mathematics

MDPI

Article

Dimension Reduction of Machine Learning-Based Forecasting Models Employing Principal Component Analysis

Yinghui Meng [1], Sultan Noman Qasem [2,3], Manouchehr Shokri [4] and Shahab S [5,*]

[1] School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China; yinghuimeng@zzuli.edu.cn

[2] Computer Science Department, College of Computer and Information Sciences, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia; SNMohammed@imamu.edu.sa

[3] Computer Science Department, Faculty of Applied Science, Taiz University, Taiz, Yemen

[4] Faculty of civil engineering, Institute of Structural Mechanics (ISM), Bauhaus-Universität Weimar, 99423 Weimar, Germany; Manouchehr.shokri@uni-weimar.de

[5] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

[*] Correspondence: shamshirbandshahaboddin@duytan.edu.vn

Received: 8 July 2020; Accepted: 23 July 2020; Published: 27 July 2020

check for updates

Abstract: In this research, an attempt was made to reduce the dimension of wavelet-ANFIS/ANN (artificial neural network/adaptive neuro-fuzzy inference system) models toward reliable forecasts as well as to decrease computational cost. In this regard, the principal component analysis was performed on the input time series decomposed by a discrete wavelet transform to feed the ANN/ANFIS models. The models were applied for dissolved oxygen (DO) forecasting in rivers which is an important variable affecting aquatic life and water quality. The current values of DO, water surface temperature, salinity, and turbidity have been considered as the input variable to forecast DO in a three-time step further. The results of the study revealed that PCA can be employed as a powerful tool for dimension reduction of input variables and also to detect inter-correlation of input variables. Results of the PCA-wavelet-ANN models are compared with those obtained from wavelet-ANN models while the earlier one has the advantage of less computational time than the later models. Dealing with ANFIS models, PCA is more beneficial to avoid wavelet-ANFIS models creating too many rules which deteriorate the efficiency of the ANFIS models. Moreover, manipulating the wavelet-ANFIS models utilizing PCA leads to a significant decreasing in computational time. Finally, it was found that the PCA-wavelet-ANN/ANFIS models can provide reliable forecasts of dissolved oxygen as an important water quality indicator in rivers.

Keywords: machine learning; dimensionality reduction; wavelet transform; water quality; principal component analysis

1. Introduction

Due to the importance of environmental issues that play a vital role in health, food supply, and in general in the ecosystem, reliable forecasting of water quality indicators is beneficial for better management and probably to mitigate risk impacts. Dissolved oxygen (DO) represents the amount of oxygen dissolved in water which is available to living aquatic organisms. These aquatic organisms are the main elements in the food supply chain as they feed other larger species. Furthermore, it is among the key variables indicating water quality. Therefore, sound forecasts of the water quality parameters such as DO can provide suitable information for environmental monitoring and assessment. Reliable forecasting models can be considered as an early warning to take serious actions in case of emergency to save aquatic life or mitigate the risks. This case may occur due to extreme weather conditions or a sudden release of effluents in upstream of rivers.

Generally, DO concentration in the water body is depended on the physiochemical and biochemical activities which are promising to developed models to forecast it as a function of other variables. Therefore, it can be formulated as a function of some other physicochemical and biochemical indicators which are easier to measure. There are different types of models and techniques to simulate DO in rivers which can be mainly categorized as analytical, numerical, and statistical approaches [1–3]. Analytical methods are based on mathematical expressions in which the exact solutions for them are difficult or in some cases impractical. On the other hand, the numerical models and also conceptual models may need a large range of variables to be introduced to the models and also tuning the variables which can differ from place to place. On the other hand, machine learning-based models as statistical approaches have shown the great capability for simulating and forecasting of complicated phenomena. These techniques have been successfully applied to forecasting purposes for a large number of real-world applications such as river flow forecasting and hydrological modeling [4–7], water quality predictions [8–10], and groundwater level estimations [11–13], etc. Detailed descriptions of big data in complex and social networks are presented in [14]. Following the literature, it can be derived that machine learning models can be considered as cost-effective and reliable techniques for simulation and forecasting of different problems. Aside from their capability for time series forecasting, they are handy to be combined with different data pre-processing techniques such as wavelet transform [15–17]. This feature enhanced their accuracy, popularity, and applications.

Wavelet transforms as a data pre-processing technique is common to remove errors and de-noise time series which can be subsequently employed as input variables for the forecasting models. It is based on the Fourier transform which decomposes time series to several low and high-frequency filters. Many research studies are demonstrating that combining the wavelet approach with the machine learning models can improve the performance of the existing models remarkably [18,19]. Efficient modeling of a target variable requires to know the effective parameters on it. Regarding DO, a wide range of biochemical and physiochemical factors may affect it. Therefore, the application of wavelet transforms to decompose the input variables may lead to a large number of subseries and subsequently generating a large number of rules which increase computational time. Moreover, considering too many input variables may deteriorate the models' efficiency due to the inter-correlation of input variables. The principal component analysis is a suitable proxy to reduce the dimension of input variables to make it applicable for further simulations. PCA has been successfully used for dimensionality reduction of different models in for many applications such as biochemical oxygen demand (BOD) and solar radiation forecasting, etc. [20,21].

In this study, an attempt was made to evaluate the applicability and suitability of principal component analysis for dimensionality reduction of the combined wavelet-ANN/ANFIS (artificial neural network/adaptive neuro-fuzzy inference system) models for dissolved oxygen (DO) forecasting. The modeling procedures are proceeded by performing PCA on decomposed time series of water surface temperature, salinity, electric conductivity, and DO in the current time step to forecast DO in the three days in advance. Prior to PCA and wavelet implementations, separate ANN and ANFIS models are developed using the original time series without any pre-processing and dimension reduction process. Afterward, these techniques are employed to establish combined models for further applications and to provide more comparisons. In Section 2, data and methodology are described while the modeling procedures are explained in Section 3. The results of different models are discussed in Section 4. Conclusions will form the last section of the manuscript.

2. Materials and Methods

2.1. Data

In this study, after considering different stations available at USGS (United States Geological Survey) data portal, water quality data of chlorophyll (Chl) in micrograms per liter (µg/L), water temperature (T) in degrees Celsius (°C), specific conductivity (SC) in micro Siemens per centimeter (µS/cm), turbidity (Tur) in Formazin Nephelometric Unit (FNU), and dissolved oxygen (DO) in milligram per liter (mg/L) for the station ID USGS 14,211,720 have been downloaded. These data are freely available from the following website: https://waterdata.usgs.gov/usa/nwis/uv?site_no=14211720. The datasets are related to water quality measurements for the Willamette River at Portland, Oregon State recorded from 2018 to 2020. The geographical coordinates of the station in terms of longitude and latitude are 122°40′09′′ and 45°31′ 03′′. Table 1 gives a statistical analysis of the data including minimum ("Min"), maximum ("Max"), average ("Mean"), skewness ("Skew"), coefficient of variation (CV) in percent and correlation between the input variables with the target variable here means DO. Some of these data along with the data of phycocyanin pigment concentration for the year 2015 have been formerly employed by Heddam, et al. [22]. However, they only compared the performance of the feed-forward neural network, ANFIS, and gene expression programming models without wavelet applications but still concluded slightly outperformance of ANFIS models over the other models.

Table 1. Data Statistical analysis.

Variable	Min	Max	Mean	Skew	CV (%)	CC
Chl (µg/L)	0.52	10.37	1.81	2.36	73	0.75
T (°C)	4.45	24.87	13	0.45	47	0.82
SC (µS/cm)	53.17	106.17	80.94	−0.29	11	0.97
Turbidity (FNU)	1.00	60.57	6.50	3.10	122	0.70
DO (mg/L)	6.91	14.30	11.00	−0.23	17	1

Following Table 1, it can be found that there is a relatively high correlation between the input variables and the DO concentration for the study area. Moreover, the data related to turbidity, chlorophyll, and water temperature have relatively higher variation than specific conductivity and DO. It is pointed out that the data have been recorded in a 30-min interval while this study employs daily data. Therefore, they have been averaged to convert on a daily scale.

