Journal of Environmental Treatment Techniques **2014, Volume 2, Issue 4, Pages: 176-183**

Environ. Treat. Tech.
ISSN: 2309-1185
Journal web link: http://www.jett.dormaj.com

Evaluation of Selected Formulas and Neural Network Model for Predicting the Longitudinal Dispersion Coefficient in River

Abbas Parsaie [1]*, Amir Hamzeh Haghiabi [2]

1- Ph.D. student of hydro structures, Department of water Engineering, Lorestan University, KhoramAbad, Iran.
2- Associate professor of water Engineering, Lorestan University, KhoramAbad, Iran.

Received: 07/09/2014 Accepted: 22/10/2014 Published: 30/10/2014

Abstract

Longitudinal dispersion coefficient (LDC) is one of the most important parameters in the river water quality management. Several ways as empirical formulas and artificial intelligent techniques are proposed for predicting the LDC and it is necessary to evaluate the performance of them. In this study, a Multilayer perceptron neural network (MLP) model has been developed and 12 formulas empirical formulas were collected. To assess the performance of these formulas and MLP model in a case study problem, calculating the LDC by dispersion routing method for Severn River in UK was considered. Results shows that the best accuracy is related to the Tavakollizadeh and Kashefipour formula ($R^2 \approx 0.45$) based on data set and for Severn River, its accuracy is $R^2 \approx 0.4$. the (MLP) model has acceptable accuracy ($R^2 \approx 0.83$) to predict the LDC in Severn River.

Key words: Longitudinal dispersion coefficient; empirical formula; Multilayer Perceptron (MLP) neural network; dispersion routing method.

1 Introduction

Rivers are one of the most vulnerable environments so study on the water quality of this source of water is extremely important. Recently, river pollution has become one of the most important problems in the environment knowledge, [1, 2]. The mechanism of Pollutant transmission in river is a more complex phenomenon so the major part of the Knowledge of Environmental Engineering is related to this issue. Study on the mechanism of pollution transmission in rivers with various condition leads to improve the public human's health. [3].

When a source of pollution released into the river, Due to molecular motion, turbulence and non-uniform velocity in flow cross section, quickly spreads at the section and moved along the river with the flow [1, 4].

The governing equation of the pollution transmission in river is Advection Dispersion Equation (ADE). This equation is a partial differential equation and named convection equation in general. The ADE has many applications in simulating the hydraulic phenomenon's in water and environmental engineering Such as simulation of sediment transport and pollution transmission in rivers and groundwater pollution modeling [5-7].

Computer modeling the pollution transmission in river included two parts. One- Selecting the Numerical methods with reasonable accuracy (Fortunately, recently techniques that have been provided are appropriate). Two- estimating the Longitudinal Dispersion coefficient, the main disadvantage of numerical modeling is related to the infirmity of the LDC prediction. Increasing the accuracy of estimating this parameter leads to increase the computer modeling. So the major part of studies on the river water quality is related to measurement, calculating and estimating of this parameter [8-10]

LDC is a function of the hydraulic and geometry characteristics of river and also its path, so it is very variable and susceptible. Thus, with minimal changes in the hydraulic conditions and river geometries, the value fluctuations of this parameter will be very much. Many empirical formulas have been proposed for calculating the LDC. These formulas are usually obtained by classical regression and usually obtained by study investigators on the one or more rivers. Some of these formulas are collected in Table (1) [11-17]

To calculate the LDC usually a tracer injects in the river and some stations along the river for sampling and measuring the tracer concentration in river water should be considered. To this purpose, the first station should be considered after the completely spreading the tracer in all the flow cross section and recording the concentration of tracer, the LDC will be calculate by a method such as Dispersion Routing Method (DRM). A brief instruction of DRM is given in the methodology section. Some researcher uses of the more powerful tools such as image processing, Acoustic Doppler Current Profiler (ADCP) and so on technique to obtain more details of dispersing process in the cross section of river [18]. These tools helped to the researcher for obtaining the effect of river geometric such

---

Corresponding author: Abbas Parsaie, Ph.D. student of hydro structures, Department of water Engineering, Lorestan University, KhoramAbad. Iran.Abbas_Parsaie@yahoo.com, phone: +989163685867

as meandering and dead zone area on the dispersion process in the river. Studies are based on the measuring the concentrations of tracer are suitable references to evaluate the performance of empirical formulas. Due to the high cost of laboratory and field studies and also based on some researches that reported the lack of empirical formulas, researchers have turned to the development and applying the artificial intelligence techniques for predicting the LDC [3].

