# Support Vector Regression Integrated with Fruit Fly Optimization Algorithm for River Flow Forecasting in Lake Urmia Basin

**Saeed Samadianfard** [1], **Salar Jarhan** [2], **Ely Salwana** [3], **Amir Mosavi** [4,5,6],
**Shahaboddin Shamshirband** [7,8,*] and **Shatirah Akib** [9]

[1]  Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Bahman Boulevard 29, Iran; s.samadian@tabrizu.ac.ir
[2]  Department of Hydrosciences, Technische Univeristät Dresden, 01069 Dresden, Germany; salar.jarhan@mailbox.tu-dresden.de
[3]  Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia; elysalwana@ukm.edu.my
[4]  School of the Built Environment, Oxford Brookes University, Oxford OX30BP, UK; amirhosein.mosavi@qut.edu.au
[5]  Institute of Automation, Kando Kalman Faculty of Electrical Engineering, Obuda University, 1034 Budapest, Hungary
[6]  Queensland University of Technology, Brisbane, QLD 4059, Australia
[7]  Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chí Minh, Vietnam
[8]  Faculty of Information Technology, Ton Duc Thang University, Ho Chí Minh, Vietnam
[9]  School of Architecture, Design and the Built Environment, Nottingham Trent University, 50 Shakespeare St, Nottingham NG1 4FQ, UK; shatirah.akib@ntu.ac.uk
[*]  Correspondence: shahaboddin.shamshirband@tdtu.edu.vn

**Abstract:** Advancement in river flow prediction systems can greatly empower the operational river management to make better decisions, practices, and policies. Machine learning methods recently have shown promising results in building accurate models for river flow prediction. This paper aims to identify models with higher accuracy, robustness, and generalization ability by inspecting the accuracy of a number of machine learning models. The proposed models for river flow include support vector regression (SVR), a hybrid of SVR with a fruit fly optimization algorithm (FOA) (so-called FOASVR), and an M5 model tree (M5). Additionally, the influence of periodicity ($\pi$) on the forecasting enactment was examined. To assess the performance of the proposed models, different statistical meters were implemented, including root mean squared error (RMSE), mean absolute error (MAE), correlation coefficient (R), and Bayesian information criterion (BIC). Results showed that the FOASVR with RMSE (4.36 and 6.33 m$^3$/s), MAE (2.40 and 3.71 m$^3$/s) and R (0.82 and 0.81) values had the best performance in forecasting river flows at Babarud and Vaniar stations, respectively. Also, regarding BIC parameters, $Q_{t-1}$ and $\pi$ were selected as parsimonious inputs for predicting river flow one month ahead. Overall findings indicated that, although both the FOASVR and M5 predicted the river flows in suitable accordance with observed river flows, the performance of the FOASVR was moderately better than the M5 and periodicity noticeably increased the performance of the models; consequently, FOASVR can be suggested as the most accurate method for forecasting river flows.

**Keywords:** river flow forecasting; stream flow; hybrid machine learning; M5 model tree; fruit fly optimization algorithm (FOA); support vector regression; big data; deep learning; hydro-informatics

## 1. Introduction

Dependable approximation of discharge is imperative in water resources management [1]. River flow prediction has emerged from hydrological modeling and transformed into a dynamic and active research area [2,3]. Studying the river flow and stream flow is fundamental to flood protection, sustainable irrigation, and urban development [4,5]. Due to the uncertainties in atmospheric behavior associated with change climate, the dynamic and data-driven methods for hydrological modeling of catchments have become popular more than ever [6]. Data mining and machine learning methods have brought novelty in producing insight from big data [7,8]. These tools forecast forthcoming trends using knowledge-driven decisions resulting from enormous input-output data. The literature includes reviews of the latest machine learning models and comparative studies of the models in river and stream flow forecasting [8–13]. Among the machine learning methods used for river flow prediction, machine learning models presented higher performance with better accuracy and generalization ability for river flow as well many hydrological applications [14]. However, there are many machine learning methods that have never been applied in this realm, which presents a gap in the literature. This paper, consequently, aims at introducing new models and evaluating their performance.

The M5 model tree (M5), as a sub-technique of data mining, constructs tree-based linear models for continuous data. Lately, the implementation of M5 as a decision tree-based regression method has been used for hydrological and water-related studies [15–22]. Londhe and Dixit [23] implemented M5 to estimate the river flow at two stations in India. The models to predict the next day's river flow were established by the preceding day's gauged river flow and rainfall. Sattari et al. [24] inspected the proficiencies of the support vector machine (SVM) and M5 model tree in forecasting the flows of the Sohu River. They revealed that M5 provided more precise predictions when compared with SVM.

SVM is a technique in which the strong points of traditional statistical methods, which are theory-oriented and analytically simple, are utilized. The SVM approach has been frequently implemented in the areas of hydrology and time series forecasting. Liong and Sivapragasam [25] applied the method to foresee floods. Yu et al. [26] suggested a method for forecasting the daily runoff by combining Chaos Theory and the SVM method. Recently, the support vector regression (SVR) method has been developed based on SVM and shows superiority in the prediction of hydrologic processes. Kalteh [27], by applying an Artificial Neural Network (ANN) and SVR to monthly streamflow recorded at two different stations, revealed that both models coupled with wavelet transformation produced more accurate outcomes than the regular models. Also, the results specified that SVR models had enhanced performance in comparison to ANN models. Wu et al. [28] used a genetic algorithm to optimize the SVR model, and the result established that the suggested model could anticipate river flow precisely in comparison with other models. Londhe and Gavraskar [29] utilized the SVR model to forecast river flow one day ahead in two studied locations. The model results were favorable according to the low values of the evaluating metrics.

On the other hand, Cao and Wu [30] coupled the Fruit Fly Optimization algorithm (FOA) with SVR (the combination was named FOASVR) to optimize the parameters of SVR for seasonal electricity consumption forecasting. The results showed that applying FOASVR had a significant role in increasing the prediction accuracy. Lijuan and Guohua [31] used FOASVR to estimate monthly incoming tourist flow, and it was reported that the suggested FOASVR is a viable option for touristic applications. To the best knowledge of the authors, FOA has not previously been integrated with SVR in river flow forecasting.

The primary objective of this research is to present models with higher accuracy, robustness, and generalization ability for river flow prediction. M5, support vector regression, and optimized SVR with FOA were used to forecast river flow at the Vaniar and Babarud stations on the Aji Chay and the Barandouz rivers, respectively, located in the Lake Urmia Basin of Iran. Some evaluation parameters for error estimation are utilized to assess the enactment of the considered models.

## 2. Study Area

The current study used data on the monthly river flow for the Vaniar station on the Aji Chay River and the Babarud station on the Barandouz River, both located at Lake Urmia Basin of Iran (Figure 1). The observed data includes 780 monthly river flow measurements (taken over 65 years from 1952 to 2017) at Babarud station and 744 monthly records (over 62 years from 1952 to 2014) at Vaniar station. There is no basic and technical way of separating training and testing data. For example, the study by Kurup and Dudani [32] used 63% of total data for model development, whereas Pal [33] used 69%, Samadianfard et al. [20,21,34] used 67% of total data, and Deo et al. [35] and Samadianfard et al. [36] used 70% of the total data to develop their models. Thus, to develop the studied models, the data are divided into training (70%) and testing (30%). Table 1 displays the statistics of implemented data for both stations. The observed data confirms high positive values of skewness ($C_{sx}$ = 2.13 and 3.19). Furthermore, the low auto-correlations demonstrate the low persistence for both mentioned stations. It should be noted that Lake Urmia is currently in a drought crisis, with the amount of precipitation and consequently river flow having decreased in recent years; therefore, this may cause some difficulties in forecasting river flows.



