 doi:10.20944/preprints201905.0110.v1

Estimating longitudinal dispersion coefficient in natural streams using empirical models and machine learning algorithms

Katayoun Kargar[a], Saeed Samadianfard[b], Javad Parsa[b], Narjes Nabipour[c], Shahaboddin Shamshirband[d,e], Timon Rabczuk[f]

[a] *Department of Civil Engineering, Urmia University, Urmia, Iran.*

[b] *Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran*

[c] *Independent researcher, Mazabdaran, Sari, Iran*

[d] *Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam; shahaboddin.shamshirband@tdtu.edu.vn*

[e] *Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

[f] *Institute of Structural Mechanics, Bauhaus University Weimar, 99423 Weimar, Germany*

Abstract

Longitudinal dispersion coefficient (LDC) plays an essential role in modeling the transport of pollution and sediment in the natural rivers. As a result of transportation processes, the concentration of pollution changes along the river. Different studies have been conducted to provide simple equations for estimating LDC. In this study, Support Vector Regression (SVR), Gaussian Process Regression (GPR), M5P and Random Forest (RF) examined to predict the LDC in the natural streams. The hydraulic and geometric features of different rivers gathered for developing the mentioned models for LDC estimation and various statistical criteria were utilized to scrutinize of the models. Furthermore, the Taylor chart was used to evaluate the models and achieved results showed that among machine learning models, M5P displayed the superior performance with CC of 0.823, SI of 0.812, NS of 0.577 and WI of 0.879. As well, S-D model with CC of 0.795 SI of 0.827, NS of 0.558 and WI of 0.890 had more precise results than other empirical models. The results indicates that the developed M5P model with simple formulations was superior to other machine learning models and empirical models and therefore, it can be used as a proper tool for estimating LDC in natural rivers.

Keywords: longitudinal dispersion coefficient, machine learning algorithms, rivers, statistical parameters.

Introduction

In the last decades, variations in quantity and quality of water resource systems have attracted the attention all around the world. Due to creating the communal health, the most significant and vital areas are cities where rivers provide drinking water and factories are located adjacent these streams (Tayfour

& Singh, 2005; Pourabadei and Kashefipour, 2007). Therefore, flow estimation as well as pollution transportation in natural streams are really important in water resources management. This necessitates accurate knowledge and data on the transmission and mixing of contaminations in the rivers and the ability to transport these materials by river flow (Pourabadei & Kashefipour, 2007). The dispersion issue is used to the mixing in natural rivers as well as in open channels. When pollutants and sewage are discharged into natural streams, they move with flow and mixing is occurred in three stages (Jirka, 2004). In the first step, the pollutant is rapidly mixed in the vertical direction. The lateral mixing is done at second stage and the pollutant is distributed sporadically. In the last step, the pollutant is dispersed longitudinally. It is created due to the lateral variation of the longitudinal velocity. For water quality analysis, one-dimensional model is used, that includes the last stage, and its severity can be determined by LDC, which is a key factor in modeling and estimating the distribution of sediment and pollution in water (Kashefipour and Falconer, 2002).

In many hydraulic problems, such as environmental engineering and pollutants in river flows, it is really important to estimate LDC accurately. When real data of this process are available in the river, data sets such as mean (U) and shearing velocity (U*), channel width (W), depth of water (H), channel slope (S) and etc., LDC can be determined readily. Several methods have been evolved to estimate the LDC value. Julínek and Říha (2017) used fluorescein color as a tracer in an open channel for determining LDC value. Results from this study were compared with values gained by formerly empirical formula and it showed well agreement with aforementioned studies. An artificial neural network (ANN) model was established by Sahay (2011) for predicting LDC in rivers. The result of ANN demonstrated that it had high performance than other methods. Noori et al (2015) utilized three methods namely, ANN, adaptive neuro fuzzy inference system and support vector machine, for LDC estimation in natural streams. High degree of doubt was found in the models, while LDC which estimated by SVM method had lower error in comparison with ANN and ANFIS models. Azamathulla and Ab.Ghani (2011) used genetic programming (GP) method to estimate LDC in streams and it was revealed that GP had provided more accurate predictions than empirical models. For prediction of LDC, ANFIS method was used with Riahi-Madvar et al (2009) who implemented several statistical methods for scrutinizing the model. The result showed that ANFIS is superior in estimating LDC in comparison with empirical models.

