International Journal of Waste Resources
ISSN: 2252-5211
Parsaie and Haghiabi, Int Int J Waste Resour 2015, 5:4
DOI: 10.4172/2252-5211.1000186
Research Article Open Access

Principle Component Analysis of Longitudinal Dispersion Coefficient Parameters

Abbas Parsaie* **and Amir Hamzeh Haghiabi**

Department of water Engineering, Lorestan University, Khorram Abad, Iran

Abstract

Study on the river water quality is a main part of environmental engineering. Longitudinal dispersion coefficient (D_L) is one of the main important parameters in the river water quality studies. Several parameters such as hydraulically and morphological rivers are affective on the D_L, whereas the mount effectiveness of some of them such as riverbed form cannot be measured. So, researchers proposed that the D_L is proportional to the flow velocity, channel width, river flow depth, and shear velocity. Defining the most influence parameters on the D_L leads to develop an optimal structure for empirical formulas and soft computing techniques which will be propose for estimation of D_L .In this paper the principle component analysis (PCA) was used to define the most affective parameters on the D_L. The PCA results indicated that the width of the river, flow depth, and flow velocity are the most important parameters on the D_L. Evaluating the performance of empirical formulas with considering the PCA results showed that the Tavakollizadeh and Kashefipur formula is accurate among the empirical formulas.

Keywords: Longitudinal dispersion coefficient; Water quality; River pollution; Soft computing; Principle component analysis

Introduction

River pollution has become one the main problem in the environmental health monitoring. Study on the river water quality is one of the main parameters in the part of the environmental engineering. Longitudinal dispersion coefficient definition is the important parameter in the river pollution studies. As shown in the Figure 1 when a contamination poured in the river flow rapidly emissions and moved to the downstream through the river. Figure 1 shows that at the first the dispersion mechanism is 3D dimensionality whereas by passing the time and moving through the river the dispersion mechanism is one dimensional which named in general longitudinal dispersion coefficient [1-3]. Study on the longitudinal dispersion coefficient in rivers usually conducted by field studies. Flied studies are usually conducted by injection a tracer which has not interaction with the water in the river and some station along the river is considered for sampling from the river water. The main notes that are more important in the field studies included the material of the tracer which should not interaction with water and has any destructive effect on the environment and another note is related to the place of the sampling station, the location of the first station should be considered after complete mixing the trace in the cross section of the river flow [4]. Several field and laboratory studies have been conducted on the mechanism of the D_{Lc} in the rivers. In this regard, the studies which conducted by [5-12] can be mentioned. Due to high cost of the field studies and laboratory equipment recently researcher welcomed to use the numerical approaches. In the field of the numerical modelling the government equation which usually is differential equation solves by numerical method such as finite difference, finite volume and finite element another numerical method which are widely uses in the environmental studies are artificial neural networks (ANN) such as Multilayer perceptron neural network (MLP), adaptive neuro fuzzy inference system (ANFIS) ,genetic programing(GP) and support vector machine. [13-23] The soft computing technique can be uses with numerical methods to increase the accuracy of the numerical simulation [24,25]. Developing the soft computing models are based on the data set, it means that the important parameters which are influence on the phenomena should be measured in the past. In this paper the principle component analysis (PCA) are used to derived most important parameters on the D_L. Developing the empirical formulas and soft computing models based on the PCA leads to prepare an optimal structure of the model by more reliability.

