

Article

Improving Daily Peak Flow Forecasts Using Hybrid Fourier-Series Autoregressive Integrated Moving Average and Recurrent Artificial Neural Network Models

Mohammad Ebrahim Banihabib ^{1,*}, Reihaneh Bandari ² and Mohammad Valipour ^{3,4,*}

¹ Department of Irrigation and Drainage Engineering, College of Aburaihan, University of Tehran, Pakdasht, Tehran 3391653755, Iran

² Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Melbourne 3125, Australia; rbandari@deakin.edu.au

³ Center of Excellence for Climate Change Research/Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁴ Department of Civil and Environmental Engineering and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI 96822, USA

* Correspondence: banihabib@ut.ac.ir (M.E.B.); valipour@hawaii.com (M.V.)

Received: 27 March 2020; Accepted: 2 June 2020; Published: 7 June 2020

Abstract: In multi-purpose reservoirs, to achieve optimal operation, sophisticated models are required to forecast reservoir inflow in both short- and long-horizon times with an acceptable accuracy, particularly for peak flows. In this study, an auto-regressive hybrid model is proposed for long-horizon forecasting of daily reservoir inflow. The model is examined for a one-year horizon forecasting of high-oscillated daily flow time series. First, a Fourier-Series Filtered Autoregressive Integrated Moving Average (FSF-ARIMA) model is applied to forecast linear behavior of daily flow time series. Second, a Recurrent Artificial Neural Network (RANN) model is utilized to forecast FSF-ARIMA model's residuals. The hybrid model follows the detail of observed flow time variation and forecasted peak flow more accurately than previous models. The proposed model enhances the ability to forecast reservoir inflow, especially in peak flows, compared to previous linear and nonlinear auto-regressive models. The hybrid model has a potential to decrease maximum and average forecasting error by 81% and 80%, respectively. The results of this investigation are useful for stakeholders and water resources managers to schedule optimum operation of multi-purpose reservoirs in controlling floods and generating hydropower.

Keywords: ARIMA; autoregressive hybrid model; daily flow; forecasting; long horizon; recurrent artificial neural network; Dez reservoir

1. Introduction

Stream flow forecasting plays an important role in environmental and hydrological research and disaster management. Nowadays, mathematical models are emerging for more precise and longer-horizon daily reservoir inflow forecasting that can be beneficial for reservoir operation, flood warning, and optimal water allocation for various water users. Predicting peak flows can help us to control floods and hydropower operation.

One of the successful mathematical forecasting models are autoregressive models such as Autoregressive Integrated Moving Average (ARIMA) and multilayer perceptron Artificial Neural Network (ANN). Autoregressive models forecast future values of time series based on the

identification of past temporal patterns of the records. The main hypothesis of ARIMA models, which are conventional forecasting models, assumes a linear relationship between historical time series for the forecasting of future variables [1]. Hence, ARIMA models cannot forecast the nonlinear pattern of inflow time series. On the other hand, ANN models, as well-known forecasting models, are able to model nonlinear time series [1]. Autoregressive forecasting models can be categorized based on their time horizon as short-term, mid-term, and long-term which are from a part of a day to a week, weeks up to a month, and months to years, respectively. Applied time scales for inflow forecasting are hourly, daily, weekly, monthly, quarterly, and yearly [2–7]. One of the advantages of machine learning models is to forecast time series for long-term horizons. Recently, many studies reported the need for forecasting daily hydrological data at a long-term horizon (at least for one year ahead) [8–12].