Dealing with any data-driven models and factor analysis, it is recommended to normalize or standardize the data into a range where all data have a roughly similar scale for a perfect training purpose and to restrict the search space to decrease computational time. For PCA implementation, data standardization is a common process before its application. In this study, the following relationship was applied to standardize the data with mean value ($\overline{x}$) of zero and standard deviation (Std) of 1.

$$x_{i,s} = \frac{x_i - \overline{x}}{Std} \tag{1}$$

where x_i and $x_{i,s}$ are the original and standardized data, respectively.

2.2. ANN

Artificial neural networks (ANNs) inspiring the biological process of animal brains are computational systems that learn by considering the examples. Generally, the network is established from a collection of connections of neurons or nodes. The nodes are aggregated to a layer in which each neuron in the layer is connected to the other nodes in the subsequent layer transferring signal or information. The importance of each node or connection is recognized by assigning weight in which higher weights show more importance or stronger connections and vice versa. Based on the structure and learning algorithms, different types of ANNs have been developed such as radial basis function

Mathematics 2020, 8, x 4 of 15

(RBF), multi-layer perceptron (MLP), etc. Feedforward backpropagation (FFBP) networks of MLP are amongst the most common type ANNs in use. It gains a backpropagation algorithm to decrease the error by getting back to modify the weights toward more accurate predictions. Input variables are processed in the input layers in a way each variable is represented by a node or neuron. Afterward, these data through connections are transmitted to the neurons in the hidden layer in which main computations are performed therein. Data transformation from input to hidden and hidden to the output layers are carried out utilizing log-sigmoid and linear activation functions, respectively. Finally, they are transformed into the output layer which provides the model predictions. In brief, the model can be formulated as follows.

$$o_j = \sum_{i=1}^{n} \beta_i g\left(w_i x_j + b_i\right) \tag{2}$$

where o_j is the predicted value given in the output at node j, x_j is the input value to node j, g is the activation function transforming the input data to the hidden layer, biases in the hidden layer are represented with b_i, and n denotes the number of the hidden layer neurons. Weights in the input to hidden layer connections are represented by w_i. Similarly, the weights of the connections from the hidden layer to the output layer is denoted by β_i. A detailed description of neural networks can be found in Zurada [23] and Beale, et al. [24].

function (RBF), multi-layer perceptron (MLP), etc. Feedforward backpropagation (FFBP) networks of MLP are amongst the most common type ANNs in use. It gains a backpropagation algorithm to decrease the error by getting back to modify the weights toward more accurate predictions. Input variables are processed in the input layers in a way each variable is represented by a node or neuron. Afterward, these data through connections are transmitted to the neurons in the hidden layer in which main computations are performed therein. Data transformation from input to hidden and hidden to the output layers are carried out utilizing log-sigmoid and linear activation functions, respectively. Finally, they are transformed into the output layer which provides the model predictions. In brief, the model can be formulated as follows.

$$o_j = \sum_{i=1}^{n} \beta_i g(w_i x_i + b_i) \tag{2}$$

where o_j is the predicted value given in the output at node j, x_j is the input value to node j, g is the activation function transforming the input data to the hidden layer, biases in the hidden layer are represented with b_i, and n denotes the number of the hidden layer neurons. Weights in the input to hidden layer connections are represented by w_i. Similarly, the weights of the connections from the hidden layer to the output layer is denoted by β_i. A detailed description of neural networks can be found in Zurada [23] and Beale, et al. [24].

2.3. ANFIS

2.3. ANFIS

After introducing the theory of fuzzy logic by Zadeh [25], thoughts to combine its features for predicting applications have been established. It was shown that fuzzy logic can describe complicated systems. However, it does not have the learning ability which is not possible to directly employ it for numeric predictions. In this regard, a combined system of ANN with the capability of learning (numeric power) and fuzzy logic to gain its reasoning features led to an adaptive neuro-fuzzy inference system (ANFIS) that can map the input space to the output space. A fuzzy system is if-then rule-based approach which for a first-order Sugeno fuzzy model with two variables of x and y, the following rules can be extracted as [26]:

$$If\ x\ is\ A_1\ and\ y\ is\ B_1\ then\ z_1 = p_1x + q_1y + r_1 \tag{3}$$

$$If\ x\ is\ A_2\ and\ y\ is\ B_2\ then\ z_2 = p_2x + q_2y + r_2 \tag{4}$$

where A and B are the membership functions and p, q, and r represent parameters of the Sugeno fuzzy model. Generally, five layers can be illustrated to describe the characteristics of the ANFIS model (Figure 1).

After introducing the theory of fuzzy logic by Zadeh [25], thoughts to combine its features for predicting applications have been established. It was shown that fuzzy logic can describe complicated systems. However, it does not have the learning ability which is not possible to directly employ it for numeric predictions. In this regard, a combined system of ANN with the capability of learning (numeric power) and fuzzy logic to gain its reasoning features led to an adaptive neuro-fuzzy inference system (ANFIS) that can map the input space to the output space. A fuzzy system is if-then rule-based approach which for a first-order Sugeno fuzzy model with two variables of x and y, the following rules can be extracted as [26]:

$$If\ x\ is\ A_1\ and\ y\ is\ B_1\ then\ z_1 = p_1x + q_1y + r_1 \tag{3}$$

$$If\ x\ is\ A_2\ and\ y\ is\ B_2\ then\ z_2 = p_2x + q_2y + r_2 \tag{4}$$

where A and B are the membership functions and p, q, and r represent parameters of the Sugeno fuzzy model. Generally, five layers can be illustrated to describe the characteristics of the ANFIS model (Figure 1).

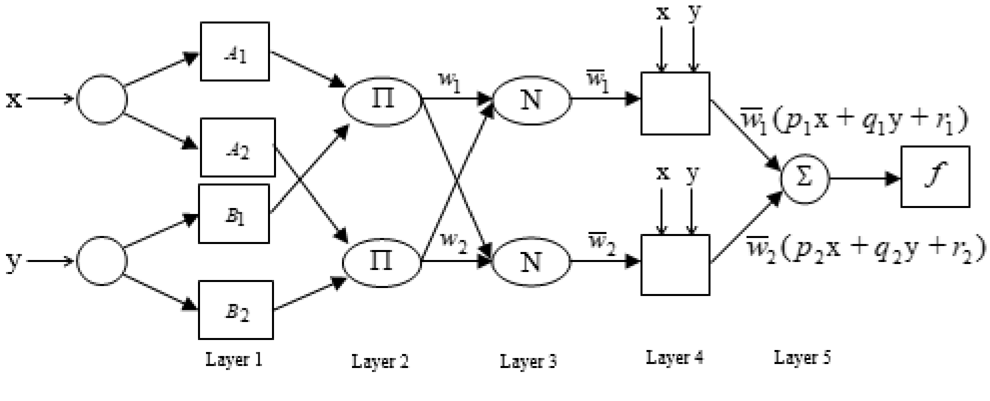

Figure 1. Layouts of the ANFIS model.

Figure 1. Layouts of the ANFIS model.

These five layers can be sequentially named as input nodes, rule nodes, average nodes, consequent nodes, and output nodes. In the first layer, membership grades are produced from the input variable.