**Figure 1.** Babarud and Vaniar stations, located at Lake Urmia Basin [37].

**Table 1.** Statistical parameters of the implemented data ($X_{mean}$, $X_{max}$, $X_{min}$, Sx, Csx, $a_1$, $a_2$, $a_3$ denote the overall mean, maximum, minimum, standard deviation, skewness, lag-1, lag-2, lag-3 auto-correlation coefficients, respectively).

| Station | Data Set | $X_{mean}$ $(m^3/s)$ | $X_{max}$ $(m^3/s)$ | $X_{min}$ $(m^3/s)$ | Sx $(m^3/s)$ | Csx $(m^3/s)$ | $r_1$ | $r_2$ | $r_3$ |
|---------|----------|------|------|------|------|------|------|------|------|
| | Training data | 8.75 | 66.50 | 0.00 | 9.63 | 2.05 | 0.70 | 0.25 | −0.07 |
| Babarud | Testing data | 4.71 | 43.27 | 0.00 | 7.37 | 2.54 | 0.59 | 0.14 | −0.12 |
| | Entire data | 7.74 | 66.50 | 0.00 | 9.28 | 2.13 | 0.69 | 0.25 | −0.05 |
| | Training data | 14.28 | 178.29 | 0.00 | 21.35 | 2.94 | 0.62 | 0.15 | −0.11 |
| Vaniar | Testing data | 5.66 | 65.30 | 0.00 | 10.50 | 3.02 | 0.50 | 0.11 | −0.05 |
| | Entire data | 12.13 | 178.29 | 0.00 | 19.58 | 3.19 | 0.63 | 0.18 | −0.07 |

## 3. Techniques Applied in Modeling

### 3.1. M5 Model Tree

With a constant value at their leaves, model trees are based on regression trees [38]. In this regard, M5, as one of the versions of model trees, has a high capability to forecast continuous numerical attributes [39]. Two different steps are necessary to develop tree models. First, a splitting principle should be utilized to create a decision tree. This criterion is constructed using the standard deviation (*SD*) of the class values which reach a node as a size of the error. So, the standard deviation reduction (*SDR*) is given by:

$$SDR = SD(T) - \sum \frac{|T_i|}{|T|} \times SD(T_i) \tag{1}$$

where $T$ is a set of data that reaches the node and $T_i$ is the $i$th subset of data. After the first step, data in the secondary nodes have lower *SD* as compared with the initial nodes, so M5 selects the split which expects to maximize error reduction. The main drawback of this step is the production of a large tree, which may cause an overfitting problem. Pruning techniques should be employed in order to fix this problem and to avoid overfitting. Therefore, the second step for developing M5 involves these techniques and substitution of subtrees with linear functions. By applying these two steps, M5 develops a linear model for each subspace.

### 3.2. Support Vector Regression (SVR)

SVM is a recognized technique for classification and regression [40]. Generally, regression-based SVM is called SVR. To solve complex problems effectively, SVR is constructed based on minimizing the structural risk. The insensitive loss function ($\varepsilon-$) is identified as the model tolerating errors up to in the training data. Thus, the SVR $\varepsilon$ pursues a linear function as follows:

$$F(x) = w^T x + b \tag{2}$$

where $w$ and $b$ represent the coefficients of the weight vector. This can be clarified as the following problem:

$$Min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\left(\xi_i + \xi_i^*\right) \qquad Subject\ to \begin{cases} F(x) - y_i \leq \varepsilon + \xi_i^* \\ y_i - F(x) \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \ldots, N \end{cases} \tag{3}$$

where $C > 0$ is a penalty parameter which has to be selected earlier. The constant $C$ can grade the experimental error. Moreover, $\xi_i$ and $\xi_i^*$, which are known as slack variables, indicate the distance between real values and the corresponding boundary values of $\varepsilon$-tube. Hence, in order to minimize Equation (2) subject to Equation (3), the function is given by [41,42]:

$$f(x) = \sum_{i=1}^{N}\left(\alpha_i^* - \alpha_i\right)K(x, x_i) + B \tag{4}$$

where $K(x, x_i)$ is the kernel function, $a_i, a_i^* \geq 0$ are the Lagrange multipliers and B is a bias term. The kernel trick is an approach which is used to solve this problem by SVR [43]. In this study, the widely implemented kernel named radial basis function (RBF) is utilized for building an optimum SVR model. Converging fast, working well in high dimensional spaces, and being simple are some advantages of the selected kernel [44]. The kernel function equation is as follows:

$$k(x, x_i) = \exp\left(-\gamma\|x - x_i\|^2\right) \tag{5}$$

where $\gamma$ is the bandwidth of the kernel function and *C*, $\gamma$, and $\varepsilon$ are three predefined parameters. In this research, 1, 0.01, and 0.001 were selected as default amounts which are used in WEKA software. Figure 2 indicates the schematic configuration of the SVR model.
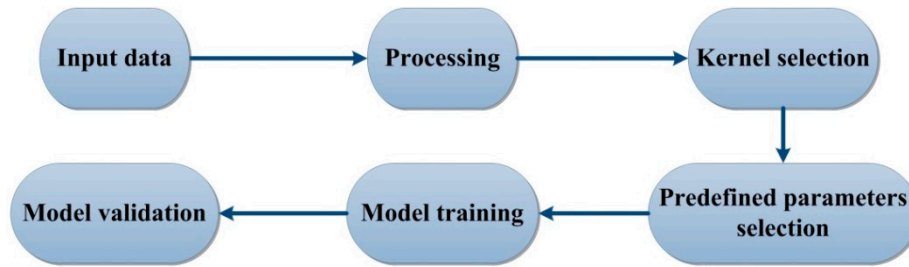


**Figure 2.** Schematic configuration of the SVR model.

### 3.3. Fruit fly Optimization Algorithm (FOA)

FOA, which was introduced by Pan [45], is a swarm intelligence optimization algorithm that imitates the activities of fruit flies to search for the global optimum. Fruit flies can identify a smell from a distance of as much as 40 km and fly all the way to it. Figure 3 displays the food searching progression utilized by the fruit fly iteratively. According to Pan [45], the following equations are exploited to acquire the initial swarm location of a fruit fly (*LR*):

$$X_0 = rand(LR)$$
$$Y_0 = rand(LR) \tag{6}$$

where *LR* is the location range of the accidental initial fruit fly swarm. Subsequently, unexpected search direction and distance for foraging of the fruit flies are given by:

$$X_i = X_0 + rand(FR)$$
$$Y_i = Y_0 + rand(FR) \tag{7}$$

where *FR* is the random flight range, so, smell concentration judgment value (*S*) can be computed by:
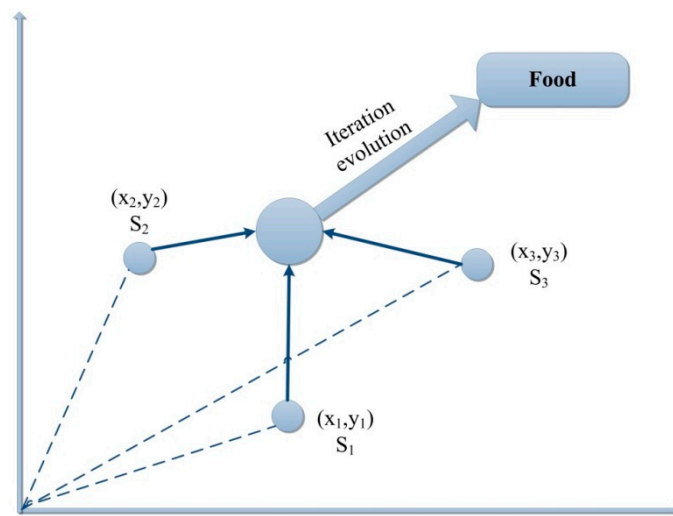
$$S_i = 1/\sqrt{x_i^2 + y_i^2} \tag{8}$$



