Simulating the Discharge Coefficient of an Improved Triangular Side Weir using a Novel Hybrid Method of Firefly Optimization based Support Vector Regression

Amir Hossein Zaji1 and Hossein Bonakdari2*

1Department of Civil Engineering, Razi University, Kermanshah, Iran
2Department of Civil Engineering, Razi University, Kermanshah, Iran

Submission: December 07, 2017; Published: March 19, 2018

*Corresponding author: Amir Hossein Zaji, Department of Civil Engineering, Razi University, Kermanshah, Iran, Tel: +98 831 427 4537; Fax: +98 831 428 3264; E-mail: bonakdari@yahoo.com

Abstract

The discharge coefficient of side weirs is the most important required information in designing side weirs and tributary channels. An accurate discharge coefficient method needs to do a trustable design. The aim of this paper is to introduce and compare the Firefly optimization-based Support Vector Regression (FFSVR) with other powerful artificial intelligence approaches, including Artificial Neural Network (ANN), Genetic Programming (GP) and Support Vector Regression (SVR) to predict the discharge coefficient of an improved shape, triangular side weir. According to the results, the FFSVR model with RMSE of 0.035 performed the best among ANN, GP and SVR, with RMSE of 0.036, 0.040 and 0.039 in the test dataset, respectively. It can be concluded that combining the Firefly optimization algorithm with SVR can produce a higher-performance model.

Keywords: Artificial neural network, Discharge coefficient, Firefly optimization algorithm, Genetic programming, Support vector regression, Triangular side weir

Abbreviations: MLPNN: Multi-Layer Perceptron Neural Network; RBNN: Radial Basis Neural Network; PSO: Particle Swarm Optimization; GA: Genetic Algorithm

Introduction

Side weirs are always used in a wide range of hydraulic structures, including irrigation and drainage systems, diverting and dividing flow structures, as well as flood control utilities. The first study to mathematically describe side weir flow was undertaken by De Marchi [1]. The author presented the following equation using the constant specific energy assumption for the upstream and downstream of the side weir.

\[ \frac{dQ}{dx} = \frac{2}{3} C_d \sqrt{2g} (y - w)^{1.5} \]  

(1)

Where \( C_d \) is the discharge coefficient, \( w \) weir height, \( y \) flow depth and \( \frac{dQ}{dx} \) variation of discharge with respect to distance along the main channel. The first side weirs that were practically employed had a rectangular shape. Determining the discharge of such side weirs is a topic of interest in many studies [2-7].

For a side weir to pass additional flow, the one option is to increase the side weir length. Increasing the length of the side weir can be done in two ways. One way is to increase the tributary channel width, or second, the side weir shape can be modified. Obviously, the second way is much more economical than the first. Therefore, numerous studies have focused on modifying the shape of side weirs [8-12]. Geometric modification can help increase side weir performance by 1.5 to 4.5 times [13].

On account of the complex nature and various geometric and hydraulic dependencies of discharge coefficient conditions, different soft computing methods are often used as a fast and accurate means of predicting the discharge coefficient [13-16]. The discharge coefficient of the triangular side weir that is considered in the present study is modelled in various studies. Zaji and Bonakdari [17] modelled the present side weir discharge coefficient using Multi-Layer Perceptron Neural Network (MLPNN), Radial Basis Neural Network (RBNN), and Particle Swarm Optimization (PSO) based equations. Zaji et al. [18] modelled the considered side weir discharge coefficient using a hybrid method of PSO based RBNN. The authors concluded that PSO improved the RBNN prediction performance. By using ANFIS there are two studies in simulating the discharge coefficient of the present side weir [19,20]. The
authors concluded that ANFIS has a high capability in simulating the present side weir discharge coefficient.

To enhance neural network performance, one possibility is to combine the neural networks with optimization algorithms [21-25]. In this study a hybrid combination of Support Vector Regression method with Firefly optimization algorithm (FFSVR) is introduced and compared with three well-known artificial intelligence techniques, namely ANN, GP and SVR. To compare the models, the discharge coefficients of the improved side weir were predicted and verified using a published experimental dataset [26].