The importance of LDC on the transport of pollutants along the rivers and its dependence on the hydrodynamic and geometric parameters have caused the necessity of many studies to estimate the LDC (McQuivey and Keefer 1974; Fischer 1979; Bencala and Walters 1983; Rutherford 1994; Seo and Cheong 1998; Kashefipour and Falconer 2002; Etemad-Shahidi and Taghipour 2012; Wang et al. 2017). In other words, LDC is a measure of intensity of pollutants mixing and plays a crucial role in the modeling of water quality in rivers and decisions made by water authorities as well. In order to provide a satisfactory estimation of LDC, different empirical formulae have been presented. The accurate estimation of LDC leads to accurate modeling of pollutants concentration along the rivers and streams

(Kashefipour and Falconer 2002). Derivation of empirical formulae for LDC is based on Π-Buckingham theory (Seo and Cheong, 1998). This classic procedure is mainly utilized for most complex hydraulic problems when the theory is incomplete to accurate and/or analytical study. In the current work attempts have been made to predict LDC using the non-dimensional parameters obtained by Π-Buckingham theory and machine learning algorithms. In this regard, the main purpose of the current research was utilizing the machine learning and data driven algorithms for improving the precision of LDC estimation. So, the applicability of GPR, SVR, M5P, and RF were examined and their results were compared with outputs of common empirical models. To the best of our knowledge, the applications of GPR and RF have not been reported in the literature for estimating LDC in natural streams.

Material and methods

Theory of dispersion

Water quality of rivers is affected by pollutions and their distributions along the riverine flows. Non-uniformity in the geometry of natural streams along with the effects of shear stresses and flow turbulence result in the complex flow field (Wang et al., 2017). After the completion of cross-sectional mixing, the following one-dimensional unsteady advection-dispersion equation is extensively used to predict the water quality in rivers.

$$\frac{\partial C}{\partial t}+U\frac{\partial C}{\partial x}=K\frac{\partial^2 C}{\partial x^2}+S \tag{1}$$

in which C is the cross-sectional averaged concentration, U is the cross-sectional averaged velocity, K is LDC, S is the source term, x and t are the mean flow direction and time, respectively. According to equation (1), main transport processes are advection and dispersion. In the mixing process, the pollutant is diffused due to the velocity differences over the cross-section. By transporting the pollutant towards the downstream, turbulent diffusion causes to mix pollutant completely and then the concentration of pollutants along the river depends mainly on the LDC. Hence, the LDC is essential parameter in predicting the solute concentration in the flow direction (Fischer et al. 1979). Based on the Taylor (1954) study, shear velocity and turbulence have main effects on the mixing intensity and the combination of longitudinal advection and longitudinal mixing can be resulted in LDC. The effect of hydrodynamic and geometric parameters on the LDC shows its variability in different streams and rivers.

Experimental data

The utilized data in the current research were measured data for above 50 rivers of USA and UK, which gathered from different studies (Fischer 1968; McQuivey and Keffer 1974; Nordin and Sabol 1974; Rutherford 1994; Graf 1995). So, 147 sets of data, which their statistical characteristics are presented in Table 1, were used in the current study. In the mentioned table, W, H, U, U_*, W/H, U/U_*, K and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2019 doi:10.20944/preprints201905.0110.v1

K/(HU*) denotes channel width, depth of water, mean velocity, shear velocity, ratio of channel width to the depth of water, longitudinal dispersion coefficient (LDC) and non-dimensional LDC, respectively.

Table 1 Statistical characteristics of used data

Variable	mean	minimum	maximum	standard deviation	coefficient of variation	skewness
W (m)	60.021	1.400	711.200	91.753	1.529	4.582
H (m)	1.550	0.140	19.940	2.131	1.375	5.243
U (ms^{-1})	0.492	0.029	1.730	0.339	0.689	1.474
U_* (ms^{-1})	0.089	0.002	0.553	0.081	0.910	3.760
W/H	43.452	2.200	156.500	29.719	0.684	1.472
U/U_*	6.954	0.770	20.770	4.651	0.669	1.186
K (m^2s^{-1})	84.228	0.200	1486.500	180.816	2.147	4.777
$K/(HU_*)$	785.110	3.080	7692.000	1119.676	1.426	3.336

Empirical models

In this section, six empirical models for estimation of LDC in rivers were presented in Table 2. These empirical models can be just within a range of specific flow and geometry and may not have appropriate results for other ranges.

Table 2 Empirical models for estimation of LDC values

Method	Equation	Notation
Fisher et al (1979)	$\frac{K}{HU_*} = 0.011\,(\frac{W}{H})^2\,(\frac{U}{U_*})^2$	F
Seo and Cheong (1998)	$\frac{K}{HU_*} = 5.915\,(\frac{W}{H})^{0.62}\,(\frac{U}{U_*})^{1.428}$	S-C
Kashefipour and Falconer (2002)	$\frac{K}{HU_*} = 10.612\,(\frac{U}{U_*})^2$	K-F
Sahay and Dutta (2009)	$\frac{K}{HU_*} = 2\,(\frac{W}{H})^{0.96}\,(\frac{U}{U_*})^{1.25}$	S-D
Wang and Huai (2016)	$\frac{K}{HU_*} = 17.648\,(\frac{W}{H})^{0.3619}\,(\frac{U}{U_*})^{1.16}$	W-H
Li et al (2013)	$\frac{K}{HU_*} = 2.828\,(\frac{W}{H})^{3.7613}\,(\frac{U}{U_*})^{1.4713}$	L

