# Natural Rivers Longitudinal Dispersion Coefficient Simulation Using Hybrid Soft Computing Model

1st Sadeq Oleiwi Sulaiman
Dams and Water Resources Engineering Department, College of Engineering
University of Anbar
Anbar, Iraq
sadeq.sulaiman@uoanbar.edu.iq

2nd Ghassan Al-Dulaimi
Civil Engineering Department, College of Engineering
Philadelphia University
Amman, Jordan

3rd Hayder Al Thamiry
Water Resources Engineering Department, College of Engineering
University of Baghdad
Baghdad, Iraq

***Abstract*— The determination of longitudinal dispersion coefficient (LDC) of pollutants in stream contributes to several environmental and hydraulic engineering practices. Hence, providing an accurate and reliable methodology for predicting LDC is an essential process required water resources engineers. In this research, new hybrid soft computing model called deep neural network (DNN) coupled with genetic algorithm (GA), is developed to predict LDC using historical information attained from published researches in the literature. The GA is established as an evolutionary modeling phase to define the highly influencing hydraulic variables as an input combination attributes to compute the LDC. The hydraulic variables belonged to various stream all around the world, are utilized to build the modeling structure. The developed prediction model assessed using various statistical metrics to visualize its predictability. The proposed coupled predictive model validated with the core established research on the same application. In general, the model exhibited an excellent methodology for the environmental and hydraulic engineering aspects. Most importantly, the proposed model fulfilled the contribution of river engineering sustainability.**

***Keywords*—*longitudinal dispersion coefficient, hybrid model, prediction problem, natural rivers.***

## I. Background

Both international and national agencies responsible for water preservation have stated the need for river protection and sustainability to ensure the preservation of water quality and water resources [1], [2]. This has become necessary owing to the need to protect the health of the public who feed and drink from these water bodies. In water resource management, the simulation and estimation of water flow, sediments and contaminants are very significant and necessary [3]. A precise estimation can minimize the risk of environmental pollution both short- and long-term scale, and further enhance the impact of quality environmental engineering projects on the quality of water and its recourses. Although there are three main aspects of mixing and dispersion of water pollutants (i.e., lateral, longitudinal, and vertical), longitudinal dispersion predominantly dominates in water bodies [4].

## II. Research significant

An effective solving of most hydraulic problems demands an accurate longitudinal dispersion coefficient (LDC) estimation. Some of the problems that demand an accurate LDC estimation include problems in environmental engineering, estuaries, river engineering, intake designs, and during the assessment of the risk of discharging dangerous contaminants into river bodies [5], [6]. The process of estimating this coefficient is complex owing to the irregularities of natural channels in terms of configuration and shape, as well as their hydraulic conditions. Due to the limited availability river data, the quantification of these bathymetric parameters is usually a difficult task.

There is a need for a combined mixing and attenuating processes in natural rivers management due to the processes that usually leads to surface water pollution. Until now, the commonest and most effective method of managing river environments is the improvement of the river's ability to auto or self-clean. A usual practice in environmental engineering in recent times is the sinking of several agricultural and industrial remainders into natural rivers to oxidize and eliminate organic river contaminants. The controlling of surface water resource quality via sinking of pollutants into river bodies demands a logical and concise approach and demands a detailed information on the ways of discharging pollutants into river bodies, as well as a knowledge of how rivers mix, transport, and self-clean pollutants. Solid contaminants are subjected to several stages of mixing as river flow and are dispersed vertically, longitudinally and transversely via advection and dispersion processes.