These five layers can be sequentially named as input nodes, rule nodes, average nodes, consequent nodes, and output nodes. In the first layer, membership grades are produced from the

The following outputs denoted by O_i^1 can be obtained for layer 1 assuming the bell-shaped membership function [27]:

$$O_i^1 = \mu_{Ai\ (x)} = \frac{1}{1+\left|(x-c_i)/a_i\right|^{2b_i}} \tag{5}$$

where $\{a_i,\ b_i,\ c_i\}$ are called premise parameters. In layer 2, firing strength determining the degree to which the antecedent part of a fuzzy rule is satisfied. The output of this layer is computed by multiplying the input signals as:

$$O_i^2 = w_i = \mu_{Ai\ (x)} \cdot \mu_{Bi\ (y)},\ i = 1,2 \tag{6}$$

In the third layer, the firing strength is normalized ($\overline{w}_i$) as:

$$O_i^3 = \overline{w}_i = \frac{w_i}{\sum_{i=1}^{2} w_i} \tag{7}$$

In layer 4, the contribution of each rule (i) toward the total output is defined as:

$$O_i^4 = \overline{w}_i z_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{8}$$

The overall output is calculated in layer with single node by summing all coming signals (O_i^5). Finally, the defuzzification process is conducted to obtain the transformed values resulted from rules.

$$O_i^5 = \sum_{i=1}^{2} \overline{w}_i z_i = \frac{\sum_{i=1}^{2} w_i z_i}{\sum_{i=1}^{2} w_i} \tag{9}$$

2.4. Wavelet Transform

Wavelet transform (WT) is to some extent similar to Fourier transform (FT) or a generalization of FT but with the advantage of time-frequency localization. Through a wavelet decomposition process, a signal or image is decomposed into a sequence of sub-signals or new images to extract more information. Generally, WT basis function consisted of two main factors of scaling and shifting in which the earlier one represents shrinking or stretching of a signal in time and the latter one means delaying or advancing the wavelet along the length of the signal. Regarding the discretization of the scale and translation parameters of a signal, two types of WT including continuous (CWT) wavelet transform and discrete wavelet transform (DWT) can be employed. CWT is appropriate for time-frequency analysis and filtering of time localized frequency components. On the other hand, DWT is suitable for de-noising and compression of signals and images when simultaneous time-frequency analysis of a signal is required. Due to the nature of the CWT, it suffers from two drawbacks of redundancy and impracticality due to the nature of the WT and since both transform parameters are continuous [28]. Therefore, for time series forecasting (our case), DWT can be sought as the appropriate type of the transform. Simply, a DWT can be mathematically expressed as:

$$\psi_{m,n}(t) = a_0^{-\frac{m}{2}} \psi\left(a_0^{-m} t - n b_0\right) \tag{10}$$

where ψ is the wavelet function, t is the time, m and n are integers to control the wavelet dilation and transformation, respectively. a_0 and b_0 are called fined dilation step (greater than 1) and location parameter (greater than zero). However, the wavelet function must satisfy the admissibility condition (Equation (11)).

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{11}$$

Finally, using DWT, the original signal of the series are decomposed using different wavelet functions at different levels. The DWT decomposes the signal into a low scale (high frequency) and low frequency (high scale) components in which they are called as details and approximation elements, respectively. For a signal with DWT at decomposition level of k, $2*k$ sub-signals are obtained. Dealing with wavelet transformation, selection of appropriate wavelet function and decomposition level can play an important role in the time series analysis. For more details, one can refer to the related citations [28,29].

2.5. PCA

The principal component analysis is a suitable approach to reduce the dimension of input data by deleting some trivial information where the data are to some extent correlated. In other words, PCA is mainly performed for dimensionality reduction to project high dimension data into smaller ones but with keeping as much information as possible. Assuming a 2-dimensional scatter plot, PCA finds the best fitting line by maximizing the sum of the squared (SS) distances from the projected points to the origin. The models try with many different fitting lines in which the line with the largest SS is recognized as PC1 which is a linear combination of the variables. The number of PCs in a dataset is equal to the number of variables but it may be the same as the sample number if the sample number is smaller than the number of variables. Mathematically, factor analysis gains concepts of eigenvalues and eigenvectors in which for each PC, the sum of squared distances are eigenvalues of that PC, and the singular value for the PC is computed as the squared root of the eigenvalue of the PC. The eigenvalues and eigenvectors give the magnitude and direction of transformation carried out on the data matrix, respectively. Moreover, the proportion of variation of each PC can be determined when the corresponding eigenvalue of the PC is divided by the sample size minus 1 (i.e., $n - 1$). In this way, the contribution of each component can be figured out and the main components associated with higher variances can be selected for further implications. On the other hand, the components with lower variances indicate that their contribution can be neglected and they do not catch much information. Therefore, for dimensionality reduction purposes, only the PCs with higher contributions or variances are employed for the modeling procedures. However, the applicability of PCA and correlation among the data matrix should be examined before performing factor analysis [30,31]. Considering time series forecasting already decomposed with wavelet transform, a high dimension of input data is expected in which PCA can be served as a suitable proxy for dimensionality reduction of the input variables.

3. Modeling Procedures and Error Measures

The modeling procedure is established by selecting input variables to predict dissolved oxygen in time step ($t + 3$). Since the earlier studies have shown that the models can be efficiently employed for short term forecasting and suitable results were reported for one and two-time step ahead forecasting. Therefore, an effort was made to develop the models in a way to be applicable for longer time series forecasting. In this regard, measurements of the chlorophyll, DO, water temperature, turbidity, and specific conductivity in the current day (t) were used as input variables. Primarily, ANN and ANFIS models with the mentioned input variables are developed. Afterward, the original time series are decomposed by WT to feed the input structure of the models. Finally, PCA is performed for dimensionality reduction since the ANFIS models with too many input variables do not work due to the generation of too many rules. Moreover, for the ANN models, the effect of PCA on the performance and run-time are investigated and compared with those of the wavelet-ANN and ANN models. Dealing with the ANN models (ANN, wavelet-ANN, and wavelet-PCA-ANN), the number of iterations and neurons in the hidden layer were set as 120 and 10 respectively. These numbers have been obtained through a trial and error process indicating the appropriate performance of the models. For the ANFIS models, type and number of membership functions can affect the performance of the models remarkably which should be chosen accordingly. For this study, the Gaussian membership function with many two has been selected for the modeling process. Generally speaking, 70 percent of the data has been used for training of either ANN or ANFIS model or 30% left for the testing. It is noticed that in the ANN models, extra data are required for the validation period to prevent overfitting but to keep similar conditions, these data were selected from the training data. Therefore, for the ANN models, 55% of the data were used for training, and 15% for validation which includes 70% of the whole dataset. Dealing with the discrete wavelet transform, appropriate wavelet function, and also suitable decomposition level are principal steps toward sufficient outputs in which for this study, Meyer function ("dmey") with decomposition level of 3 have been employed to decompose original time series. It is noteworthy the decomposition level for a given wavelet function and time series are depended on the length of the time series. Finally, PCA was performed on the decomposed time series. Considering five input variables with three details and one approximation sub-signals as the effective sub-signals, the initial dimension of the input variables is 20 for four time series for each original time series. However, the dimension was reduced to four based on calculated variances for the principal components. Dealing with PCA, two metrics of eigenvalue and variance as well as break-in screen plot are three common criteria to extract the best components. The first criterion represents the components with eigenvalues higher than 1 as the suitable components. The variance factor shows the contribution of each component. Finally, the graphical method shows a breakpoint as the stop point or component. In this study, the variance metric with over 90% contribution of the whole variation is considered for component selection. The calculated variance indicating the contribution of each component is illustrated in Figure 2.