M5 model tree (M5P)

The M5P algorithm was first introduced by Quinlan (1992) and it is an extended version of the M5 algorithm. Model trees can consider a set of data with a large number of features and sizes and can work with high degree of efficiency. The M5P algorithm contains four stages. The first stage is building a tree by dividing the input space into numerous subspaces. The variation in intra-space from root to node is lessened by using some attributes and division criterion. To measure the subspaces variability, the values of standard deviation for each node were utilized. By using standard deviation reduction factor, the tree is built. This method remarkably reduces the expected errors in the node using following equation.

$$SDR= sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \quad (2)$$

In this equation, sd denotes the standard deviation, T is a set of examples which reach the node and T_i is the outcomes of the node division pursuant to the attributes (Wang & Witten, 1997). The linear regression model is advanced in each of the sub-spaces for each node, which is done in the second step. Then, the pruning method is used to overcome the problem of over-training which happens when the correspondent SDR value of the linear model becomes lower than the predetermined error. The adjacent linear model can show severe disturbances in the result of pruning and it can happen mostly for some models made from a smaller amount of training data but this can be balanced by smoothing in the last step. In the smoothing procedure, to create the last model of the leaf, all models from the leaf to the root are combined.

Support Vector Regression (SVR)

Support vector regression is evolved from support vector machines (SVM), which was provided by (Vapnik, 1995, 1998) and has been used in many hydrological applications. This approach is a data-based method and it deals with the predicted problems and the structural risk minimization principle (Pai and Hong, 2007). Achieving a regression model with suitable predictive performance is the main goal of the SVR. Variables in the SVR model is $(x_i, y_i)_{i=1}^{N}$ in which x_i is input parameter, y_i is output parameter and the total number of data is represented by N. The SVR is expressed as:

$$f(x) = w\varphi(x) + b \quad (3)$$

where w is a weight vector, b is a threshold value and $\varphi(x)$ is a non-linear mapping variable. Input patterns are designed in a large space, therefore, in the mapped space it can be linearly regressed. In the SVR model, the optimal amount of w and b computed by the below formula:

$$\min\left\{\frac{1}{2}||w||^2 + C\sum_{l=1}^{n}|y_t - f(x_t)|\right\} \quad (4)$$

where C demonstrate the penalty parameter, and n is the sample size.

Gaussian Process Regression (GPR)

It is defined as a set of random variables in which each of them has a common Gaussian distribution. To represent the relation between inputs and outputs, the f function is modeled. By a precise f model, for each possible entry, an output is predicted. In the GPR process, in the first step the training samples should be determined. In GPR, *x* is input and *y* is output parameter. To achieve a model between *x* and *y*, a GPR model is made as the regression function and the noise term ($\varepsilon \sim N(0, \sigma_n^2)$) is used in this function:

$$y = f(x) + \varepsilon \tag{5}$$

where σ_n is the standard deviation of the noise.

This can be completely determined by a mean $m(x)$ and a covariance $k(x, x')$.

$$f(x) \sim GP(m(x), k(x, x')) \tag{6}$$

$m(x)$=0 is assumed to facilitate the computation and there are different choices for k. Covariance function, which is known as kernel function, is a linear separator and is used to obtain the connection between input and output of the model. If points are moved to higher spaces, their internal multiplication (*k*) will be changed too. Selecting a suitable kernel function based on assumptions such as smoothness and possible patterns in the data is really significant. The kernel functions which used in this study are, the polynomial kernel, the Normalized Poly Kernel, the radial basis function or the Gaussian kernel (RBF), and the Pearson universal kernel (PUK). In this section GPR modeling method was introduced briefly, for more detailed explanation see Rasmussen and Williams (2006).

Random Forest (RF)

This method was a series of relatively complex relationships that were able to take into account the interaction between predictors and multiplex forms of responses without any relations between substrate values. The RF model includes a number of uncomplicated decision trees (Breiman, 2001). By using a sample subset of the available data, each of the component trees forming an RF. The same example can happen in several subsets because these subsets are independent. For each tree, a subset of predictions is chosen with the same chance. By combining and averaging the single predictions of all compounding trees, the predictive output is achieved. The RF algorithm consist of two random levels in each tree. The first step is bagging and the second one is selection of the features randomly, these are indicated that the performance of this model is superior to other models (Archer & Kimes, 2008). RF, without any assumptions about an independent or dependent variable, explains both linear and nonlinear relationships.