## III. Literature review

The determination of LDC in natural rivers had been calculated using theoretical, experimental and soft computing models. The first theoretical procedure proposed by [7]. The research followed by implementing the proposal of Taylor on uniform flow by [8]. Another development established by [9] through combining a linear flow with the empirical formulation of the LDC. Based on the suggested detailed theory of [4], the simulation of LDC had been introduced empirically, in which it relies on fluid, hydraulic and geometry characteristics. In more descriptive manners, LDC is a function for channel depth (H), width (B), shear velocity (u*), water velocity (U), width depth ratio (B/H), relative shear velocity (U/u*), channel shape parameter (β), channel sinuosity (σ) it can be formulated as follows:

$$LDC = f(H, B, u^*, U, {}^{B}\!/_{H}, {}^{U}\!/_{u^*}, \beta, \sigma) \quad (1)$$

978-1-5386-6712-5/18/$31.00 ©2018 IEEE
DOI 10.1109/DeSE.2018.00056

The interest of applying soft computing models in solving the LDC problem has been the motivation of hydraulic engineering since early of the 21th century. This is with enthusiasm of establishing more reliable and cohesion modeling strategies for this natural problem that characterized with highly non-linear pattern. The core study of this manuscript is the one developed by Prof. Vijay [10]. The potential of ANN is proposed to predict the LDC. Note that the motivation that encourage the main authors is to develop more advance version of artificial neural network that can enhance the predictability accuracy. In another attempt, Prof. Vijay modeled the LDC using the applicability of fuzzy logic [11]. In 2009, three different optimization algorithm tried to configure the ANN model including feed forward backpropagation, radial basis function and generalized neural network, established for predicting natural stream LDC [12]. The feasibility of evolutionary programming models investigated for LDC prediction by [13]. Modern machine learning based on decision tree modeling proposed for predicting the same targeted variable [14]. Another study conducted on what is the reliability of various machine learning models including artificial neural network, support vector machine and adaptive neuro-fuzzy inference system for establishing an accurate prediction model for the LDC [15]. Most recently, The implementation of complementary modeling based on tuning the application of ANN via bee algorithm [16]. The research pattern since 2005 up to date indicated the magnificent of exploring new models for solving this problem where attract the author(s) of this work to bring new knowledge by integrating predictive model with variable optimization approach.

## IV. RESEARCH AIMS

In this manuscript, the author attempts to introduce new hybrid soft computing predictive model called GA-DNN. The developed DNN model coupled with nature-inspired optimization algorithm called genetic algorithm (GA). The concept of coupling the GA optimizer is to abstract the most correlated variables influencing the LDC magnitude. The employed engineering information collected from published research established by [10]. The results of the modeling are validated against the outcome of the traditional artificial neural network (ANN) model conducted on the same inspected problem conducted by [10].

## V. DEEP LEARNING NEURAL NETWORK

DNN model first proposed by [17]. The method proposed through the pretraining procedure as an advance step to the layer-by-layer training. The dimension of the data can be reduced by extracting features from the inputs to obtain a compact representation. The second step, samples are labeled by exporting the features to the next layer. Finally, the labeled data deployed to fine tune the network. There are two reasons attributable to the popularity of deep learning methods: (i) the issue of data overfitting can be addressed by the development of the big data analysis techniques; (ii) non-random initial values will be assigned to the network during the pre-training procedure prior to the unsupervised learning. Therefore, a faster coverage rate and a better local minimum can be achieved after the training process.

Classical DNN model included of closely embedded input, output, and several hidden layers. The input and hidden layers are directly connected and function together to weigh the input values to produce a new set of real numbers that will be transmitted to the output layer (Figure 1). Finally, the output layer, based on the transmitted values, classify or predict the outcome of the process.

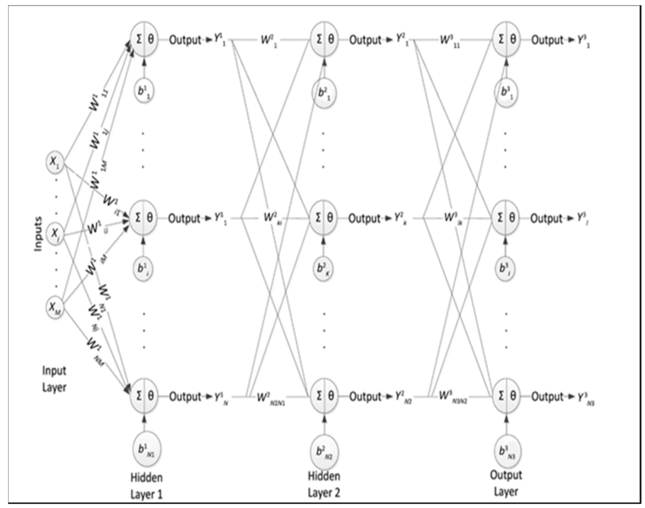


Fig. 1. The slandered architecture of deep neural network description.