Materials and Methods

Experimental dataset

Borghei and Parvaneh [26] experimental dataset for an improved triangular side weir was used to verify and compare the FFSVR, SVR, GP and ANN methods. The authors introduced a new side weir shape that could increase the pass discharge ability up to two times (Figure 1). The experiments were done in a glass channel with 11m length, 0.4m width and 0.66m height.

Table 1: Various geometric and hydraulic parameters of a modified labyrinth side weir [26].

<table>
<thead>
<tr>
<th>$\theta/2(\circ)$</th>
<th>$L$(m)</th>
<th>$w$(mm)</th>
<th>$w/Y_i$</th>
<th>$Q_1$(m$^3$/s)</th>
<th>$F_i$</th>
<th>Numb of Runs</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.3</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>50</td>
</tr>
<tr>
<td>45</td>
<td>0.3</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>5,01,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td>60</td>
<td>0.3</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>50, 100,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>5,01,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td>70</td>
<td>0.3</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>5,07,51,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>5,01,00,150</td>
<td>0.46-0.83</td>
<td>0.019-0.030</td>
<td>0.19-0.96</td>
<td>55</td>
</tr>
</tbody>
</table>

Numerical models

Artificial neural network (ANN): The back propagation learning algorithm artificial neural network has been widely used in numerous studies. An ANN model consists of three major layers: input, hidden and output layers. Each layer comprises neurons. The neurons of the input and output layers are the input and output variables of the problem, respectively. The back propagation algorithm attempts to determine the weight and bias of each neuron. The neurons in each layer calculate the weighted summation of the last layer’s neurons and bias values and place this summation in the transfer function. Neurons can use many transfer functions with the hyperbolic transform function being the most common [14,27,28] and defined as follows:

$$\text{tansig}(n) = \frac{2}{1+e^{-2n}} - 1 \quad (2)$$

In this study, the input layer consists of the considered non-dimensional parameters of the improved triangular side weir and the output layer is the discharge coefficient. Details of the ANN algorithm are given in [29,30].

Genetic programming (GP): Genetic Programming [31], a subset of the Genetic Algorithm (GA) [32], is a program-based method that uses the symbolic form of programs as chromosomes. The initial population is selected randomly. Each individual of the population is a randomly generated program made from random input combinations, random arithmetic (+, -, $\times$, $\div$) and mathematical functions (sin, cos, etc.), and random logical functions (NOT, AND, OR, etc.). By using the fitness function, each individual cost is calculated. Subsequently, the individuals (programs) with the best costs are selected to perform the crossover and mutation process. Crossover is meant to change the parts of two programs with each other and mutation is to randomly change a program to create new
programs. This evolution is repeated to achieve the defined stop criterion. The output of the program is a program that could predict the discharge coefficient of improved triangular side weir by using the non-dimensional input parameters. The GP procedure is presented in more detail in [33,34].

**Support vector regression (SVR):** Support vector regression is a high-performance method capable of handling regression and classification problems much better than other neural networks [35,36]. Most of the other machine learning methods are performed from local training errors in their training process, but SVR uses upper band generalization error minimization in the training process [37]. SVR is described in detail in [38,39]. In this study, an SVR model with RBF kernel function was used to model the discharge coefficient of the improved triangular side weir. RBF kernel function can handle complex problems more accurately than other kernel functions [37,40]. An RBF kernel function is defined as:

\[
K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|) \tag{3}
\]

**Firefly optimization-based support vector regression (FFSVR):** The accuracy of SVR model prediction is directly related to the correct specification of model variables. In this study, the optimal values of SVR variables were determined using the Firefly optimization algorithm [39]. The Firefly algorithm can perform much better in finding the optima compared to other evolutionary optimization algorithms [40-43]. The basic steps of the Firefly algorithm are summarized in Figure 2.