In the artificial intelligence techniques usually provide a network instead of the classical regression. The accuracy of the AI models, based on the research conducted, are much more from empirical formula. In the field of AI model using the Multilayer Neural network (MLP), Adaptive Neuro Fuzzy Inference System (ANFIS), M5 algorithm, Support Vector Machine (SVM) and Genetic programing(GP) can be mentioned [3, 19-23].

In this paper after assessing the accuracy of empirical formulas by data set which range of them given in table (2), the LDC in the Severn River was calculate by DRM, empirical formulas and MLP model.

2 Methodology

As shown in the figure (1) when a pollutant releases in the river, it rapidly disperses in all the flow cross section and transports to the downstream by the river flow. After covering the pollution all the flow cross section, the transmission of pollutant along the river is one dimensional. The governing equation of pollutant transmission in river is introduced to show the important role of LDC in computer modeling of pollution transmission in river. To extract the governing equation of pollution transmission in river, it is enough to consider an element of river and by using the continuity equation for balancing the inputs and outputs of the pollution discharge in the element of the fluid and by aid of Fick lows, the one dimensional ADE (eq.1) will be extracted [4]. in this equation, C (ppb) is concentration, u(m/s) is the mean flow velocity and x(m) is the distance from pollution injection and D_L (m^2/s) is the LDC. To calculate the dispersion coefficient, several ways as empirical formulas and artificial intelligent techniques has been proposed and developed. All of them are based on dimensionless parameters that extracted by using the Buckingham theory on influence parameters on LDC and will be explained in the next section. in this study some selected empirical formulas and AI model has been assessed to calculate or predict the LDC and to increase the accuracy of LDC prediction in a case study, a novel approach has been proposed and it is related to the calculating the Longitudinal dispersion coefficient by powerful method such as neural network models for Severn River. In this paper, the MLP model was considered because it's suitable accuracy, automatic developing without any human making in parameters setting and easily in development. About 100 data related to these parameter was collected and the range of them given in the table (2)

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}=D_L\frac{\partial^2 C}{\partial x^2} \qquad (1)$$

Fig 1: shamanic figure of pollutant transmission in rivers

2.1 Longitudinal Dispersion coefficient

The longitudinal dispersion coefficient is related to the properties of Fluid, Hydraulic condition and the river geometry (Cross sections and path line). For fluid properties the density and dynamic viscosity and for hydraulic condition, the velocity, flow depth share velocity and energy gradient slope and for river geometry the width of cross section and longitudinal slope can be mentioned. Several other parameters that are influences on the LDC, but cannot be clearly measurement such as sinuosity path (S_f)and bed form of river(S_n). All the influences parameter can be written as below function:

$$D_L=f_1(\rho,\mu,u,u_*,h,w,s_f,s_n) \qquad (2)$$

where ρ is fluid density; μ is dynamic viscosity; w is the width of cross section; h is flow depth; u_* is share velocity, S_f is longitudinal bed shape and S_n is sinuosity.to extract the dimensionless parameter on the LDC, the Buckingham theory was considered and dimensionless parameter will be extracted as below[12].

Always the flow in the nature spatially in the river is turbulent so the Reynolds number ρ×(uh/μ) can be ignored and the bed form and sinusitis path parameters also cannot be measuring clearly so the effect of them can be considered as flow resistant and seems in the flow depth. The dimensionless parameters that can be clearly measurement give as below[11, 12].

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

These dimensionless parameters are the base for the each empirical equations and development of the AI models also are based on these parameters. For developing the AI models such as MLP, ANFIS SVM the data that are related to these parameters must be considered.

Table 1: Empirical equations for estimating the longitudinal dispersion coefficient

