Granular computing-neural network model for prediction of longitudinal dispersion coefficients in rivers

Behzad Ghiasi, Hossein Sheikhian, Amin Zeynolabedin
and Mohammad Hossein Niksokhan

ABSTRACT

Successful application of one-dimensional advection-dispersion models in rivers depends on the accuracy of longitudinal dispersion coefficient (LDC). In this regards, this study aims to introduce an appropriate approach to estimate LDC in natural rivers that is based on a hybrid method of granular computing (GRC) method and artificial neural network (ANN) model (GRC-ANN). Also, Adaptive Neuro-Fuzzy Inference System (ANFIS) and ANN models were developed to investigate the accuracy of three credible Artificial Intelligence (AI) models and the manner of these models in different LDC values. By Comparing with empirical models developed in other studies, the results revealed the superior performance of GRC-ANN for LDC estimation. The sensitivity analysis of the three intelligent models developed in this study is done to determine the sensitivity of each model to its input parameters specially the most important ones. The sensitivity analysis results showed that the W/H parameter has the most significant impact on the output of all three models in this research.

Key words | ANFIS, ANN, artificial intelligence, GRC-ANN, longitudinal dispersion coefficient

Behzad Ghiasi
Mohammad Hossein Niksokhan (corresponding author)
School of Environment, College of Engineering,
University of Tehran,
Iran
E-mail: niksokhan@ut.ac.ir

Hossein Sheikhian
Department of Geospatial Information Systems,
College of Engineering,
University of Tehran,
Iran

Amin Zeynolabedin
School of Civil Engineering, College of Engineering,
University of Tehran,
Iran

INTRODUCTION

Dispersion is a dominant process for mixing pollutants in narrow water bodies such as natural rivers. In natural rivers, due to lower width and depth compared to length, usually – except for near-field regions close to source of pollution – hydrodynamic and water quality variations occur along the length of the river. In this regard, flow and mixing in rivers in the far-field (after cross-sectional mixing) can be considered as one-dimensional (1D). Therefore, for river water quality modeling in the far field, lateral and vertical mixings are less relevant than longitudinal dispersion (Deng et al. 2001). The lateral difference of longitudinal velocity causes longitudinal dispersion and the intensity of this mixing is determined by longitudinal dispersion coefficient (LDC) (Seo & Cheong 1998). Due to the importance of LDC, many studies have been done on this parameter.

Many researches focused on the estimation of LDC, have led to introducing different solutions which are based on integral methods (Fischer 1967; Fischer et al. 1979; Deng et al. 2001), empirical formula (Seo & Cheong 1998; Kashefipour & Falconer 2002; Azamathulla & Ghani 2011; Zeng & Huai 2014; Noori et al. 2017a), dye tracing

measurements (Clark et al. 1996), and also more sophisticated methods based on artificial intelligence (AI) methods (Riahi-Madvar et al. 2009; Sahay & Douta 2009; Parsaie & Haghiabi 2015; Noori et al. 2017b; Balf et al. 2018).

Generally, the estimation of LDC in natural rivers is very complicated due to high variations of hydraulic and geometric characteristics, which significantly affect this parameter. Regarding the application of basic method proposed by Fischer (1979) to determine LDC, it is required to have a detailed transverse profile of velocity and geometry of the river. Otherwise, the lateral profile of longitudinal velocity should be assumed (Zeng & Huai 2014). Also, many researchers have demonstrated that some tracers (e.g., chemical or fluorescent tracers) are non-conservative in natural waters (Smart & Laidlaw 1977). Toxicity of by-products and the high cost of soluble tracers are other flaws related to the application of dye tracing measures (Clark et al. 1996). The mentioned shortcomings have limited the integral-empirical-dye tracer methods for LDC determination in rivers. Regarding AI methods, although they have recorded acceptable performance for LDC estimation, their uncertainty can limit the application of results for

water quality management and control strategies in rivers (Noori *et al.* 2015).

In this study, Considering the disadvantages and limitations of proposed methods for the prediction of LDC, an alternative approach was introduced that is based on a hybrid method of granular computing (GRC) approach and artificial neural network (ANN) model. GRC is a general theory of information granulation and processing, acting upon different levels of information discretization (Bargiela & Pedrycz 2003). Yao (2001) defined the GRC as a natural data-mining tool to solve real-world problems by analyzing, presenting and decision-making based on available information about the problem. GRC model was first introduced by Zadeh (1996), focusing on interval-based information granulation and was developed to conceptualize granulation by Lin *et al.* (2002), which operates based on the information table. Pedrycz *et al.* (2008) discussed different combinations of fuzzy sets and GRC, proving that fuzzy sets can be applied to the presentation of GRC outputs, or as the input granules of the GRC model. Qian *et al.* (2010) extended Pawlak's rough set model (Pawlak 1982) to a multi-granulation rough set model (MGRS), where granules are extracted in a hierarchical structure. GRC approach has been applied in the fields of computer science and engineering such as machine learning (Yao & Yao 2002a, 2002b), knowledge reduction (Wu *et al.* 2009), intelligent data analysis (Chen & Yao 2006) and intelligent social networks (Yager 2008). However, different types of GRC have been proposed to improve its capabilities including GRC based interval analysis (Moore 1966), integration of rough set theory and GRC approach (Pawlak 1982) and embedding fuzzy sets into GRC (Zadeh 1997). Integration of GRC with interval analysis and rough set theory is mostly used for simplifying the problem, which is not the matter of interest in this research and the theory of fuzzy-granulation is a more advanced and flexible method. However, it suffers from a shortage of transparency and validity of the results. In order to overcome the limitations of the mentioned methods, integration of ANN and GRC approaches (GRC-ANN) is proposed here. This approach provides a model with unique capabilities in comparison to fuzzy, rough set theory and interval information granulations. Reliability, transparency, efficiency, and robustness of the model can be considered as the advantages of GRC-ANN approach application, which can lead to the more reliable and accurate estimation of LDC.

The main goal of this study is to provide a new model to accurately estimate the LDC by GRC-ANN model and comparing GRC-ANN, ANN, ANFIS models with other studies.

Finally, sensitivity analysis is done to check models performances and how parameters affect the results.

METHODOLOGY

GRC-ANN model

GRC model is a method of data processing based on multiple levels of data granularity (Pawlak 1982; Zadeh 1997). In this method, the whole data is divided into granules and clusters or subsets as fundamental elements of granular computing (Yao 2001, 2008). Granulation categorizes individual elements of the whole universe of data into clusters, based on relevant information, in the form of an Information Table (Yao 2008). These granules of data are formed according to the existing similarity between objects to put them in different granules, and then extract the prediction rules (Hobbs 1990; Yao 2001; Bargiela & Pedrycz 2003; Keet 2008; Yao 2008). Also, Information Table is the presentation of information in the form of objects described by a set of valid attributes, defined by Equation (1) (Yao & Zhong 2002):

$$S = (U, A_t, L, \{V_a | a \in A_t\}, \{F_a | a \in A_t) \tag{1}$$

where U are the whole objects used to solve the problem, A_t is a set of attributes for the mentioned objects, L is a language of defining object attributes, V_a is the set of valid values for the attribute a and the F_a is the function relating the objects to their attributes.

In GRC approach, for prediction, a set of rules is extracted in the form of IF-THEN: 'if an object satisfies ϕ, then the object satisfies ψ.', Where concepts ϕ and ψ are sets of attribute-values for a set of objects and assigned output value, respectively. In the process of rule extraction, GRC algorithm applies measurements on granules formed in the procedure in order to select the best set of possible rules and to prioritize the rules. Generality (G), absolute support (AS), coverage (CV) and conditional entropy (CE) are four parameters used to extract the optimum rules and to construct the best model.