![Firefly algorithm flowchart.](image)

**Results**

Improved triangular side weir discharge coefficient is related to various geometric and hydraulic situations, such as length (L), height (W), weir apex angle (θ), flow depth (y), upstream Froude number (Fr) and upstream discharge (Q). To run a numerical model to be used in practical and scientific processes without concern about the geometric and hydraulic scales, it is ideal to use non-dimensional input variables. Thus, ten non-dimensional input variables were performed using the efficient variables. Here, the numerical methods FFSVR, SVR, ANN and GP were used to model the discharge coefficient of an improved triangular side weir with the input variables: \( \frac{W}{L}, \frac{y}{L}, \frac{Fr}{L}, \frac{Q}{L^3}, \sin^2(\theta), \frac{w}{L}, \frac{1}{\sin(\theta)} \) and \( \frac{w}{L}, \frac{1}{\sin(\theta)} \).

Figure 3 shows the scatter plot and statistic errors of the investigated model for training and testing. From this figure, it can be concluded that the FFSVR model with RMSE, MAE and of 0.035, 0.026 and 4.218, respectively, performed best among the investigated models.

The trend line equation of each model is shown in Figure 3. The form of a linear trend line equation is \( y = C_0x + C_1 \). It is
obvious that as $C_1$ approaches one and $C_2$ approaches zero, the model is more fit to the experimental results. The figure shows that the FFSVR model with $C_1$ equaling 0.976 and 0.912 in the training and testing data, respectively, could predict the improved triangular side weir more accurately.

<table>
<thead>
<tr>
<th>Model</th>
<th>Train dataset</th>
<th>Test dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firefly Support Vector Regression (FFSVR)</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>$y = 0.9762x + 0.0165$</td>
<td>$R^2 = 0.965$</td>
<td>$y = 0.9124x + 0.0547$</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.037</td>
<td>0.035</td>
</tr>
<tr>
<td>MAE</td>
<td>0.026</td>
<td>0.030</td>
</tr>
<tr>
<td>$%$</td>
<td>4.088</td>
<td>4.211</td>
</tr>
<tr>
<td>Support Vector Regression (SVR)</td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td>$y = 0.9694x + 0.0265$</td>
<td>$R^2 = 0.9609$</td>
<td>$y = 0.8818x + 0.0692$</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.039</td>
<td>0.039</td>
</tr>
<tr>
<td>MAE</td>
<td>0.030</td>
<td>0.030</td>
</tr>
<tr>
<td>$%$</td>
<td>4.211</td>
<td>4.833</td>
</tr>
<tr>
<td>Genetic Programming (GP)</td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>$y = 0.9287x + 0.0485$</td>
<td>$R^2 = 0.9569$</td>
<td>$y = 0.8874x + 0.076$</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.041</td>
<td>0.040</td>
</tr>
<tr>
<td>MAE</td>
<td>0.033</td>
<td>0.030</td>
</tr>
<tr>
<td>$%$</td>
<td>4.789</td>
<td>4.773</td>
</tr>
<tr>
<td>Artificial Neural Network (ANN)</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
<tr>
<td>$y = 0.9404x + 0.0442$</td>
<td>$R^2 = 0.9435$</td>
<td>$y = 0.8148x + 0.1166$</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.047</td>
<td>0.036</td>
</tr>
<tr>
<td>MAE</td>
<td>0.035</td>
<td>0.023</td>
</tr>
<tr>
<td>$%$</td>
<td>4.982</td>
<td>4.515</td>
</tr>
</tbody>
</table>

*Figure 3: Scatter plot and performance evaluation of SVR-FF, SVR, ANN and GP methods for the testing and training datasets*
Figure 3 indicates that for all four models, the performance of training and testing data is very similar. This is indicative that the models were not trapped in over-fitting. Over-fitting occurs when the instruments used to design the model are more than enough. Thus, in an over-fitting situation, the performance of the training dataset prediction is significantly better than the testing dataset.

Conclusion

In this study, the Firefly optimization algorithm was combined with an SVR model to achieve a new, powerful hybrid method, the FFSVR. The performance of the FFSVR was then investigated and compared with the three most applied artificial intelligence methods: SVR, ANN and GP. The output of the models was the discharge coefficient of an improved triangular side weir, and the inputs were ten non-dimensional parameters obtained from the main parameters: and . The results signify that the FFSVR with RMSE of 0.035 in the test dataset achieved the highest performance among all investigated models. It can be concluded that the Firefly optimization algorithm can be successfully combined with the SVM method to increase the model’s ability to predict accurately.

References