**Figure 3.** Food searching process utilized by the fruit fly iteratively.

To improve the performance of the river flow forecasting, FOA was implemented for choosing optimized values of three SVR parameters including $C$, $\varepsilon$, $\gamma$, which are connected to $(S_{ci}, S_{\varepsilon i}, S_{\gamma i})$ (i.e., $C = S_{ci}$, $\gamma = S_{\gamma i}$, and $\varepsilon = S_{\varepsilon i}$). It should be noted that the values of LR and FR were selected based on a trial and error procedure for minimizing the prediction errors. The flowchart of the mentioned procedure (FOASVR) is displayed in Figure 4.

The differences between the predicted and the actual values were evaluated by mean squared error (*MSE*) as presented in the equation below:

$$MSE = \frac{\sum_{i=1}^{n} (p_i - o_i)^2}{n} \tag{9}$$

where $p_i$ and $o_i$ are the $i^{th}$ predicted and observed values and n is the entire number of data. The fruit fly saves the finest smell concentration value and the corresponding coordinate among the swarms, then flies toward the next place. When the new result is not superior to the previous iteration, the iteration number reaches its maximum, or the error of the prediction reaches the predefined value, this process will stop. Therefore, optimal values are acquired, and the model has the best performance with these values.
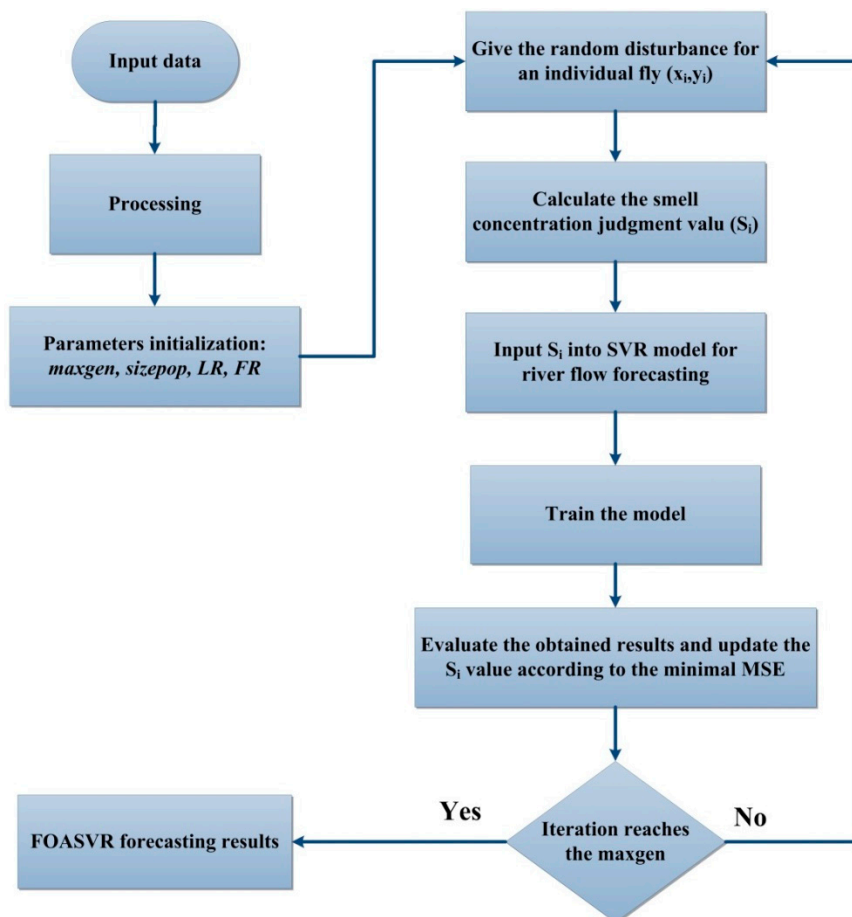


**Figure 4.** FOASVR flowchart.

In this research, data were normalized to be between 0 and 1 because this helps to increase the accuracy of the model and to predict performance [46]. Additionally, LR and FR were chosen to be included at [0, 10] and [−1, 1], respectively; also, the maximum iteration number (maxgen) was equal to 100, and a population size (sizepop) of 20 was selected in order to have reasonable efficiency. Libsvm toolbox was used to run SVR in this article.

## 4. Evaluation Parameters

In this study, different evaluation parameters were considered for scrutinizing the precision of the examined models for river flow forecasting.

As one of the widely-used statistical parameters, the root mean squared error (RMSE) measures the average amount of error (the difference between predicted and observed flows) appropriately, and it can be determined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_p(i) - Q_o(i)\right)^2} \tag{10}$$

where $Q_p(i)$, $Q_o(i)$, and $n$ represent the predicted river flow, the observed river flow, and the number of observations, respectively.

The bias in the predicted river flow is calculated by the mean absolute error (MAE) which measures the closeness of the predictions to the actual flows. Lower MAE values represent more precise predictions of river flow, either equal or close to the observed values. It is calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Q_p(i) - Q_o(i)\right| \tag{11}$$

The correlation coefficient ($R$), which describes the amount of linearity among simulated and observed values of river flows, ranges from −1 to 1 and is described as follows:

$$R = \frac{\left(\sum_{i=1}^{n} Q_o(i)Q_p(i) - \frac{1}{n}\sum_{i=1}^{n} Q_o(i)\sum_{i=1}^{n} Q_p(i)\right)}{\left(\sum_{i=1}^{n} Q_o(i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Q_o(i)\right)^2\right)\left(\sum_{i=1}^{n} Q_p(i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Q_p(i)\right)^2\right)} \tag{12}$$

Also, the Bayesian information criterion ($BIC$) was utilized in order to specify the best model parsimoniously, which means that the model with fewer input parameters would have better performance in comparison to others. $BIC$ measures models relative to each other; in fact, the model with the best performance has the smallest quantity of the BIC [47]. It is given as follows:

$$BIC = n \times \ln\left(\frac{RSS}{n}\right) + K \times \ln(n) \tag{13}$$

where $K$ indicates the number of input parameters and residual sum of squares ($RSS$) can be determined as follows:

$$RSS = \sum_{i=1}^{n}\left(Q_p(i) - Q_o(i)\right)^2 \tag{14}$$

Furthermore, a Taylor diagram (TD), which is a graphical illustration of the observed and forecasted data, was applied to inspect the precision of models [48]. The TD is able to encapsulate some characteristics of the predicted and observed flows at the same time. For instance, this diagram can illustrate RMSE, R, and SD between the forecasted and actual data simultaneously. In the TD, the azimuth angle, the radial distance from the origin, and radial distance from the observed data point denote the R-value, the ratio of the normalized SD, and the RMSE value of the prediction, respectively.

