



# Article Novel Ensemble Forecasting of Streamflow Using Locally Weighted Learning Algorithm

Rana Muhammad Adnan <sup>1</sup>, Abolfazl Jaafari <sup>2,\*</sup>, Aadhityaa Mohanavelu <sup>3</sup>, Ozgur Kisi <sup>4,\*</sup>, and Ahmed Elbeltagi <sup>5</sup>

- State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China; rana@hhu.edu.cn
- <sup>2</sup> Forest Research Division, Research Institute of Forests and Rangelands, Agricultural Research, Education and Extension Organization (AREEO), Tehran 1496813111, Iran
- <sup>3</sup> Department of Civil Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amritanagar, Coimbatore 641 112, India; aadhityaa65@gmail.com
- <sup>4</sup> Civil Engineering Department, Ilia State University, 0162 Tbilisi, Georgia
- <sup>5</sup> Agricultural Engineering Department, Faculty of Agriculture, Mansoura University, Mansoura 35516, Egypt; ahmedelbeltagy81@mans.edu.eg
- \* Correspondence: jaafari@rifr-ac.ir (A.J.); ozgur.kisi@iliauni.edu.ge (O.K.)

Abstract: The development of advanced computational models for improving the accuracy of streamflow forecasting could save time and cost for sustainable water resource management. In this study, a locally weighted learning (LWL) algorithm is combined with the Additive Regression (AR), Bagging (BG), Dagging (DG), Random Subspace (RS), and Rotation Forest (RF) ensemble techniques for the streamflow forecasting in the Jhelum Catchment, Pakistan. To build the models, we grouped the initial parameters into four different scenarios (M1-M4) of input data with a fivefold cross-validation (I–V) approach. To evaluate the accuracy of the developed ensemble models, previous lagged values of streamflow were used as inputs whereas the cross-validation technique and periodicity input were used to examine prediction accuracy on the basis of root correlation coefficient (R), root mean squared error (RMSE), mean absolute error (MAE), relative absolute error (RAE), and root relative squared error (RRSE). The results showed that the incorporation of periodicity (i.e., MN) as an additional input variable considerably improved both the training performance and predictive performance of the models. A comparison between the results obtained from the input combinations III and IV revealed a significant performance improvement. The cross-validation revealed that the dataset M3 provided more accurate results compared to the other datasets. While all the ensemble models successfully outperformed the standalone LWL model, the ensemble LWL-AR model was identified as the best model. Our study demonstrated that the ensemble modeling approach is a robust and promising alternative to the single forecasting of streamflow that should be further investigated with different datasets from other regions around the world.

**Keywords:** ensemble modeling; additive regression; bagging; dagging; random subspace; rotation forest

## 1. Introduction

To understand the current state, potential, and prospects of water availability, systematic studies on all aspects of basin hydrology (e.g., precipitation, surface, and sub-surface water) and investigation of all indicators are required [1–3]. Streamflow is one such indicator that has a direct influence on local drinking water supply and the quantity of water available for irrigation, hydro-electricity generation, and other needs [4]. Indeed, projections have shown that 20% of the river discharge is controlled by human interventions [5]. Changes in land use and land cover over time, glaciers, snowfields, topographic boundaries, dams, and reservoir management are some of the key factors influencing



Citation: Adnan, R.M.; Jaafari, A.; Mohanavelu, A.; Kisi, O.; Elbeltagi, A. Novel Ensemble Forecasting of Streamflow Using Locally Weighted Learning Algorithm. *Sustainability* 2021, *13*, 5877. https://doi.org/ 10.3390/su13115877

Academic Editor: Giuseppe Barbaro

Received: 21 April 2021 Accepted: 20 May 2021 Published: 24 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). streamflow trends [6]. Streamflow data are a very valuable asset if available over a long period of years. Advantages of streamflow forecasting include early flood warning and mitigation, reservoir planning and management, quantification of available water resources for water supply projects, etc. [7]. Accurate forecasting of streamflow is crucial for the efficient management of water reservoir systems, such as dams, under competing demand for water for irrigation, domestic use, and hydro-power generation activities while at the same time maintaining an adequate environment in the river (or stream) system [8]. In addition, both short-term and long-term streamflow forecasting is necessary pertaining to the optimization of the hydrological components of water resource systems mainly during flood or drought periods [9]. Early prediction of streamflow could provide an imminent warning to disaster management organizations to prepare in response to floods quite early thus preventing the costly socio-economic losses incurred from such extreme events [10].

Since streamflow is a derivative of a complex physical system, the predictions of streamflow using physical-based models generally have significant, inherent uncertainty caused by inaccurate or simple representation of hydrological processes, incomplete or incorrect antecedent conditions, bias or errors in the input variables, or uncertainty in the model parameters. In addition, the requirement for big data (n number of parameters) to simulate the hydrological process also restricts the application of physical models [11]. The application of statistics-based time series models such as the autoregressive integrated moving average (ARIMA) model and its derivatives such as periodic or seasonal ARIMA models and more complex multivariate models such as transfer function-noise (TFN) models have been particularly widely used in forecasting monthly streamflow [12]. However, these models are mostly built upon the assumption that the process follows a normal distribution, however the streamflow process is generally non-linear and stochastic in nature [13]. Machine learning (ML) models, which have been widely used in recent decades to model many real-world problems [14-20], have the unique ability to identify the complex non-linear relationships between the predictors (inputs) and targets (outputs) without the need for the physical characterization of the system or the requirement of making any underlying assumptions. Many hybrid ensemble ML models with the integration of different data preprocessing techniques such as wavelet transformations, empirical mode decomposition, etc. have very high efficiency in accurately forecasting the future streamflow using only antecedent streamflow time series data as input [12,21,22]. Examples of the most recent works on streamflow forecasting can be found in Adnan, Liang, Heddam, Zounemat-Kermani, Kisi and Li [2], Ferreira, et al. [23], Piazzi, et al. [24], Saraiva, et al. [25], and Tyralis, et al. [26].

Although several ML-derived models have been suggested and used to forecast streamflow, there is no model that can forecast streamflow without any biases or with utmost certainty based on the time series of antecedent streamflow values. While literature shows evidence that some single and hybrid ML models, such as OSELM, BGWO-RELM, MLR—KNN, RMGM-BP, RBF-ANN, and MARS-DE, are very effective in forecasting streamflow in river basins across the world, none of these models have been proven to forecast streamflow without any biases or with utmost certainty based on the time series of antecedent streamflow values [2,9,12,24,27]. Hence, the development and application of novel and sophisticated machine learning algorithms for streamflow forecasting are critical to overcoming such limitations in favor of improving the overall forecasting accuracy and model performance. Locally Weighted Learning (LWL) is one such novel machine learning algorithm that has proven efficient for modeling environmental problems. Recently LWLbased ensemble models have been successfully used to model groundwater potential [28] and forest fire susceptibility [29]. One unique advantage of LWL is that for each point of interest a local model is created based on neighboring data of the query point instead of building a whole global model for the entire functional space. Based on this strategy, data points closer to the query point receive a higher weight that can control overprediction. In this study, we combine the LWL algorithm with five ensemble learning techniques, that is, Additive Regression (AR), Bagging (BG), Dagging (DG), Random Subspace (RS), and

Rotation Forest (RF), to develop five ensemble models for a novel ensemble forecasting of streamflow. We apply the models to the lagged streamflow time-series input derived from the antecedent streamflow data. To the best of our knowledge, the LWL technique has not yet been investigated for streamflow forecasting and this study is the first to use and compare different versions of the LWL-based ensemble models for this purpose.

