Intelligent Modeling of Discharge Coefficient of Lateral Intakes

A. H. Haghiabi¹, A. Parsaie¹,.*, Z. Shamsi²

¹ Water Engineering Department, Lorestan University, Khorramabad, Iran
² Department of Water and Soil Conservation, Ministry of Agriculture Jihad, Kerman, Iran

ABSTRACT: Intake structures have been widely used for flow diversion in the irrigation and drainage networks. In this paper, the multivariate adaptive regression splines (MARS), artificial neural network (ANN), and support vector machine (SVM) techniques were utilized for prediction of discharge coefficient \( C_d \) of lateral intakes. The experimental data pertaining to dimensionless parameters on \( C_d \) were collected to develop the models. The results indicated that the best performance in modeling is related to the MARS model with \( R^2=0.98 \) and \( \text{RMSE}=0.023 \) and the MARS model outperforms the ANN and SVM models. The tangent sigmoid and radial basic functions were found to be the most efficient transfer and kernel functions for ANN and SVM, respectively. Moreover, Froude number \( (Fr) \) and the ratio of the weir height to the upstream flow depth \( (P/d) \) were the most effective factors for predicting \( C_d \). Evaluation of the performance of applied models in term of developed discrepancy ratio \( (\text{DDR}) \) index shows that the minimum data dispersivity is related to the MARS model.

Keywords: Flow Measurement, Multivariate Adaptive Regression Splines, Neural Networks, Support Vector Machine

1- Introduction

Lateral intakes (LIs) have been widely used in water engineering projects such as irrigation and drainage networks. A LI sets at the side wall of open channels or riverine. Weir is the main part of lateral intake. The main task of the LIs are diverting a certain volume of flow [1]. Study on the hydraulic properties of LIs initially was started by experimental studies conducted by De Marchi [2] and till to now, many investigators have investigated the hydraulic properties of this structure and they have proposed and tested many theories tho this end [3-5]. Discharge capacity of LI is proportional of \( C_d \) and length of crest of weir. To increase the discharge capacity, it is possible to increase the length of crest and improve the discharge coefficient. By considering these points, researchers has proposed some palns for shape of crest and improving the discharge coefficient. in this regard proposing the labyrinth, oblique, semi-elliptical, curved plan-form have been proposed for increasing the length of crest [6-8] and circulating the crest and using the guide vanes have been proposed for increasing the \( C_d \) [6, 9-12]. Based on reports, increasing the length of crest leads to improve the discharge capacity about three or four times [13-16]. Associated laboratory studies, numerical methods have been used for simulation of flow through the LIs [17-19]. In the field of numerical modeling, water surface profile over the LIs has been simulated using the classical numerical approaches such as Runge–Kutta method and advanced numerical methods such as computational fluid dynamic techniques. The aim of those studies in addition to modeling of water surface profile, were characterizing the flow properties such as flow pattern, distribution of velocity and pressure along the LIs [20]. In the numerical approaches, modeling of \( C_d \) also has been taken into consideration. To this end, soft computing techniques have been applied [21]. Using the artificial neural network has been reported by [Bilhan, Emin Emiroglu and Kisi [22], Bilhan, Emiroglu and Kisi [23]]; Based on these reports the performance of the ANN are so suitable for predicting the \( C_d \). Emiroglu and Kisi [24] have stated that the Neuro-Fuzzy method has suitable performance for prediction of discharge coefficient of the labyrinth LIs. Successfully Using the GMDH for predicting the \( C_d \) of LIs was reported by Etehaj, Bonakdari, Zaji, Azimi and Khoshbin [25]. This paper considers intelligent modeling of \( C_d \) of LIs using MARS method as new soft computing approach in hydraulic engineering. Performance of MARS model is compared with other types of soft computing techniques including ANN and SVM. The experimental data pertaining to dimensionless parameters on \( C_d \) were collected from the literature.