The main merit of the DNN is that the deep multi-layer neural network is made up of several levels of nonlinearities which made them applicable in the representation of highly non-linear and/or highly-varying functions.

## VI. GENETIC ALGORITHM OPTIMIZER

GA is very well-known optimization technique that can classified as an evolutionary method based on biological process. The effectiveness of this optimization approach discussed comprehensibly in term of solving the nonlinearity and stochasticity by [18]. The main processes are involved the implementation of the heuristic GA are including reproduction of chromosomes, crossover, and mutation. Note these processes are applied to satisfied the probability of the discretization of the input variables that are coded into binary strings [19]. The GA processes integrated with the DNN predictive model are presented in Figure 2.

## VII. MODELING DEVELOPMENT

The hydraulic natural streams data are organized to be supplied for the coupled prediction model. The eight input parameters including $LDC = f(H, B, u^*, U, {}^{B}/_{H}, {}^{U}/_{u^*}, \beta, \sigma)$ are defined to determine the LDC. The aim of the GA optimizer is applied to abstract the highly influencing parameters starting from two parameters until including all the eight parameters. Table 1 presented the optimization selection procedure. It can be notice, seven input combinations are prepared. These input combinations are provided for the prediction matric as input attributes; whereas, all the attributes are corresponding to the LDC. The modeling of the presented regression problem validated via multiple statistical formulations including root mean square error (RMSE), mean absolute error (MAE), and the correlation coefficient ($R^2$).

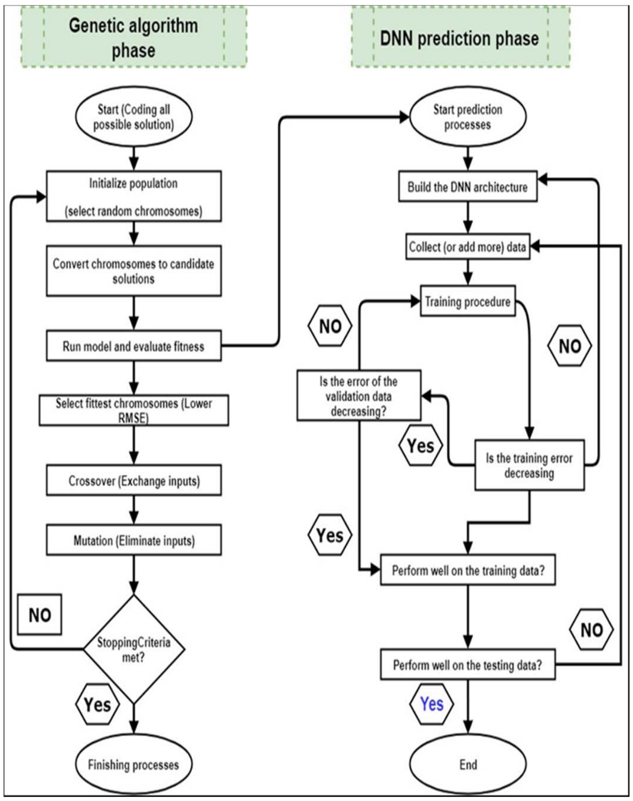


Fig. 2. The proposed hybrid genetic algorithm deep neural network (GA-DNN) predictive model.