The contingency table in Table 1 represents the quantitative information about the rule $\varphi \Rightarrow \psi$. The values in the four cells are not independent. They are linked by the constraint $a + b + c + d = n$. The 2×2 contingency table has been used by many authors for representing information of rules (Gaines 1991; Silverstein *et al.* 1998).

The generality of concept ϕ displays the relative size of the constructive granule of the concept ϕ, as defined in

Table 1 │ Quantitative information of the $\varphi \Rightarrow \psi$ rule

	ψ	$\neg\psi$	Totals
Φ	$\|m(\varphi) \cap m(\neg\psi)\|$	$\|m(\varphi) \cap m(\neg\psi)\|$	$\|m(\varphi)\|$
$\neg \varphi$	$\|m(\neg\varphi) \cap m(\psi)\|$	$\|m(\neg\varphi) \cap m(\neg\psi)\|$	$\|m(\neg\varphi)\|$
Totals	$\|m(\psi)\|$	$\|m(\neg\psi)\|$	$\|U\|$

Equation (2) (Pawlak 1982), where $|m(\phi)|$ is the size of granule forming concept ϕ and $|U|$ is the size of the whole universe.

$$G(\phi) = \frac{|m(\phi)|}{|U|} = \frac{\text{number of data records included in } \phi}{\text{all records count}} \quad (2)$$

A concept is more general if it covers more instances of the universe. If $g(\varphi) = \alpha$, then $(100\alpha)\%$ of objects in U satisfy φ. The quantity may be viewed as the probability of a randomly selected element satisfying φ. Obviously, we have $0 \leq g(\varphi) \leq 1$.

AS which is the conditional probability where a randomly selected object satisfies both ϕ and ψ (Yao & Yao 2002a, 2002b), can be achieved through Equation (3).

$$AS(\phi \rightarrow \psi) = \frac{|m(\phi \wedge \psi)|}{|m(\phi)|}$$
$$= \frac{\text{number of data records included in both } \phi \text{ and } \psi}{\text{number of data records included in } \phi} \quad (3)$$

It may be interpreted as the degree to which φ implies ψ. If $as(\psi|\varphi) = \alpha$, then $(100\alpha)\%$ of objects satisfying φ also satisfy ψ. It is in fact the conditional probability of a randomly selected element satisfying ψ given that the element satisfies φ. In set-theoretic terms, it is the degree to which $m(\varphi)$ is included in $m(\psi)$. Clearly, $AS(\psi|\varphi) = 1$, if and only if $m(\varphi) \subseteq m(\psi)$.

CE, represented by $H(\psi|\phi)$ that reveals the uncertainty of formula based on the formula is defined by Equation (4) (Yao & Yao 2002a, 2002b).

$$H(\psi|\phi) = -\sum_{i=1}^{n} p(\psi_i|\phi) \log(p(\psi_i|\phi)) \quad (4)$$

where $p(\psi_i|\phi) = \frac{|m(\phi \cap \psi_i)|}{|m(\psi)|}$. *CV* denotes the conditional probability of a randomly selected object to satisfy ϕ while

satisfying ψ (Yao & Yao 2002a, 2002b).

$$CV(\phi \rightarrow \psi) = \frac{|m(\phi \cap \psi)|}{|m(\psi)|}$$
$$= \frac{\text{number of data records included in both } \phi \text{ and } \psi}{\text{number of data records included in } \psi} \quad (5)$$

Unlike the absolute support, the change of support varies from -1 to 1. One may consider g(ψ) to be the prior probability of ψ and as ($\psi|\varphi$) the posterior probability of ψ after knowing φ. The difference of posterior and prior probabilities represents the change of our confidence regarding whether φ actually confirms ψ. For a positive value, one may say that φ confirms ψ; for a negative value, one may say that φ does not confirm ψ.

Rule extraction starts with extracting and grouping instances possessing the same set of attribute values in the information table to construct concepts. Decision attributes are then grouped in the form of concepts showing the output classes. GRC applies the same measures to construct high-quality classification rules in the form of IF-THEN statements (Yao 2001). GRC approach extracts the rules from the dataset based on *CE* and *AS* so that rules with minimum *CE* and maximum *AS* values are extracted. To form a granular decision tree, the priority of rules in the tree is determined based on higher *G* and *CV*. The rule extraction procedure using GRC for LDC modelling has been illustrated in first section of Figure 1.

This paper proposes an integrated model of GRC rule generation and ANN (GRC-ANN model). It considers the relationship among extracted rules from GRC rule extraction algorithm in the previous section, by applying a group of rules satisfied using the input pattern of data, ranked by the rule importance measurements undertaken by the GRC. This approach allows the ANN to use the mentioned rule quality parameters to construct the approximator structure, instead of time-consuming iterative learning procedure.

To integrate GRC and ANN, existing relations in the data are extracted in the form of classification rules. A classifier network is then constructed from the rules considering their quality parameters computed during the rule extraction procedure. In this way, the network is constructed using the rule quality parameters.

GRC-ANN classifier addresses the drawbacks and limitations of both granular computing and neural network methods. To do so, it embeds the rules extracted by granular computing into a feed forward multi-layer artificial neural network structure. It significantly contributes to the
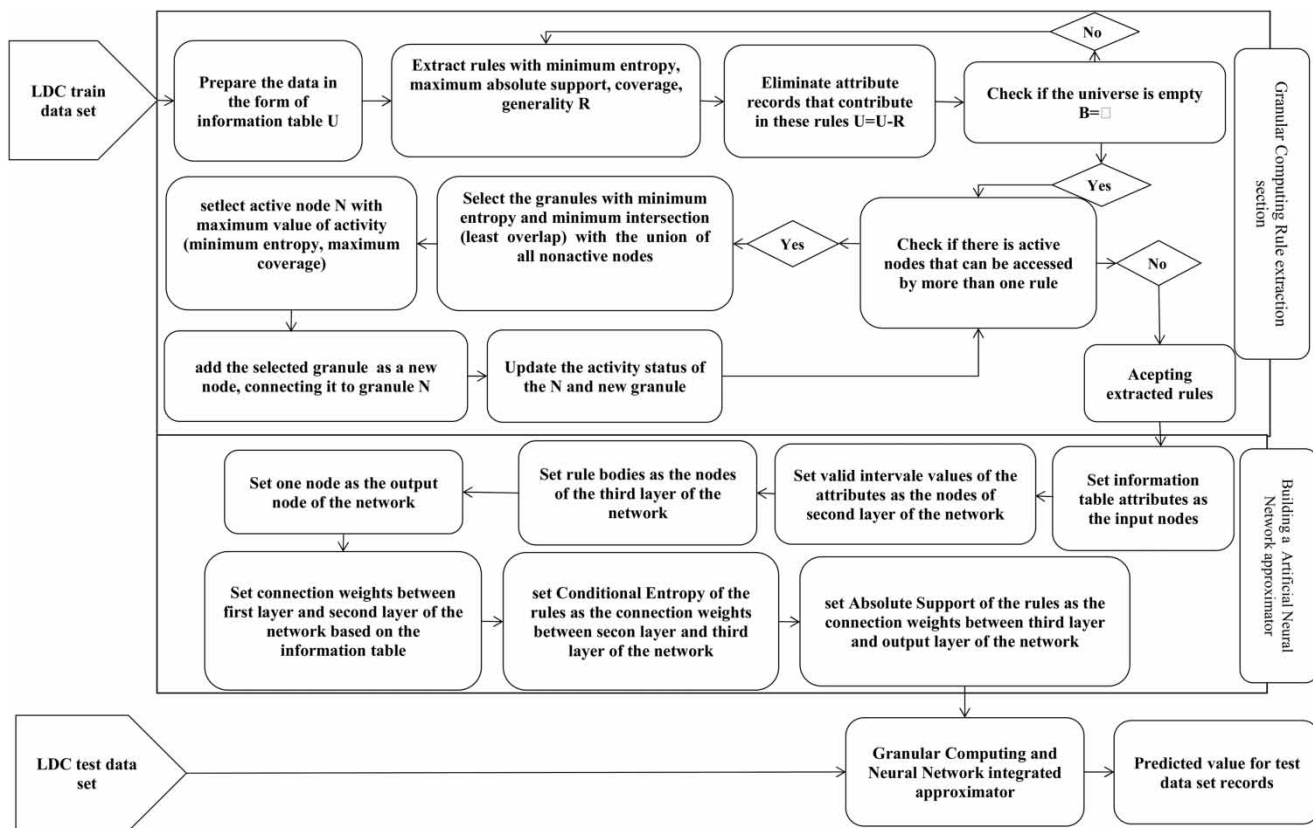
Figure 1 | Procedure of structuring GRC-ANN from the LDC data.