Mathematics **2020**, *8*, x 7 of 15

sufficient outputs in which for this study, Meyer function ("dmey") with decomposition level of 3 have been employed to decompose original time series. It is noteworthy the decomposition level for a given wavelet function and time series are depended on the length of the time series. Finally, PCA was performed on the decomposed time series. Considering five input variables with three details and one approximation sub-signals as the effective sub-signals, the initial dimension of the input variables is 20 for four time series for each original time series. However, the dimension was reduced to four based on calculated variances for the principal components. Dealing with PCA, two metrics of eigenvalue and variance as well as break-in screen plot are three common criteria to extract the best components. The first criterion represents the components with eigenvalues higher than 1 as the suitable components. The variance factor shows the contribution of each component. Finally, the graphical method shows a breakpoint as the stop point or component. In this study, the variance metric with over 90% contribution of the whole variation is considered for component selection. The calculated variance indicating the contribution of each component is illustrated in Figure 2.

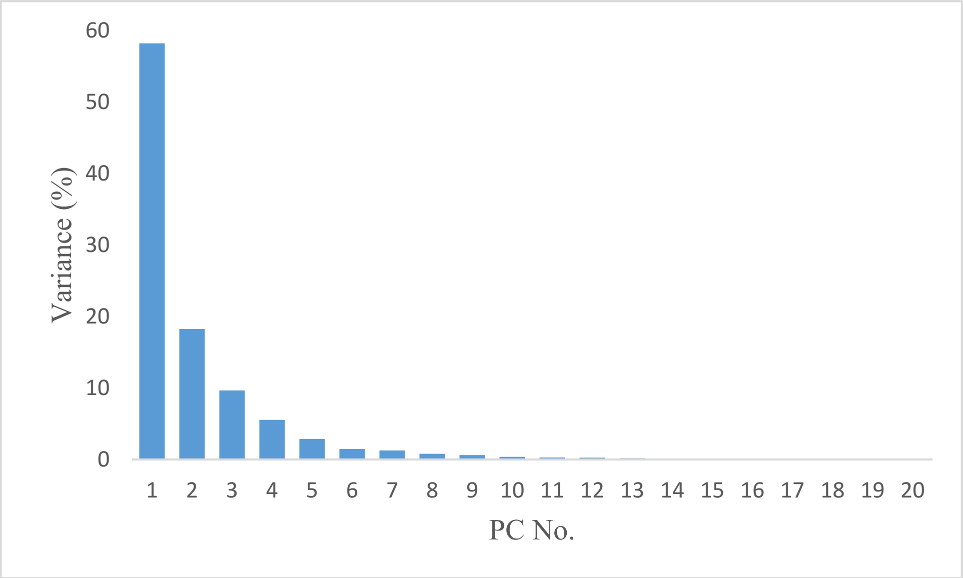

Figure 2. The variance of each PC obtained from 20 sub-signals as input variables.

Figure 2. The variance of each PC obtained from 20 sub-signals as input variables.

Regarding Figure 2, it can be obtained that the first four PCs have the most contribution associating with more than 90% of the whole variance of the 20 components. Therefore, only three to four PCs are selected to feed the input structure of the wavelet-ANN/ANFIS models. However, to evaluate the efficiency of different models, two error measures of the coefficient of determination (R^2) and root mean square error (RMSE) are employed. Moreover, the effect of dimensional reduction approach on the computational time is compared with those of the models with the whole signal/sub-signals. The error measured used in this study are defined as follows:

Regarding Figure 2, it can be obtained that the first four PCs have the most contribution associating with more than 90% of the whole variance of the 20 components. Therefore, only three to four PCs are selected to feed the input structure of the wavelet-ANN/ANFIS models. However, to evaluate the efficiency of different models, two error measures of the coefficient of determination (R^2) and root mean square error (RMSE) are employed. Moreover, the effect of dimensional reduction approach on the computational time is compared with those of the models with the whole signal/sub-signals. The error measured used in this study are defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i(measured)} - y_{i(forecasted)}\right)^2}{\sum_{i=1}^{n}\left(y_{i(measured)} - \overline{y}_{(measured)}\right)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i(measured)} - y_{i(forecasted)}\right)^2}{\sum_{i=1}^{n}\left(y_{i(measured)} - \overline{y}_{(measured)}\right)^2} \tag{13}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i(measured)} - y_{i(forecasted)}\right)^2}{\sum_{i=1}^{n}\left(y_{i(measured)} - \overline{y}_{(measured)}\right)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i(measured)} - y_{i(forecasted)}\right)^2}{\sum_{i=1}^{n}\left(y_{i(measured)} - \overline{y}_{(measured)}\right)^2} \tag{13}$$

where y represents the target variable either observed or predicted one and n denote the sample number. To easily remind the main steps of the study, Figure 3 presents a flowchart of the study.

where y represents the target variable either observed or predicted one and n denote the sample number. To easily remind the main steps of the study, Figure 3 presents a flowchart of the study.

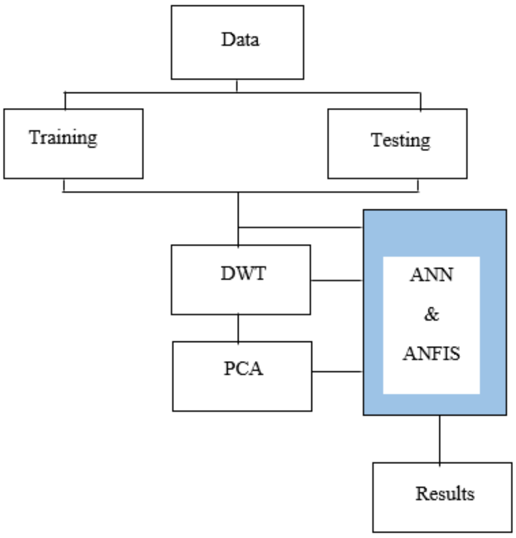

Figure 3. Flowchart of the study toward forecasting DO.

Figure 3. Flowchart of the study toward forecasting DO.

4. Results and Discussion

4. Results and Discussion

The results are presented in three sub-sections in which firstly, the original models without any pre-processing techniques are discussed. In the next sub-section, the results of the models fed by the outputs of the DWT are given. Finally, the results of the wavelet-PCA-ANN/ANFIS models are explained. It is noteworthy that all the models are calibrated and developed with similar conditions but with a difference in the input structure.

The results are presented in three sub-sections in which firstly, the original models without any pre-processing techniques are discussed. In the next sub-section, the results of the models fed by the outputs of the DWT are given. Finally, the results of the wavelet-PCA-ANN/ANFIS models are explained. It is noteworthy that all the models are calibrated and developed with similar conditions but with a difference in the input structure.

4.1. ANN and ANFIS Models

4.1. ANN and ANFIS Models

As mentioned earlier, the ANN and ANFIS models have been developed using five input variables of DO, turbidity, specific conductivity, chlorophyll, and water temperature in the current time to predict DO value three days ahead. This is due to the ability of the available models for shorter period forecasting since there is available literature indicating the capability of the machine learning-based models for one and two days ahead forecasting. However, here, the models were organized to forecast the DO for three days in advance and in case of reliable results for the models, the methodology can be extended for longer periods. Table 2 presents the results of the ANN and ANFIS models in terms of R^2, RMSE, and the required run time as well. The results are given for training and testing periods separately.

As mentioned earlier, the ANN and ANFIS models have been developed using five input variables of DO, turbidity, specific conductivity, chlorophyll, and water temperature in the current time to predict DO value three days ahead. This is due to the ability of the available models for shorter period forecasting since there is available literature indicating the capability of the machine learning-based models for one and two days ahead forecasting. However, here, the models were organized to forecast the DO for three days in advance and in case of reliable results for the models, the methodology can be extended for longer periods. Table 2 presents the results of the ANN and ANFIS models in terms of R^2, RMSE, and the required run time as well. The results are given for training and testing periods separately.

Table 2. Results of the ANN and ANAFIS model for DO forecasting.