Author	Equation
Elder (1959)	$D_L = 5.93hu_*$
McQuivey and Keefer (1974)	$D_L = 0.58\left(\frac{h}{u_*}\right)^2 uw$
Fisher (1967)	$D_L = 0.011\frac{u^2w^2}{hu_*}$
Li et al. (1998)	$D_L = 0.55\frac{wu_*}{h^2}$
Liu (1977)	$D_L = 0.18\left(\frac{u}{u_*}\right)^{0.5}\left(\frac{w}{h}\right)^2 hu_*$
Iwasa and Aya (1991)	$D_L = 2\left(\frac{w}{h}\right)^{1.5} hu_*$
Seo and Cheong (1998)	$D_L = 5.92\left(\frac{u}{u_*}\right)^{1.43}\left(\frac{w}{h}\right)^{0.62} hu_*$
Koussis and Rodriguez-Mirasol (1998)	$D_L = 0.6\left(\frac{w}{h}\right)^2 hu_*$
Li et al. (1998)	$D_L = 5.92\left(\frac{u}{u_*}\right)^{1.2}\left(\frac{w}{h}\right)^{1.3} hu_*$
Kashefipur and Falconer (2001)	$D_L = 2\left(\frac{u}{u_*}\right)^{0.96}\left(\frac{w}{h}\right)^{1.25} hu_*$
Tavakollizadeh and Kashefipour (2007)	$D_L = 7.428 + 1.775\left(\frac{u}{u_*}\right)^{1.752}\left(\frac{w}{h}\right)^{0.62} hu$
Rajeev and Dutta (2009)	$D_L = 10.612\left(\frac{u}{u_*}\right) hu$

2.2 Artificial Neural Network (ANN)

ANN is a nonlinear mathematical model that is able to simulate arbitrarily complex nonlinear processes that relate the inputs and outputs of any system. In many complex mathematical problems that lead to solve complex nonlinear equations, Multilayer Perceptron 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 4 demonstrates a three-layer neural network consisting of inputs layer, hidden layer (layers) and outputs layer. As shown in Fig. 4. 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 [24, 25].

Table 2: Range of LDC collected data

	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
Stdev	106.9	2.3	0.4	0.1	218.7

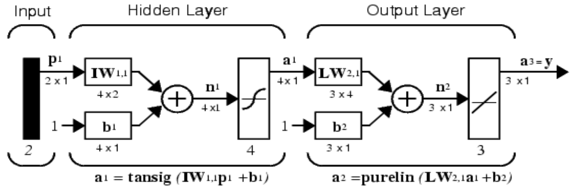

Fig. 2 A three-layer ANN architecture

2.3 Dispersion Routing Method

As mentioned in the past session, calculating the LDC is more important, so firstly, the longitudinal dispersion was calculated from the concentration profile by Dispersion Routing Method (DRM). The basic formula used in the dispersion routing method given in the equation (5). Using the DRM included the four stage. 1-considering of initial value for LDC. 2- Calculating the concentration profile at the downstream station by using the upstream concentration profile and LDC. 3- Done a comparison between the calculating profile and measurement profile. 4- If the calculating profile doesn't a suitable cover the measurement profile the process will be repeated until the calculating profile has a good covering on the measurement profile. In the equation (5), $\bar{t}$ is the average time of measured concentration and index one and two is considered for upstream and downstream sampling stations [26].

$$C(x_2,t) = \int_{-\infty}^{+\infty} uC(x_1,t)\frac{\exp\left\{\frac{-\left[\bar{t_2}-\bar{t_1}-t+\tau\right]^2}{4D_L\left(\bar{t_2}-\bar{t_1}\right)}\right\}}{\sqrt{4\pi D_L\left(\bar{t_2}-\bar{t_1}\right)}}d\tau \tag{4}$$

$$\bar{t} = \frac{\int_0^\infty tC(x,t)}{\int_0^\infty C(x,t)} \tag{5}$$

2.4 Case study

To assess the performance of empirical formulas in differ river, the field study that did by Atkinson and Davis [27] on pollutant transmission mechanism in Severn River has been considered. They selected about 14 km length of the river to study and showing the effect of some hydraulic and geometry condition of river such as bed form, revers and dead zone on the pollution concentration profile. They used of RhodamineWT 20% as tracer and considered six stations after the place of injection to take the samples from the water of river. Figure (3) shows a schematic map of the river and location of the sampling stations. Universal Coordinates of sampling along the river in each station was given in the table (3) [28, 29].figure (4) shows the values of measurement concentrations at sampling station.