## 5. Results and Discussion

For evaluating the effects of previous monthly flows, three input combinations were established (Table 2). Moreover, the periodicity effect was inspected by appending a component $\pi$ (1 to 12 for each month).

**Table 2.** Input parameters of the established models.

| Model | Input Parameters | Output Parameters |
|-------|------------------|-------------------|
| 1 | $Q_{t-1}$ | $Q_t$ |
| 2 | $Q_{t-1}, \pi$ | $Q_t$ |
| 3 | $Q_{t-1}, Q_{t-2}$ | $Q_t$ |
| 4 | $Q_{t-1}, Q_{t-2}, \pi$ | $Q_t$ |
| 5 | $Q_{t-1}, Q_{t-2}, Q_{t-3}$ | $Q_t$ |
| 6 | $Q_{t-1}, Q_{t-2}, Q_{t-3}, \pi$ | $Q_t$ |

The results of the statistical parameters for studied techniques in the test phase for the Babarud station are given in Table 3. From these tables, it is clear that the periodicity considerably increased each model's accuracy. For the FOASVR model, R increased from 0.63 (for input combination (1)) to 0.82 (for input combination (2)) and similarly, RMSE and MAE indices decreased from 5.74 to 4.36 and from 3.29 to 2.40, respectively. Regarding two previous cases, by adding the periodicity component, R increased from 0.70 to 0.80 and RMSE and MAE decreased from 5.33 to 4.50 and from 2.90 to 2.67, respectively. Finally, in the case of three previous discharge inputs, R increased from 0.67 to 0.79 and RMSE and MAE decreased from 5.69 to 4.58 and from 3.20 to 2.67, respectively. Comparison of FOASVR, M5, and SVR models indicated that the FOASVR-2 model, whose inputs are $Q_{t-1}$, and $\pi$, had better accuracy than the M5 and SVR models. M5 also performed better than the SVR model. Overall, FOASVR performed better than SVR and M5. Also, FOA increased the accuracy of SVR by approximately 27% for RMSE and 38% for MAE in the second scenario, which performed roughly (4% RMSE and 14% MAE) better than M5. Without periodicity, FOASVR-3 indicated a 6% better performance than M5-3, and both models performed better than the SVR-5 model. The relative RMSE and MAE differences between the optimal FOASVR-3 model without periodicity and FOASVR-3 model with periodicity input were 18.2% and 17.2%, respectively. From the BIC point of view, FOASVR-2, M5-2, and SVR-4, with the values of 597.55, 581.85, and 701.18, respectively, had better performance in comparison with other models, which means that these scenarios had parsimonious inputs (accurate results with fewer input parameters). So, for this station input combination, (2) was a reasonable choice. The time variation of observed and predicted river flows by the optimal periodic and non-periodic FOASVR, M5, and SVR models is illustrated in Figures 5–8. It can be comprehended from the Figures 5 and 6 that all three periodic and non-periodic models considerably underestimate some peak flows. It seems that the precision of these models decreases with increasing flow rate. However, the superior accuracy of FOASVR and M5 to the SVR model can also be comprehended from these figures. Comparison of Figures 5 and 6 visibly indicate that the periodic models better approximate the observed river flows than the non-periodic models. Figure 9 displays the scattered diagrams of the observed and predicted monthly river flows for each method. It is noticeable from the graphs that the SVR model performs worse than the other two methods, especially in the prediction of peak river flows. Comparison of the two figures reveals that the periodic models produce more accurate estimates than non-periodic models. Also, this figure indicates that all models (periodic and non-periodic) overestimate some low flows, and the periodic models perform worse than periodic ones in estimating peak flows. This may be because the peak flows do not have any high correlation with the time of the year (i.e., the periodicity value).

The test statistics of the FOASVR, M5, and SVR models for the Vaniar station are provided in Table 4. The encouraging influence of the periodicity component on models' precision is clearly seen for this station. For the FOASVR model, R increased from 0.57 (for input combination (1)) to 0.79 (for input combination (2)) and similarly, RMSE and MAE values decreased from 8.78 to 6.58 and from 4.77 to 3.86, respectively. In the case of two previous discharge inputs, when the periodicity component was added, R increased from 0.55 to 0.80 and RMSE and MAE decreased from 8.88 to 6.48 and from 4.97 to 3.75, respectively. Finally, in the three previous discharge inputs case, R increased from 0.55 to 0.81 and RMSE and MAE values decreased from 8.99 to 6.33 and from 5.53 to 3.71, respectively. Comparison of the three models reveals that the optimal FOASVR-6 model whose inputs are $Q_{t-1}, Q_{t-2}, Q_{t-3}$, and $\pi$

performed better than optimal M5-2 comprising $Q_{t-1}$ and $\pi$ inputs, and both performed better than the optimal SVR-6 model whose inputs are the same as FOASVR-6. Generally, FOASVR performed better than SVR and M5 models; moreover, the accuracy of SVR was increased by 29.7% and 30.4% related to RMSE and MAE, respectively, in the optimal scenario (FOASVR-6) by applying FOA; also, FOASVR showed 16.8% and 19.7% better performance than M5 in terms of RMSE and MAE, respectively, for this scenario. Without the periodicity component, the optimal FOASVR-1 model performed better than the optimal M5-1 and SVR-3 models. The relative RMSE and MAE differences between the optimal FOASVR-1 model without periodicity and FOASVR-1 model with periodicity input were 25.1% and 19.1%, respectively. The best values for BIC in this station were related to FOASVR-6 with 703.64, M5-2 with 740.34, and SVR-2 with 825.05. In light of the fact that FOASVR-6 was closely followed by FOASVR-4 with the value of 707.09 and FOASVR-2 with the value of 707.53, it is better to choose a combination with fewer input parameters. Thus, the input parameters of $Q_{t-1}$ and $\pi$ were selected as a parsimonious scenario for this station, which is similar to the previous station. Figures 7 and 8 demonstrate the time variation of observed and predicted river flows by the optimal periodic and non-periodic FOASVR, M5, and SVR models. As found for the Vaniar station, here also the three periodic and non-periodic models underestimate some peak flows. Comparison of Figures 7 and 8 confirms that appending the periodicity component as the input increases the estimation capacity of the models. The scatter plots of the observed and predicted monthly river flows by each method are shown in Figure 9. As with the previous station, the FOASVR and M5 perform better than the SVR model, especially in the prediction of peak river flows. This figure indicates that the estimates of periodic models are more accurate. According to Figure 9, as at Babarud station, the models overestimate the low flows at the Vaniar station, thereby forecasting shifts from overestimation to underestimation with increasing flow rate.