# 2. Case Study

For this study, the Jhelum Catchment located in the western Himalayas in the north part of Pakistan was selected. This catchment originates from India and drains the southern slope of the Greater Himalayas and the northern slope of the Pir Punjal Mountains. The upstream side of the basin located in India is occupied with great glaciers. Due to climate change in recent years, this transboundary river in Pakistan side is greatly affected by glacier melt. Pakistan has a key reservoir (i.e., Mangla Reservoir) downstream of this basin. This reservoir is the second biggest reservoir in Pakistan with an installed capacity of 1000 MW and fulfills 6% of the electricity generation demand of the country. Therefore, precise estimation of this key catchment is very crucial for the economy and sustainability of water resources in Pakistan. This catchment mainly consists of two main sub-basins, that is, the Naran and Neelum basins. The catchment covers a drainage area of 33,342 km<sup>2</sup> up to Mangla Dam with an elevation variation of 200 m to 6248 m. For accurate estimation of streamflow in this basin, the key hydraulic station, that is, Kohala station, at the main river Jhelum streamline after the confluence of both key tributaries (Neelum and Naran) was selected as shown in Figure 1. For model development, the monthly streamflow data of the selected station were obtained from the Water and Power Development Authority (WAPDA) of Pakistan for the duration of 1965 to 2012. For a robust data analysis with the models, a cross-validation scheme was applied. Therefore, data were divided into four equal datasets where each dataset was used for model testing whereas the other three datasets were set aside for model training.



Figure 1. Location map of study area.

#### 3. Methods

### 3.1. Locally Weighted Learning (LWL) Algorithm

The locally weighted learning (LWL) algorithm is motivated by the classification of example-based approaches [30]. In this algorithm, the regression model is not processed unless the output value of the new vector is presented. This is required to correctly execute all learning at the prediction moment. LWL is an advanced type of M5 method in a way that suits both linear and non-linear regression in space for the unique fields of example [31]. Based on the weighted results, distance according to the questionnaire was used to allocate the weights to the training datasets and a regression equation is produced. There is a wide range of methods of distance-based weighting that can be used on the basis of the problem preference in LWL [32]. The statistical model for basic linear regression and the linear model of the multiple regression are presented, respectively, in Equations (1) and (2):

$$y_i = \beta_0 + \beta_{1xi} + \varepsilon_i \ i = 1, \ 2, \ 3....n$$
 (1)

$$y_i = \beta_0 + \beta_{1xi1} + \beta_{2xi2} + \dots + \beta_{kxik} + \varepsilon_i \tag{2}$$

where *y* is the response (dependent variable), *x* is the predictor (independent variable),  $y_i$  and  $\varepsilon_i$  represent random variables, and *xi* is constant. The linear existence of the model is due to  $\beta$ -parameters. The LWL objective function of squared error is expressed as follows:

Minimize 
$$F = \frac{1}{2N} \sum_{K=1}^{N} w_k \left( \propto_{K0} + \sum_{n=1}^{M} \alpha_{kn} x_{kn} + \varepsilon_k - y_k \right)^2$$
 (3)

where *F* is the function of objective, *w* is the weight function matrix, *M* is total variables number,  $\varepsilon_k$  is the random error, and  $\alpha_{K0} \dots \alpha_{kn}$  are regression coefficients.

#### 3.2. Bagging

Bagging or "Bootstrap Aggregating" is a method composed of two major steps for getting more stable, robust, and precise models [33,34]. Bagging is one of the stable ensemble learning techniques used for resampling the training dataset. The first phase consists of bootstrapping the raw data samples that make up the various sets of training data. From these training datasets, multiple models are created. Prediction is generated from the continuous training processes for datasets and multiple models. The underlying notion of the Bagging technique is straightforward. Instead of generating predictions from a standalone model that is appropriate for the actual data, the relationship between the input-output variables is defined by multiple models generated. Then using the weighted average in the Bagged algorithm, various models are coupled to form a single output [35,36]. This strategy can effectively reduce the possible uncertainties in the modeling process. Previous works prove that Bagging is a favorable choice for ensemble modeling of many environmental problems [29].

#### 3.3. Additive Regression

Additive Regression was first developed by Stone [37] as a nonparametric method to approximate a multivariate function by using multiple unary functions. For the dependent variable Y and the independent variables  $X_1, X_2, ..., X_p$ , the nonparametric additive model can be given by:

$$E(Y|X_1, X_2, \dots, X_p = \alpha + \sum_{i=1}^{p} f_i(X_i)$$
 (4)

where  $f_i(X_i)$  is a unary nonparametric function. To satisfy the identifiable conditions, it is generally required that  $f_i(X_i) = 0$ , i = 1, 2, ..., p. Compared to traditional linear models, the nonparametric regression model does not pre-suppose the relationship between variables and the form of the regression function. Further, it is an adaptable and robust data-driven model that can yield a better approximation for nonlinear nonhomogeneous problems [38]. Given these advantages, many researchers applied this technique to study the linear and nonlinear relationships in environmental problems [39].

#### 3.4. Random Subspace (RS)

Random Subspace (RS) was developed by Ho [40] as a new ensemble learning technique for resolving real-world problems. The numerous classifiers of this technique are combined and trained on an altered feature space to generate multiple training subsets for the classifiers, which are the training bases. RS applies multiple samples on function space, as opposed to the example space as in other ensemble models, as stated by Havlíček, et al. [41]. This strategy takes advantage of bootstrapping and grouping. The RS inputs are the training set (x), the base-classifier (w), and the subspaces number (L) [42]. It is strongly recommended by Pham, et al. [43] that this approach be used to prevent over-fitting issues and to cope with the most unnecessary datasets.

#### 3.5. Dagging

Ting and Witten [44] pioneered the Dagging algorithm as a resampling ensemble technique that uses most votes to combine various classifiers to get improved prediction accuracy for the base classifier. Dagging generates multiple different samples instead of producing the bootstrap samples to acquire the base classifier. In recent years, it has been considered a promising machine learning algorithm for classification problems. In the real world, the Dagging ensemble technique has been applied to solve different classification problems. The development of an *M* dataset can occur with a specific training dataset containing *N* samples which may come from the existing training datasets [45,46]. There are many n (n < N) samples in any dataset that are distinct from each other. In the particular training datasets, the variables are not replaced and can be chosen as a part of the dataset specified where the size of sample datasets. Ultimately, depending on the training dataset, many classifiers can be acquired. The capability of Dagging has been frequently proven for obtaining improved predictive modeling of different classification problems [29,47].

#### 3.6. Rotation Forest

Rotation Forest (RF) is an ensemble learning technique that independently trains *L* decision trees using, for each tree, a different set of extracted features. Suppose the  $x = (x_1, \dots, x_n)$  T represents an example defined by n characteristics (attributes) and let *X* be an  $N \times n$  matrix including examples of the training process. We assume that the actual class labels of all instances of training are also given. Let go of  $D = \{D_1, \dots, D_L\}$  is the set of classifiers for *L* and *F* is the set of characteristics. The purpose of Rotation Forest is to create precise and diverse classifiers. As in Bagging, bootstrap samples are taken as the training collection for the individual classifiers. The key heuristic is to introduce extraction of features and to recreate a complete feature set for each classifier in the ensemble afterward [48]. The feature collection is randomly divided into *K* subsets to do this. The principal component analysis (PCA) is run on each subset separately, and a new set of n linear extracted features is constructed by pooling all main components. The data are translated into the new space of the function linearly. With this data collection, classifier  $D_i$  is educated. Multiple splits of the collection of features will contribute to various extracted features, thereby leading to the diversity of the bootstrap sampling implemented.