Performance criteria

To statistically examine the performance of models created in the current study, correlation coefficient (CC), Scattered Index (SI), Nash-Sutcliffe (NS), Willmott's Index (WI) were used. The mathematical representations are cited as follows.

$$CC = \frac{\left(\sum_{i=1}^{N} O_i P_i - \frac{1}{N}\sum_{i=1}^{N} O_i \sum_{i=1}^{N} P_i\right)}{\left(\sum_{i=1}^{N} O_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} O_i\right)^2\right)\left(\sum_{i=1}^{N} P_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} P_i\right)^2\right)} \tag{7}$$

$$SI = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}}{\overline{O_i}} \tag{8}$$

$$NS = 1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \overline{O_i})^2}\right], 0 \leq NS \leq 1 \tag{9}$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}\left(|P_i' - \overline{O_i}| + |O_i' - \overline{O_i}|\right)^2}\right], 0 \leq WI \leq 1 \tag{10}$$

in above-mentioned formulas O_i and P_i are the measured and estimated value of the dispersion coefficient, $\bar{O}$ is the mean of measured O and the number of data is represented by N.

Additionally, Taylor diagrams (Taylor, 2001) were utilized for checking the precision of the implemented models and empirical models for LDC estimation in natural rivers. It is notable that in the mentioned diagram, measured and some correspondent statistical parameters are presented, simultaneously. Moreover, different points on a polar plot are used in Taylor diagrams for investigating the differences between measured and estimated values. Also, the CC and normalized standard deviation are specified by azimuth angle and radial distances from the base point, respectively (Taylor, 2001).

Results and Discussion

The capability of machine learning and data driven algorithms such as GPR, SVR, M5P and RF in estimating LDC values in different streams were compared with the potential of common empirical models. For this purpose, the hydraulic parameters of several streams in different geographical locations, including channel width, depth of water, mean velocity, shear velocity and LDC, were gathered. In the current study, the whole data includes the hydraulic and geometric properties of 147 streams, which randomly partitioned into training (67%) and testing data (33%). In the other word, $\frac{W}{H}$ and $\frac{U}{U_*}$ were used for estimating LDC values. It should be noted that four kernel functions including polynomial, normalized polynomial, Pearson VII function-based and radial basis function were investigated for GPR (GPR-1, GPR-2, GPR-3, GPR-4) and SVR (SVR-1, SVR-2, SVR-3, SVR-4) models. The results of statistical parameters including CC, SI, NS and WI in LDC estimation for considered models and mentioned common empirical models are displayed in Table 3. It is obvious from this table that among GPR models, GPR-3 with CC of 0.679, SI of 1.053, NS of 0.288 and WI of 0.766 had the best performance. Moreover, SVR-3 with CC of 0.788, SI of 0.822, NS of 0.566 and WI of 0.818 estimated LDC values with lower errors comparing with different SVR models. It is notable that Pearson VII function-based GPR and SVR models (GPR-3 and SVR-3) had more accuracy in comparison with other GPR and SVR models. In the other words, Pearson VII function-based kernel had more applicability in LDC estimation than other mentioned kernel functions. Furthermore, it can be comprehended from Table 3 that M5P with CC of 0.823, SI of 0.812, NS of 0.577 and WI of 0.879 had the precise prediction between machine learning and data driven algorithms. Unlike GPR, SVR and M5P models, RF with CC of 0.482, SI of 1.532, NS of -0.506 and WI of 0.666 had an unacceptable accuracy and it is not recommended for LDC estimation. Additionally, among considered empirical models, S-D with CC of 0.795 SI of 0.827, NS of 0.558 and WI of 0.890 had high precision in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2019 doi:10.20944/preprints201905.0110.v1

comparison with other empirical models. In the other words, errors generated by S-D model are only lower than GPR-3 and RF. This implies that the SVR-3 and M5P are able to attain more accurate performances compared to empirical models. Evident by statistical metrics presented in Table 3, it can be concluded that the accuracy of M5P far exceeds the mentioned models and empirical models. In terms of practical value of the M5P model, the resulted explicit formulations can be a powerful and easy tool for accurate estimation of LDC values.

Scatter plot of the observed values of LDC and the corresponding values estimated by the studied methods and empirical models are shown in Figure 1. Also, it is clear from Fig 1 that the estimate of M5P are less scattered across the bisection line. So, the estimates of M5P are much closer to the bisection line than data driven algorithms and empirical models. Moreover, Fig 2 presented the measured and estimated values of LDC in testing phase. In accordance to the obtained concluding remarks from scatter plots, it is obvious that the estimates of M5P are in better agreement with measured LDC values. Besides, an additional evaluation of measured and estimated LDC values by GPR-3, SVR-3, M5P, RF and S-D were accomplished and Fig 3 presented the obtained results. This figure is an instrumental tool to understand the different potential of studied models. In the Taylor diagram, the best model is clarified by the point with the lower RMSE and higher CC values. According to the lower distance from measured point (the point with green color) to the correspondent point of the M5P (the point with cyan color), the estimates of M5P were more accurate than the results of other models.