**Table 3.** The evaluation parameters of studied models in the test period for Babarud Station.

| Model Input | Model | Statistical Parameters | | | |
|---|---|---|---|---|---|
| | | RMSE (m³/s) | MAE (m³/s) | R | BIC |
| $Q_{t-1}$ | SVR-1 | 6.10 | 4.10 | 0.59 | 706.88 |
| | M5-1 | 5.94 | 3.62 | 0.61 | 696.57 |
| | FOASVR-1 | 5.74 | 3.29 | 0.63 | 683.28 |
| $Q_{t-1}, \pi$ | SVR-2 | 5.97 | 3.88 | 0.61 | 703.79 |
| | M5-2 | 4.54 | 2.73 | 0.80 | 597.55 |
| | FOASVR-2 | 4.36 | 2.40 | 0.82 | 581.85 |
| $Q_{t-1}, Q_{t-2}$ | SVR-3 | 5.98 | 4.04 | 0.62 | 704.44 |
| | M5-3 | 5.79 | 3.49 | 0.68 | 691.92 |
| | FOASVR-3 | 5.33 | 2.90 | 0.70 | 659.80 |
| $Q_{t-1}, Q_{t-2}, \pi$ | SVR-4 | 5.85 | 3.83 | 0.64 | 701.18 |
| | M5-4 | 4.55 | 2.83 | 0.80 | 603.67 |
| | FOASVR-4 | 4.50 | 2.67 | 0.80 | 599.39 |
| $Q_{t-1}, Q_{t-2}, Q_{t-3}$ | SVR-5 | 5.91 | 3.90 | 0.62 | 705.14 |
| | M5-5 | 5.79 | 3.50 | 0.68 | 697.18 |
| | FOASVR-5 | 5.69 | 3.20 | 0.67 | 690.42 |
| $Q_{t-1}, Q_{t-2}, Q_{t-3}, \pi$ | SVR-6 | 5.82 | 3.77 | 0.64 | 704.46 |
| | M5-6 | 4.54 | 2.84 | 0.80 | 608.09 |
| | FOASVR-6 | 4.58 | 2.67 | 0.79 | 611.49 |

**Table 4.** The evaluation parameters of studied models in the test period for Vaniar Station.

| Model Input | Model | Statistical Parameters | | | |
|---|---|---|---|---|---|
| | | RMSE (m³/s) | MAE (m³/s) | R | BIC |
| $Q_{t-1}$ | SVR-1 | 9.33 | 5.60 | 0.50 | 831.52 |
| | M5-1 | 9.57 | 5.44 | 0.54 | 840.91 |
| | FOASVR-1 | 8.78 | 4.77 | 0.57 | 809.04 |
| $Q_{t-1}, \pi$ | SVR-2 | 9.04 | 5.31 | 0.52 | 825.05 |
| | M5-2 | 7.19 | 4.46 | 0.77 | 740.34 |
| | FOASVR-2 | 6.58 | 3.86 | 0.79 | 707.53 |
| $Q_{t-1}, Q_{t-2}$ | SVR-3 | 9.21 | 5.57 | 0.53 | 831.95 |
| | M5-3 | 9.80 | 5.46 | 0.59 | 854.92 |
| | FOASVR-3 | 8.88 | 4.97 | 0.55 | 818.45 |
| $Q_{t-1}, Q_{t-2}, \pi$ | SVR-4 | 8.96 | 5.33 | 0.54 | 826.99 |
| | M5-4 | 7.58 | 4.64 | 0.76 | 765.10 |
| | FOASVR-4 | 6.48 | 3.75 | 0.80 | 707.09 |
| $Q_{t-1}, Q_{t-2}, Q_{t-3}$ | SVR-5 | 9.22 | 5.73 | 0.52 | 837.57 |
| | M5-5 | 9.79 | 5.55 | 0.60 | 859.76 |
| | FOASVR-5 | 8.99 | 5.53 | 0.55 | 828.22 |
| $Q_{t-1}, Q_{t-2}, Q_{t-3}, \pi$ | SVR-6 | 9.01 | 5.53 | 0.53 | 834.27 |
| | M5-6 | 7.61 | 4.62 | 0.75 | 771.78 |
| | FOASVR-6 | 6.33 | 3.71 | 0.81 | 703.64 |
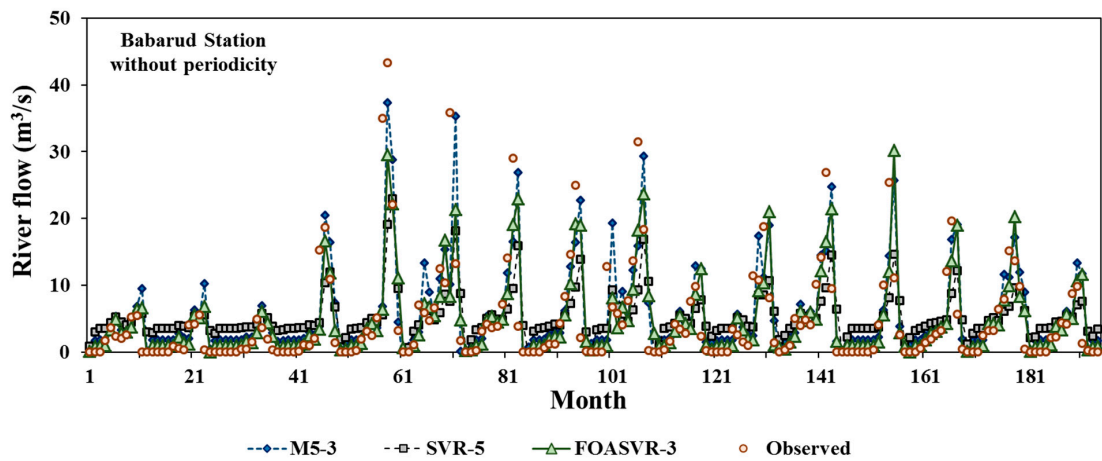


**Figure 5.** The observed and forecasted monthly river flows for Babarud station without periodicity.
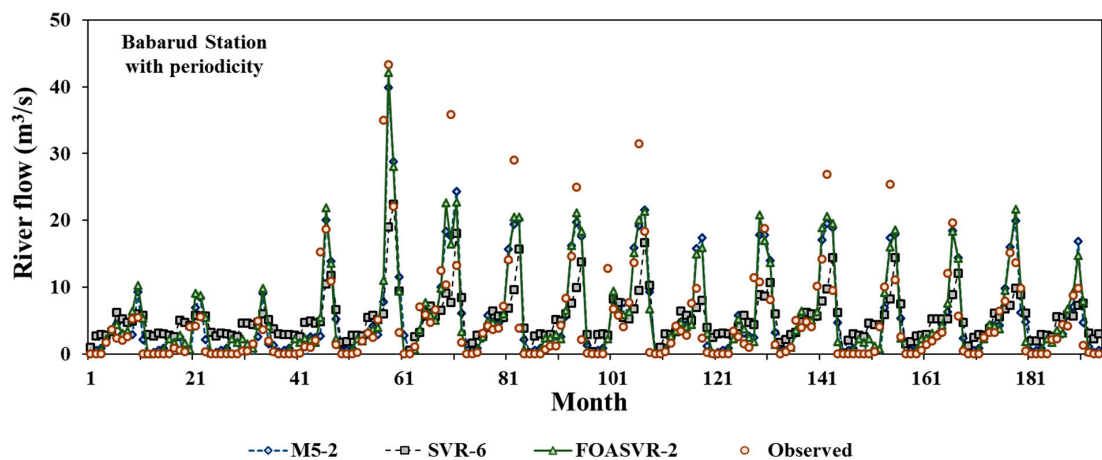