TABLE I. THE OPTIMIZED INPUT VARIABLES USING GA OPTIMIZER FOR BUILDING THE DNN MODEL.

| *No. Inputs* | *Selected Parameters* |
|---|---|
| *2-input* | B (m), U (m/s), |
| *3-input* | B (m), U (m/s), $\sigma$ |
| *4-input* | B (m), U (m/s), $\beta, \sigma$ |
| *5-input* | B (m), H (m), U (m/s), B / H, $\sigma$ |
| *6-input* | B (m), H (m), U (m/s), u* (m/s), B / H, U / u* |
| *7-input* | *B (m), H (m), U (m/s), u* (m/s), B / H,* $\beta, \sigma$ |
| *8-input* | B (m), H (m), U (m/s), u* (m/s), B / H, U / u*, $\beta, \sigma$ |

## VIII. APPLICATION AND DISCUSSION

The aim of this research is to provide a reliable predictive model based on coupled machine learning model with variable input optimization. The modeling assessed using natural stream information presented the hydraulic properties. The archived results of the current research validated against the published research by [10]. In this way, the benchmark research is more reliable and trustful to be justify the current research. Based on the allocated correlated parameters to predict the LDC, the width (B) and water velocity (U) are the principal parameters that are affecting the LDC amount. This is totally agreed and reliable based on the conclusion of the based study conducted by professor Seo Won, [20]. The followed input combinations included various hydraulic characteristics. The optimal solution for the correlated parameters determined based on the accurate prediction performance indicators tabulated in Table 3. It is clearly can be observed that including four parameters (i.e., B (m), U (m/s), β,and σ) attained the perfect prediction accuracy. This is also reflected on the scatter plot illustrated in Figure 3. The ideal data collection around the fit line 45o realized for the third input combination, in which including four parameters. The best values of the (RMSE and MAE) archived (47.967 and 32.326). In comparison with the published research of Tayfur and Vijay, the RMSE improved by (64%). The result augmentation is highly acceptable and contribute to the knowledge of the hydraulic and environmental engineering aspect.

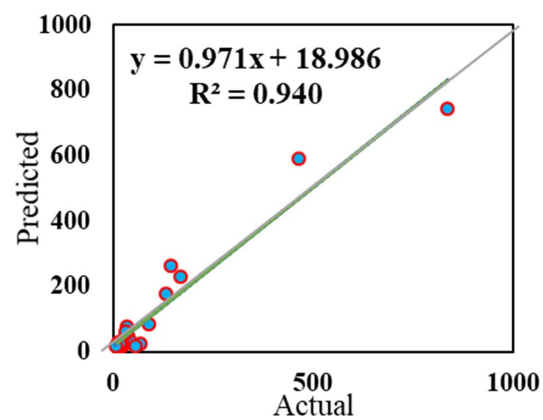


Fig. 3. The scatter plots including the correlation formulation between the actual and predicted LDC using four input variables

TABLE II. The modeling performance of predicting the LDC using all the determined input combinations.

| ***Models*** | ***RMSE*** | ***MAE*** | ***R2*** |
|---|---|---|---|
| *2-input* | 134.946 | 76.385 | 0.498 |
| *3-input* | 136.901 | 98.488 | 0.714 |
| *4-input* | 47.967 | 32.326 | 0.940 |
| *5-input* | 79.659 | 37.215 | 0.911 |
| *6-input* | 102.403 | 68.993 | 0.745 |
| *7-input* | 111.178 | 89.636 | 0.823 |
| *8-input* | 142.456 | 98.495 | 0.495 |

## IX. Conclusions

Performing a robust predictive model for LDC in natural streams can contributes to various environmental and river engineering aspects. The core obstacle in this engineering problem is the complexity of the correlation of the targeted variable with numerous river characteristics including stream velocity, physical properties, and hydrological features. In which make solving this problem required a reliable intelligence system to comprehend the internal mechanism perfectly. The current study dedicated on implementing relatively new soft computing hybrid model called GA-DNN for predicting LDC. The modeled data gathered from published studies over the past decades. The main merit distinguished this study is the simulation of real world information characterized by natural phenomena. The current research results evidenced the capability of the applied modeling approach over the existed literature methods. As future research to be devoted, different data set can be implemented to validate the generalization capacity of the applied model.

## References


[1] G. Lollino, M. Arattano, M. Rinaldi, O. Giustolisi, J. C. Marechal, and G. E. Grant, Engineering geology for society and territory – volume 3: River basins, reservoir sedimentation and water resources. 2015.