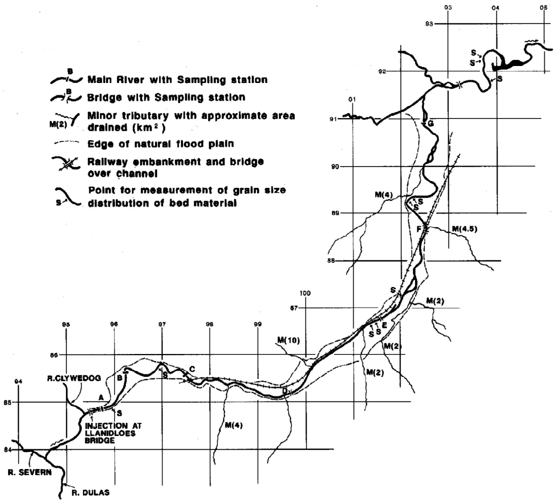

Fig. 3: Schematic map of the Severn River and sampling stations

3 Results and Discussion

Initially assessing the empirical formulas was done with data set that the range of them given in the table (2) and indices error(eq.6) for result of them for calculating the LDC was determined and gives in table(4). As given in the table (4), Tavakollizadeh and Kashefipour (2007) formula has the best accuracy through the empirical formulas and the performance of other formulas is not suitable. Then to assessing the performance of all empirical formulas, the LDC in the Severn River has been calculated by DRM. All the stage of calculating the LDC for Severn River shows in the figures 4 to 9 and gives in the table (5). Again the LDC was calculating by empirical formulas for Severn River and result of them given in the table (6).

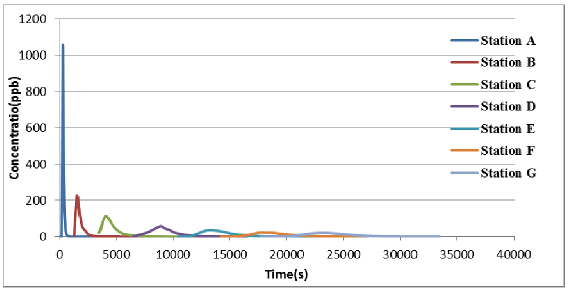

Fig. 4: concentrations value measurement at Severn River sampling stations

As given in the table (6), Tavakollizadeh and Kashefipour (2007) has the best accuracy for calculating the LDC for Severn River through the empirical formulas. After calculating the LDC from the dispersion routing method for each station, again, the LDC was calculated by MLP model and the results of this model also given in the table (6).

Table 3: Universal coordinate of Sampling Stations

Site	UK (Grid reference)	Distance(m)
Injection	SN 9549 8479	0
A	SN 9570 8488	210
B	SN 9621 8561	1175
C	SN 9748 8558	2875
D	SN 9969 8518	5275
E	SO 0160 8677	7775
F	SO 0252 8858	10275
G	SO 0220 9090	13775

With a review of the table (6) it is clear Almost of the empirical formulas has not suitable ability to calculate the LDC. To reach more accuracy in calculating the LDC; the MLP model was developed. As shown in the figure (10), The MLP model contains two hidden layers; the hidden layer contains eighteen (18) neurons in the first hidden layer and five (5) in the second hidden layer and transfer functions were tangent sigmoid (tansig). The training of the MLP model was performed with levenberg_marquat technique. 70 % of data is used for training, 15 % for validation and 15 % for testing the model. The performance of MLP model in each stage of development (training, validation and testing) is shown in the Figures 11, 12 and 13. As shown in the figures 11 to 13, the accuracy of the MLP model is more suitable than the empirical equations. The dimensionless parameter that extracted in the dimensional analysis stage was considered as input parameters to the MLP model. To predict the LDC for Severn River, its dimensionless was extract for each sampling station and gives to the MLP model as inputs parameters and the LDC was predicted. The result of MLP model gives in the table (6). The MLP model has acceptable accuracy to predict the LDC. In final conclusion it seems to have a powerful tool to predict the LDC in rivers it is good to prepare free codes or commercial software's that they are based on AI model with a suitable GUI.

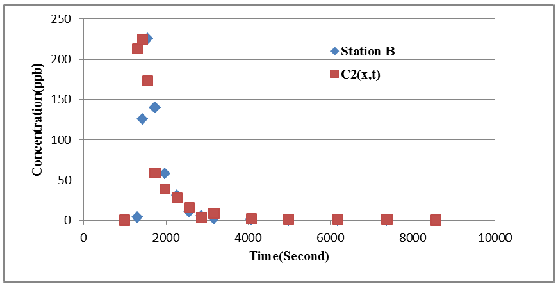

Fig 4: Station B:Comparisons of measured and calculated concentration profiles

Table (4): Error indexes for Empirical Equation

Author Equation	R^2	RMSE
Elder (1959)	0.1	245.6
McQuivey and Keefer (1974)	0	5485096.39
Fisher (1967)	0.29	1470.9
Li et al. (1998)	0.07	245.7
Liu (1977)	0.22	353.46
Iwasa and Aya (1991)	0.21	198.78
Seo and Cheong (1998)	0.38	378.94
Koussis and Rodriguez-Mirasol (1998)	0.17	335.96
Li et al. (1998)	0.33	186.17
Kashefipur and Falconer (2001)	0.32	230.94
Tavakollizadeh and Kashefipour (2007)	0.45	3388.41
Rajeev and Dutta (2009)	0.36	314.03