**Figure 6.** The observed and forecasted monthly river flows for Babarud station with periodicity.
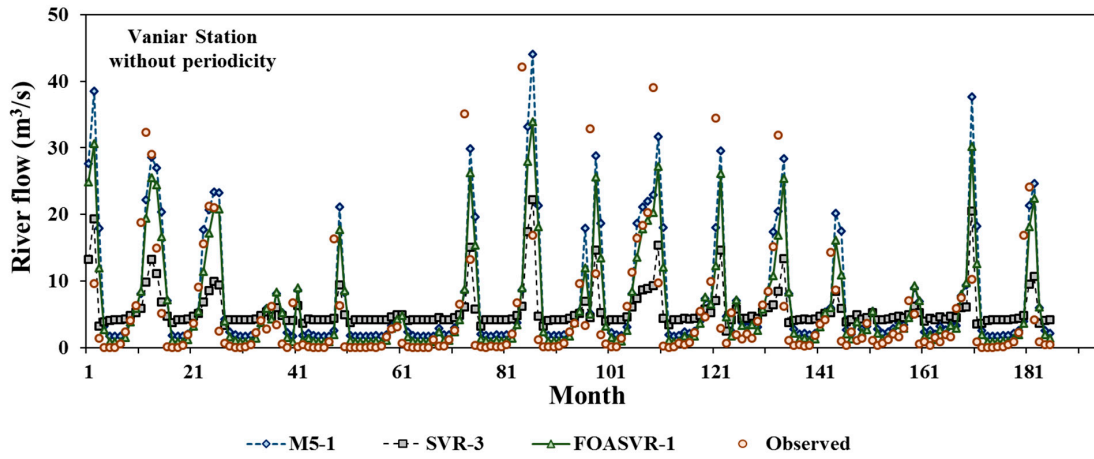
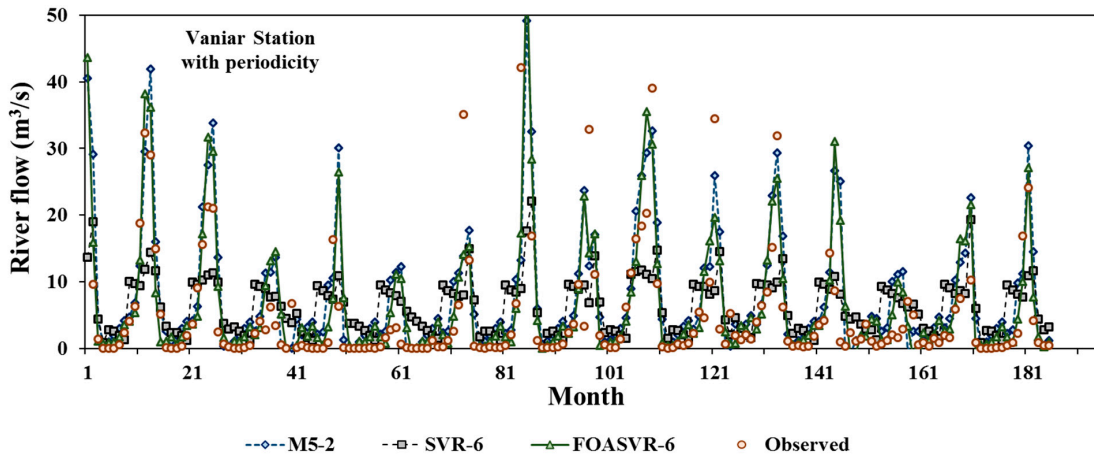**Figure 7.** The observed and forecasted monthly river flows for Vaniar station without periodicity.



**Figure 8.** The observed and forecasted monthly river flows for Vaniar station with periodicity.
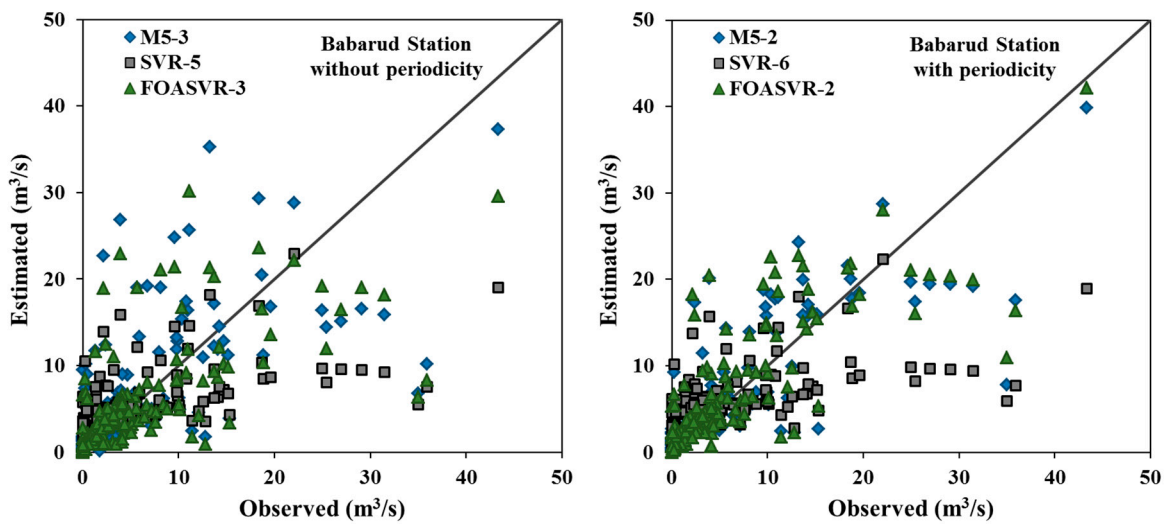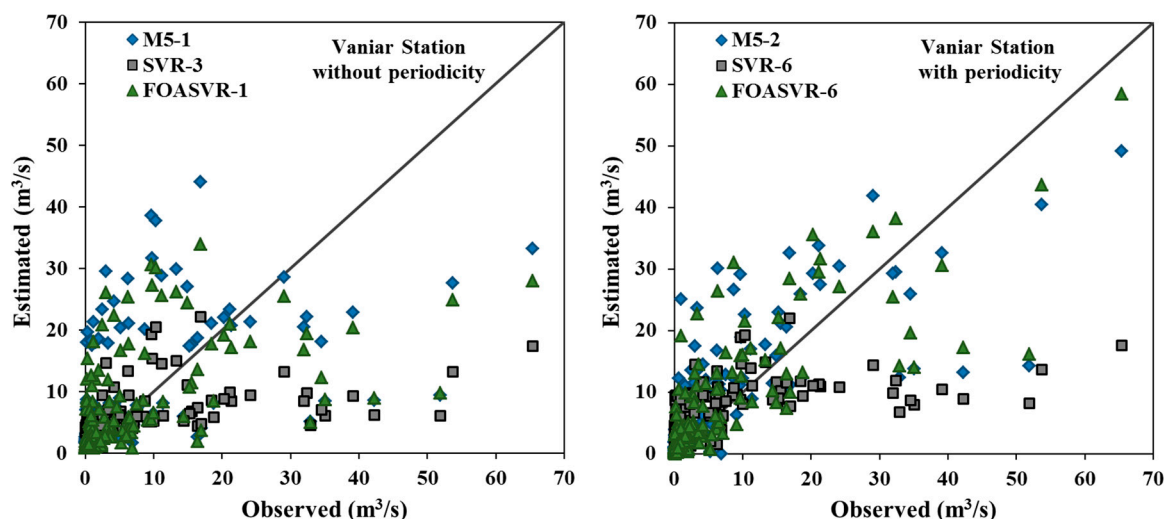


**Figure 9.** *Cont.*

**Figure 9.** The scatter plots of the observed and forecasted monthly river flows. (**a**) Babarud station without periodicity, (**b**) Babarud station with periodicity, (**c**) Vaniar station without periodicity, (**d**) Vaniar station with periodicity.