Table 3 performance of various models based on different statistical parameters

Methods	Statistical parameters			
	CC	SI	NS	WI
GPR-1	0.078	1.447	-0.345	0.358
GPR-2	0. 645	1.623	-0.690	0.487
GPR-3	0.679	1.053	0.288	0.766
GPR-4	0.493	1.389	-0.238	0.403
SVR-1	0.089	1.260	-0.020	0.221
SVR-2	0.651	1.008	0.348	0.625
SVR-3	0.788	0.822	0.566	0.818
SVR-4	0.547	1.154	0.145	0.326
M5P	0.823	0.812	0.577	0.879
RF	0.482	1.532	-0.506	0.666
F	0.809	3.894	-8.789	0.563
S-C	0.734	1.033	0.312	0.852
K-F	0.411	1.438	-0.334	0.661
S-D	0.795	0.827	0.558	0.890
W-H	0.643	0.996	0.359	0.747
L	0.768	0.942	0.427	0.873

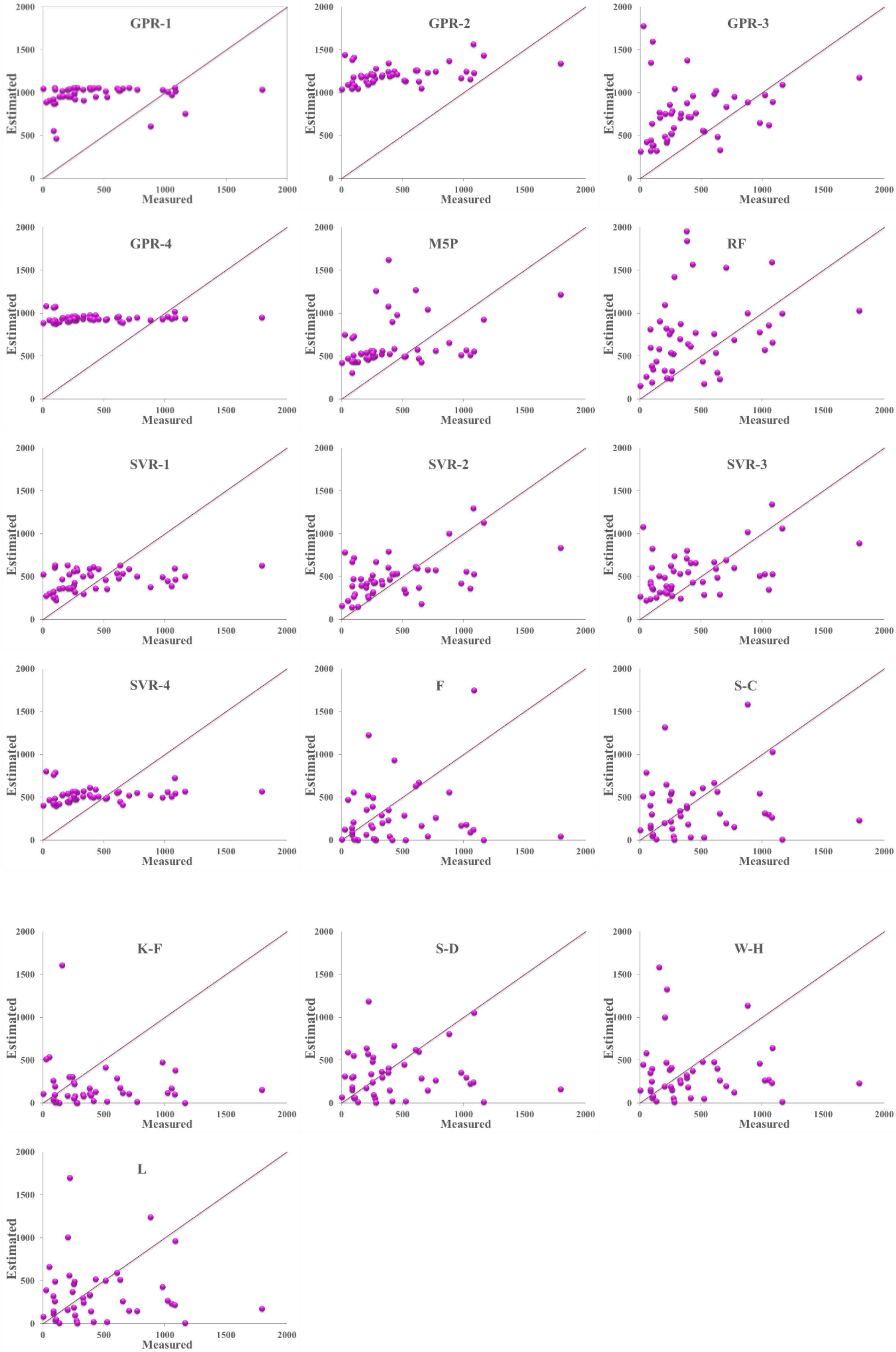
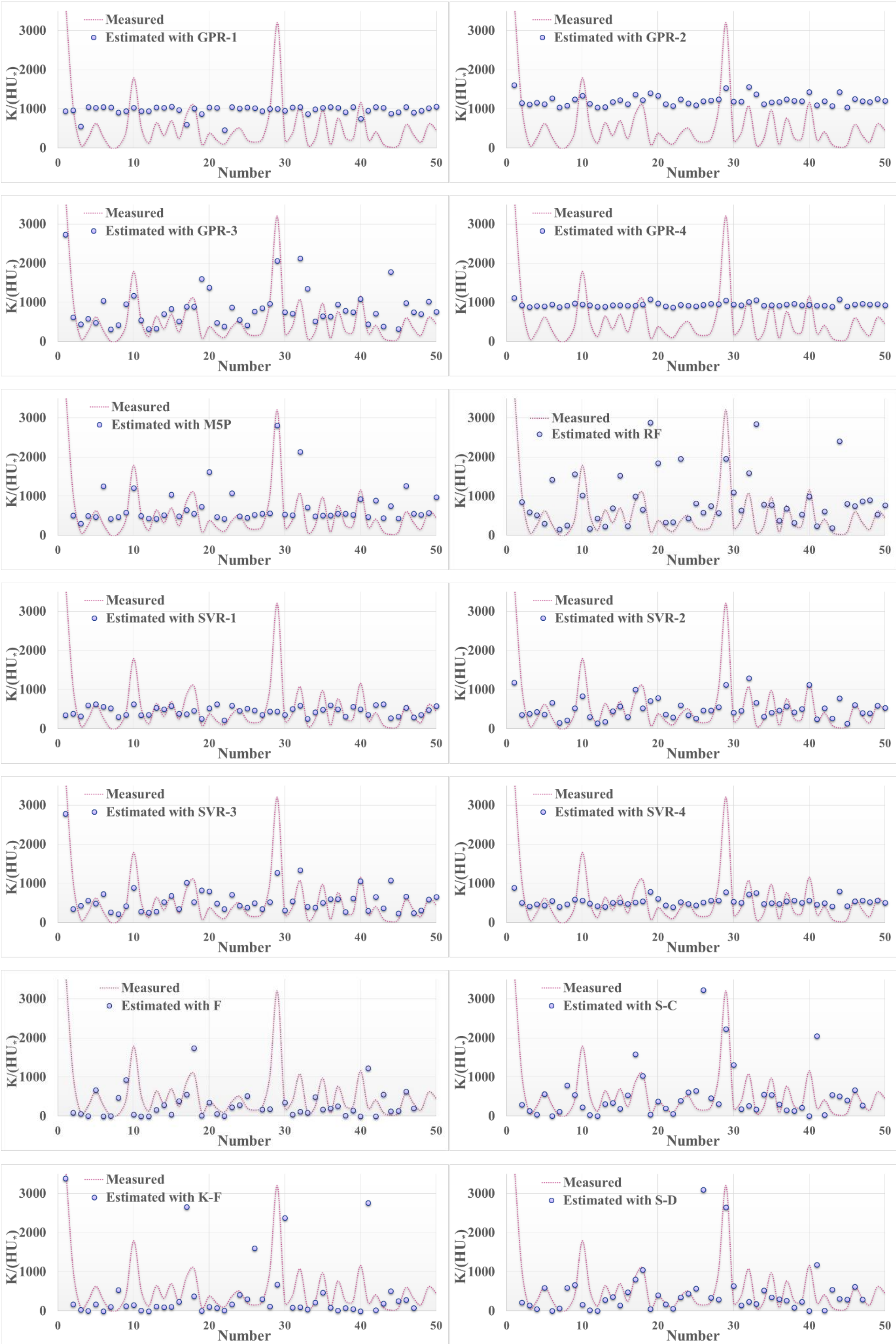