transparency of the resulting structure. Layers, neurons, and connection weights are all accompanied by physical descriptions. The network uses the information provided by a set of rules to decide about the presented pattern and take advantage of all available information. By using this approach, no input pattern remains unclassified. In this way, learning is done by granular computing rather than by the network itself. The constructed network does not contain any connection or node without a clear description, despite the conventional neural networks containing hidden neurons and connection weights obtained from a black-box learning algorithm. (Sheikhian *et al.* 2017). The proposed method gives an exact and robust prediction, despite neural networks employing a learning algorithm which relies on the initial weights, causing different network structures after each learning instance. In the proposed model, the number of inputs, outputs, hidden neurons and the pattern of the connections is determined automatically from the data set.

The second section of Figure 1 algorithm is suggested to integrate GRC rule generation and ANN model.

The quality of prediction depends on the rule set quality and accuracy, including both test dataset adequacy and

quality, and rule extraction algorithm reliability and robustness. The structure of the proposed procedure is illustrated in Figure 1.

The ANN used in this paper comprises four layers integrated with rule parameters, including input layer, pattern layer, rule firing layer and output layer (Figure 2). The number of nodes used in input layer equals to attributes of data records. Pattern layer nodes contain normalized and quantized valid values of the criteria in input layer. The classification rules are embedded in the rule-firing layer and the output node is one node which contains value of the output parameter.

In this model, the connections among nodes in the input and pattern layer are determined based on valid values of the parameters in the Information Table. The connection weights would be considered one if the value of the node in the pattern layer is valid for the node in the input layer and otherwise it would be zero. The connection weights between the pattern layer and rule-firing layer depend on whether the nodes in the pattern layer are related to a specified rule or not. It is based on the *CE* of the rules which shows the degree of certainty in the utilized rules. The
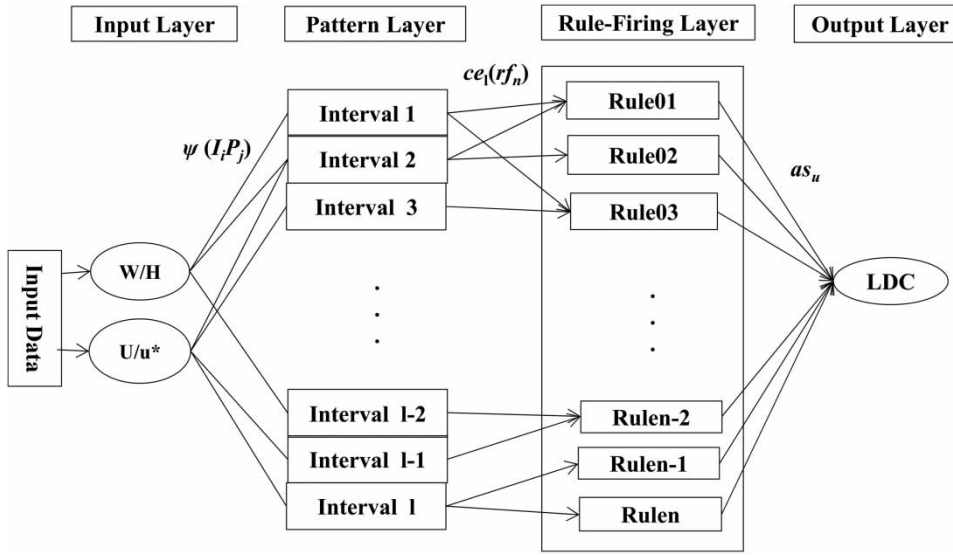
Figure 2 | Structure of the integrated GRC-ANN approach.

links between the rule-firing layer and output nodes are defined by *AS* of the rules that shows the degree of confidence provided by the supporting rules for the output class or value. The output of the rule-firing layer are used to compute the overall output of the network as follow:

$$LDC = \frac{\sum_{i=1}^{n} F(rf_i)w_i}{n_{fired}} = \frac{as_1 F(rf_1) + as_2 F(rf_2) + \cdots + as_n F(rf_n)}{n_{fired}}$$

(6)

where LDC is the predicted value for the input pattern, as_i is the support of the ith fired rule used as the weight of connection between the rule-firing layer and the output layer, n_{fired} is the number of rules fired and $F(rf_i)$ is the output of rule-firing layer nodes obtained through Equation (7).

$$F(rf_i) = ce_1(rf_i)\psi(I_1P_1)O_i + ce_2(rf_i)\psi(I_2P_2)O_i + \cdots + ce_l(rf_i)\psi(I_mP_l)O_i$$

(7)

In Equation (7), $ce_i(rf_i)$ is the weight of connection between the nodes in the pattern layer and the ith node in the rule-firing layer, which is the *CE* of the rule corresponding to that node. $\psi(I_iP_j)$ is the weight of connection between the ith node in the input layer and jth node in the pattern layer, which is Boolean. Thus, $\psi(I_iP_j)$ is unit if the jth node in the pattern layer is a valid interval for the value of the ith node in the input layer, and it is zero otherwise. Finally, O_i is the output of the ith rule fired by the presented pattern of the dataset. The final output value would be

obtained as:

$$LDC = \frac{\begin{aligned}&as_1\{ce_1(rf_1)\psi(I_1P_1)O_1 + ce_1(rf_1)\psi(I_2P_1)O_1 + \cdots \\ &+ce_l(rf_1)\psi(I_mP_l)O_1\} \\ &+as_2\{ce_1(rf_2)\psi(I_1P_1)O_2 + ce_2(rf_2)\psi(I_2P_1)O_2 + \cdots \\ &+ce_l(rf_2\psi)(I_mP_l)O_2\} + \cdots \\ &+as_n\{ce_1(rf_n)\psi(I_1P_1)O_n + ce_2(rf_n)\psi(I_2P_1)O_n + \cdots \\ &+ce_l(rf_n\psi)(I_mP_l)O_n\}\end{aligned}}{n_{fired}}$$

(8)

In the process of prediction, a set of rules satisfied by an input pattern of the data, are triggered that is followed by grouping these rules according to their output partition classes. For each output partition class, an activation strength value is calculated based on its triggered rules and their parameters computed by GRC. The output will be the average of the center of partitions with maximum activation values. In this way, no object remains unclassified. Quality of the classification depends upon the accuracy and adequacy of the extracted and pruned rule set, quality of the data set and reliability and accuracy of the rule extraction method.

Considering the provided description, it is evident that the network is independent of the number of input criteria and the number of extracted rules.

