Eng* 138(6):542–554
- Fisher BH (1967) The mechanics of dispersion in natural streams. *J Hydraul Div* 93(6):187–216
- Iwasa Y, Aya S (1991) Predicting longitudinal dispersion coefficient in open channel flows. In: *Proceedings of international symposium on environmental hydraulics*, Hong Kong, pp 505–510
- Kashefipour SM, Falconer A (2002) Longitudinal dispersion coefficients in natural channels. In: *Proceedings of the fifth international hydro informatics conference*, 1–5 July. Cardiff University, pp 95–102
- Koussis AD, Rodriguez-Mirasol J (1998) Hydraulic estimation of dispersion coefficient for streams. *J Hydraul Eng* 124:317–320
- Li ZH, Huang J, Li J (1998) Preliminary study on longitudinal dispersion coefficient for the gorges reservoir. In: *Proceedings of the seventh international symposium environmental hydraulics*, 16–18 December, Hong Kong, China
- Liu H (1977) Predicting dispersion coefficient of stream. *J Environ Eng Div* 103(1):59–69
- McQuivey RS, Keefer TN (1974) Simple method for predicting dispersion in streams. *J Environ Eng Div* 100(4):997–1011
- Najafzadeh M, Sattar AA (2015) neuro-fuzzy GMDH approach to predict longitudinal dispersion in water networks. *Water Resour Manage* 29(7):2205–2219
- Noori R, Karbassi A, Farokhnia A, Dehghani M (2009) Predicting the longitudinal dispersion coefficient using support vector machine and adaptive neuro-fuzzy inference system techniques. *Environ Eng Sci* 26(10):1503–1510
- Noori R, Karbassi AR, Mehdizadeh H, Vesali-Naseh M, Sabahi MS (2011) A framework development for predicting the longitudinal dispersion coefficient in natural streams using an artificial neural network. *Environ Prog Sustain Energy* 30(3):439–449
- Noori R, Deng Z, Kiaghadi A, Kachooosangi FT (2015) How reliable are ANN, ANFIS, and SVM techniques for predicting longitudinal dispersion coefficient in natural rivers? *J Hydraul Eng* 0(0):04015039
- Palancar MC, Aragón JM, Sánchez F, Gil R (2003) The determination of longitudinal dispersion coefficients in rivers. *Water Environ Res* 75(4):324–335
- Parsaie A, Haghiabi A (2015a) The effect of predicting discharge coefficient by neural network on increasing the numerical modeling accuracy of flow over side weir. *Water Resour Manage* 29(4):973–985
- Parsaie A, Haghiabi A (2015b) Computational modeling of pollution transmission in rivers. *Appl Water Sci* 1–10
- Parsaie A, Yonesi H, Najafian S (2015) Predictive modeling of discharge in compound open channel by support vector machine technique. *Model. Earth Syst. Environ.* 1(2):1–6
- Piotrowski AP, Rowinski PM, Napiorkowski JJ (2012) Comparison of evolutionary computation techniques for noise injected neural network training to estimate longitudinal dispersion coefficients in rivers. *Expert Syst Appl* 39(1):1354–1361
- Rajeev RS, Dutta S (2009) Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrol Res* 40(6):544–552
- Riahi-Madvar H, Ayyoubzadeh SA, Khadangi E, Ebadzadeh MM (2009) An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS. *Expert Syst Appl* 36(4):8589–8596
- Seo I, Baek K (2004) Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams. *J Hydraul Eng* 130(3):227–236
- Seo I, Cheong T (1998) Predicting longitudinal dispersion coefficient in natural streams. *J Hydraul Eng* 124(1):25–32
- Tavakollizadeh A, Kashefipour SM (2007) Effects of dispersion coefficient on quality modeling of surface waters. In: *Proceedings of the sixth international symposium river engineering*, Ahwaz, Iran, 16–18 October 2007, pp 67–78
- Toprak ZF, Cigizoglu HK (2008) Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrol Process* 22(20):4106–4129
- Toprak ZF, Şen Z, Savci ME (2004) Comment on “Longitudinal dispersion coefficients in natural channels”. *Water Res* 38(13):3139–3143
- Toprak ZF, Hamidi N, Kisi O, Gerger R (2014) Modeling dimensionless longitudinal dispersion coefficient in natural streams using artificial intelligence methods. *KSCE J Civ Eng* 18(2):718–730
- Wang Z-Y, Lee JW, Melching C (2015) *Water quality management. river dynamics and integrated river management*, edited. Springer, Heidelberg, pp 555–631
- Zhang W (2010) *Computational ecology: artificial neural networks and their applications*. World Scientific Publishing Company Pte Limited