#### 4. Ensemble Forecasting

Ensemble forecasting of the monthly streamflow was performed using the LWL algorithm that was used as the base model and was combined with the Additive Regression (AR), Bagging (BG), Dagging (DG), Random Subspace (RS), and Rotation Forest (RF) ensemble techniques. This combination resulted in five ensemble models, namely the ensemble LWL-AR, LWL-BG, LWL-DG, LWL-RS, and LWL-RF models. In each model, the

ensemble learning technique performs resampling of the training dataset to train the base LWL algorithm. Table 1 details the summary of statistical characteristics of the data used in this study. To build the models, we grouped the initial input parameters into four different scenarios of input data. They include:

- (i) Qt-1
- (ii) Qt-1, Qt-2
- (iii) Qt-1, Qt-2, Qt-3
- (iv) Qt-1, Qt-2, Qt-3, MN

where Qt-1 is the streamflow at 1 previous month and vice versa and MN is the month number of the streamflow.

	Table 1. An overview	of the statistical	characteristics	of the data used.
--	----------------------	--------------------	-----------------	-------------------

Statistics	Whole Dataset (m <sup>3</sup> /s) 1965 to 2012	M1 Dataset (m <sup>3</sup> /s) 2001 to 2012	M2 Dataset (m <sup>3</sup> /s) 1989 to 2000	M3 Dataset (m <sup>3</sup> /s) 1977 to 1988	M4 Dataset (m <sup>3</sup> /s) 1965 to 1976
Mean	772.9	794.0	783.7	835.8	678.0
Min.	110.7	112.3	134.9	127.0	110.7
Max.	2824	2824	2426	2773	2014
Skewness	0.886	0.931	0.716	0.845	0.888
Std. dev.	609.2	645.1	600.6	651.7	514.1
Variance	371,069	416,106	360,780	424,712	264,330

In a cross-validation approach, data were divided into four equal sets such that three sets were used for model training and the remaining set was used for validation [49–51]. We used several performance metrics to measure the performance of the models during both training and validation phases. These metrics include: correlation coefficient (R) (Equation (5), root mean square error (RMSE) (Equation (6), mean absolute error (MAE) (Equation (7), relative absolute error (RAE) (Equation (8), and root-relative square error (RRSE) (Equation (9). A full description of these metrics can be found in the corresponding literature [2,24,52–55].

$$R = \frac{\sum (P_i - \overline{P}) (T_i - \overline{T})}{\sqrt{\sum (P_i - \overline{P})^2 \sum (T_i - \overline{T})^2}}$$
(5)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(P_{ij} - T_{j}\right)^{2}}{N}}$$
(6)

$$MAE = \frac{\sum_{i=1}^{n} |P_{ij} - T_j|}{N}$$
(7)

$$RAE = |\frac{P_{ij} - T_j}{T_j}| \times 100 \tag{8}$$

$$RRSE = \sqrt{\frac{\sum_{i=1}^{n} (P_{ij} - T_{j})^{2}}{\sum_{i=1}^{n} (T_{j} - \overline{T}_{j})^{2}}}$$
(9)

where *P* is the value predicted, *T* is the target value,  $\overline{P}$  and  $\overline{T}$  are the mean predicted and target values.

We developed the models using the open-source Weka software on an HP Laptop with an Intel(R) Core (TM) i3-3110M CPU @ 2.40GHz, 4 GB of RAM, an x64-based processor, and the Microsoft Windows 8.1 operating system. The optimum value for each model parameter was identified via a trial-and-error process. To do so, we arbitrarily entered different values until the best model performance was achieved [36,56]. Table 2 details the optimum parameter setting of each model.

<b>D</b> (			Mod	el		
Parameter	LWL	AR	BG	DG	RS	RF
Debug	False	False	False	False	False	False
Search algorithm	Linear NN search	-	-	-	-	-
Weighting kernel	0	-	-	-	-	-
Number of iterations	-	14	12	10	10	11
Shrinkage	-	0.1	-	-	-	-
Bag size percent	-	-	100	-	-	-
Seed	-	-	1	1	1	1
Number of folds	-	-	-	10	-	-
Verbose	-	-	-	False	-	-
Number of boosting iterations	-	30	-	-	-	-
Subspace size	-	-	-	-	0.5	-
Max group	-	-	-	-	-	3
Min group	-	-	-	-	-	3
Number of groups	-	-	-	-	-	False
Projection filter	-	-	-	-	-	PCA
Removed percentage	-	-	-	-	-	50

Table 2. Optimum parameter setting of the models.

#### 5. Results

Table 3 shows the results of the single LWL model with different input combinations and datasets. Given the mean values of each metric obtained from each input combination and dataset, the model with input combination IV performed the best and achieved RMSE = 244.6 m<sup>3</sup>/s, MAE = 175 m<sup>3</sup>/s, RAE = 34.47 m<sup>3</sup>/s, RRSE = 40.90 m<sup>3</sup>/s, and R = 0.834 in the training phase and RMSE = 274.8 m<sup>3</sup>/s, MAE = 199.2 m<sup>3</sup>/s, RAE = 38.70 m<sup>3</sup>/s, RRSE = 44.08 m<sup>3</sup>/s, and R = 0.809 in the testing phase. Importing periodicity (i.e., MN) as an additional input variable into the model considerably improved both the training performance and prediction performance. A comparison between the results obtained from the input combinations III and IV revealed a significant performance improvement, that is, RMSE, MAE, RAE, and RRSE decreased up to 10.12, 14.59, 15.41, and 10.69% in the training phase and 6.17, 9.41, 8.40, and 9.56% in the testing phase, respectively. In terms of the R metric, the results showed 5.3 and 6.1% training and testing improvements when we used input combination IV. Further, the results revealed that the best and worst predictive performance (i.e., testing performance) was obtained with the datasets M3 and M2, respectively.

			Trai	ning		Testing					
Metric	Data Set		Input Cor	nbination			Input Combination				
		Ι	II	III	IV	Ι	II	III	IV		
	M1	358.6	300.3	295.5	255.9	365.8	308.6	311.4	295.4		
	M2	358.7	303.7	275.5	242.1	397.0	370.2	369.5	328.0		
RMSE	M3	358.8	283.8	271.5	244.0	382.1	303	292.9	274.8		
	M4	362.3	306.5	300.4	252.3	397.9	342.2	312.4	277.7		
	Mean	359.6	298.6	285.7	248.6	385.7	331.0	321.6	294.0		
	M1	282.6	227.7	226.0	183.5	271.0	231.3	241.4	207.7		
MAE	M2	279.9	227.0	210.1	178.8	306.9	263.9	265.0	222.2		
	M3	274.4	213.8	204.9	175.0	291.5	228.8	219.9	199.2		
	M4	281.5	230.8	229.0	183.5	309.8	257.2	240.9	200.0		
	Mean	279.6	224.8	217.5	180.2	294.8	245.3	241.8	207.3		
	M1	52.24	42.09	41.78	35.92	57.57	49.12	51.27	44.12		
	M2	55.67	44.14	41.39	35.57	56.68	48.74	48.95	41.03		
RAE	M3	53.35	42.51	40.75	34.47	56.01	43.95	42.25	38.70		
	M4	55.47	45.47	44.53	33.67	57.56	47.79	44.75	38.16		
	Mean	54.18	43.55	42.11	34.91	56.96	47.40	46.81	40.50		
	M1	58.80	47.42	46.64	39.94	65.81	57.8	58.32	55.32		
	M2	60.51	49.62	46.19	40.85	63.60	56.34	56.24	49.91		
RRSE	M3	58.63	47.88	45.80	40.90	60.42	50.43	48.74	44.08		
	M4	60.74	51.38	49.09	41.72	61.62	52.99	48.38	46.24		
	Mean	59.67	49.08	46.93	40.85	62.86	54.39	52.92	48.89		
	M1	0.659	0.776	0.783	0.841	0.594	0.672	0.676	0.746		
	M2	0.642	0.750	0.792	0.834	0.612	0.687	0.694	0.759		
R	M3	0.658	0.773	0.792	0.834	0.629	0.746	0.762	0.809		
	M4	0.634	0.736	0.759	0.826	0.619	0.723	0.757	0.789		
	Mean	0.648	0.759	0.782	0.834	0.614	0.707	0.722	0.776		

Table 3. Results of the single LWL model.