Sensitivity analysis

A sensitivity analysis for the prediction of longitudinal dispersion coefficient in dimensionless format is carried out

Uncorrected Proof

6   B. Ghiasi *et al.* │ GRC-ANN to predict longitudinal dispersion coefficient   **Water Science & Technology** │ in press │ 2020

for the mean values of input and output parameters. The analysis is based on the assumption that errors are independent of changes in inputs. Two indices are used in literature (Sattar & Gharabaghi 2015) named S_c which is marginal sensitivity coefficient and S_n which is normalized sensitivity coefficient and are defined by below equations:

$$S_c = \frac{\Delta X}{\Delta Y} \tag{9}$$

$$S_n = S_c \frac{X}{Y} \tag{10}$$

where output ΔY is the deviation in Y caused by the deviation in the input X. Investigation of the sensitivity of the error is carried out by incrementing each input parameter by 10%.

Models evaluation

Three common indices including determination of coefficient (R^2), root mean square error (RMSE) and mean absolute error (MAE) were used to evaluate the models' performance. The best value for R^2 is equal to one and for the two other indices is zero. However, these criteria show the average error in models and do not give any information about the error distribution. Hence, it is important to test the models' performance using some other criteria such as threshold statistics (*TS*) (Equation (11)) (Jain & Indurthy 2003) and developed discrepancy ratio (DDR) (Equation (12)) (Noori *et al.* 2010). In the DDR index, the higher crowns of each model show its better performance in estimating.

$$TS_x = \frac{n_x}{N} \times 100\% \tag{11}$$

where n_x is the total number of points that the absolute value of the relative error is less than x%.

$$DDR = \left(\frac{Predicted_{Value}}{Observed_{Value}} \right) - 1 \tag{12}$$

Dataset

In this research, a dataset including 100 patterns consisting of hydraulic and geometry characteristics measured in rivers (Seo & Cheong 1998; Noori *et al.* 2011) was used to calibrate and verify the GRC-ANN model (The details of dataset is available in Appendix B, Table 1). These data sets were collected from published works of Seo & Cheong (1998), Kashefipour & Falconer (2002), Tayfur (2006), Toprak & Cigizoglu (2008), and Etemad-Shahidi & Taghipour (2012). Hydraulic characteristics were the cross-sectional average velocity (U) and shear velocity (u*) and also geometric characteristics included channel width (W) and the flow depth (H). To eliminate the possible inconsistency due to different data dimensions, the dimensionless parameters W/H, U/u* and LDC/Hu* were used for calibration and verification of GRC-ANN model. From the existing dataset, 70 data randomly selected for calibration and remaining were used for validation of the results. In this dataset, W and H parameters had more variability than U and u*. LDC measurement data presented higher variability. Therefore, it can be concluded that the LDC treat as a nonlinear and complicated parameter because of its high variability.

RESULT AND DISCUSSION

GRC-ANN model evaluation

The data used in GRC-ANN model must be ranged between (0,1). Using $x = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})}$ the input and outputs data were rescaled in which $x_{\min}$ and $x_{\max}$ are the minima and maximum values in the datasets. Normalized calibration data were used as the input of the GRC-ANN model in the training procedure. The training step is the best network structure describing the relations between the input data was determined based on quality measurements applied by the model such as generality and absolute support. Network design starts with rule induction through GRC algorithm which consists of two main sections: extracting the set of possible rules and selecting the high-quality rules. Quality of the rules is assessed based on the measurements undertaken by GRC which includes Generality, Absolute Support, Coverage, and Entropy. Higher values of Absolute Support, Coverage, and Generality indicated that the rules are more desirable, while higher Entropy values show the opposite effect. The GRC algorithm used in this research applies the threshold of maximum 0.5 for entropy and minimum 0.7 for Absolute Support to achieve high-quality rules. In the case of Generality and Coverage, high values cannot be expected except in some cases, but rules having higher values of these parameters gain a priority in the set of rules. Four out of 43 extracted rules are shown in Table 2.

Table 2 | GRC rules

Rule	Generality	Absolute Support	Coverage	Entropy
If $0.41 < U/u^* < 0.78$ and $0.08 < W/H < 0.21$ Then $0.01 < LDC/Hu^* < 0.06$	0.08	1	0.6	0.3
If $0.25 < U/u^* < 0.41$ and $0.31 < W/H < 0.49$ Then $0.021 < LDC/Hu^* < 0.43$	0.06	0.9	0.25	0.1
If $0.25 < U/u^* < 0.33$ and $0.07 < W/H < 0.21$ Then $0.01 < LDC/Hu^* < 0.06$	0.11	0.91	0.1	0.06
If $0.008 < U/u^* < 0.12$ and $0.16 < W/H < 0.36$ Then $0.006 < LDC/Hu^* < 0.024$	0.09	0.77	0.16	0.32

The induction process can be briefly described as follows. Based on a measure of association between two partitions, the universe is divided into partitions. If a partition is not a subset of a user defined class or data-driven class, it is further divided by using another measure. The process continues until a set of rules are found that correctly classify all records. In a granule network, each node contains a subset of objects. A larger granule is divided into smaller granules by an atomic formula. In other words, the smaller granule is obtained by selecting those objects of the larger granule that satisfy the atomic formula. The set of the smallest granules thus form a conjunctively definable coverage of the universe. Atomic formulas define basic granules which are the basis for the granule network constructed from the data. Each node in the granule network is a conjunction of some basic granules, and thus a conjunctively definable granule. The algorithm is a heuristic search algorithm. The measures discussed can be used to define different fitness functions for the algorithm. After forming the initial network, a pruning process is undertaken to reduce the number of rules while maintaining useful information. In this process, granular computing measures are used to prune the rules not satisfying minimum requirements of a rule to be included in the classification rule set. For example, in this table GRC-ANN structure is constructed based on the set of qualified rules with the constraint of covering the data set values. The acquired structure was examined against the expected values to ensure that the existing relations between the data patterns are perfectly reflected in the structure of the data. In this regard, the best network structure, which can be built upon the GRC quality measurements, were formed and tested against the 30 normalized instances and were used as the test data set.

Figure 3 shows the quality measurements for the extracted rules. Because of the considered thresholds for AS and CE parameters, they fluctuate between the defined range. In the case of G parameter, most of the rules acquired low values because of the complexity of the problem that requires detailed rules to cover different situations.

Comparison

In order to compare acquired GRC-ANN results concerning previously implemented models, the GRC-ANN results are compared with the outputs of the empirical equations applied to test data of this research. Table 3 shows the detailed result of the comparison.

Table 3 shows that the empirical methods extravagantly lack the accuracy of the GRC-ANN model. This can be due to applying smaller data set compared to GRC-ANN model. As a result, Deng *et al.* (2002) method which showed the best performance amongst empirical methods had approximately three and four times more RMSE and MAE values compared to GRC-ANN model. Considering the R^2, empirical methods reached the value of 0.8349 in the best case, which is the Deng *et al.* (2002) method, while the GRC-ANN value of R^2 was approximately 0.9824. The GRC model has also outperformed the new methods proposed by Disley *et al.* (2015) and Sattar & Gharabaghi (2015). GRC-ANN RMSE and MAE are calculated 28.97 and 10.08 respectively which are better than mentioned methods while R^2 value is computed 0.982 which is slightly higher.

Predicted values of the mentioned models were compared with true measured LDC values. Figure 4 demonstrates this comparison for the calibration and test data. Apart from some anomalies, ANFIS model outputs were significantly closer to the desired values and ANN model was not capable of precisely pursuing the pattern of measured values. Among the implemented models, GRC-ANN perfectly estimated the measured LDC values, which verifies the accuracy and reliability of the GRC-ANN model once more.