Methods and Material

Longitudinal dispersion coefficient is a function of the river geometries, fluid properties and hydraulic condition. The main parameters that are influence on the D_c are given on the Equation (1).

$$D_c = f_1\left(\rho, \mu, u, u_*, h, w, s_f, s_n\right) \quad \text{........... (1)}$$

Where ρ is fluid density; μ is dynamic viscosity; w is the width

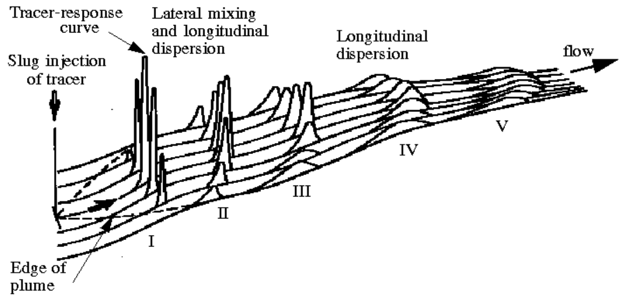

Figure 1: Schematic shape of the dispersion mechanism in rivers

***Corresponding author:** Abbas Parsaie, Ph.D. Candidate of Hydro Structures, Department of Water Engineering, Lorestan University, Khorram Abad, Iran, E-mail: Abbas_Parsaie@yahoo.com; Tel: +989163685867

Received October 10, 2015; **Accepted** November 12, 2015; **Published** November 19, 2015

Citation: Parsaie A, Haghiabi AH (2015) Principle Component Analysis of Longitudinal Dispersion Coefficient Parameters . Int J Waste Resour 5: 186. doi: 10.4172/2252-5211.1000186

Citation: Parsaie A, Haghiabi AH (2015) Principle Component Analysis of Longitudinal Dispersion Coefficient Parameters . Int J Waste Resour 5: 186. doi: 10.4172/2252-5211.1000186

of cross-section; h is flow depth; u_* is share velocity, s_f is longitudinal bed shape and s_n is sinuosity of the river. To derive the dimensionless parameter on the D_L , the Buckingham theory was considered and dimensionless parameter will be derived as shown below

$$\frac{D_L}{hu_* u} = f_1\left(\rho\frac{uh}{\mu}, \frac{u}{u_*}, \frac{w}{h}, s_f, s_n\right) \text{......... (2)}$$

The flow in the nature is always turbulent especially in the river. Therefore, the Reynolds number $\rho\frac{uh}{\mu}$ can be ignored and the bed form and sinusitis 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. Table 1 gives some of the famous empirical formula which proposed by researchers. As mentioned in past developing the ANN models is based on the data set so about 150 data set related to the D_L was collected and range of them given in the Table 2.

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

Principle Component Analysis (PCA)

The Principle Component Analysis (PCA) is an advanced categorized method in the factor analysis approaches and usually uses for data reduction in the field of engineering. The main application of the PCA is in the compression and classification of data; in the other words the main usage of this approach is to reduce the dimensionality of a data set (sample) by finding a new set of variables, smaller than the original set of variables that nonetheless retains most of the sample's information. During the PCA process because of using the all initial variable, the new variable involved all the initial variable information. The process of PCA continued two steps which explained in the follow [26].

Results and Conclusion

The accuracy of the empirical formulas was conducted by compression with measurement data and results of these are shown in the Figure 2. Figure 2 shows the results of the empirical formulas versus the measured data and also the standard error indices such as correlation coefficient (R^2) and root mean square of error (RMSE) was added to this figure. As seems in the Figure 2 the Tavakollizadeh and Kashefipur formula is the accurate among the empirical formulas by (0.45 and RMSE=5861). In overall, assessing the performance of the empirical formulas shows that these formulas have unacceptable accuracy to use in management problems. To define the most influence parameters on the D_L , the PCA technique was carried out on the collected data which these range was given in the Table 2. The result of the PCA is given in the Figure 3. As shown in the Figure 3, the channel width, flow depth, flow velocity are the most important parameters for D_L prediction. Another scenario as similar to the equation (2 &3) was considered, the result of the scenario number (2) shows the ratio of the u/u* is more important than the W/H.