Model	Training		Testing		Run Time
	R2	RMSE	R2	RMSE	
ANN	0.97	0.45	0.97	0.45	8.6
ANFIS	0.99	0.19	0.92	0.56	20.6

Table 2. Results of the ANN and ANAFIS model for DO forecasting.

Model	Training		Testing		Run Time
	R2	RMSE	R2	RMSE	
ANN	0.97	0.45	0.97	0.45	8.6
ANFIS	0.99	0.19	0.92	0.56	20.6

According to Table 2, it can be found that the ANN models outperform the ANFIS model for DO forecasting. It has a lower mean squared error, higher correlation than the ANFIS model. Moreover, it is time effective where less computational time is required for this model than the other. Generally, the ANFIS due to their identity which is rule-based models requires more memory and time than ANN models. The computational time and required memory sharply increase as the number of input

According to Table 2, it can be found that the ANN models outperform the ANFIS model for DO forecasting. It has a lower mean squared error, higher correlation than the ANFIS model. Moreover, it is time effective where less computational time is required for this model than the other. Generally, the ANFIS due to their identity which is rule-based models requires more memory and time than ANN models. The computational time and required memory sharply increase as the number of input variables increases. Therefore, to achieve reliable and desired ANFIS models, further actions needed to reduce the dimension of the input variables using data pre-processing techniques. Moreover, the performance of the ANFIS model for training and testing period differ remarkably which demonstrates the model drawback to catch necessary information. On the other hand, the variables increases. Therefore, to achieve reliable and desired ANFIS models, further actions needed to reduce the dimension of the input variables using data pre-processing techniques. Moreover, the performance of the ANFIS model for training and testing period differ remarkably which demonstrates the model drawback to catch necessary information. On the other hand, the performance of the ANN model for both periods is roughly the same indicating its efficiency for DO forecasting. Figure 4 illustrates a scatter plot for the ANN and ANFIS models during the testing stage.

Mathematics **2020**, *8*, x 9 of 15

performance of the ANN model for both periods is roughly the same indicating its efficiency for DO forecasting. Figure 4 illustrates a scatter plot for the ANN and ANFIS models during the testing stage.

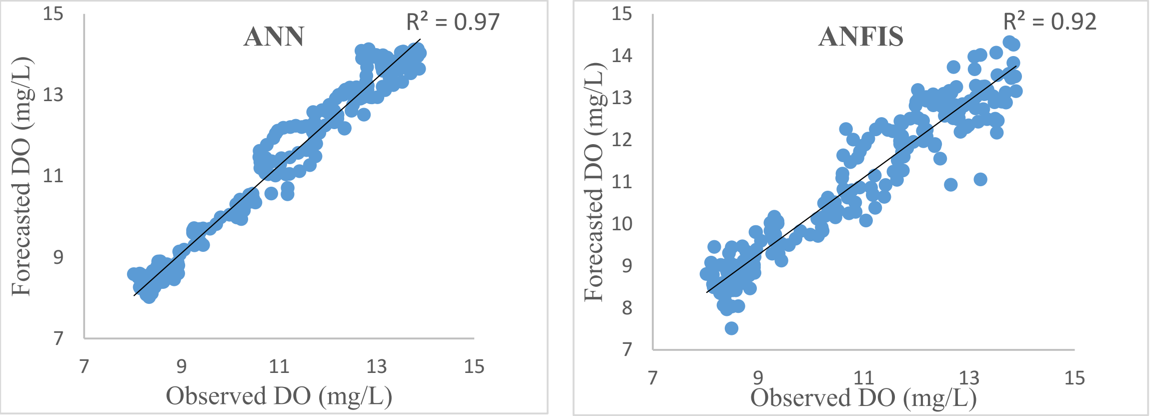

Figure 4. Results of the ANN and ANFIS models for the testing period.

Figure 4. Results of the ANN and ANFIS models for the testing period.

As observed from Figure 4, there is a good agreement of similarity between the model forecasts and observed values of DO. Generally, both models provide high correlated outputs with their corresponding values measures at the river station. However, the ANN model gives more accurate and reliable forecasts which indicate its outperformance for this case. As the study was mainly organized to deal with the dimensionality problem of the machine learning-based models, these two models were main prepared to provide more comparisons for the models combined with the pre-processing technique.

As observed from Figure 4, there is a good agreement of similarity between the model forecasts and observed values of DO. Generally, both models provide high correlated outputs with their corresponding values measures at the river station. However, the ANN model gives more accurate and reliable forecasts which indicate its outperformance for this case. As the study was mainly organized to deal with the dimensionality problem of the machine learning-based models, these two models were main prepared to provide more comparisons for the models combined with the pre-processing technique.

4.2. Wavelet-ANN and -ANFIS Models

4.2. Wavelet-ANN and -ANFIS Models

Discrete wavelet transforms are suitable tools for de-noising and extracting information from time series. They are being increasingly used for forecasting purposes and linking to machine learning-based models. In this section, the original has been decomposed into three details and three approximations sub-signals for each variable. Afterward, the three details and only the last approximation sub-signal were used as effective series to feed the ANN and ANFIS models. Therefore, the models were fed with a totally 20 input sub-signals. Performance of the combined models of wavelet-ANN (WANN) and wavelet-ANFIS (WANFIS) are presented in Table 3.

Discrete wavelet transforms are suitable tools for de-noising and extracting information from time series. They are being increasingly used for forecasting purposes and linking to machine learning-based models. In this section, the original has been decomposed into three details and three approximations sub-signals for each variable. Afterward, the three details and only the last approximation sub-signal were used as effective series to feed the ANN and ANFIS models. Therefore, the models were fed with a totally 20 input sub-signals. Performance of the combined models of wavelet-ANN (WANN) and wavelet-ANFIS (WANFIS) are presented in Table 3.

Table 3. Results of the WANN and WANFIS models.

Model	Training		Testing		Run time
	R2	RMSE (mg/L)	R2	RMSE (mg/L)	(s)
WANN	0.97	0.43	0.97	0.52	9.0
WANFIS	NAN	NAN	NAN	NAN	NAN

Table 3. Results of the WANN and WANFIS models.

Model	Training		Testing		Run Time (s)
	R2	RMSE (mg/L)	R2	RMSE (mg/L)	
WANN	0.97	0.43	0.97	0.52	9.0
WANFIS	NAN	NAN	NAN	NAN	NAN

Considering the performance of the WANN model, it can be found that its performance can be evaluated satisfactory since it has high and low values of the coefficient of determination and root mean square error, respectively. Moreover, it sounds to be economic in terms of computational cost as well. However, a comparison between the results of the WANN model and the ANN model, it can be found that the WANN model takes more time to be implemented, more expensive in terms

Considering the performance of the WANN model, it can be found that its performance can be evaluated satisfactory since it has high and low values of the coefficient of determination and root mean square error, respectively. Moreover, it sounds to be economic in terms of computational cost as well. However, a comparison between the results of the WANN model and the ANN model, it can be found that the WANN model takes more time to be implemented, more expensive in terms of computational effects and complexity but still lower performance. However, it should not be misunderstood or mislead that the DWT does not improve the performance of the existing ANN model. On the other hand, it is mainly due to feeding the WANN model with too many input variables (20 sub-signals) which are inter-correlated. Therefore, it should be manipulated to only

of computational effects and complexity but still lower performance. However, it should not be misunderstood or mislead that the DWT does not improve the performance of the existing ANN model. On the other hand, it is mainly due to feeding the WANN model with too many input variables (20 sub-signals) which are inter-correlated. Therefore, it should be manipulated to only select the appropriate sub-signals toward increasing the accuracy and reliability of the model outputs. In this regard, the principal component analysis was performed to derive the most efficient sub-signals to be used in the input structure of the models. Similarly, for the WANFIS model, using all the 20 sub-signals as the model input leads to the generation of too many rules which are not possible to be efficiently executed by the usual CPUs. In this regard, its performance in terms of the error measures cannot be assessed. Therefore, this model with the current input structure requires more amendments to be employed for forecasting purposes. In Table 3, NAN is used to represent that the WANFIS model due to too many rules cannot be implemented.