2- Method and Materials

Most important geometrical and hydraulic parameters effect the \( C_d \) are shown in Figure 1.
As seen from Figure 1, $d_1$ is the flow depth at the beginning of the weir, $y$ is the flow depth on the weir. $P$ is the weir height, $B$ is channel width, $b$ is the length of weir. Using the Buckingham theory the dimensionless parameters related to $C_d$ can be written as Equation 1 [26].

$$C_d = f \left( \frac{Fr_d}{b} \frac{b}{d_1} \frac{P}{d_1} \right)$$  \hspace{1cm} (1)

In which $Fr_d$ is the Froude number. Equation 1 is fundamental for developing the empirical formula and soft computing models to predict and mathematical model of $C_d$. Developing of soft computing techniques are based on the dataset, therefore 169 datasets with regarding to the Equation 2 were collected from Jalili and Borghei [27]. The histograms of collected dataset are shown in Figure 2.

$|x-t|_+ = \max(0, x-t) = \begin{cases} x-t & x > t \\ 0 & x \leq t \end{cases}$  \hspace{1cm} (2.a)

$|t-x|_+ = \max(0, t-x) = \begin{cases} t-x & x < t \\ 0 & x \geq t \end{cases}$  \hspace{1cm} (2.b)

where $t$ denotes the knot. Basic functions some time calls as mirrored pair functions. These functions are defined for each input variables such as $X_j$ at observed dataset related of them. Sets of basic functions are defined as

$$C = \left\{ (x_1-t_1), (t-x_i) \right\} \times \{x_1, x_2, \ldots, x_p\}, \, j = 1, \ldots, p$$ \hspace{1cm} (3)

The general form of function derived from MARS model is written as an adaptive function as bellow.

$$y = \beta_0 + \sum_{i=1}^{M} \beta_i BF_i(X)$$ \hspace{1cm} (4)

where $\beta$ is the constant value, $BF_i(X)$ known is basic functions and $\beta_i$ are the coefficients of the as basic functions. The constant and coefficient of derived function in MARS model are justified using least square error technique. The $M$ is number of basic functions derived from the final stage of model development. Developing of MARS model has two stages. One forward stage in this stage number of basic function is increased to decrease difference between the results of model and observed data. In the next step of model development to avoid over parameterization and over fitting pruning the some of the basic functions are considered. In this stage with regarding to cross-validation (GCV) criteria that are given as bellow the basics function are pruned.

$$GCV = \frac{\text{SSE}}{n \left( 1 - \frac{C(B)}{n} \right)^2}$$ \hspace{1cm} (5.a)

$$C(B) = (B+1) + dB$$ \hspace{1cm} (5.b)

Where, SSE, is the sum of square of residuals, $n$ denotes the number of records and $C(B)$ defined a difficulty criteria, which increases by the number of basic functions [32-35].

2-2- Artificial neural networks (ANNs)

ANN is an advance mathematical method that has ability for mapping complexes systems which are based dataset. Common type of ANN is MLP are that are widely used in the researches. To use MLP model, definition of appropriate transfer functions, designing a suitable structure in term of computational cost should be considered. Different transfer function can be tested. An ANN maybe has one or more hidden layers. Figure 3 demonstrates a neuron consisting of inputs, weight and output. As shown in Figure 3, $w_i$ is the weight and $b_i$ is the bias for each neuron. After designing the structure of MLP (number of hidden layer and number of neurons in each hidden layer), definition of weight and biases’ should considered. This stage named model training. Several
methods, whether classical or modern have been suggested. In classical approach, Levenberg–Marquardt technique and modern ways such as modern optimization approach such as PSO, GA can be stated [21, 36-39].

After selecting the activation appropriate function, to improve the precision of MLP model, number of hidden layer(s) and number of neurons in hidden layer may be increased step by step. The hyperbolic tangent sigmoid (tansig) transfer function was considered for neurons in the first and second hidden layer respectively. The Levenberg–Marquardt technique was used for MLP model learning.

To provide a good simulation, it is better that range of training and testing dataset would be close together. Designing of MLP model include some steps such as 1- considering the number of the neurons in each layer, 2- defining of suitable transfer function of the hidden layer(s), 3- considering the number of the input features is considered. Next, different type of transfer functions such as logsig, tansig, purelin are tested. The value of the regularization parameter is defined during the calibration process. For gave more information about the SVM are presented in the literature [21, 36-39].