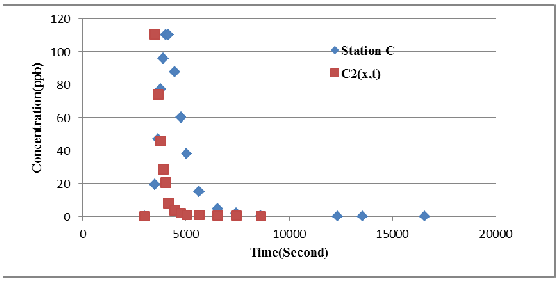

Fig 5: Station C: Comparisons of measured and calculated concentration profiles

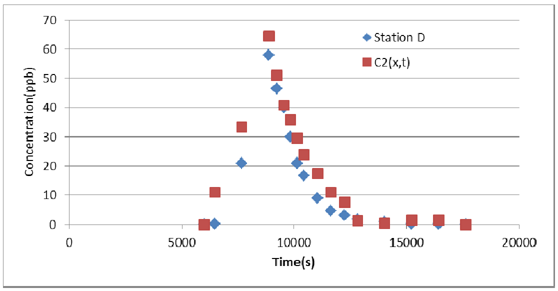

Fig 6: Station D:Comparisons of measured and calculated concentration profiles

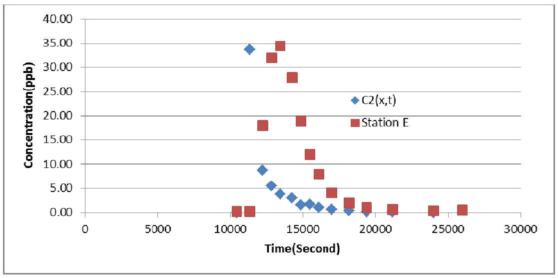

Fig 7: Station E:Comparisons of measured and calculated concentration profiles

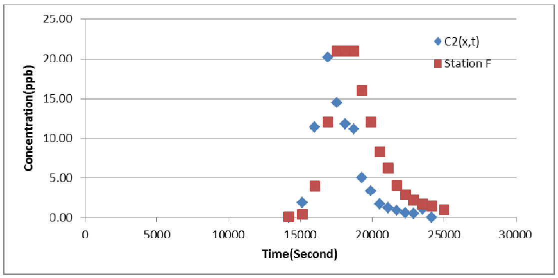

Fig 8: Station F:Comparisons of measured and calculated concentration profiles

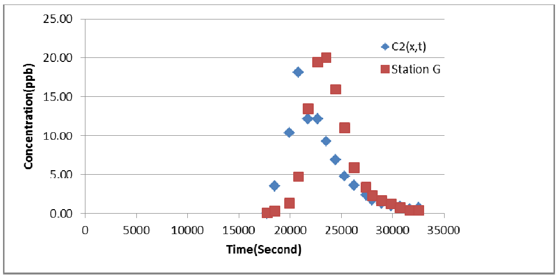

Fig 9: Station G:Comparisons of measured and calculated concentration profiles

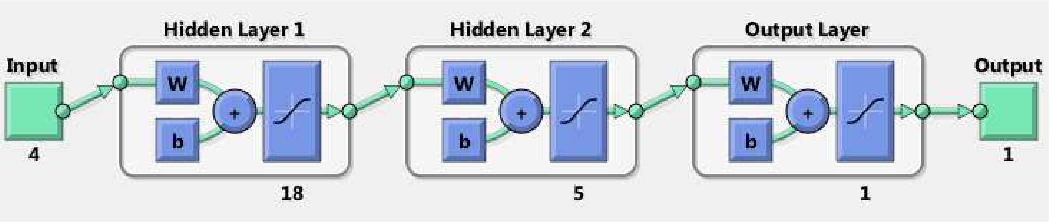

Fig. 10: structure of MLP model

Table 5: Result of Dispersion routing Method for Severn River

River	Station	Δx	Δt	D_L
Severn River	A	4	2	41.5
	B	4	2	26.5
	C	4	2	12.5
	D	4	2	26.5
	E	4	2	37.5
	F	4	2	29.5
	G	4	2	