4.3. PCA-Wavelet-ANN and -ANFIS Models

Using high dimension machine learning-based models may misconduct the training process since the input variables can be inter-correlated. Moreover, the model fed by many input variables may sharply increase computational costs and complexity as well. Therefore, dimensionality reduction utilizing factor analysis can be performed to improve the efficiency of the models. In this regard, the input structure of the already WANN and WANFIS models have been re-organized through factor analysis, and the most effective components (PC1, PC2, PC3, and PC4) were chosen as the tentative input variables for the model development. In other words, the 20 sub-signals obtained from DWT have been processed by PCA and more important components were employed to feed the ANN and ANFIS models. These models are represented by PCA-WANN and PCA-WANFIS models as their results are presented in Table 4. It is noticed that in Table 4, either for ANN or ANFIS models, two sets of results are given which means two models for each technique are developed. In the first step, PCs number 1 to 3 have been selected as the efficient components in which the results are shown with ($PC_{1\text{-}3}$-WANN/WANFIS). Next, the fourth component was added to the input variables presented as ($PC_{1\text{-}4}$-WANN/WANFIS).

Table 4. Results of the models using different PCs.

Model	Training		Testing		Run Time
	R2	RMSE	R2	RMSE	
PC1-3-WANN	0.96	0.40	0.92	0.58	9.3
PC1-4-WANN	0.98	0.25	0.97	0.36	9.6
PC1-3-WANFIS	0.98	0.25	0.97	0.36	6.5
PC1-4-WANFIS	0.99	0.18	0.88	1.017	8.8

Regarding Table 4, it can be derived that all the models developed by PCs as the input variables have suitable performance. They are efficient models in terms of performance evaluation criteria and run time as well. Considering the WANN models, the model with four PCs ($PC_{1\text{-}4}$-WANN) outperforms the model with the same conditions only with one less component ($PC_{1\text{-}3}$-WANN). Therefore, for the WANN models, feeding the input structure with the first four components leads to more accurate predictions than the model with three PCs. In other words, adding the fourth component to the input structure of the WANN model increase its performance remarkably while the run time for that slightly increases. On the other hand, the WANFIS model with three PCs provides better performance than the model with four PCs. It can be concluded that for the ANFIS models, training the model with fewer signals are desired toward achieving reliable predictions were in the $PC_{1\text{-}4}$-WANFIS model, adding a new input variable (PC) increase the error of the forecasts remarkably. Besides, it requires less computational time. The $PC_{1\text{-}3}$-WANFIS simultaneously improves the performance of the existing WANFIS models and decrease computational time. Overall, the $PC_{1\text{-}3}$-WANFIS has the highest performance compared with the other models. It has a very high value of the coefficient of determination and also very low values of RMSE for both periods. Similar performance for training and testing stages also confirms the suitability of the proposed model. Figure 5 illustrates scatter plots

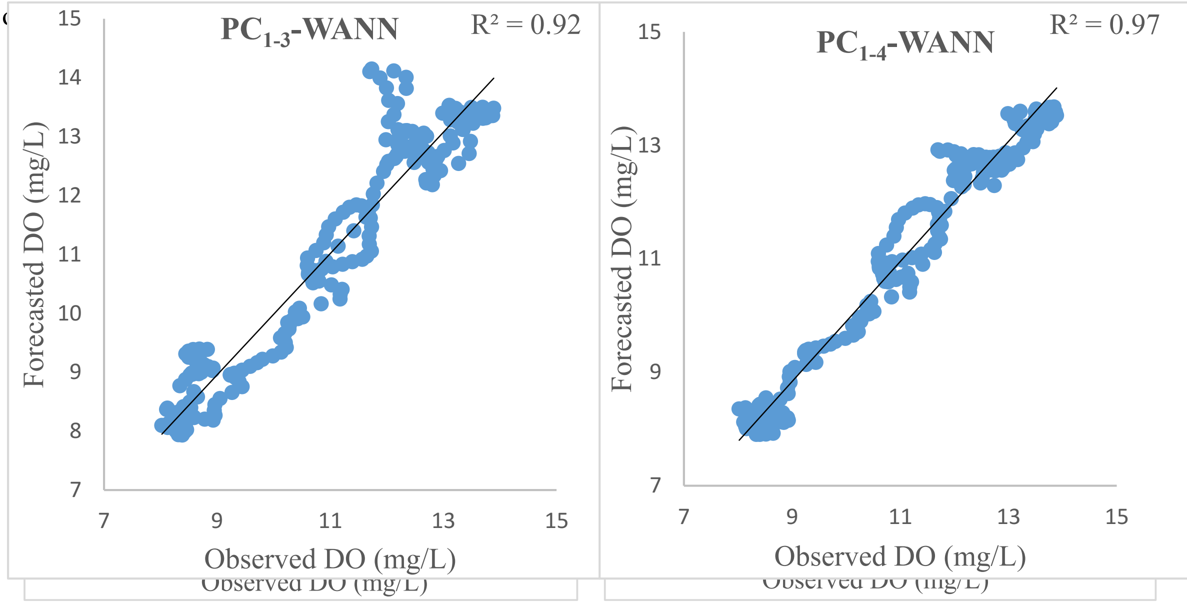

Figure 5. Results of the WANN models with two different input combinations for Table 4.

It is derived that both models generally provide sound forecasts of DO in which good agreement between the model outputs and field measurements can be found. However, for the measurements in a range of 11 to 13 mg/L, the estimated values are to some extend deviated from the correlation line. This deviation is more obvious for the left-hand side graph than the other one. Therefore, the correlation between model outputs of $PC_{1\text{-}4}$-WANN and the observed values are stronger than the other WANN model outputs. It shows that adding the fourth PC to the input variables can provide useful information for the model to train better and consequently to provide more accurate forecasts. In a similar way to the WANN models, the outputs of the WANFIS models during the testing period are depicted in Figure 6.

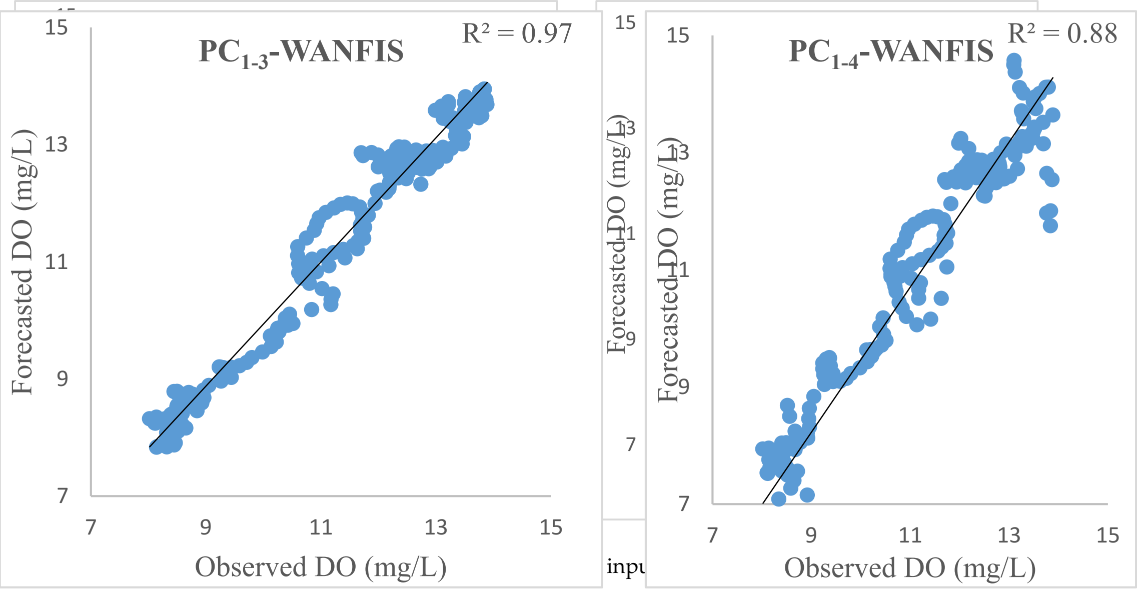

Figure 6. Results of the WANFIS models with two different input combinations for the testing period.