The best performance is shown in bold.

The results of the five ensemble models, that is, LWL-AR, LWL-BG, LWL-DG, LWL-RS, and LWL-RF, are summed up in Tables 4–8. Similar to the single LWL model, the performance of the ensembles models was predominantly influenced by the input combination and dataset. For example, RMSE of the testing phase ranged from 223.9 m<sup>3</sup>/s (M3-IV) to 407.8 m<sup>3</sup>/s (M2-I) for LWL-AR, from 255.3 m<sup>3</sup>/s (M3-IV) to 345.2 m<sup>3</sup>/s (M2-III) for LWL-BG, from 233.5 m<sup>3</sup>/s (M3-IV) to 390.6 m<sup>3</sup>/s (M2-I) for LWL-DG, from 242.8 m<sup>3</sup>/s (M3-IV) to 397.2 m<sup>3</sup>/s (M4-I) for LWL-RS, and from 229.4 m<sup>3</sup>/s (M3-IV) to 397 m<sup>3</sup>/s (M2-I) for LWL-RF. Given these values and also the values of other performance metrics, it is evident that the best performance of all models was achieved by the dataset M3 and the input combination IV (i.e., M3-IV).

Table 4. Results of the ensemble LWL-AR model.

			Trai	ning		Testing					
Metric	Dataset		Input Cor	nbination		Input Combination					
		Ι	II	III	IV	Ι	II	III	IV		
	M1	321.0	184.5	162.4	143.7	327.5	292.8	293.3	261.9		
	M2	310.0	183.8	170.3	128.3	407.8	334.4	315.2	273.5		
RMSE	M3	306.9	174.0	152.1	138.4	373.6	264.4	258.8	223.9		
	M4	314.2	193.3	169.1	139.1	377.6	294.6	284.2	242.9		

			Trai	ning			Test	ing	
Metric	Dataset		Input Cor	nbination			Input Con	nbination	
		Ι	II	III	IV	I	II	III	IV
	Mean	313.0	183.9	163.5	137.4	371.6	296.6	287.9	250.6
	M1	248.1	135.8	115.8	97.47	247.2	199.8	195.9	171.1
	M2	241.6	130.2	120	88.47	310.9	224.2	209.4	168.6
MAE	M3	232.2	125.1	104.8	95.72	292.1	191.7	183.8	150.7
	M4	242.6	136.6	117.8	95.72	295.1	200.4	198.6	156.0
	Mean	241.1	131.9	114.6	94.30	286.3	204.0	196.9	161.6
	M1	45.87	25.11	21.43	17.70	52.51	42.44	41.61	33.90
RAE	M2	48.04	25.31	23.32	17.60	57.43	41.41	38.67	31.13
	M3	45.15	24.87	20.83	18.95	56.11	36.82	35.31	28.95
	M4	47.79	26.92	23.20	18.60	54.83	37.23	36.89	31.80
	Mean	46.71	25.55	22.20	18.21	55.22	39.48	38.12	31.45
	M1	50.69	29.13	25.64	21.86	61.33	54.83	54.93	45.49
	M2	52.29	30.03	27.83	21.64	62.06	50.90	47.96	41.62
RRSE	M3	50.15	29.36	25.65	23.48	62.18	44.00	43.08	23.26
	M4	52.67	32.4	28.34	23.31	58.47	45.63	44.02	40.57
	Mean	51.45	30.23	26.87	22.57	61.01	48.84	47.50	37.74
	M1	0.743	0.916	0.935	0.953	0.621	0.740	0.733	0.823
	M2	0.728	0.910	0.924	0.953	0.612	0.743	0.773	0.828
R	M3	0.750	0.914	0.935	0.945	0.616	0.808	0.821	0.867
	M4	0.723	0.903	0.922	0.947	0.658	0.794	0.806	0.835
	Mean	0.736	0.911	0.929	0.950	0.627	0.771	0.783	0.838

Table 4. Cont.

The best performance is shown in bold.

# Table 5. Results of the ensemble LWL-BG model.

			Trai	ning	Testing				
Metric	Dataset		Input Cor	nbination	L	I	nput Con	nbinatio	n
		Ι	II	III	IV	Ι	II	III	IV
	M1	363.6	290.0	272.9	240.9	345.5	294.0	274.3	261.2
	M2	352.3	285.8	266.5	237.9	398.1	340.6	345.2	306.5
RMSE	M3	336.6	262.0	250.7	229.4	359.4	292.8	276.1	255.3
	M4	342.8	289.5	272.6	250.5	376.4	319.1	294.7	258.8
	Mean	348.8	281.8	265.7	239.7	369.9	311.6	297.6	270.5
	M1	284.4	224.2	212.6	177.4	269.0	226.6	218.5	187.7
	M2	273.6	217.4	202.1	174.8	310.9	243.5	247.3	208.0
MAE	M3	265.9	202.0	191.0	169.8	279.4	223.6	209.5	191.7
	M4	270.6	220.8	205.1	183.0	300.5	248.6	225.3	182.8
	Mean	273.6	216.1	202.7	176.3	290.0	235.6	225.2	192.6
	M1	52.58	41.45	39.31	32.90	57.13	48.14	46.40	38.83
	M2	53.19	42.27	39.83	34.45	57.42	44.97	45.69	38.43
RAE	M3	52.88	40.16	37.98	33.77	53.68	42.95	40.24	34.88
	M4	53.32	43.50	39.88	35.57	55.83	46.19	41.86	36.83
	Mean	52.99	41.85	39.25	34.17	56.02	45.56	43.55	37.24
	M1	57.42	45.79	43.10	38.04	64.71	55.05	51.37	48.46
	M2	57.56	46.71	44.67	39.88	60.58	51.83	52.54	46.65
RRSE	M3	56.77	44.19	42.29	38.70	59.82	48.73	45.96	40.45
	M4	57.47	48.53	44.54	40.94	58.30	49.42	45.64	42.48
	Mean	57.31	46.31	43.65	39.39	60.85	51.26	48.88	44.51
	M1	0.672	0.794	0.817	0.859	0.590	0.694	0.743	0.781
	M2	0.669	0.783	0.803	0.845	0.627	0.736	0.731	0.796
R	M3	0.679	0.808	0.824	0.854	0.646	0.762	0.789	0.845
	M4	0.671	0.766	0.803	0.834	0.661	0.760	0.799	0.821
	Mean	0.673	0.788	0.812	0.848	0.631	0.738	0.766	0.811

The best performance is shown in bold.