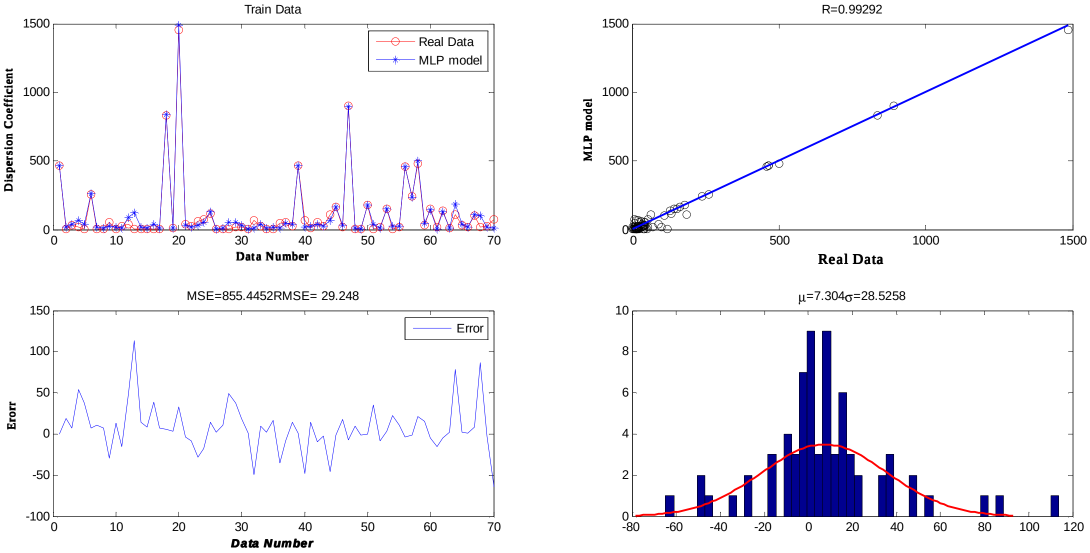

Fig. 11: Performance of MLP models during the training Stage.

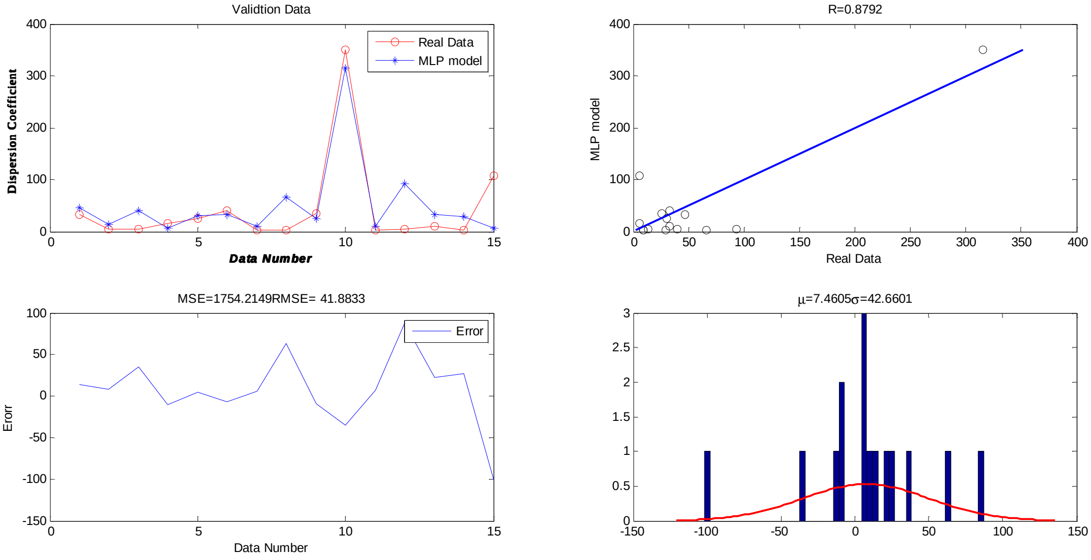

Fig. 12: Performance of MLP models during the Validation Stage.