By attention to the PCA results and reviewing the structure of empirical formulas which was given in the table 1 and evaluating the results of the empirical formula which was given in the Figure 2 it could be found that the empirical formula such as Tavakollizadeh and Kashefipur which considered more weight for the parameters such as W

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

Table 1: Empirical equations for calculating the longitudinal dispersion coefficient

Citation: Parsaie A, Haghiabi AH (2015) Principle Component Analysis of Longitudinal Dispersion Coefficient Parameters . Int J Waste Resour 5: 186. doi: 10.4172/2252-5211.1000186

Data Range	W (m)	H(m)	U(m/s)	u*(m/s)	W/H	U/u*
min	1.40	0.14	0.03	0.00	2.20	1.06
max	711.20	19.94	1.74	0.55	156.54	20.77
AVG	68.23	1.72	0.53	0.09	44.57	7.54
STDEV	99.28	2.20	0.35	0.07	29.76	4.79

Table 2: Range of collected data.

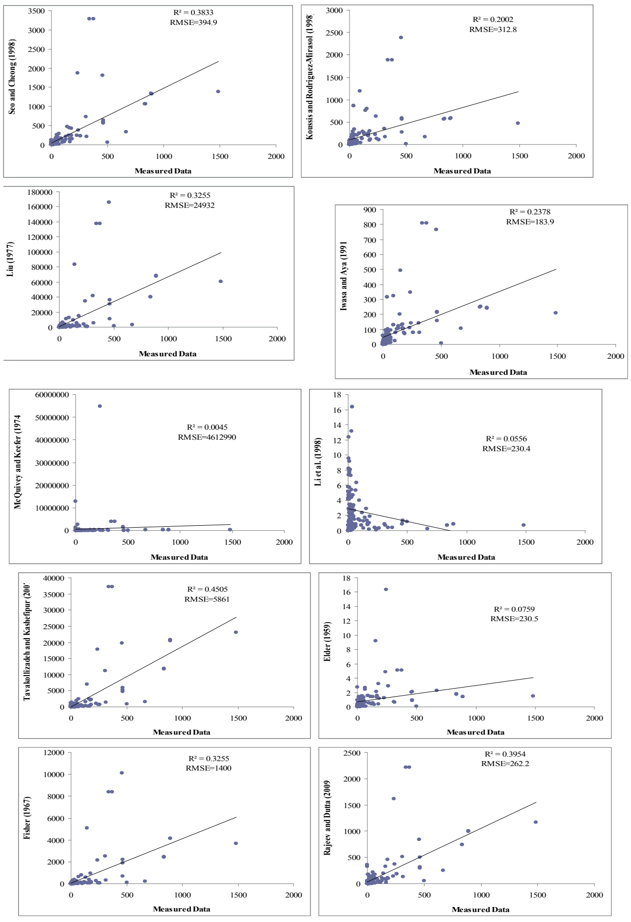

Figure 2: Performance of the empirical formulas to calculate the D_L

and H or u/u_* are more accurate in the compare to the other formulas.

Conclusion

In this study, some famous analytical approaches for calculating the longitudinal dispersion coefficient (D_L) were assessed. To this purpose, 150 experimental data which published in the credible journal was collected. Calculation the standard error indices for analytical approaches results show that the Tavakollizadeh and Kashefipur formula by correlation coefficient about 0.45 is accurate among the empirical formulas. To define the most influence parameters on the D_L, the PCA technique was used. The PCA result indicated that the channel width, flow depth, flow velocity are the most important parameters on the D_L. Using the results of principle component analysis (PCA) techniques leads to develop an optimal model structure for empirical formula or soft computing models.

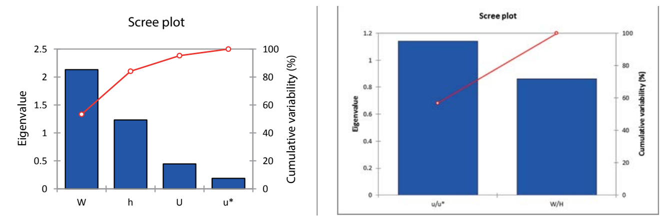

Figure 3: Results of the PCA techique, left: Scenario 1, right: Scenario 2.