2- 3- Support vector machine (SVM)
SVMs are a type of artificial intelligence method has been widely used in the field of hydrology studies. Published literatures, the SVM has high ability for pattern recognition. Developing the SVM is based on the data set. Therefore, collected data set has been separated into two categories as calibration and valuation. For preparation of SVM for modeling and predicting the problems two steps should be considered. Choosing the regularization parameter (C) and kernel function, and learning algorithm has high effect on increasing the SVM performance as well. Various types of kernel functions such as linear, polynomial, Gaussian radial basis function have been proposed for the kernel functions. Selecting the type of kernel function is a trial and error process and for specific problem a types of functions should be tested. The performance of the SVM for predicting the C_d as output parameters. As seems from the Figure 4, the MLP model contain two hidden layers which five and three neurons are located in the first and second hidden layer respectively. The hyperbolic tangent sigmoid (tansig) transfer function was considered for neurons in the first and second hidden layers, respectively. The Levenberg–Marquardt technique was used for MLP model learning.

The performance of the MLP model during the training and testing stages are shown in Figure 6. In this figure, the predicted C_d plotted versus the observed C_d. Moreover, in this figure, the results of error indices such as such as R^2 (Equation 7), and RMSE (Equation 8) have been presented, as well. In overall, these figure shows that the performance of the MLP for predicting the C_d is suitable.

![Figure 3. Sketch of three-layer ANN architecture](image)

![Figure 4. Architect of the developed MLP model](image)

\[ R^2 = \left( \frac{\sum_{i=1}^{n} (O_i - \bar{O})(P_i - \bar{P})^2}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}} \right)^2 \]  \hspace{1cm} (6)

\[ RMSE = \frac{\sqrt{\sum_{i=1}^{n} (O_i - P_i)^2}}{n} \]  \hspace{1cm} (7)

Developing the SVM model is similar to the ANN. This means that the same dataset which had been used for training and testing the ANN was applied for preparation of SVM. The structure of developed SVM model is shown in Figure 5. The results of SVM in stages of model preparation are shown in Figure 6. For preparation of SVM, radial basis function (RBF) and polynomial kernel function were...
assessed. Testing of both kernel functions show that the RBF kernel function has better performance in compare with the polynomial function. During the preparation of SVM, the value of internal coefficient (γ) and constant (σ) of radial basis function were found equal to 10931.07 and 51.24 respectively. Assessing the performance of SVM model in both stages of development (training and testing) shows that this model with $R^2 = 0.98$ and RMSE=0.013 for training stage and $R^2 = 0.96$ and RMSE=0.046 for testing stage is more then the MLP model with $R^2 = 0.87$ and RMSE=0.035 for training stage and $R^2 = 0.86$ and RMSE=0.040.

Equation 9 can be used for predicting the $C_d$. As seen from Table 1, $F_{r1}$, $P/d_1$ and $L/b$ have been appeared in the almost of the basic functions. It means that these three parameters in compare to other parameters are more affective on the $C_d$. This result of MARS model upholds the results of the MLP model sensitivity analysis were reported by Azamathulla, Haghiabi and Parsaie [1]. Moreover, the performances of the MARS model during the development process (training and testing stages) are given in the Figure 6. Observing Figure 6, it was found that the performance of the MARS model for prediction the $C_d$ is so suitable.

Preparation of MARS model as similar to ANN and SVM is based on the data set. To this purpose, collected data set which had used for developing the MLP and SVM are used for preparation of MARS model. During the MARS model development, at the first step thirty basic functions was considered and at the second step (pruning step) five basic functions was pruned and at the end, the optimal MARS model with twenty five basic functions was derived. The general form of the obtained model MARS is given in the Equation 9. The extended form of the MARS model is given in Table 1.

$$C_d = 0.6502 + \sum_{m=1}^{25} \beta_m h_m(x)$$  \hspace{1cm} (8)

Equation 9 can be used for predicting the $C_d$. As seen from Table 1, $F_{r1}$, $P/d_1$ and $L/b$ have been appeared in the almost of the basic functions. It means that these three parameters in compare to other parameters are more affective on the $C_d$. This result of MARS model upholds the results of the MLP model sensitivity analysis were reported by Azamathulla, Haghiabi and Parsaie [1]. Moreover, the performances of the MARS model during the development process (training and testing stages) are given in the Figure 6. Observing Figure 6, it was found that the performance of the MARS model for prediction the $C_d$ is so suitable.