Most of the AI methods are unstable in estimating low or high values of LDC (Noori *et al.* 2017a). Some of these models ignore over 100 LDC values to exaggerate about the results, where in some others, estimated LDC least value is extravagantly higher than the measured least value. To sort this out, stability of the three implemented models were examined in estimating LDC values of higher than 200. Evidence for this can be found in Table 4.
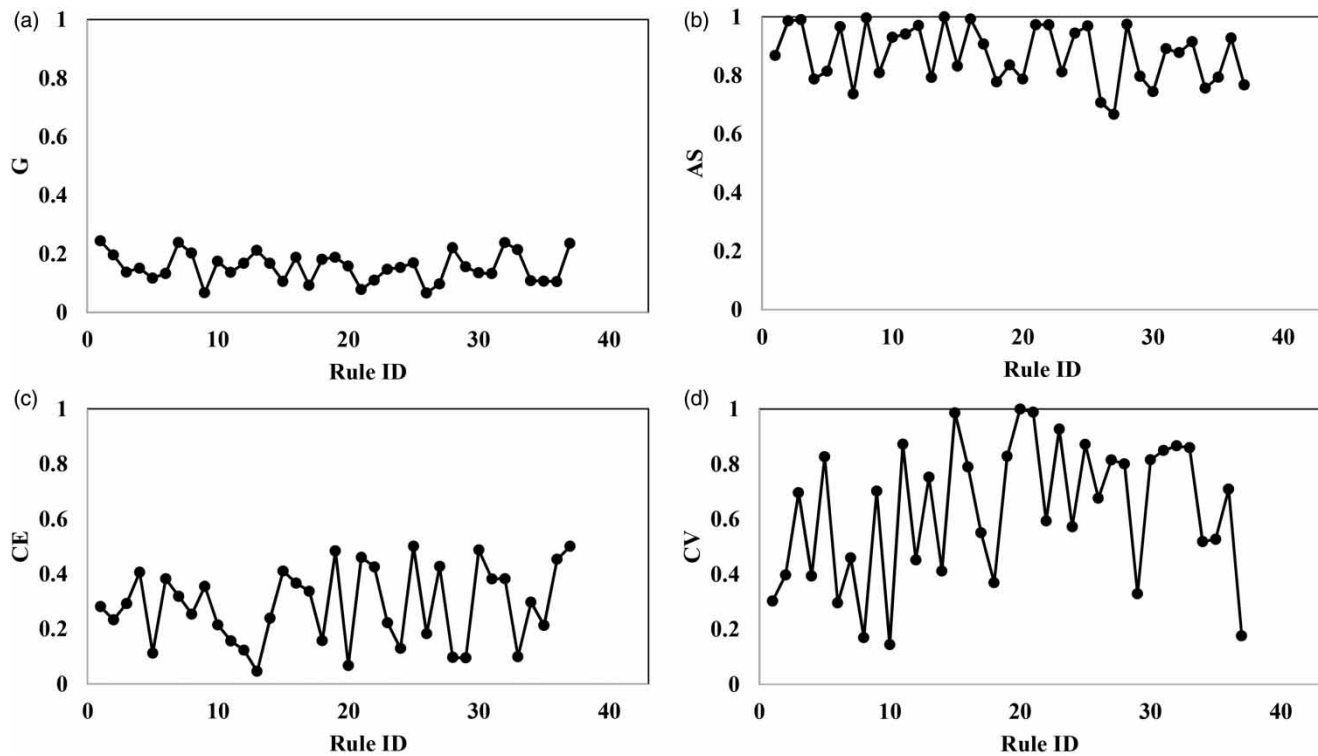
Uncorrected Proof

8   B. Ghiasi *et al.* │ GRC-ANN to predict longitudinal dispersion coefficient   **Water Science & Technology** │ in press │ 2020

Figure 3 │ Quality parameters of the extracted rules from GRC-ANN model.

Table 3 │ Comparison with empirical models

| Reference	Equation	RMSE	MAE	R²
Fischer *et al.* (1979)	$\frac{LDC}{Hu^*} = 0.011\left(\frac{W}{H}\right)^2\left(\frac{U}{u^*}\right)^2$	456.46	179.23	0.6185
Seo & Chang (1998)	$\frac{LDC}{Hu^*} = 5.195\left(\frac{W}{H}\right)^{0.620}\left(\frac{U}{u^*}\right)^{1.428}$	94.68	55.28	0.8180
Deng *et al.* (2001)	$\frac{LDC}{Hu^*} = 0.15\left(\frac{1}{8\varepsilon}\right)\left(\frac{W}{H}\right)^{1.667}\left(\frac{U}{u^*}\right)^2$ $\varepsilon = 0.145 + \left(\frac{1}{3520}\right)\left(\frac{W}{H}\right)^{1.38}\left(\frac{U}{u^*}\right)$	82.83	42.31	0.8349
Kashefipoor & Falconer (2002)	$\frac{LDC}{Hu^*} = 10.612\left(\frac{U}{u^*}\right)^2$	97.47	46.44	0.7957
Sahay & Douta (2009)	$\frac{LDC}{Hu^*} = 2\left(\frac{W}{H}\right)^{0.96}\left(\frac{U}{u^*}\right)^{1.25}$	92.06	45.36	0.7897
Zeng & Huai (2014)	$\frac{LDC}{HU} = 5.4\left(\frac{W}{H}\right)^{0.7}\left(\frac{U}{u^*}\right)^{0.13}$	105.65	48.10	0.8320
Disley *et al.* (2015)	$\frac{LDC}{Hu^*} = 3.563\left(\frac{W}{H}\right)^{0.6776}\left(\frac{U}{u^*}\right)^{1.0132}(Fr)^{-0.4117}$	108.79	51.20	0.9103
Sattar & Gharabaghi (2015)	$\frac{LDC}{Hu^*} = \frac{2.9\times4.6^{\sqrt{Fr}}}{\sqrt{Fr}}\left(\frac{W}{H}\right)^{0.5-Fr}\left(\frac{U}{u^*}\right)^{1+\sqrt{Fr}}$	87.17	44.47	0.8891
GRC-ANN		28.97	10.08	0.9824
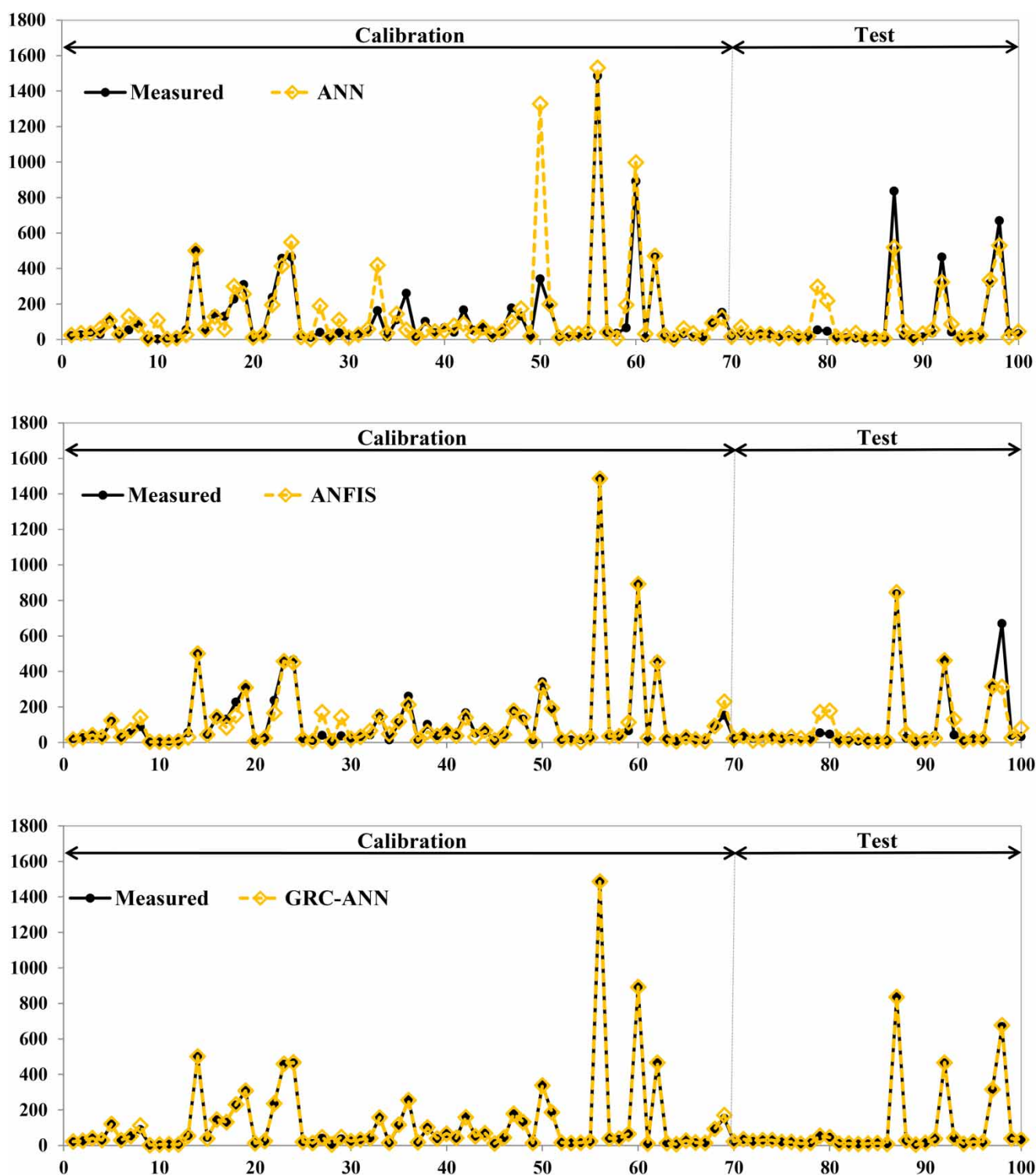
Figure 4 │ comparison between model and measured data for GRC-ANN, ANN, and ANFIS models for the calibration and test data.