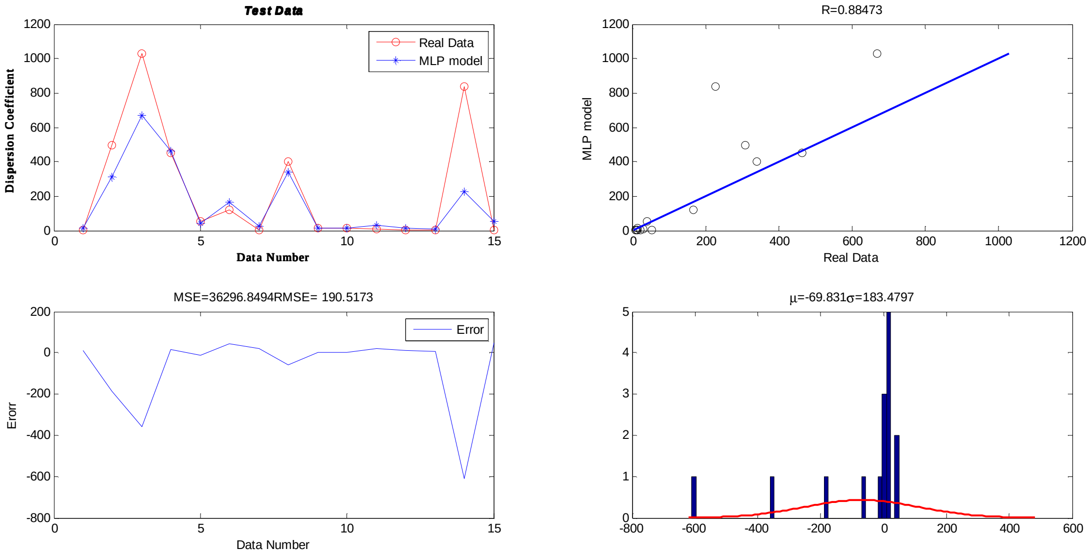

Fig. 13: Performance of MLP models during the Testing Stage.

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{N} \left(D_{L_{i\,(\text{model})}} - D_{L_{i\,(Actual)}} \right)^2}{\sum_{i=1}^{N} \left(D_{L_{i\,(Actual)}} \right)^2} \right], RMSE = \left[\frac{\sum_{i=1}^{N} \left(D_{L_{i\,(\text{model})}} - D_{L_{i\,(Actual)}} \right)^2}{N} \right]^{1/2} \quad (6)$$

Table 6: Calculating the DL by Empirical formulas and MLP model

model	St.(A-B)	St.(B-C)	St.(C-D)	St.(D-E)	St.(E-F)	St.(F-G)	
$D_{L-DRM}\,(m^2/s)$	41.5	26.5	12.5	38.5	37.5	29.5	R^2
$Eq(1)$	0.11	0.12	0.14	0.14	0.12	0.15	0.21
$Eq(2)$	1069.3	905.96	1071.9	1069	957	1166.76	0.00
$Eq(3)$	150.1	132.5	124.8	116.4	147.9	117.7	0.17
$Eq(4)$	2.81	2.7	2.38	2.43	3.06	2.35	0.11
$Eq(5)$	12.92	34.6	34.8	35.9	43.1	37.8	0.7
$Eq(6)$	74.03	13.8	14.7	15.3	16.8	16.5	0.38
$Eq(7)$	74.4	69.07	72.53	66.24	67.35	69.7	0.01
$Eq(8)$	27.1	29.04	29.8	31.8	37.5	34	0.02
$Eq(9)$	18.26	17.42	17.73	16.93	18.85	17.71	0.05
$Eq(10)$	3.49	3.39	3.77	3.45	3.15	3.71	0.4
$Eq(11)$	586.1	490.8	490.5	416.2	441.1	424.4	0.00
$Eq(12)$	56.13	53.29	55.25	52	55	54.47	0.01
ANN(MLP)	38.2	29.8	14.5	49.7	40.4	32.6	0.83

References

1. Holzbecher, E., *Environmental Modeling: Using MATLAB*. 2012: Springer.
2. Benedini, M. and G. Tsakiris, *Water Quality Modelling for Rivers and Streams*. 2013: Springer.
3. Riahi-Madvar, H., et al., *An expert system for*

predicting longitudinal dispersion coefficient in natural streams by using ANFIS. Expert Systems with Applications, 2009. 36(4): p. 8589-8596.