Assessing the results of utilized models (MARS, SVM and MLP) with regarding to error indices (R and RMSE) donanestated that the MARS model was the most accurate espaciflay in testing stage. They indices provide only an average value and does not provide any information about the error distribution. Hence, in this study another error index including developed discrepancy ratio (DDR) was used to present more information about the error properties. The DDR index is calculcated as Equation 9. As presented in Equation 9, the DDR index is defined as ratio of predicted valued to observed value. It is notable that the predicted values are the output of applied models in training and testing stages. DDR index in addition to provide more information about the error distribution, it is charercterized the over and lower propertices of models. Results of DDR index for both stage of model development are shown in Figure 7 and their histograms are shown in Figure 8. Reviewing these figures shows that in training stage, the DDR values of SVM model varies between the -0.08 and 0.24, the DDR values of MLP model varies between -0.4 and 0.4 and DDR values of MARS model vaies between 0.08 and 0.1. Comparison the DDR of models in training stage shows that the minimum range of DDR value in training stage is related MARS model and maximum range of DDR is related to MLP model. These results regareding the DDR index for testing stage were almost repeated. Evalution the result of applied models in term of DDR index shows that the results of MARS model are more relaible. Another finding from the DDR index is defining the over and lower estimation propertices of models. Histogram of DDR index shows that they models has not significant over-lowe estimation propricetie.
Table 1. The Basic function and related coefficient of the MARS model

<table>
<thead>
<tr>
<th>Basic function</th>
<th>equation</th>
<th>Coefficient (β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$ (X)</td>
<td>BF1 = max(0, 0.508 - b/B)</td>
<td>-0.9929</td>
</tr>
<tr>
<td>$h_2$ (X)</td>
<td>BF2 = max(0, P/d_1 - 0.37)</td>
<td>-3.2360</td>
</tr>
<tr>
<td>$h_3$ (X)</td>
<td>BF3 = max(0, 0.37 - P/d_1)</td>
<td>0.6680</td>
</tr>
<tr>
<td>$h_4$ (X)</td>
<td>BF4 = BF1 × max(0, Fr_1 - 0.294)</td>
<td>2.1341</td>
</tr>
<tr>
<td>$h_5$ (X)</td>
<td>BF5 = BF1 × max(0, 0.294 - Fr_1)</td>
<td>-2.4153</td>
</tr>
<tr>
<td>$h_6$ (X)</td>
<td>BF6 = max(0, Fr_1 - 0.453)</td>
<td>-0.8700</td>
</tr>
<tr>
<td>$h_7$ (X)</td>
<td>BF7 = max(0, 0.453 - Fr_1)</td>
<td>-19.5510</td>
</tr>
<tr>
<td>$h_8$ (X)</td>
<td>BF8 = BF7 × max(0, 4.167 - b/d_1)</td>
<td>-1.1161</td>
</tr>
<tr>
<td>$h_9$ (X)</td>
<td>BF9 = BF2 × max(0, b/B - 0.508)</td>
<td>2.6338</td>
</tr>
<tr>
<td>$h_{10}$ (X)</td>
<td>BF10 = BF7 × max(0, b/B - 1.26)</td>
<td>89.0015</td>
</tr>
<tr>
<td>$h_{11}$ (X)</td>
<td>BF11 = BF7 × max(0, 1.26 - b/B)</td>
<td>-91.9290</td>
</tr>
<tr>
<td>$h_{12}$ (X)</td>
<td>BF12 = BF7 × max(0, 1.5 - b/B)</td>
<td>90.9047</td>
</tr>
<tr>
<td>$h_{13}$ (X)</td>
<td>BF13 = max(0, b/B - 0.508) × max(0, Fr_1 - 0.558)</td>
<td>0.8923</td>
</tr>
<tr>
<td>$h_{14}$ (X)</td>
<td>BF14 = max(0, b/B - 0.508) × max(0, 0.558 - Fr_1)</td>
<td>-1.1671</td>
</tr>
<tr>
<td>$h_{15}$ (X)</td>
<td>BF15 = max(0, b/B - 0.508) × max(0, b/d_1 - 4.167)</td>
<td>-0.1619</td>
</tr>
<tr>
<td>$h_{16}$ (X)</td>
<td>BF16 = BF2 × max(0, 4.167 - b/d_1)</td>
<td>0.8684</td>
</tr>
<tr>
<td>$h_{17}$ (X)</td>
<td>BF17 = BF7 × max(0, 3.409 - b/d_1)</td>
<td>1.1468</td>
</tr>
<tr>
<td>$h_{18}$ (X)</td>
<td>BF18 = BF2 × max(0, b/d_1 - 3.495)</td>
<td>0.6231</td>
</tr>
<tr>
<td>$h_{19}$ (X)</td>
<td>BF19 = BF2 × max(0, 3.495 - b/d_1)</td>
<td>-0.8032</td>
</tr>
<tr>
<td>$h_{20}$ (X)</td>
<td>BF20 = max(0, b/d_1 - 0.953) × max(0, P/d_1 - 0.88)</td>
<td>2.4854</td>
</tr>
<tr>
<td>$h_{21}$ (X)</td>
<td>BF21 = BF2 × max(0, b/B - 1.508)</td>
<td>-61.3033</td>
</tr>
<tr>
<td>$h_{22}$ (X)</td>
<td>BF22 = BF2 × max(0, 1.508 - b/B)</td>
<td>2.0168</td>
</tr>
<tr>
<td>$h_{23}$ (X)</td>
<td>BF23 = max(0, L/d_1 - 0.953) × max(0, 0.156 - Fr_1)</td>
<td>0.8152</td>
</tr>
<tr>
<td>$h_{24}$ (X)</td>
<td>BF24 = max(0, 4.46 - b/d_1)</td>
<td>-0.0328</td>
</tr>
<tr>
<td>$h_{25}$ (X)</td>
<td>BF25 = BF7 × max(0, b/d_1 - 2.788)</td>
<td>-0.5373</td>
</tr>
</tbody>
</table>