Although ANFIS model showed an acceptable performance in estimating higher than 200 LDC values, its accuracy was not convincing because of having RMSE index of about 2 times worse compared to the GRC-ANN model. However, GRC-ANN model provided the best results amongst all of the methods in extremely high LDC values which certify

Uncorrected Proof

10   B. Ghiasi *et al.* │ GRC-ANN to predict longitudinal dispersion coefficient   Water Science & Technology │ in press │ 2020

Table 4 │ Errors result for GRC-ANN, ANN and ANFIS models for LDC > 200 m^2/s

	RMSE	MAE	R^2
ANN	281.58	150.66	0.5186
ANFIS	97.18	41.63	0.9315
GRC-ANN	**39.86**	**15.11**	**0.9867**

the good performance of the GRC-ANN model in prediction LDC measurements.

Natural rivers have widespread W/H ratio affecting the LDC. To investigate this effect in models predictions, low and high values of W/H ratio were considered separately. The results are demonstrated in Table 5.

The Table 5 illustrates that ANN and ANFIS models are more affected by W/H ratio than GRC-ANN. Eliminating this range, the ANN and ANFIS models accuracy improves because these models have better performance in W/H > 100. Despite the ANN and ANFIS models showed weak performance in low and high values of W/H, the GRC-ANN model has an excellent performance. The R^2 value of GRC-ANN models is close to 1.000 in W/H > 100 and close to 0.9 in W/H < 25. Therefore, W/H changes slightly affected the GRC-ANN and this model is applicable to every natural river.

In the case of U/u*, GRC-ANN model outperformed the other models, as in the case of W/H.

Table 6 indicates that the ANN model showed the worst performance with low U/u* values. Although GRC-ANN performance in U/u* < 3 is worse than U/u* > 15, it has better performance than other models since RMSE of the ANN and ANFIS models is worse than GRC-ANN about 4 and 3 times when U/u* < 3 and 140 and 3 times when U/u* > 15.

According to Tables 5 and 6, models can be ranked based on their performance in determining low and high values of W/H and U/u. The performance of the three models is prioritized in Table 7.

In this Table, numbers 1, 2 and 3 indicate the order of choosing models for rivers which based on W/H or U/u*

Table 5 │ Errors result for GRC-ANN, ANN, and ANFIS models for W/H < 25 and W/H >100

	RMSE		MAE		R^2	
Model	**W/H <25**	**W/H >100**	**W/H <25**	**W/H >100**	**W/H <25**	**W/H >100**
ANN	88.08	15.70	60.98	8.28	0.0003	0.9993
ANFIS	58.00	**0.2819**	34.87	**0.12**	0.3821	**0.9999**
GRC-ANN	**21.53**	0.3729	**16.13**	0.23	**0.8864**	**0.9999**

Table 6 │ Errors result for GRC-ANN, ANN and ANFIS models for U/u* <3 and U/u* >15

	RMSE		MAE		R^2	
Model	**U/u* <3**	**U/u* >15**	**U/u* <3**	**U/u* >15**	**U/u* <3**	**U/u* >15**
ANN	36.92	332.18	22.00	140.28	0.3682	0.6970
ANFIS	24.77	9.78	11.00	3.38	**0.9862**	0.9996
GRC-ANN	**9.65**	**2.24**	**7.27**	**1.04**	0.9619	**0.9999**

Table 7 │ Rating of models in errors comparison for W/H <25, W/H >100, U/u* <3 and U/u* >15

	ANN	ANFIS	GRC-ANN
W/H < 25	3	2	1
W/H >100	3	**1**	2
U/u* < 3	3	2	1
U/u* > 15	3	2	1

special range. Rank 1 represents that the model has the lowest error among others presented in Tables 5 and 6 (it is considered as the best model) while rank 3 indicates that the model has the highest error. This table indicates that GRC-ANN model can be chosen in about all of the natural rivers with best performance and least error. ANFIS model showed a better performance than ANN model. So, after GRC-ANN, ANFIS model can be chosen in special conditions. ANN model in 3 categories was determined as the last priority, so in the special condition, ANN performance is the worst.

In Figure 5, more tendencies to the centreline in the error distribution graph and larger values of the maximum QDDR denote more accuracy. By comparing DDR results
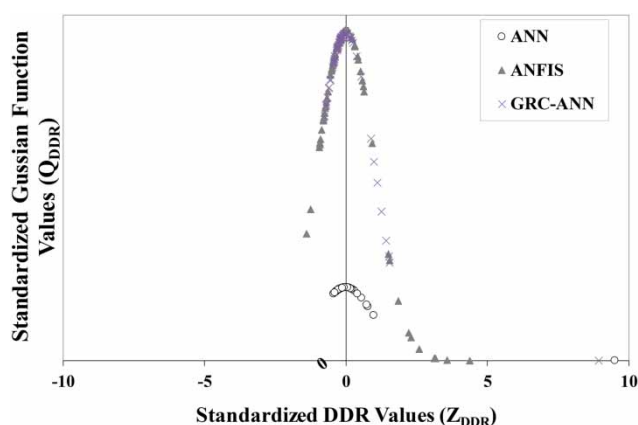

Figure 5 │ DDR results of GRC-ANN, ANN, and ANFIS models.