			Trai	ning		Testing				
Metric	Dataset		Input Cor	nbination			Input Cor	nbination		
		Ι	II	III	IV	Ι	II	III	IV	
	M1	369.2	310.1	279.1	241.0	320.3	274.1	259.0	249.4	
	M2	355.4	296.3	270.9	264.8	390.6	335.8	326.4	298.8	
RMSE	M3	338.3	285.6	262.6	234.5	349.6	288.2	253.9	233.5	
	M4	346.6	299.3	271.7	239.5	337.6	324.9	286.0	247.1	
	Mean	352.4	297.8	271.1	245.0	349.5	305.8	281.3	257.2	
	M1	286.0	236.6	211.7	171.3	248.0	217.4	206.0	191.1	
	M2	275.2	225.8	201.7	194.0	299.3	246.8	229.3	209.4	
MAE	M3	262.2	220.1	200.1	166.9	274.3	227.2	197.8	181.3	
	M4	267.5	225.2	207.3	181.4	298.8	252.2	220.8	179.4	
	Mean	272.7	226.9	205.2	178.4	280.1	235.9	213.5	190.3	
	M1	52.87	43.75	39.14	31.67	52.67	46.18	43.76	39.85	
	M2	53.51	43.89	39.21	37.71	55.28	45.59	42.36	38.68	
RAE	M3	52.15	43.77	39.80	32.89	52.69	43.66	37.99	33.34	
	M4	52.70	44.37	40.84	36.07	55.52	46.85	41.02	36.72	
	Mean	52.81	43.95	39.75	34.59	54.04	45.57	41.28	37.15	
	M1	58.30	48.97	44.07	38.06	59.98	51.32	48.50	44.56	
	M2	58.07	48.42	44.27	43.28	59.44	51.10	49.67	45.48	
RRSE	M3	57.06	48.18	44.30	39.30	58.18	47.96	42.26	38.26	
	M4	58.09	50.18	45.54	40.40	58.64	50.32	44.29	41.52	
	Mean	57.88	48.94	44.55	40.26	59.06	50.18	46.18	42.46	
	M1	0.663	0.766	0.814	0.867	0.623	0.724	0.753	0.803	
	M2	0.663	0.771	0.806	0.815	0.643	0.753	0.778	0.797	
R	M3	0.676	0.774	0.815	0.848	0.663	0.774	0.824	0.847	
	M4	0.663	0.753	0.797	0.841	0.659	0.764	0.821	0.828	
	Mean	0.666	0.766	0.808	0.843	0.647	0.754	0.794	0.819	

 Table 6. Results of the ensemble LWL-DG model.

The best performance is shown in bold.

Table 7.	Results of	the ensemble	LWL-RS model.

			Trai	ning			Tes	ting	
Metric	Dataset		Input Cor	nbination			Input Cor	nbination	
	-	Ι	II	III	IV	Ι	II	III	IV
	M1	371.3	329.9	287.0	270.4	351.4	302.1	274.1	248.7
	M2	358.7	319.8	301.1	282.0	397.0	362.1	371.6	345.5
RMSE	M3	359.8	317.2	279.3	268.2	382.1	326.8	302.7	242.8
	M4	362.3	296.7	319.8	280.4	397.9	344.3	319.2	302.5
	Mean	363.0	315.9	296.8	275.3	382.1	333.8	316.9	284.9
	M1	282.6	261.1	225.2	200.3	271.0	248.1	221.8	192.7
	M2	279.9	238.0	231.8	207.2	306.9	295.5	274.1	243.0
MAE	M3	274.4	245.1	215.7	205.9	291.5	263.8	238.5	191.4
	M4	281.5	248.6	238.0	219.3	309.8	276.7	240.7	228.3
	Mean	279.6	248.2	227.7	208.2	294.8	271.0	243.8	213.9
	M1	52.25	46.76	41.64	37.09	57.56	52.68	47.11	40.66
	M2	55.68	49.13	45.06	40.29	56.68	53.43	50.63	44.88
RAE	M3	53.35	50.64	42.51	40.57	56.01	51.27	45.81	37.03
	M4	55.47	48.78	49.13	43.62	57.56	54.41	44.73	42.41
	Mean	54.19	48.83	44.59	40.39	56.95	52.95	47.07	41.25
	M1	58.63	49.88	45.32	42.70	65.81	61.71	51.32	45.47
	M2	60.51	55.97	49.20	46.08	60.42	58.54	56.56	52.58
RRSE	M3	58.80	54.89	46.82	44.96	63.60	57.68	50.39	41.39
	M4	60.74	54.83	55.97	47.30	61.62	57.72	49.43	46.86
	Mean	59.67	53.89	49.33	45.26	62.86	58.91	51.93	46.58
	M1	0.659	0.676	0.806	0.837	0.594	0.637	0.736	0.796
	M2	0.642	0.714	0.769	0.814	0.629	0.659	0.702	0.773
R	M3	0.659	0.682	0.790	0.821	0.612	0.676	0.750	0.848
	M4	0.634	0.679	0.714	0.792	0.619	0.671	0.769	0.815
	Mean	0.649	0.688	0.770	0.816	0.614	0.661	0.739	0.808

The best performance is shown in bold.

			Trai	ning		Testing				
Metric	Dataset		Input Cor	nbination	L	I	nput Con	nbinatio	ı	
		Ι	II	III	IV	Ι	II	III	IV	
	M1	371.3	259.9	261.7	225.4	351.4	289.8	278.3	232.4	
	M2	359.8	271.7	269.4	229.1	397.0	307.4	336.3	300.2	
RMSE	M3	358.7	242.9	253.0	213.8	382.1	265.6	266.3	229.4	
	M4	362.3	271.0	311.5	230.6	397.9	297.3	311.4	266.4	
	Mean	363.0	261.4	273.9	224.7	382.1	290.0	298.1	257.1	
Metric RMSE MAE RAE RRSE	M1	282.6	196.2	200.0	167.4	271.0	218.0	212.8	195.7	
	M2	274.4	203.4	205.2	165.0	306.9	223.1	237.4	201.7	
MAE	M3	279.9	184.2	193.0	160.0	291.5	195.5	195.6	173.0	
	M4	281.5	204.1	237.8	167.3	309.8	227.8	237.8	175.7	
	Mean	279.6	197.0	209.0	164.9	294.8	216.1	220.9	186.5	
RAE	M1	52.25	36.27	36.97	30.96	57.57	46.30	45.19	37.31	
	M2	53.35	39.55	39.90	32.07	56.68	41.20	43.86	37.25	
	M3	55.68	36.62	38.38	31.36	56.01	37.56	37.57	33.23	
	M4	55.47	40.22	44.18	32.96	57.56	42.32	44.18	36.37	
	Mean	54.19	38.17	39.86	31.84	56.96	41.85	42.70	36.04	
	M1	58.63	41.03	41.33	35.59	65.81	54.27	52.11	42.96	
	M2	58.80	44.41	44.03	37.44	60.42	46.78	51.17	45.68	
RRSE	M3	60.51	40.98	42.68	36.06	63.60	44.21	44.32	38.68	
IUCE	M4	60.74	45.42	48.24	38.66	61.62	46.04	48.24	41.26	
	Mean	59.67	42.96	44.07	36.94	62.86	47.83	48.96	42.15	
	M1	0.659	0.834	0.830	0.88	0.594	0.714	0.753	0.821	
	M2	0.659	0.805	0.806	0.869	0.629	0.787	0.750	0.806	
R	M3	0.642	0.835	0.819	0.882	0.612	0.808	0.805	0.858	
	M4	0.634	0.796	0.771	0.856	0.619	0.799	0.771	0.846	
	Mean	0.649	0.818	0.807	0.872	0.614	0.777	0.770	0.833	

Table 8. Results of the ensemble LWL-RF model.

The best performance is shown in bold.

A comparison between the results obtained from the single LWL model and its ensembles clearly indicates that the ensemble learning techniques considerably improved the training and testing performances of the base LWL algorithm. The ensemble models achieved greater training performance than the single LWL model by about 44.7, 44.7, 47.8, 44.7, and 13.9% in terms of the RMSE, MAE, RAE, RRSE, and R metrics, respectively. In the case of the testing performance, LWL-AR showed 53.3, 54.5, 55, 53.8, and 22.4% improvements. Similarly, testing performance improvements in the corresponding metrics are 8, 7.1, 8, 9, and 4.5% by applying LWL-BG, 12.5, 8.2, 8.3, 13.2, and 5.5% by applying LWL-DG, 3.1, 3.2, 1.9, 4.7, and 4.1% by applying LWL-RS, 12.6, 10, 11, 13.8, and 7.3% by applying LWL-RF, respectively.