Fig 1 Scatter plots of measured and estimated dispersion coefficient

 doi:10.20944/preprints201905.0110.v1



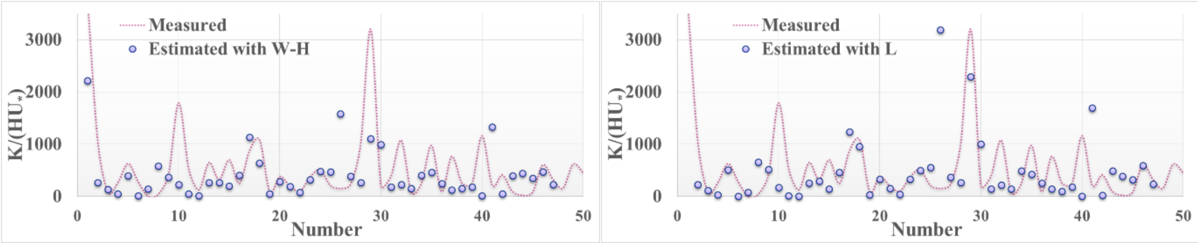


Fig 2 Series plots of measured and estimated dispersion coefficient values

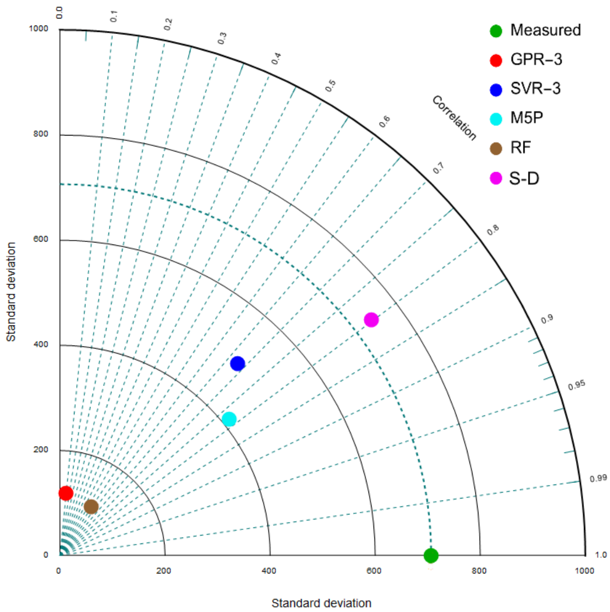
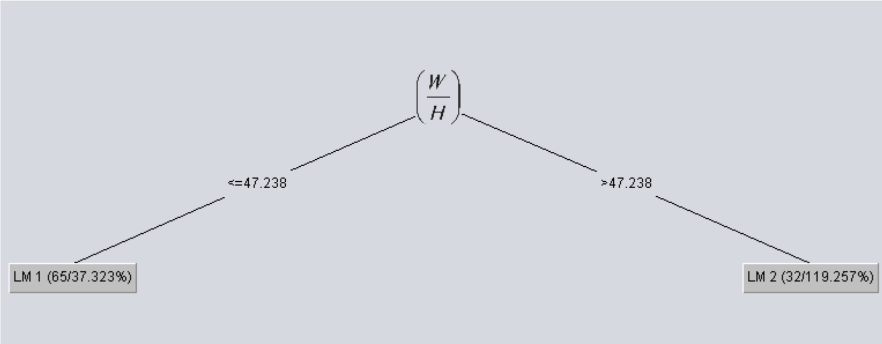


Fig 3 Taylor diagram of estimated dispersion coefficient values

Also, resulted tree structure from M5P model as the best accurate model is displayed in Fig 4. This illustrated tree model is based on the characteristics of used streams, and two linear equations are applied for LDC estimation, while other studied data driven models do not have such a precision and capability.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2019 doi:10.20944/preprints201905.0110.v1



LM number	Resulted linear equation
LM 1	$\left(\frac{K}{HU_*}\right) = 1.6896 \times \left(\frac{W}{H}\right) + 20.0124 * \left(\frac{U}{U_*}\right) + 393.3346$
LM 2	$\left(\frac{K}{HU_*}\right) = 2.8759 \times \left(\frac{W}{H}\right) + 181.7915 * \left(\frac{U}{U_*}\right) + 339.5557$

Fig 4 M5P obtained tree with two different linear rules

Conclusion

In this study, various machine learning algorithms including GPR, SVR, M5P and RF were utilized to estimate LDC values in the natural streams and rivers. LDC values can be estimated using flow variables and channel geometric characteristics. So, in the current research, $\frac{W}{H}$ and $\frac{U}{U_*}$ were considered as input parameters and $\frac{K}{U_* H}$ as an output parameter. The performances of the models were evaluated based on error measures of CC, SI, NS, WI and Taylor diagram as well. The results indicated that although, GPR-3 and SVR-3 using Pearson VII function-based kernel and M5P had satisfactory performances, but M5P provided the most accurate estimations of LDC. Furthermore, among six common empirical models, which were implemented in the current study, S-D model had the best result. In conclusion, the developed M5P model outperformed others in terms of accuracy and it is recommended for LDC estimation.

References

Archer, K. J., Kimes, R.V. (2008). Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52 (4), 2249–2260.

Azamathulla, H. M., & Ghani, A. A. (2011). Genetic programming for predicting longitudinal dispersion coefficients in streams. Water resources management, 25(6), 1537-1544.

Bencala, K. E., & Walters, R. A. (1983). Simulation of solute transport in a mountain pool-and-riffle stream: A transient storage model. Water Resource Research. 19(3), 718-724.