Furthermore, TDs were utilized to examine SD and R values for the FOASVR, M5, and SVR models. Figure 10 exhibits TDs for all models, where the space from the reference green point is an amount of the centered RMSE. It can be comprehended from Figure 10 that FOASVR (a point with yellow color) provided relatively precise predictions of river flow at both stations.
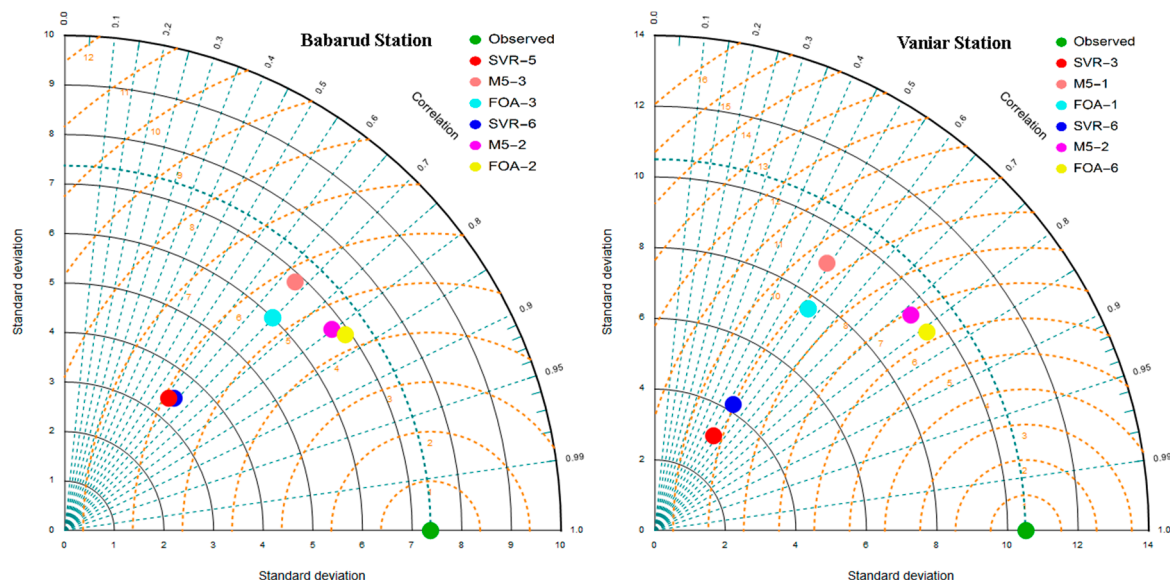


**Figure 10.** Taylor diagrams (TDs) of the monthly predicted river flow. (**a**) Babarud station, (**b**) Vaniar station.

## 6. Conclusions

In the current study, three different data-driven techniques, FOASVR, M5, and SVR, were compared for one month of river flow forecasting at two stations located in the Lake Urmia Basin of Iran. Comparison of three periodic models revealed that the periodic FOASVR model had better accuracy than the periodic M5 and SVR models. M5 was also found to achieve more suitable results than the SVR model. Similarly, the comparison of non-periodic models showed that the optimal FOASVR also had better performance than the M5 and SVR models. It was proved that appending a periodicity component significantly increases models' accuracy in forecasting monthly river flows for both stations. For the Babarud station, the relative RMSE and MAE differences between the optimal periodic and

non-periodic FOASVR models were found to be 18.2% and 17.2%, respectively. For the Vaniar station, the periodicity component decreased the RMSE and MAE values of the optimal FOASVR models by 27.9 and 22.2%, respectively. According to BIC, the second input combination ($Q_{t-1}$ and $\pi$) were used as parsimonious inputs for FOASVR with values of 581.85 and 707.53 for Babarud and Vaniar stations, respectively. Generally, the FOASVR models performed better than the other two methods in forecasting monthly river flows. However, all methods had some difficulties in forecasting peak river flows, while the FOASVR models provided a better forecast in this case. The presented advancement in river flow prediction can highly empower operational river management to make better decisions and policies. The hybrid model of FOASVR shows promising results in building accurate models for river flow prediction.

**Author Contributions:** S.S. (Saeed Samadianfard) and S.S. (Shahaboddin Shamshirband) designed the research; S.J. and E.S. collected hydrological data and interpreted the results, S.S. (Saeed Samadianfard), S.J. and A.M. wrote the initial draft; S.S. (Saeed Samadianfard), S.S. (Shahaboddin Shamshirband) and S.A. revised the manuscript; all authors contributed to the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Onyari, E.; Ilunga, F. Application of MLP neural network and M5P model tree in predicting streamflow: A case study of Luvuvhu catchment, South Africa. In Proceedings of the International Conference on Information and Multimedia Technology (ICMT), Hong Kong, China, 28–30 December 2010; pp. 156–160.

2. Harrigan, S.; Prudhomme, C.; Parry, S.; Smith, K.; Tanguy, M. Benchmarking ensemble streamflow prediction skill in the UK. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2023–2039. [CrossRef]

3. Muhammad, A.; Stadnyk, T.; Unduche, F.; Coulibaly, P. Multi-model approaches for improving seasonal ensemble streamflow prediction scheme with various statistical post-processing techniques in the Canadian Prairie region. *Water* **2018**, *10*, 1604. [CrossRef]

4. Bou-Fakhreddine, B.; Mougharbel, I.; Faye, A.; Chakra, S.A.; Pollet, Y. Daily river flow prediction based on Two-Phase Constructive Fuzzy Systems Modeling: A case of hydrological–meteorological measurements asymmetry. *J. Hydrol.* **2018**, *558*, 255–265. [CrossRef]

5. Baydaroğlu, Ö.; Koçak, K.; Duran, K. River flow prediction using hybrid models of support vector regression with the wavelet transform, singular spectrum analysis and chaotic approach. *Meteorol. Atmos. Phys.* **2018**, *130*, 349–359. [CrossRef]

6. Fernando, A.K.; Shamseldin, A.Y.; Abrahart, B.J. River Flow Forecasting Using Gene Expression Programming Models. In Proceedings of the 10th International Conference on Hydroinformatics HIC 2012, Hamburg, Germany, 14–18 July 2012.

7. Rehana, S. River Water Temperature Modelling Under Climate Change Using Support Vector Regression. In *Hydrology in a Changing World*; Springer: Cham, Switzerland, 2019; pp. 171–183.