A comparison of the models' outcomes also reveals that the ensemble LWL-AR model performed better than the other models in both training and testing phases of the monthly streamflow modeling. The LWL-DG and LWL-RF models showed similar performance and ranked as the second-best models, followed by the LWL-RS model that was identified as the least effective ensemble model.

To further compare the models' performance, we used time variation, scatter plots, and Taylor and violin diagrams to visualize the results obtained from the best input combination (i.e., M3-IV). Figure 2 shows that LWL-AR predictions are much closer to the observed values compared to the other models. Figure 3 reveals that the ensemble LWL-AR model performed better compared to other models in catching the extreme streamflow values (minimum and maximum), which is an important indicator in water resource management and for the evaluation of extreme events such as drought and flood.



**Figure 2.** Time variation graphs of the observed and predicted streamflow by different LWL-based ensemble models in the testing phase using the best input combination (M3-IV).



**Figure 3.** Comparison of different LWL-based ensemble models for forecasting peak streamflow in the testing phase using the best input combination (M3-IV).

Figure 4 compares the single LWL model with its ensemble models in low streamflow (i.e., lower than 500 m<sup>3</sup>/s) prediction and clearly demonstrates the superiority of LWL-AR in catching the minimums of streamflow. Figure 5 shows the scatter plots of the observed and predicted monthly streamflow for the best input combination (i.e., M3-IV). While the single LWL model resulted in a highly scattered prediction with R<sup>2</sup> = 0.809, the LWL-AR ensemble model produced a fit line equation (y = 0.9401x + 56.669) close to the exact line (y = x) with the highest R<sup>2</sup> value (0.867) compared to the other models.



**Figure 4.** Comparison of different LWL-based ensemble models for forecasting low streamflow in the testing phase using the best input combination (M3-IV). Notice that only the streamflow values lower than 500 m<sup>3</sup>/s are shown.

Figure 6 shows the Taylor diagram of the models and indicates how well the models match each other in terms of their standard deviation and correlation difference. Among the different models, LWL-AR achieved a closer standard deviation to the observed data with the lowest square error and highest correlation, which is followed by the LWL-BG and LWL-DG models. Figure 7 shows the violin graph of the models and indicates that LWL-AR achieved a data distribution similar to the observed data, which is followed by the LWL-DG model.

Overall, our case study demonstrated that the ensemble models successfully outperformed the single LWL model and provided promising accuracy for streamflow forecasting. Due to the non-linear nature of many environmental processes and phenomena (e.g., streamflow), hybrid ensemble models that benefit from the advantages of multiple methods/models can better capture the complexity of these phenomena and often yield more accurate results than single simple models.



**Figure 5.** Scatterplots of the observed and predicted streamflow by (**a**) LWL, (**b**) LWL-AR, (**c**) LWL-BG, (**d**) LWL-DG, (**e**) LWL-RS, (**f**) LWL-RF ensemble models in the testing phase using the best input combination (M3-IV).



**Figure 6.** Taylor diagram of different LWL-based ensemble models in the testing phase using the best input combination (M3-IV).



**Figure 7.** Violin diagrams of different LWL-based ensemble models in the testing using the best input combination (M3-IV).

#### 6. Discussion

In all ensemble models, considering periodicity (i.e., MN) as an additional input variable substantially improved both the training performance and predictive performance. During the testing phase, for the LWL-AR model, the improvements in RMSE, MAE, RAE, RRSE, and R were up to 13, 17.9, 17.5, 20.5, and 7%, respectively. For the LWL-BG model, the metrics improved up to 9.1, 14.5, 14.5, 8.9, and 5.9%, respectively. For the LWL-DG model,

the metrics improved up to 8.6, 10.9, 10, 8.1, and 3.1%, respectively. For the LWL-RS model, the metrics improved up to 10.1, 12.3, 12.4, 10.3, and 9.3%, respectively. For the LWL-RF model, the metrics improved up to 13.8, 15.6, 15.6, 13.9, and 8.2%, respectively. These results are in agreement with the previous studies that reported on the improvement of predictive accuracy using the periodicity variable. For example, Kişi [57] demonstrated the improved performance of the three types of ANN models using the periodicity variable for the prediction of monthly streamflow of the Canakdere and Goksudere rivers, Turkey. Adnan, et al. [58] used the periodicity variable to improve the predictive capability of the FFNN, RBNN, GRNN, and ANFIS models for the prediction of the monthly streamflow of the Gilgit River, Pakistan. In a recent study, Adnan, Zounemat-Kermani, Kuriqi and Kisi [53] achieved an improved performance of the long short-term memory (LSTM), extreme learning machines (ELM), and random forest (RF) models for the monthly streamflow of the Kohala and Garhi Habibullah stations in Pakistan. They showed that the inclusion of the periodicity component (MN) decreased the RMSE of the optimal LSTM, ELM, and RF models by 11.9%, 6.9%, and 1% for the Garhihabibullah Station and by 20.8%, 20.5%, and 3.7% for the Kohala Station, respectively.

A comparison of the models' outcomes revealed that the ensemble LWL-AR model performed better than the other models in both training and testing phases of the monthly streamflow modeling. The LWL-DG and LWL-RF models showed similar performance and ranked as the second-best models, followed by the LWL-RS model that was identified as the least effective ensemble model. The results of other modeling studies support our findings that the application of the ensemble learning techniques can considerably improve the capability of the base models for modeling different environmental problems [26,29,47,59]. Overall, our case study demonstrated that the ensemble models successfully outperformed the single LWL model and provided promising accuracy for streamflow forecasting. Due to the non-linear nature of many environmental processes and phenomena (e.g., streamflow), hybrid ensemble models that benefit from the advantages of multiple methods/models can better capture the complexity of these phenomena and often yield more accurate results than single simple models.

#### 7. Conclusions

This study investigated the capability of five ensemble models, that is, LWL-AR, LWL-BG, LWL-DG, LWL-RS, and LWL-RF, for monthly streamflow forecasting. The results were validated using several performance metrics and compared to those of a single LWL model. Based on the results obtained, we conclude that:

- The ensemble models are predominantly superior to the single LWL model for monthly streamflow forecasting.
- Among the ensemble methods, the LWL-AR model surpasses the other models in both training and testing performances.
- The most accurate models are developed when the periodicity variable (MN, month number) is incorporated into the modeling process.
- Ensemble forecasting is a robust and promising alternative to the single forecasting of streamflow.

Although the developed ensemble models were verified using a regional-scale dataset from Pakistan, they are sufficiently general to be applied in any other region around the world with minor adjustments in the variables relative to local conditions. Future research can extend this ensemble forecasting approach by using other ensemble learning techniques (e.g., AdaBoost, MultiBoost, LogitBoost, Decorate, etc.) and, perhaps even more interesting, by testing various types of state-of-the-art machine learning methods as the base classifier. The idea of coupling machine learning methods with ensemble learning techniques with the aim of enhancing the computational performance and improving the predictive accuracy can be extended beyond forecasting monthly streamflow to solve many other complex geo-hydrology problems. In this study, previous streamflow values and periodicity information were considered as inputs to the ensemble models. In future works, streamflow forecasting considering the flood mitigation capacity of Mangla Dam can be investigated using ensemble models. Furthermore, by taking into account the landforms (the digital terrain model) and the dimensions of the river basin as inputs, the implemented methods may provide more accurate forecasting results.

**Author Contributions:** Conceptualization: R.M.A., O.K., and A.J.; formal analysis: A.J., A.E., O.K., and R.M.A.; validation: R.M.A., A.J., A.M., A.E., and O.K.; supervision: A.J. and O.K.; writing original draft: R.M.A., A.J., A.M., A.E., and O.K.; visualization: R.M.A., A.J., and A.E.; investigation: R.M.A., A.J., and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key R&D Program of China. (2016YFC0402706).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The data presented in this study will be available on interested request from the corresponding author.