Following Figure 6, it can be found the results projected in the left panel are more accurate and reliable than those illustrated in the right panel demonstrating that the WANFIS model with three PCs is superior to the WANFIS model with four PCs. Results of the $PC_{1\text{-}3}$-WANFIS model reveal that

Following Figure 6, it can be found the results projected in the left panel are more accurate and reliable than those illustrated in the right panel demonstrating that the WANFIS model with three PCs is superior to the WANFIS model with four PCs. Results of the $PC_{1\text{-}3}$-WANFIS model reveal that

Following Figure 6, it can be found the results projected in the left panel are more accurate and reliable than those illustrated in the right panel demonstrating that the WANFIS model with three PCs is superior to the WANFIS model with four PCs. Results of the $PC_{1\text{-}3}$-WANFIS model reveal that the model forecasts are of great consistency with those of the observed values for a wide range of variety of the target variable. The models' forecasts even for the extreme values are close to the real values which are promising to employ such an approach for extreme value analysis and forecasting. Reliable estimation of the extreme values is of great interest and benefit for practical applications to prevent or mitigate impacts of a rapid deterioration of water quality parameters. To figure out how well the model forecasts DO values for a wide range from low to medium and medium to high values, Figure 7 presents a time series of the models' outputs versus the real values.

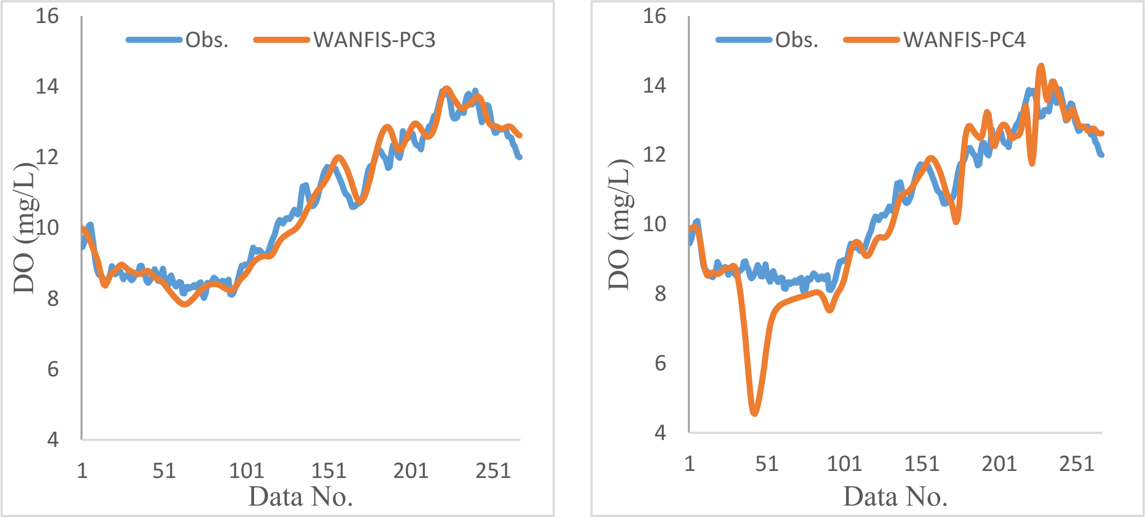

Figure 7. Time series of the forecasted and observed DO for the WANFIS models.

The time series illustrations in Figure 7 confirms that the $PC_{1\text{-}3}$-WANFIS model can be successfully applied for DO forecasting from low to high values. The model simulations roughly overlap observation demonstrating model efficiency. On the other hand, the $PC_{1\text{-}4}$-WANFIS model has some drawbacks for DO forecasting, especially low values. There are remarkable inconsistencies between the model outputs and those of the corresponding field measurements. Thus, dealing with ANFIS models, only the model with three PCs can estimate DO concentration in the river accordingly.

Generally speaking, in this study, it was found that principal component analysis can be employed as a suitable tool for dimensionality reduction but still keeping main features of the time series in which the models fed by the components obtained from the factor analysis showed a great performance. Through this study, it can be derived that the proposed combination of the discrete wavelet transform, PCA and ANFIS can improve performance of an existing ANFIS model with original time series of input variables about 6%, 30% and 70% in terms of R^2, RMSE, and run time, respectively. Similarly, the proposed approach can be successfully to improve the efficiency of the ANN and WANN models already developed. Dealing with ANN models, it was found that feeding the model with a large number of decomposed time series may mislead the training process of the ANN model due to the inter-correlation of the available data. Therefore, finding suitable sub-series to eliminate this problem can be fulfilled via PCA.

The results obtained through this study are comparable with those of literature where demonstrated PCA and wavelet transform as suitable proxies can be linked to ANN and ANFIS models to improve their forecasting performance. For instance, Solgi, Pourhaghi, Bahmani and Zarei [21] concluded that wavelet and PCA combinations with support vector regression model can enhance the R^2 of the model during the testing period about 10% when the model is applied for BOD forecasting. Similarly, Sahoo, et al. [32] applied the PCA-ANFIS model for forecasting of water quality index River Brahmani, India. The results indicated the efficiency of the proposed model for modeling the index. However, the application of the proposed model is not limited to the abovementioned fields and they can be considered as a suitable approach for forecasting different atmospherics, hydrologic and other processes as well.

Sahoo, et al. [32] applied the PCA-ANFIS model for forecasting of water quality index River Brahmani, India. The results indicated the efficiency of the proposed model for modeling the index. However, the application of the proposed model is not limited to the abovementioned fields and they can be considered as a suitable approach for forecasting different atmospherics, hydrologic and other processes as well.

5. Conclusions

The main focus of this study was to manipulate machine learning-based forecasting models with a large number of input data. In this regard, five different input variables have been decomposed through discrete wavelet transform to generate high dimension input data. Afterward, artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) as machine learning models have been employed for long term forecasting of dissolved oxygen (DO) in Willamette River, Oregon State, USA. Finally, principal component analysis (PCA) was considered for dimensionality reduction purposes to improve the accuracy of the available models and also to decrease computational time. To provide more comparisons, different models of ANN and ANFIS with different combinations of input variables from the original time series to models with efficient principal components were developed. Dealing with wavelet transform, Meyer wavelet function at level 3 was employed to decompose original time series. The main findings of the current study can be summarized as the following concluding remarks.

- The models fed by the PCs have the highest performance among the other models demonstrating the PCA approach to catch suitable information from time series as well as reducing the dimension of the input variables.
- The wavelet-ANN model with the first four PCs has a better performance than the model with three PCs while for the ANFIS model the results were conversely. Therefore, more PCs for ANN and fewer PCs for the ANFIS models lead to the desired outputs.
- Using factor analysis improved the performance of the existing wavelet-ANN and ANFIS models while decreased computational time and complexity. Therefore, the proposed approach can be employed for forecasting of other time series as well.
- Among different models examined in this study, the $PC_{1\text{-}3}$-WANFIS model indicating a wavelet-ANFIS model using three principal components from the decomposed time series has the best performance. The proposed models perform fast with accurate forecasts for a wide range of variation for the DO. Moreover, the model has an excellent performance to forecast extreme values which are of great performance for environmental management and planning.
- Results of this study that the factor analysis is a suitable proxy for dimensionality reduction of the forecasting models which improves the performance in terms of computational time and reliability of the outputs. The PCA has a great capability to detect the inter-correlation among time series which may lead to model misconduct if it does not manipulate accordingly.