References

1. Chanson H (2004) Environmental Hydraulics for Open Channel Flows.
2. Holzbecher E (2012) Environmental Modeling: Using MATLAB.
3. Sahin S (2014) An Empirical Approach for Determining Longitudinal Dispersion Coefficients in Rivers. Environ Process 1: 277-285.
4. Atkinson TC, Davis PM (1999) Longitudinal dispersion in natural channels: I. Experimental results from the River Severn, U.K. Hydrol Earth Syst Sci 4: 345-353.
5. Baek KO, Seo IW (2010) Routing procedures for observed dispersion coefficients in two-dimensional river mixing. Advances in Water Resources 33: 1551-1559.
6. Davis PM, Atkinson TC (1999) Longitudinal dispersion in natural channels: 3. An aggregated dead zone model applied to the River Severn, U.K. Hydrol Earth Syst Sci 4: 373-381.
7. Kashefipour SM, Falconer RA (2002) Longitudinal dispersion coefficients in natural channels. Water Res 36: 1596-1608.
8. Nordin CF, Troutman BM (1980) Longitudinal dispersion in rivers: The persistence of skewness in observed data. Water Resources Research 16: 123-128.
9. Papadimitrakis I, Orphanos I (2004) Longitudinal Dispersion Characteristics of Rivers and Natural Streams in Greece. Water Air & Soil Pollution: Focus 4: 289-305.
10. Seo I, Park S, Choi H (2009) A Study of Pollutant Mixing and Evaluating of Dispersion Coefficients in Laboratory Meandering Channel. Advances in Water Resources and Hydraulic Engineering 485-490.
11. Azamathulla H, Ghani A (2011) Genetic Programming for Predicting Longitudinal Dispersion Coefficients in Streams. Water Resources Management 25: 1537-1544.
12. Azamathulla HM, Wu FC (2011) Support vector machine approach for longitudinal dispersion coefficients in natural streams. Applied Soft Computing 11: 2902-2905.
13. Benedini M, Tsakiris G (2013) Water Quality Modelling for Rivers and Streams.
14. Etemad-Shahidi A, Taghipour M (2012) Predicting Longitudinal Dispersion Coefficient in Natural Streams Using M5′ Model Tree. Journal of Hydraulic Engineering 138: 542-554.
15. Fuat Toprak Z, Savci ME (2007) Longitudinal Dispersion Coefficient Modeling in Natural Channels using Fuzzy Logic." CLEAN – Soil, Air, Water 35: 626-637.
16. Li X, Liu H, Yin M (2013) Differential Evolution for Prediction of Longitudinal Dispersion Coefficients in Natural Streams. Water Resour Manage 27: 5245-5260.
17. Mirbagheri S, Abaspour M, Zamani K (2009) Mathematical modeling of water quality in river systems.
18. 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 Systems with Applications 36: 8589-8596.
19. Sahay R (2011) Prediction of longitudinal dispersion coefficients in natural rivers using artificial neural network. Environ Fluid Mech 11: 247-261.
20. Szymkiewicz R (2010) Numerical Modeling in Open Channel Hydraulics.
21. Tayfur G, Singh V (2005) Predicting Longitudinal Dispersion Coefficient in Natural Streams by Artificial Neural Network. Journal of Hydraulic Engineering

131: 991-1000.

22. 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: 718-730.

23. Parsaie A, Haghiabi A (2015) The Effect of Predicting Discharge Coefficient by Neural Network on Increasing the Numerical Modeling Accuracy of Flow Over Side Weir. Water Resour Manage 29 973-985.

24. 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: 1-6.

25. Alvarez PA (2011) Exploratory Data Analysis with MATLAB, (2nd Edn). International Statistical Review 79: 492-492.

26. Camacho J, Pérez-Villegas A, Rodríguez-Gómez RA, Jiménez-Mañas E (2015) Multivariate Exploratory Data Analysis (MEDA) Toolbox for Matlab. Chemometrics and Intelligent Laboratory Systems 143: 49-57.