4. Chanson, H., *19 - Design of weirs and spillways*, in *Hydraulics of Open Channel Flow (Second Edition)*, H. Chanson, Editor. 2004, Butterworth-Heinemann: Oxford. p. 391-430.
5. Chau, K.W., *Modelling for Coastal Hydraulics and Engineering*. 2010: Taylor & Francis.
6. Portela, L. and R. Neves, *Numerical modelling of suspended sediment transport in tidal estuaries: A comparison between the Tagus (Portugal) and the Scheldt (Belgium-the Netherlands).* Netherland Journal of Aquatic Ecology, 1994. 28(3-4): p. 329-335.
7. Julien, P.Y., *Erosion and Sedimentation*. 1998: Cambridge University Press.
8. Kashefipour, S.M. and R.A. Falconer, *Longitudinal dispersion coefficients in natural channels.* Water Research, 2002. 36(6): p. 1596-1608.
9. Deng, Z., et al., *Longitudinal Dispersion Coefficient in Single-Channel Streams.* Journal of Hydraulic Engineering, 2002. 128(10): p. 901-916.
10. Deng, Z., V. Singh, and L. Bengtsson, *Longitudinal Dispersion Coefficient in Straight Rivers.* Journal of Hydraulic Engineering, 2001. 127(11): p. 919-927.
11. Seo, I. and K. Baek, *Estimation of the Longitudinal Dispersion Coefficient Using the Velocity Profile in Natural Streams.* Journal of Hydraulic Engineering, 2004. 130(3): p. 227-236.
12. Seo, I. and T. Cheong, *Predicting Longitudinal Dispersion Coefficient in Natural Streams.* Journal of Hydraulic Engineering, 1998. 124(1): p. 25-32.
13. Seo, I. and T. Cheong, *Moment-Based Calculation of Parameters for the Storage Zone Model for River Dispersion.* Journal of Hydraulic Engineering, 2001. 127(6): p. 453-465.
14. Shen, C., et al., *Estimating longitudinal dispersion in rivers using Acoustic Doppler Current Profilers.* Advances in Water Resources, 2010. 33(6): p. 615-623.
15. Seo, I.W. and K.O. Baek, *Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams.* Journal of Hydraulic Engineering, 2004. 130(3): p. 227-236.
16. West, J.R. and J.S. Mangat, *The determination and prediction of longitudinal dispersion coefficients in a narrow, shallow estuary.* Estuarine, Coastal and Shelf Science, 1986. 22(2): p. 161-181.
17. Seo, I., S. Park, and H. Choi, *A Study of Pollutant Mixing and Evaluating of Dispersion Coefficients in Laboratory Meandering Channel*, in *Advances in Water Resources and Hydraulic Engineering*. 2009, Springer Berlin Heidelberg. p. 485-490.
18. Carr, M. and C. Rehmann, *Measuring the Dispersion Coefficient with Acoustic Doppler Current Profilers.* Journal of Hydraulic Engineering, 2007. 133(8): p. 977-982.
19. Tayfur, G. and V. Singh, *Predicting Longitudinal Dispersion Coefficient in Natural Streams by Artificial Neural Network.* Journal of Hydraulic Engineering, 2005. 131(11): p. 991-1000.
20. Toprak, Z.F. and H.K. Cigizoglu, *Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods.* Hydrological Processes, 2008. 22(20): p. 4106-4129.
21. Azamathulla, H. and A. Ghani, *Genetic Programming for Predicting Longitudinal Dispersion Coefficients in Streams.* Water Resources Management, 2011. 25(6): p. 1537-1544.
22. Sahay, R., *Prediction of longitudinal dispersion coefficients in natural rivers using artificial neural network.* Environmental Fluid Mechanics, 2011. 11(3): p. 247-261.
23. Etemad-Shahidi, A. and M. Taghipour, *Predicting Longitudinal Dispersion Coefficient in Natural Streams Using M5′ Model Tree.* Journal of Hydraulic Engineering, 2012. 138(6): p. 542-554.
24. Aleksander, I. and H. Morton, *An Introduction to Neural Computing*. 1995: Internat. Thomson Computer Press.
25. Sivanandam, S.N. and S.N. Deepa, *Introduction to Neural Networks Using Matlab 6.0*. 2006: Tata McGraw-Hill.
26. Baek, K.O. and I.W. Seo, *Routing procedures for observed dispersion coefficients in two-dimensional river mixing.* Advances in Water Resources, 2010. 33(12): p. 1551-1559.
27. Atkinson, T.C. and P.M. Davis, *Longitudinal dispersion in natural channels: l. Experimental results from the River Severn, U.K.* Hydrol. Earth Syst. Sci., 1999. 4(3): p. 345-353.
28. Davis, P.M. and T.C. Atkinson, *Longitudinal dispersion in natural channels: 3. An aggregated dead zone model applied to the River Severn, U.K.* Hydrol. Earth Syst. Sci., 1999. 4(3): p. 373-381.
29. Davis, P.M., T.C. Atkinson, and T.M.L. Wigley, *Longitudinal dispersion in natural channels: 2. The roles of shear flow dispersion and dead zones in the River Severn, U.K.* Hydrol. Earth Syst. Sci., 1999. 4(3): p. 355-371.