In brief, the results of this study were promising to apply PCA for dimensionality reduction and eliminate the inter-correlation of variables. It can successfully derive the most important input variables to be subsequently employed in the forecasting models. The proposed model provides reliable forecasts of DO for three days in advance. The combined model of PCA, wavelet, and ANN/ANFIS can be efficiently used for the forecasting of other water quality indicators or environmental indicators with different time steps in advance. Furthermore, it can be generalized and adopted for forecasting of several different hydrological variables such as flow discharge, suspended sediment load, rainfall, and groundwater level among the others.

Author Contributions: Data curation, S.S.; Formal analysis, Y.M. and S.N.Q.; Investigation, T.M.; Methodology, S.N.Q. and M.S.; Supervision, S.S.; Validation, M.S. and S.S.; Visualization, T.M. and S.N.Q.; Writing—original draft, Y.M, S.N.Q. and S.S.; Writing—review & editing, T.M. and M.S. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We acknowledge the support of the German Research Foundation (DFG) and the Bauhaus-Universität Weimar within the Open-Access Publishing Programme.

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

Nomenclature

PCA	Principal Component Analysis
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
FT	Fourier Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
RMSE	Root Mean Square Error
CV	Coefficient of Variation
DO	Dissolved Oxygen
BOD	Biochemical Oxygen Demand
Chl	Chlorophyll
SC	Specific Conductivity
Tur	Turbidity

References

1. Cox, B. A review of dissolved oxygen modelling techniques for lowland rivers. *Sci. Total Environ.* **2003**, *314*, 303–334. [CrossRef]
2. Phelps, E.B.; Streeter, H. *A Study of the Pollution and Natural Purification of the Ohio River*; US Department of Health, Education, & Welfare: Washington, DC, USA, 1958.
3. Bennett, J.P.; Rathbun, R. *Reaeration in Open-Channel Flow*; US Government Printing Office: Washington, DC, USA, 1971; Volume 737.
4. Ahani, A.; Shourian, M.; Rad, P.R. Performance assessment of the linear, nonlinear and nonparametric data driven models in river flow forecasting. *Water Res. Manag.* **2018**, *32*, 383–399. [CrossRef]
5. Anusree, K.; Varghese, K. Streamflow prediction of Karuvannur River Basin using ANFIS, ANN and MNLR models. *Proc. Technol.* **2016**, *24*, 101–108. [CrossRef]
6. Dastorani, M.T.; Moghadamnia, A.; Piri, J.; Rico-Ramirez, M. Application of ANN and ANFIS models for reconstructing missing flow data. *Environ. Monit. Assess.* **2010**, *166*, 421–434. [CrossRef]
7. Nourani, V.; Kisi, Ö.; Komasi, M. Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.* **2011**, *402*, 41–59. [CrossRef]
8. Maier, H.R.; Dandy, G.C. The use of artificial neural networks for the prediction of water quality parameters. *Water Resourc. Res.* **1996**, *32*, 1013–1022. [CrossRef]
9. Sarkar, A.; Pandey, P. River water quality modelling using artificial neural network technique. *Aquat. Proc.* **2015**, *4*, 1070–1077. [CrossRef]
10. Heddam, S.; Kisi, O. Extreme learning machines: A new approach for modeling dissolved oxygen (DO) concentration with and without water quality variables as predictors. *Environ. Sci. Pollut. Res.* **2017**, *24*, 16702–16724. [CrossRef]
11. Daliakopoulos, I.N.; Coulibaly, P.; Tsanis, I.K. Groundwater level forecasting using artificial neural networks. *J. Hydrol.* **2005**, *309*, 229–240. [CrossRef]
12. Li, H.; Lu, Y.; Zheng, C.; Yang, M.; Li, S. Groundwater level prediction for the arid oasis of Northwest China based on the artificial bee colony algorithm and a back-propagation neural network with double hidden layers. *Water* **2019**, *11*, 860. [CrossRef]
13. Gong, Y.; Wang, Z.; Xu, G.; Zhang, Z. A comparative study of groundwater level forecasting using data-driven models based on ensemble empirical mode decomposition. *Water* **2018**, *10*, 730. [CrossRef]
14. Thai, M.T.; Wu, W.; Xiong, H. *Big Data in Complex and Social Networks*; CRC Press: London, UK, 2016.
15. Hadi, S.J.; Tombul, M. Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination. *J. Hydrol.* **2018**, *561*, 674–687. [CrossRef]

16. Nourani, V.; Parhizkar, M. Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall–runoff modeling. *J. Hydroinform.* **2013**, *15*, 829–848. [CrossRef]
17. Pramanik, N.; Panda, R.K.; Singh, A. Daily river flow forecasting using wavelet ANN hybrid models. *J. Hydroinform.* **2011**, *13*, 49–63. [CrossRef]
18. Adamowski, J.; Chan, H.F. A wavelet neural network conjunction model for groundwater level forecasting. *J. Hydrol.* **2011**, *407*, 28–40. [CrossRef]
19. Sharghi, E.; Nourani, V.; Molajou, A.; Najafi, H. Conjunction of emotional ANN (EANN) and wavelet transform for rainfall-runoff modeling. *J. Hydroinform.* **2019**, *21*, 136–152. [CrossRef]
20. Zhang, X.; Wei, Z. A hybrid model based on principal component analysis, wavelet transform, and extreme learning machine optimized by Bat algorithm for daily solar radiation forecasting. *Sustainability* **2019**, *11*, 4138. [CrossRef]
21. Solgi, A.; Pourhaghi, A.; Bahmani, R.; Zarei, H. Improving SVR and ANFIS performance using wavelet transform and PCA algorithm for modeling and predicting biochemical oxygen demand (BOD). *Ecohydrol. Hydrobiol.* **2017**, *17*, 164–175. [CrossRef]
22. Heddam, S.; Sanikhani, H.; Kisi, O. Application of artificial intelligence to estimate phycocyanin pigment concentration using water quality data: A comparative study. *Appl. Water Sci.* **2019**, *9*, 164. [CrossRef]
23. Zurada, J.M. *Introduction to Artificial Neural Systems*; West Group: West St. Paul, MN, USA, 1992; Volume 8.
24. Beale, H.D.; Demuth, H.B.; Hagan, M. *Neural Network Design*; PWS: Boston, MA, USA, 1996.
25. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]
26. Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *1*, 116–132. [CrossRef]
27. Chang, F.-J.; Chang, Y.-T. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Adv. Water Resour.* **2006**, *29*, 1–10. [CrossRef]
28. Akansu, A.N.; Haddad, P.A.; Haddad, R.A.; Haddad, P.R. *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*; Academic Press: Cambridge, MA, USA, 2001.
29. Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: San Diego, CA, USA; London, UK; Boston, MA, USA; New York, NY, USA; Sydney, NSW, Australia; Tokyo, Japan; Toronto, ON, Canada, 1998.
30. Cattell, R. The scree test for the number of factors. *Multivar. Behav. Res.* **1996**, *1*, 629–637. [CrossRef] [PubMed]
31. Crane, D.R.; Busby, D.M.; Larson, J.H. A factor analysis of the Dyadic Adjustment Scale with distressed and nondistressed couples. *Am. J. Fam. Ther.* **1991**, *19*, 60–66. [CrossRef]
32. Sahoo, M.M.; Patra, K.; Khatua, K. Inference of water quality index using ANFIA and PCA. *Aquat. Proc.* **2015**, *4*, 1099–1106. [CrossRef]