Uncorrected Proof

11   B. Ghiasi *et al.* │ GRC-ANN to predict longitudinal dispersion coefficient   **Water Science & Technology** │ in press │ 2020

Table 8 │ ARE results for ANN, ANFIS and GRC-ANN models

Percent of datasets (%)	ANN	ANFIS	GRC-ANN
10	5.27	**0.04**	0.06
25	16.82	3.07	**0.33**
50	39.40	16.26	**1.18**
75	91.20	49.16	**4.98**
90	167.38	99.97	**13.53**
95	289.48	216.06	**22.43**
100	3674.66	387.8	**67.89**

illustrated in Figure 5, we can conclude that the GRC-ANN and ANFIS models significantly outperformed the ANN model.

In order to compare the TS analysis, error distribution of testing steps for the ANN, ANFIS and GRC-ANN models is computed by means of TS analysis and absolute relative error (ARE) of 10, 25, 50, 90 and 100 percent of datasets demonstrated in Table 8.

Table 8 demonstrates that although ANFIS model has the lowest error in initial datasets until 10% but it's ARE in 10%–100% datasets is about 5–17 times larger than GRC-ANN model. According to Table 8, it is clear that the GRC-ANN model had better performance compared to others.

Sensitivity analysis

Sensitivity analysis of the model has been made for three different methods named GRC-ANN, ANFIS, and ANN. X and ΔX are the inputs of the models and in each method, Y is computed and is compared with the new Y which is the result of sensitivity analysis. Table 9 shows Computed Y (Y_1), new Y (Y_2), S_c and S_n for the aforementioned methods.

Table 10 │ R^2 and RMSE of different calibration data size steps

	R^2		RMSE	
Size of Calibration	Calibration	Test	Calibration	Test
40%	0.95	0.92	52	80
50%	0.98	0.93	32	46
60%	0.99	0.98	19	32
70%	0.99	0.99	15	22

Table 9 shows that the input parameter of W/H has a greater effect on the output of the model in all three models in this research. The importance of W/H has also been confirmed by Kashefipour & Falconer (2002), Tayfur & Singh (2005) and Sattar & Gharabaghi (2015).

However, the W/H parameter has a greater impact on all three models, but its impact varies in three models. The values of S_n indicate that the W/H parameter has the highest effect in the ANFIS model and has the least effect in the GRC-ANN model. Also, the difference between the values of S_n in two input parameters demonstrates that the impact difference between the two parameters in the GRC-ANN model is the lowest.

Robustness

In order to investigate the model robustness, the GRC-ANN model rerun three times with different calibration data size from 40% to 60% (The base run was 70%) and the model performance has been investigated in this condition. The R^2 and RMSE of these runs in calibration and test steps have been shown in Table 10.

By using 40% of the dataset for calibration, the RMSE is 52 and 80 in calibration and test steps that this value is less than the best empirical model, Deng *et al.* (2001) with RMSE

Table 9 │ Computed S_c and S_n of three GRC-ANN, ANFIS and ANN methods

Method	GRC-ANN		ANFIS		ANN	
	W/H	U/u*	W/H	U/u*	W/H	U/u*
X (input)	51.46	7.75	51.46	7.75	51.46	7.75
ΔX (input)	5.14	0.77	5.14	0.77	5.14	0.77
Y_1 (computed output of model)	695.97	695.97	775.25	775.25	793.21	793.21
Y_2 (sensitivity analysis output)	780	812	803.87	1840.75	856.65	953.35
ΔY ($Y_2 - Y_1$)	84.03	116.03	28.62	1065.49	63.44	160.14
S_c (marginal sensitivity coefficient)	0.0612	149.67	5.56	1374.45	12.33	206.58
S_n (normalized sensitivity coefficient	0.83	1.67	0.37	13.74	0.8	2.02

Uncorrected Proof

12   B. Ghiasi *et al.* | GRC-ANN to predict longitudinal dispersion coefficient   **Water Science & Technology** | in press | 2020

equal to 82. Therefore, the results of Table 10 demonstrated that the GRC-ANN is a reliable and robustness model in different size of calibration dataset.

CONCLUSION

Many researchers have tried to predict the LDC more accurately through experimental methods and intelligence models. These efforts have led to numerous intelligence models and empirical equations. Although intelligent models are more accurate than empirical models, they still do not have the proper accuracy and have some weaknesses. In this study, in order to improve the performance of the GRC model, it was combined with ANN model. Also, to test the performance of other intelligence models, ANN and ANFIS models were also used to investigate the results of intelligent models accurately. The results of this research can be summarized as follows:

- Compared to empirical relationships, the GRC-ANN model has the best performance, and the error value of the Deng *et al.* (2002) equation is about three times greater than GRC-ANN model.
- The GRC-ANN model makes proper allegiance with the pattern on observation values and has the best performance among ANN and ANFIS models.
- The GRC-ANN model has higher accuracy and lower error than the ANN and ANFIS models in LDC> 200.
- GRC-ANN model was more accurate than ANFIS and ANN models in W/H >100, W/H <25, U/u* <3 and U/u* >15 ranges. As a result, GRC-ANN model can be used to estimate all the LDC values in most rivers with high confidence.
- Diagrams of error distribution of the models developed in this study indicate the advantage of GRC-ANN model.
- The input parameter W/H has a more significant effect on the output of the model in all three models of this research.
- The W/H parameter has the highest effect in the ANFIS model and has the least effect on the GRC-ANN model
- Due to results of GRC-ANN model in different size of calibration dataset, this model is very reliable and robust.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/wst.2020.006.

REFERENCES

Azamathulla, H. M. & Ghani, A. A. 2011 Genetic programming for predicting longitudinal dispersion coefficients in streams. *Water Resour. Manage.* **25** (6), 1537–1544. https://doi.org/10.1007/s11269-010-9759-9.

Balf, M. R., Noori, R., Berndtsson, R., Ghaemi, A. & Ghiasi, B. 2018 Evolutionary polynomial regression approach to predict longitudinal dispersion coefficient in rivers. *J. Water Supply: Res. Technol.–AQUA.* **67** (5), 447–457. https://doi.org/10.2166/aqua.2018.021.

Bargiela, A. & Pedrycz, W. 2003 *Granular Computing: an Introduction*. Springer.

Chen, Y. & Yao, Y. 2006 Multiview intelligent data analysis based on granular computing. *GRC* 281–286.

Clark, J. F., Schlosser, P., Stute, M. & Simpson, H. J. 1996 SF6–3he tracer release experiment: a new method of determining longitudinal dispersion coefficients in large rivers. *Environ. Sci. Technol.* **30** (5), 1527–1532. https://doi.org/10.1021/es9504606.

Deng, Z., Singh, V. P. & Bengtsson, L. 2001 Longitudinal dispersion coefficient in straight rivers. *J. Hydraul. Eng.* **127** (11), 919–927. https://doi.org/10.1061/(ASCE)0733-9429(2001)127:11(919).

Deng, Z., Bengtsson, L., Singh, V. P. & Adrian, D. D. 2002 Longitudinal dispersion coefficient in single channel streams. *J. Hydraul. Eng.* **128** (10), 901–916. https://doi.org/10.1061/(ASCE)0733-9429(2002)128:10(901).

Disley, T., Gharabaghi, B., Mahboubi, A. A. & McBean, E. A. 2015 Predictive equation for longitudinal dispersion coefficient. *Hydrol. Processes* **29** (2), 161–172. https://doi.org/10.1002/hyp.10139.

Etemad-Shahidi, A. & Taghipour, M. 2012 Predicting longitudinal dispersion coefficient in natural streams using M5′ Model Tree. *J. Hydraul. Eng.* **138** (6), 542–554. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000550.