Conflicts of Interest: There is no conflict of interest in this study.

#### References

- 1. Zhang, W.; Hu, Y.; Liu, J.; Wang, H.; Wei, J.; Sun, P.; Wu, L.; Zheng, H. Progress of ethylene action mechanism and its application on plant type formation in crops. *Saudi J. Biol. Sci.* **2020**, *27*, 1667–1673. [CrossRef] [PubMed]
- Adnan, R.M.; Liang, Z.; Heddam, S.; Zounemat-Kermani, M.; Kisi, O.; Li, B. Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *J. Hydrol.* 2020, 586, 124371. [CrossRef]
- 3. Yuan, X.; Chen, C.; Lei, X.; Yuan, Y.; Adnan, R.M. Monthly runoff forecasting based on LSTM–ALO model. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 2199–2212. [CrossRef]
- 4. Liu, J.; Liu, Y.; Wang, X. An environmental assessment model of construction and demolition waste based on system dynamics: A case study in Guangzhou. *Environ. Sci. Pollut. Res.* 2019, 27, 37237–37259. [CrossRef]
- 5. Mehran, A.; AghaKouchak, A.; Nakhjiri, N.; Stewardson, M.J.; Peel, M.C.; Phillips, T.J.; Wada, Y.; Ravalico, J.K. Compounding Impacts of Human-Induced Water Stress and Climate Change on Water Availability. *Sci. Rep.* **2017**, *7*, 6282. [CrossRef]
- 6. Zhang, C.; Zhang, B.; Li, W.; Liu, M. Response of streamflow to climate change and human activity in Xitiaoxi river basin in China. *Hydrol. Process.* **2014**, *28*, 43–50. [CrossRef]
- 7. Adnan, R.M.; Liang, Z.; Parmar, K.S.; Soni, K.; Kisi, O. Modeling monthly streamflow in mountainous basin by MARS, GMDH-NN and DENFIS using hydroclimatic data. *Neural Comput. Appl.* **2021**, *33*, 2853–2871. [CrossRef]
- 8. Gibbs, M.S.; Dandy, G.C.; Maier, H.R. Assessment of the ability to meet environmental water requirements in the Upper South East of South Australia. *Stoch. Environ. Res. Risk Assess.* **2013**, *28*, 39–56. [CrossRef]
- 9. Kişi, Ö. Streamflow Forecasting Using Different Artificial Neural Network Algorithms. J. Hydrol. Eng. 2007, 12, 532–539. [CrossRef]
- 10. Yossef, N.C.; Winsemius, H.; Weerts, A.; Van Beek, R.; Bierkens, M.F.P. Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.* **2013**, *49*, 4687–4699. [CrossRef]
- 11. Aqil, M.; Kita, I.; Yano, A.; Nishiyama, S. A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *J. Hydrol.* **2007**, *337*, 22–34. [CrossRef]
- 12. Abudu, S.; Cui, C.-L.; King, J.P.; Abudukadeer, K. Comparison of performance of statistical models in forecasting monthly streamflow of Kizil River, China. *Water Sci. Eng.* 2010, *3*, 269–281.
- 13. Wang, W. Stochasticity, Nonlinearity and Forecasting of Streamflow Processes; IOS Press: Amsterdam, The Netherlands, 2006.
- 14. Rajaee, T. Wavelet and Neuro-fuzzy Conjunction Approach for Suspended Sediment Prediction. *CLEAN Soil Air Water* **2010**, *38*, 275–286. [CrossRef]
- 15. Mehdizadeh, S.; Fathian, F.; Safari, M.J.S.; Adamowski, J.F. Comparative assessment of time series and artificial intelligence models to estimate monthly streamflow: A local and external data analysis approach. *J. Hydrol.* **2019**, *579*, 124225. [CrossRef]
- Adnan, R.M.; Petroselli, A.; Heddam, S.; Santos, C.A.G.; Kisi, O. Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stoch. Environ. Res. Risk Assess.* 2021, 35, 597–616. [CrossRef]
- 17. Rahgoshay, M.; Feiznia, S.; Arian, M.; Hashemi, S.A.A. Simulation of daily suspended sediment load using an improved model of support vector machine and genetic algorithms and particle swarm. *Arab. J. Geosci.* **2019**, *12*. [CrossRef]
- Kim, C.M.; Parnichkun, M. Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of k-means clustering and adaptive neuro-fuzzy inference system. *Appl. Water Sci.* 2017, 7, 3885–3902. [CrossRef]
- Affes, Z.; Kaffel, R.H. Forecast Bankruptcy Using a Blend of Clustering and MARS Model—Case of US Banks. SSRN Electron. J. 2016, 281, 27–64. [CrossRef]