Breiman, L. (2001). Random forests. Mach. Learn. 1 (5–32), 45.

Etemad-Shahidi, A., & Taghipour, M. (2012). Predicting longitudinal dispersion coefficient in natural streams using M5' model tree. Journal of Hydraulic Engineering. 138(6), 542–554.

Fischer, H.B. (1968). Dispersion predictions in natural streams. Journal of the Sanitary Engineering Division, ASCE, 94(5), 927-943.

Fischer, H.B. (1979). Mixing in Inland and Coastal Waters. Academic Press.

Graf, B. (1995). Observed and predicted velocity and longitudinal dispersion at steady and unsteady flow, Colorado River, Glen Canyon Dam to Lake Mead. Water Resources Bulletin. 31(2), 265-281.

Jirka, G.H. (2004). Mixing and dispersion in rivers. In: Greco, A., Carravetta, Morte, R.D. (Eds.), River Flow 2004. Taylor and Francis, London, pp. 13e27.

Julínek, T., & Říha, J. (2017). Longitudinal dispersion in an open channel determined from a tracer study. Environmental earth sciences, 76(17), 592.

Kashefipour, S. M., Falconer, R. A. (2002). Longitudinal dispersion coefficients in natural channels. Water Res 36: 1596–1608.

Li, X., Liu, H., Yin, M. (2013). Differential evolution for prediction of longitudinal dispersion coefficients in natural streams. Water Resour. Manage 27 (15), 5245–5260.

McQuivey, R. S., & Keefer, T. N. (1974). Simple method for predicting dispersion in streams. Journal of Enviromental Engineering. 100(4), 997-1011.

McQuivey, R. S., & Keffer, T. N. (1974). Simple method for predicting dispersion in streams. Journal of Environmental Engineering Division, ASCE, 100(4), 997-1011.

Noori, R., Deng, Z., Kiaghadi, A., & Kachoosangi, F. T. (2015). How reliable are ANN, ANFIS, and SVM techniques for predicting longitudinal dispersion coefficient in natural rivers? Journal of Hydraulic Engineering, 142(1), 04015039.

Nordin, C. F., & Sabol, G. V. (1974). Empirical data on longitudinal dispersion in rivers. U.S. Geological Survey Water Resource Investigation 20-74, Washington, D.C.

Pai, P., Hong, W. (2007). A recurrent support vector regression model in rainfall 827, 819-827.

Pourabadei, M., & Kashefipour, S. M. (2007). Investigation of flow parameters on dispersion coefficient of pollutants in canal. In Proceedings of the 6th International Symposium River Engineering, October (pp. 16-18).

Quinlan, J. R. (1992). Learning with continuous classes, in: Proceedings of the Australian Joint Conference on Artificial Intelligence, World Scientific, Singapore, pp. 343–348.

Rasmussen, C. E., Williams, C. K. I. (2006). Gaussian Processes for Machine Learning, MIT Press, Cambridge, MA.

Riahi-Madvar, H., Ayyoubzadeh, S. A., Khadangi, E., & Ebadzadeh, M. M. (2009). An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS. Expert Systems with Applications, 36(4), 8589-8596.

Rutherford, J. C. (1994). River Mixing, Wiley, Chichester, U. K, 347

Sahay, R. R. (2011). Prediction of longitudinal dispersion coefficients in natural rivers using artificial neural network. Environmental Fluid Mechanics, 11(3), 247-261.

Sahay, R. R., Dutta, S. (2009). Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. Hydrol Res 40(6):544–552.

Seo, I.W., & Cheong, T.S. (1998). Predicting longitudinal dispersion coefficient in natural streams. Journal of Hydraulic Engineering. 124(1), 25–32.

Tayfour, G., & Singh, V. P. (2005). Predicting longitudinal dispersion coefficient in natural streams by artificial neural network. Journal of Hydraulic Engineering, 131(11), 991–1000.

Taylor, G. (1954). The dispersion of matter in turbulent flow through a pipe. In: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 223(1155), The Royal Society, pp. 446–468 (May).

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres, 106, 7183–7192.

Vapnik, V. (1995). The Nature of Statistical Learning Theory. Springer, New York 187 pp.

Vapnik, V. (1998). Statistical Learning Theory. John Wiley & Sons, New York 740 pp.

Wang, Y., & Witten, I. H. (1997). Induction of model trees for predicting continuous classes. Proceedings European Conference on Machine Learning, Prague, Czechoslovakia. Also available as Working Paper 96/23, Department of Computer Science, University of Waikato.

Wang, Y., Huai, W., 2016. Estimating the longitudinal dispersion coefficient in straight natural rivers. J. Hydraulic Eng. (04016048)

Wang, Y., Huai, W., & Wang, W. (2017). Physically sound formula for longitudinal dispersion coefficients of natural rivers. Journal of Hydrology. 544, 511-523.