Fischer, H. B. 1967 The mechanics of dispersion in natural streams. *J. Hydraul. Eng. Div. ASCE* **93** (6), 187–216.

Fischer, H. B., List, J. E., Koh, C. R., Imberger, J. & Brooks, N. H. 1979 *Mixing in Inland and Coastal Waters*. Academic, New York, pp. 104–138.

Gaines, B. R. 1991 *The Trade-Off Between Knowledge and Data in Knowledge Acquisition*. AAAI/MIT Press, pp. 491–505.

Hobbs, J. R. 1990 Granularity. In: *Readings in Qualitative Reasoning About Physical Systems*. Morgan Kaufmann, 542–545. https://doi.org/10.1016/B978-1-4832-1447-4.50053-5

Jain, A. & Indurthy, S. K. V. P. 2003 Comparative analysis of event based rainfall-runoff modeling techniques-deterministic, statistical, and artificial neural networks. *J. Hydrol. Eng. ASCE* **8**, 93. https://doi.org/10.1061/(ASCE)1084-0699(2003)8%3A2(93).

Kashefipour, M. S. & Falconer, R. A. 2002 Longitudinal dispersion coefficients in natural channels. *Water Res.* **36** (6), 1596–1608. https://doi.org/10.1016/S0043-1354(01)00351-7.

Keet, C. M. 2008 *A Formal Theory of Granularity*. PhD Thesis, KRDB Research Centre, Faculty of Computer Science, Free University of Bozen-Bolzano, Italy.

Lin, T. Y., Yao, Y. Y. & Zadeh, L. A. 2002 *Data Mining, Rough Sets, and Granular Computing*. Springer.

Moore, R. E. 1966 *Interval Analysis*. Prentice-Hall Englewood Cliffs.

Noori, R., Hoshyaripour, G., Ashrafi, K. & Araabi, B. N. 2010 Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. *Atmos. Environ.* **44** (4), 476–482. https://doi.org/10.1016/j.atmosenv.2009.11.005.

Noori, R., Karbassi, A. R., Mehdizadeh, H., Vesali-Naseh, M. & Sabahi, M. S. 2011 A framework development for predicting the longitudinal dispersion coefficient in natural streams using an artificial neural network. *Environ. Prog. Sustainable Energy* **30** (3), 439–449. https://doi.org/10.1002/ep.10478.

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? *J. Hydrol. Eng. ASCE* **142** (1), 04015039. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001062.

Noori, R., Ghiasi, B., Sheikhian, H. & Adamowski, J. F. 2017a Estimation of the dispersion coefficient in natural rivers using a granular computing model. *J. Hydrol. Eng. ASCE* **143** (5), 04017001. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001276.

Noori, R., Sheikhian, H., Hooshyaripor, F., Naghikhani, A., Adamowski, J. F. & Ghiasi, B. 2017b Granular computing for prediction of scour below spillways. *Water Resour. Manage.* **31** (1), 313–326. https://doi.org/10.1007/s11269-016-1526-0.

Parsaie, A. & Haghiabi, A. H. 2015 Predicting the longitudinal dispersion coefficient by radial basis function neural network. *Model. Earth Syst. Environ.* **1** (4), 34. https://doi.org/10.1007/s40808-015-0037-y.

Pawlak, Z. 1982 Rough sets. *Int. J. Comput. Inf. Sci.* **11** (5), 341–356. https://doi.org/10.1007/BF01001956.

Pedrycz, W., Skowron, A. & Kreinovich, V. 2008 *Handbook of Granular Computing*. John Wiley & Sons.

Qian, Y., Liang, J., Yao, Y. & Dang, C. 2010 MGRS: a multi-granulation rough set. *Inf. Sci.* **180** (6), 949–970. https://doi.org/10.1016/j.ins.2009.11.023.

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 Syst. Appl.* **36** (4), 8589–8596. https://doi.org/10.1016/j.eswa.2008.10.043.

Sahay, R. R. & Douta, S. 2009 Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrol. Res.* **40** (6), 544–552. https://doi.org/10.2166/nh.2009.014.

Sattar, A. M. & Gharabaghi, B. 2015 Gene expression models for prediction of longitudinal dispersion coefficient in streams. *J. Hydrol.* **524**, 587–596. https://doi.org/10.1016/j.jhydrol.2015.03.016.

Seo, I. W. & Cheong, T. S. 1998 Predicting longitudinal dispersion coefficient in natural streams. *J. Hydraul. Eng.* **124** (1), 25–32. https://doi.org/10.1061/(ASCE)0733-9429(1998)124:1(25).

Sheikhian, H., Delavar, M. R. & Stein, A. 2017 A GIS-based multi-criteria seismic vulnerability assessment using the integration of granular computing rule extraction and artificial neural networks. *Trans. GIS* **21** (6), 1237–1259. https://doi.org/10.1111/tgis.12274.

Silverstein, C., Brin, S. & Motwani, R. 1998 Beyond market baskets: generalizing association rules to dependence rules. *Data Min. Knowl. Discovery* **2** (1), 39–68.

Smart, P. L. & Laidlaw, I. M. S. 1977 An evaluation of some fluorescent dyes for water tracing. *Water Resour. Res.* **13** (1), 15–33. https://doi.org/10.1029/WR013i001p00015.

Tayfur, G. 2006 Fuzzy, ANN, and regression models to predict longitudinal dispersion coefficient in natural streams. *Hydrol. Res.* **37** (2), 143–164.

Tayfur, G. & Singh, V. P. 2005 Predicting longitudinal dispersion coefficient in natural streams by artificial neural network. *J. Hydrol. Eng. ASCE* **131** (11), 991–1000. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:11(991).

Toprak, Z. F. & Cigizoglu, H. K. 2008 Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrol. Processes: Int. J.* **22** (20), 4106–4129.

Wu, W. Z., Leung, Y. & Mi, J. S. 2009 Granular computing and knowledge reduction in formal contexts. *IEEE Trans. Knowl. Data Eng.* **21** (10), 1461–1474. https://doi.org/10.1109/TKDE.2008.223.

Yager, R. R. 2008 Intelligent social network analysis using granular computing. *Int. J. Intell. Syst.* **23** (11), 1197–1219. https://doi.org/10.1002/int.20314.

Yao, Y. Y. 2001 On Modeling data mining with granular computing. In: *25th Annual International Computer Software and Applications Conference*. IEEE, Chicago, IL.

Yao, Y. Y. 2008 A unified framework of granular computing. In: *Handbook of Granular Computing* (W. Pedrycz, A. Skowron & A. V. Kreinovich, eds). Wiley, Chichester, UK, pp. 401–410.

Yao, J. T. & Yao, Y. Y. 2002a A granular computing approach to machine learning. *FSKD* **2**, 732–736.

Yao, J. & Yao, Y. 2002b Induction of classification rules by granular computing. In: *International Conference on Rough Sets and Current Trends in Computing*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45813-1_43

Yao, Y. & Zhong, N. 2002 Granular computing using information tables. In: *Data Mining, Rough Sets, and Granular Computing*. Springer, 102–124. https://doi.org/10.1007/978-3-7908-1791-1_5

Zadeh, L. A. 1996 Fuzzy logic = computing with words. *IEEE Trans. Fuzzy Syst.* **4** (2), 103–111. https://doi.org/10.1007/978-3-7908-1873-4_1.

Zadeh, L. A. 1997 Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst.* **90** (2), 111–127. https://doi.org/10.1016/S0165-0114(97)00077-8.

Zeng, Y. & Huai, W. 2014 Estimation of longitudinal dispersion coefficient in rivers. *J. Hydro-Environ. Res.* **8** (1), 2–8. https://doi.org/10.1016/j.jher.2013.02.005.