8. Yaseen, Z.M.; Sulaiman, S.O.; Deo, R.C.; Chau, K.W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *J. Hydrol.* **2018**, *569*, 387–408. [CrossRef]

9. Azad, A.; Farzin, S.; Kashi, H.; Sanikhani, H.; Karami, H.; Kisi, O. Prediction of river flow using hybrid neuro-fuzzy models. *Arab. J. Geosci.* **2018**, *11*, 718. [CrossRef]

10. Mosavi, A.; Ozturk, P.; Chau, K.W. Flood prediction using machine learning models: Literature review. *Water* **2018**, *10*, 1536. [CrossRef]

11. Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An Ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, *651*, 2087–2096. [CrossRef] [PubMed]

12. Qasem, S.N.; Samadianfard, S.; Nahand, H.S.; Mosavi, A.; Shamshirband, S.; Chau, K.W. Estimating Daily Dew Point Temperature Using Machine Learning Algorithms. *Water* **2019**, *11*, 582. [CrossRef]

13. Tongal, H.; Booij, M.J. Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. *J. Hydrol.* **2018**, *564*, 266–282. [CrossRef]

14. Darwen, P.J. Bayesian model averaging for river flow prediction. *Appl. Intell.* **2019**, *49*, 103–111. [CrossRef]

15. Bhattacharya, B.; Solomatine, D.P. Neural networks and M5 model trees in modeling water level–discharge relationship. *Neurocomputing* **2005**, *63*, 381–396. [CrossRef]

16. Bhattacharya, B.; Solomatine, D.P. Machine learning in sedimentation modeling. *Neural Netw.* **2006**, *19*, 208–214. [CrossRef] [PubMed]

17. Khan, A.S.; See, L. *Rainfall-Runoff Modeling Using Data-Driven and Statistical Methods*; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2006.

18. Siek, M.; Solomatine, D.P. Tree-like machine learning models in hydrologic forecasting: Optimality and expert knowledge. *Geophys. Res. Abstr.* **2007**, *9*, 2–5.

19. Stravs, L.; Brilly, M. Development of a low flow forecasting model using the M5 machine learning method. *Hydrol. Sci.* **2007**, *52*, 466–477. [CrossRef]

20. Samadianfard, S.; Nazemi, A.H.; Sadraddini, A.A. M5 model tree and gene expression programming based modeling of sandy soil water movement under surface drip irrigation. *Agric. Sci. Dev.* **2014**, *3*, 178–190.

21. Samadianfard, S.; Sattari, M.T.; Kisi, O.; Kazemi, H. Determining flow friction factor in irrigation pipes using data mining and artificial intelligence approaches. *Appl. Artif. Intell.* **2014**, *28*, 793–813. [CrossRef]

22. Esmaeilzadeh, B.; Sattari, M.T.; Samadianfard, S. Performance evaluation of ANNs and an M5 model tree in Sattarkhan Reservoir inflow prediction. *ISH J. Hydraul. Eng.* **2017**, *23*, 283–292. [CrossRef]

23. Londhe, S.N.; Dixit, P.R. Forecasting Stream Flow Using Model Trees. *Int. J. Earth Sci. Eng.* **2011**, *4*, 282–285.

24. Sattari, M.T.; Pal, M.; Apaydin, H.; Ozturk, F. M5 Model Tree Application in Daily River Flow Forecasting in Sohu Stream, Turkey. *Water Resour.* **2013**, *40*, 233–242. [CrossRef]

25. Liong, S.Y.; Sivapragasam, C. Flood stage forecasting with support vector machines. *J. AWRA* **2002**, *38*, 173–186. [CrossRef]

26. Yu, X.Y.; Liong, S.Y.; Babovic, V. EC-SVM approach for realtime hydrologic forecasting. *J. Hydroinf.* **2004**, *6*, 209–233. [CrossRef]

27. Kalteh, A.M. Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Comput. Geosci.* **2013**, *54*, 1–8. [CrossRef]

28. Wu, C.L.; Chau, K.W.; Li, Y.S. River stage prediction based on a distributed support vector regression. *J. Hydrol.* **2008**, *358*, 96–111. [CrossRef]

29. Londhe, S.; Gavraskar, S.S. Forecasting One Day Ahead Stream Flow Using Support Vector Regression. *Aquat. Procedia* **2015**, *4*, 900–907. [CrossRef]

30. Cao, G.; Wu, L. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting. *Energy* **2016**, *115*, 734–745. [CrossRef]

31. Lijuan, W.; Guohua, C. Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. *Knowl.-Based Syst.* **2016**, *110*, 157–166. [CrossRef]

32. Kurup, P.U.; Dudani, N.K. Neural networks for profiling stress history of clays from PCPT data. *J. Geotech. Geoenviron. Eng.* **2014**, *128*, 569–579. [CrossRef]

33. Pal, M. M5 model tree for land cover classification. *Int. J. Remote Sens.* **2006**, *27*, 825–831. [CrossRef]

34. Samadianfard, S.; Delirhasannia, R.; Kisi, O.; Agirre-Basurko, E. Comparative analysis of ozone level prediction models using gene expression programming and multiple linear regression. *Geofizika* **2013**, *30*, 43–74.

35. Deo, R.C.; Ghorbani, M.A.; Samadianfard, S.; Maraseni, T.; Bilgili, M.; Biazar, M. Multi-layer perceptron hybrid model integrated with the firefly optimizer algorithm for windspeed prediction of target site using a limited set of neighboring reference station data. *Renew. Energy* **2018**, *116*, 309–323. [CrossRef]

36. Samadianfard, S.; Asadi, E.; Jarhan, S.; Kazemi, H.; Kheshtgar, S.; Kisi, O.; Sajjadi, S.; Abdul Manaf, A. Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths. *Soil Tillage Res.* **2018**, *175*, 37–50. [CrossRef]

37. Available online: https://earth.google.com/web/@32.205151,53.07029487,2852.42968574a,2667368.97567809d, 35y,0.11753984h,16.72644158t,-0r (accessed on 21 February 2019).

38. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*; Morgan Kaufmann: San Francisco, CA, USA, 2005.

39. Quinlan, J.R. Learning with continuous classes. In Proceedings of the Fifth Australian Joint Conf. on Artificial Intelligence, Hobart, Tasmania, 16–18 November 1992; Adams, A., Sterling, L., Eds.; World Scientific: Singapore, 1992; pp. 343–348.

40. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.

41. Gunn, S.R. *Support Vector Machines for Classification and Regression, Technical Report*; University of Southampton: Southampton, UK, 1998.

42. Cimen, M. Estimation of daily suspended sediments using support vector machines. *Hydrol. Sci. J.* **2008**, *53*, 656–666. [CrossRef]

43. Smola, A.J.; Scholkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]

44. Wu, C.-H.; Tzeng, G.-H.; Lin, R.-H. A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Syst. Appl.* **2009**, *36*, 4725–4735. [CrossRef]

45. Pan, W.-T. A new Fruit Fly Optimization Algorithm: Taking the financial distress model as an example. *Knowl.-Based Syst.* **2012**, *26*, 69–74. [CrossRef]

46. Chang, C.-C.; Lin, C.-J. Training v-support vector classifiers: Theory and algorithms. *Neural Comput.* **2001**, *13*, 2119–2147. [CrossRef] [PubMed]

47. Burnham, K.P.; Anderson, D.R. *Model Selection and Inference: A Practical Information-Theoretic Approach*, 2nd ed.; Springer: New York, NY, USA, 2002.

48. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]