- 20. Adnan, R.M.; Liang, Z.; Trajkovic, S.; Zounemat-Kermani, M.; Li, B.; Kisi, O. Daily streamflow prediction using optimally pruned extreme learning machine. *J. Hydrol.* 2019, 577, 123981. [CrossRef]
- 21. Zhang, X.; Peng, Y.; Zhang, C.; Wang, B. Are hybrid models integrated with data preprocessing techniques suitable for monthly streamflow forecasting? Some experiment evidences. *J. Hydrol.* **2015**, 530, 137–152. [CrossRef]
- 22. 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]
- 23. Ferreira, R.G.; da Silva, D.D.; Elesbon, A.A.A.; Fernandes-Filho, E.I.; Veloso, G.V.; Fraga, M.D.S.; Ferreira, L.B. Machine learning models for streamflow regionalization in a tropical watershed. *J. Environ. Manag.* **2021**, *280*, 111713. [CrossRef]
- 24. Piazzi, G.; Thirel, G.; Perrin, C.; Delaigue, O. Sequential Data Assimilation for Streamflow Forecasting: Assessing the Sensitivity to Uncertainties and Updated Variables of a Conceptual Hydrological Model at Basin Scale. *Water Resour. Res.* 2021, 57, 57. [CrossRef]
- 25. Saraiva, S.V.; Carvalho, F.D.O.; Santos, C.A.G.; Barreto, L.C.; Freire, P.K.D.M.M. Daily streamflow forecasting in Sobradinho Reservoir using machine learning models coupled with wavelet transform and bootstrapping. *Appl. Soft Comput.* **2021**, 102, 107081. [CrossRef]
- 26. Tyralis, H.; Papacharalampous, G.; Langousis, A. Super ensemble learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. *Neural Comput. Appl.* **2021**, *33*, 3053–3068. [CrossRef]
- 27. Zhang, K.; Ruben, G.B.; Li, X.; Li, Z.; Yu, Z.; Xia, J.; Dong, Z. A comprehensive assessment framework for quantifying climatic and anthropogenic contributions to streamflow changes: A case study in a typical semi-arid North China basin. *Environ. Model. Softw.* **2020**, *128*, 104704. [CrossRef]
- Yen, H.P.H.; Pham, B.T.; Van Phong, T.; Ha, D.H.; Costache, R.; Van Le, H.; Nguyen, H.D.; Amiri, M.; Van Tao, N.; Prakash, I. Locally weighted learning based hybrid intelligence models for groundwater potential mapping and modeling: A case study at Gia Lai province, Vietnam. *Geosci. Front.* 2021, *12*, 101154. [CrossRef]
- 29. Tuyen, T.T.; Jaafari, A.; Yen, H.P.H.; Nguyen-Thoi, T.; Van Phong, T.; Nguyen, H.D.; Van Le, H.; Phuong, T.T.M.; Nguyen, S.H.; Prakash, I.; et al. Mapping forest fire susceptibility using spatially explicit ensemble models based on the locally weighted learning algorithm. *Ecol. Inform.* **2021**, *63*, 101292. [CrossRef]
- 30. Atkeson, C.G.; Moore, A.W.; Schaal, S. Locally Weighted Learning. Artif. Intell. Rev. 1997, 11, 11–73. [CrossRef]
- Ahmadianfar, I.; Jamei, M.; Chu, X. A novel Hybrid Wavelet-Locally Weighted Linear Regression (W-LWLR) Model for Electrical Conductivity (EC) Prediction in Surface Water. J. Contam. Hydrol. 2020, 232, 103641. [CrossRef] [PubMed]
- Kisi, O.; Ozkan, C. A New Approach for Modeling Sediment-Discharge Relationship: Local Weighted Linear Regression. Water Resour. Manag. 2016, 31, 1–23. [CrossRef]
- 33. Chen, T.; Ren, J. Bagging for Gaussian process regression. Neurocomputing 2009, 72, 1605–1610. [CrossRef]
- Zhou, Y.; Tian, L.; Zhu, C.; Jin, X.; Sun, Y. Video Coding Optimization for Virtual Reality 360-Degree Source. IEEE J. Sel. Top. Signal Process. 2020, 14, 118–129. [CrossRef]
- Azhari, M.; Abarda, A.; Alaoui, A.; Ettaki, B.; Zerouaoui, J. Detection of Pulsar Candidates using Bagging Method. *Procedia* Comput. Sci. 2020, 170, 1096–1101. [CrossRef]
- 36. Xue, X.; Zhang, K.; Tan, K.C.; Feng, L.; Wang, J.; Chen, G.; Zhao, X.; Zhang, L.; Yao, J. Affine Transformation-Enhanced Multifactorial Optimization for Heterogeneous Problems. *IEEE Trans. Cybern.* **2020**, 1–15. [CrossRef] [PubMed]
- 37. Stone, C.J. Additive Regression and Other Nonparametric Models. Ann. Stat. 1985, 13, 689–705. [CrossRef]
- Piegorsch, W.W.; Xiong, H.; Bhattacharya, R.N.; Lin, L. Benchmark Dose Analysis via Nonparametric Regression Modeling. *Risk Anal.* 2013, 34, 135–151. [CrossRef]
- 39. Zhang, M.; Yang, Z.; Liu, L.; Zhou, D. Impact of renewable energy investment on carbon emissions in China—An empirical study using a nonparametric additive regression model. *Sci. Total Environ.* **2021**, *785*, 147109. [CrossRef]
- 40. Ho, T.K. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, 20, 832–844. [CrossRef]
- 41. Havlíček, V.; Córcoles, A.D.; Temme, K.; Harrow, A.W.; Kandala, A.; Chow, J.M.; Gambetta, J.M. Supervised learning with quantum-enhanced feature spaces. *Nat. Cell Biol.* **2019**, *567*, 209–212. [CrossRef]
- 42. Kuncheva, L.I.; Rodriguez, J.J.; Plumpton, C.O.; Linden, D.E.J.; Johnston, S.J. Random Subspace Ensembles for fMRI Classification. *IEEE Trans. Med. Imaging* **2010**, *29*, 531–542. [CrossRef]
- Pham, B.T.; Bui, D.T.; Prakash, I.; Dholakia, M. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena* 2017, 149, 52–63. [CrossRef]
   Ting, K.M.; Witten, I.H. *Stacking Bagged and Dagged Models*: University of Waikato: Hamilton, New Zealand, 1997.
- Ting, K.M.; Witten, I.H. *Stacking Bagged and Dagged Models*; University of Waikato: Hamilton, New Zealand, 1997.
   Yariyan, P.; Janizadeh, S.; Van Phong, T.; Nguyen, H.D.; Costache, R.; Van Le, H.; Pham, B.T.; Pradhan, B.; Tiefenbacher, J.P. Improvement of Best First Decision Trees Using Bagging and Dagging Ensembles for Flood Probability Mapping. *Water Resour. Manag.* 2020, 34, 3037–3053. [CrossRef]
- 46. Zuo, C.; Chen, Q.; Tian, L.; Waller, L.; Asundi, A. Transport of intensity phase retrieval and computational imaging for partially coherent fields: The phase space perspective. *Opt. Lasers Eng.* **2015**, *71*, 20–32. [CrossRef]
- 47. Tran, Q.C.; Minh, D.D.; Jaafari, A.; Al-Ansari, N.; Minh, D.D.; Van, D.T.; Nguyen, D.A.; Tran, T.H.; Ho, L.S.; Nguyen, D.H.; et al. Novel Ensemble Landslide Predictive Models Based on the Hyperpipes Algorithm: A Case Study in the Nam Dam Commune, Vietnam. *Appl. Sci.* 2020, *10*, 3710. [CrossRef]

- Malek, A.G.; Mansoori, M.; Omranpour, H. Random forest and rotation forest ensemble methods for classification of epileptic EEG signals based on improved 1D-LBP feature extraction. *Int. J. Imaging Syst. Technol.* 2021, 31, 189–203. [CrossRef]
- 49. Jiang, Q.; Shao, F.; Lin, W.; Gu, K.; Jiang, G.; Sun, H. Optimizing Multistage Discriminative Dictionaries for Blind Image Quality Assessment. *IEEE Trans. Multimed.* 2018, 20, 2035–2048. [CrossRef]
- Pham, B.T.; Jaafari, A.; Avand, M.; Al-Ansari, N.; Du, T.D.; Yen, H.P.H.; Van Phong, T.; Nguyen, D.H.; Van Le, H.; Mafi-Gholami, D.; et al. Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction. *Symmetry* 2020, 12, 1022. [CrossRef]
- 51. Zhang, K.; Zhang, J.; Ma, X.; Yao, C.; Zhang, L.; Yang, Y.; Wang, J.; Yao, J.; Zhao, H. History Matching of Naturally Fractured Reservoirs Using a Deep Sparse Autoencoder. *SPE J.* **2021**, 1–22. [CrossRef]
- 52. Zhao, C.; Li, J. Equilibrium Selection under the Bayes-Based Strategy Updating Rules. Symmetry 2020, 12, 739. [CrossRef]
- 53. Adnan, R.M.; Zounemat-Kermani, M.; Kuriqi, A.; Kisi, O. Machine Learning Method in Prediction Streamflow Considering Periodicity Component. In *Understanding Built Environment*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 383–403.
- Kisi, O.; Shiri, J.; Karimi, S.; Adnan, R.M. Three different adaptive neuro fuzzy computing techniques for forecasting long-period daily streamflows. In *Big Data in Engineering Applications*; Springer: Singapore, 2018; pp. 303–321.
- 55. Alizamir, M.; Kisi, O.; Muhammad Adnan, R.; Kuriqi, A. Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. *Acta Geophys.* 2020, *68*, 1113–1126. [CrossRef]
- 56. Zhao, J.; Liu, J.; Jiang, J.; Gao, F. Efficient Deployment with Geometric Analysis for mmWave UAV Communications. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 1. [CrossRef]
- 57. Kişi, Ö. River flow forecasting and estimation using different artificial neural network techniques. *Hydrol. Res.* **2008**, *39*, 27–40. [CrossRef]
- Adnan, R.M.; Yuan, X.; Kisi, O.; Yuan, Y.; Tayyab, M.; Lei, X. Application of soft computing models in streamflow forecasting. In *Proceedings of the Institution of Civil Engineers—Water Management*; Thomas Telford Ltd.: London, UK, 2019; Volume 172, pp. 123–134. [CrossRef]
- Pham, B.T.; Jaafari, A.; Van Phong, T.; Yen, H.P.H.; Tuyen, T.T.; Van Luong, V.; Nguyen, H.D.; Van Le, H.; Foong, L.K. Improved flood susceptibility mapping using a best first decision tree integrated with ensemble learning techniques. *Geosci. Front.* 2021, 12, 101105. [CrossRef]