\[
 DDR = \left( \frac{\text{Predicted Value}}{\text{Observed Value}} \right) - 1 \quad \text{(9)}
\]

4- Conclusions

Discharge coefficient of flow measurement structure is the main parameter for controlling the efficiency of hydro systems. Among the hydraulic structures, Lateral intakes are the common structure widely used in water engineering projects. Recently by advancing the soft computing techniques in most area of engineering especially in water engineering, the discharge coefficient of lateral intakes have been models using these techniques. In this study, the new soft computing model entitled multivariate adaptive regression splines (MARS) in hydraulic engineering has been used for mathematical expression for $C_d$. The performance of MARS model was compared with the SVM and ANN which were developed to this end as well. Th MARS technique with $R^2=0.98$ and RMSE=0.013 in calibration stage and $R^2=0.96$ and RMSE=0.023 in validation stage has satisfactory performance in intelligent modeling of $C_d$ of lateral intakes. Reviewing of mathematical form derived from MARS technique for $C_d$ indicated that relative upstream head, relative weir length and Froude number of flow at beginning of weir are most importact factors on $C_d$. Results of ANN and SVM indicated that although these models have suitale value for error indices in development stages ($R^2_{MLP}=0.86$ and RMSE$_{MLP}=0.040$, $R^2_{SVM}=0.86$ and RMSE$_{SVM}=0.046$), but their precisions are less then the MARS model. The high precision of MARS model is due to intelligent definition of most effective parameters are defined automatically during
the mapping the relation between the independent and dependent variables. Another benefit of the MARS model are related to programing its results for another purposes. To prived more information about the results of applied models, the DDR index was calculated. Evaluation of DDR shows that the minimum data dispersivity is related to MARS model. Therefor, the modeling by MARS model is more reliable.

Figure 7. Values of DDR for applied models in training and testing stages

Figure 8. Histogram of DDR values for applied models in training and testing stages

References