[2] J. I. Hedges and R. G. Keil, "Sedimentary organic matter preservation: an assessment and speculative synthesis," Mar. Chem., vol. 49, no. 2–3, pp. 81–115, 1995.

[3] K. Jaglal, "Contaminated Aquatic Sediments," Water Environ. Res., vol. 89, no. 10, pp. 1542–1560, 2017.

[4] Z.-Q. Deng, V. P. Singh, and L. Bengtsson, "Longitudinal Dispersion Coefficient in Straight Rivers," J. Hydraul. Eng., vol. 127, no. November, pp. 919–927, 2001.

[5] H. Riahi-Madvar, S. A. Ayyoubzadeh, E. Khadangi, and M. M. Ebadzadeh, "An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS," Expert Syst. Appl., vol. 36, no. 4, pp. 8589–8596, 2009.

[6] K. El Kadi Abderrezzak, R. Ata, and F. Zaoui, "One-dimensional numerical modelling of solute transport in streams: The role of longitudinal dispersion coefficient," J. Hydrol., vol. 527, pp. 978–989, 2015.

[7] G. Taylor, "The Dispersion of Matter in Turbulent Flow through a Pipe," Proc. R. Soc. A Math. Phys. Eng. Sci., vol. 223, no. 1155, pp. 446–468, 1954.

[8] J. W. Elder, "The dispersion of marked fluid in turbulent shear flow," J. Fluid Mech., vol. 5, no. 4, pp. 544–560, 1959.

[9] McQuivey, "Simple Method for Predicting Dispersion in Streams," Pochvozn. i Agrokhimiya, vol. 25, no. 5, pp. 62–68, 1974.

[10] G. Tayfur and S. Vijay, "Predicting Longitudinal Dispersion Coefficient in Natural Streams by Artificial Neural Network," J. Hydraul. Eng., vol. 131, no. 11, pp. 991–1000, 2005.

[11] G. Tayfur, "Fuzzy, ANN, and regression models to predict longitudinal dispersion coefficient in natural streams," Nord. Hydrol., vol. 37, no. 2, pp. 143–164, 2006.

[12] Z. F. Toprak and H. K. Cigizoglu, "Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods," Hydrol. Process., vol. 22, no. 20, pp. 4106–4129, 2008.

[13] H. M. Azamathulla and A. A. Ghani, "Genetic Programming for Predicting Longitudinal Dispersion Coefficients in Streams," Water Resour. Manag., vol. 25, no. 6, pp. 1537–1544, 2011.

[14] A. Etemad-Shahidi and M. Taghipour, "Predicting Longitudinal Dispersion Coefficient in Natural Streams Using M5´ Model Tree," J. Hydraul. Eng., vol. 138, no. 6, pp. 542–554, 2012.

[15] R. Noori, Z. Deng, A. Kiaghadi, and F. T. Kachoosangi, "How Reliable Are ANN , ANFIS , and SVM Techniques for Predicting Longitudinal Dispersion Coefficient in Natural Rivers ?," J. Hydraul. Eng., vol. 142, no. 1, pp. 1–8, 2015.

[16] M. J. Alizadeh, A. Shabani, and M. R. Kavianpour, "Predicting longitudinal dispersion coefficient using ANN with metaheuristic training algorithms," Int. J. Environ. Sci. Technol., vol. 14, no. 11, pp. 2399–2410, 2017.

[17] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006.

[18] D. E. Goldberg, B. Korb, and K. Deb, "Messy Genetic Algorithms : Motivation , Analysis , and First Results," Engineering, vol. 3, pp. 493–530, 1989.

[19] D. Golmohammadi, R. C. Creese, H. Valian, and J. Kolassa, "Supplier Selection Based on a Neural Network Model Using Genetic Algorithm," IEEE Trans. Neural Networks, vol. 20, no. 9, pp. 1504–1519, 2009.

[20] W. Seo and S. Cheong, Tae, "Predicting Longitudinal Dispersion Coefficient in Natural Streams," J. Hydraul. Eng., vol. 124, no. January, pp. 25–32, 1998.