Some ANN models are linked with conventional auto-regressive models (Auto Regressive (AR) and ARIMA) to reveal their ability in short-term horizon forecasting of daily flow. Kisi and Cigizoglu [3] used Feed-Forward Artificial Neural Network (FF-ANN) and AR models to forecast daily flow for three rivers in the USA and a river in Turkey. Three artificial neural network structures were then selected for comparison with the AR model forecasts. Given the same input data for 1-day-ahead forecasts, the results showed that ANN structures were able to produce better results than AR models. Banihabib et al. [2] forecasted the daily inflow to the Dez reservoir by using an FF-ANN model and a linear regression model based on inflow data from hydrometric stations located upstream of River Dez. The research showed that in short-term horizon forecasting, the FF-ANN performed better than linear regression models. Xie [13] used linear regression and exponential smoothing ANN models to forecast daily flow. These models performed well during the dry seasons while the non-linear ANNs were superior to the other models in forecasting flood flows of rainy seasons; however, they still had limitations in forecasting peak flows [13]. Sattari et al. [14] forecasted daily flow upstream of Eleviyan Reservoir by a Recurrent ANN (RANN) and a Back-Propagation Neural Network (BP-NN). The results suggested that both models are fair in one-day ahead forecasting of flood flow. However, both models performed well when they were applied for forecasting low flows. Consequently, these studies show that ANN models perform better for a short-term horizon forecasting of daily low flow. However, even for the short-term horizon forecasting of inflow, ANN models do not have a considerable accuracy for peak flows such as flood flows.

Several ARIMA and Artificial Neural Network (ANN) models are proposed for forecasting of inflow in monthly, quarterly, and yearly time scales [4–6,15–20], while autoregressive models are reported for successful forecasting of daily flow at short-term time horizons [2,3,14,21–26]. The forecasting horizon of most of short term autoregressive models are only one day or 3- to 7-days ahead of flow forecasting [2,3,6,21,23–25,27]. These short-term horizon flow forecasting can benefit single-purpose dam for the purpose of flood control. However, a longer horizon of forecasting is required for optimal operation of a multi-purpose dam that provides water for hydropower generation, flood control, domestic water supply, and irrigation. There are other successful investigations of daily forecasting for at least one year ahead [8,10–12]. Banihabib et al. [9] proposed a non-linear auto-regressive ANN model for forecasting daily streamflow for a long-term horizon versus an ARIMA model. The auto-regressive ANN model improved long-term forecasting daily streamflow by continuously following a daily flow pattern compared to the ARIMA. However, the proposed model has still considerable uncertainty in forecasting peak flows.

The literature review indicates that developing autoregressive models are needed to forecast daily reservoir inflow, especially peak flows for a long-term horizon, to provide the optimal operation of multi-purpose reservoirs for flood control, hydropower generation, domestic water supply, and irrigation purposes. Since forecasting by regular ANN has already been done [2,6,9], we focused on using hybrid ANN models in this study to keep a novelty aspect of the work, as well as to increase the accuracy of the predictions. Indeed, the novelty of this research compared to previous studies like Banihabib et al. [9] is developing a hybrid autoregressive model (a combination of linear and nonlinear forecasting models) to present a robust forecasting model for peak flows in long-horizon daily reservoir inflow forecasting.

The results of this study are applicable for stakeholders and water resources managers to schedule optimum operation of multi-purpose reservoirs to control floods and to generate optimum hydropower.

2. Materials and Methods

2.1. Multi-Purpose Reservoir and Reservoir Inflow Data

To examine the performance of the proposed autoregressive hybrid model, a multi-purpose reservoir (Dez Dam) has been selected as a case study. Figure 1 shows the location of the Dez basin. The reservoir aims to control floods, generate hydropower, and supply agricultural and domestic water. The Dez reservoir basin is at a latitude of 32°35'N to 34°07'N and longitude of 48°20' to 50°20'E in Southwest Iran. In this research, the daily flow records of Taleh-Zang hydrometric station located upstream of the reservoir were used to forecast inflow to the reservoir. In addition, daily discharge data from 1975 to 2011 of the hydrometric station were used for calibration/training and forecasting daily reservoir inflow. The dataset comprised 13,140 data points from 23 September 1975 to 22 September 2011. The data set was then split into two subsets. The daily stream flow from 23 September 1975 to 22 September 2010 was chosen for training (calibration), and daily reservoir inflow from 23 September 2010 to 22 September 2011 was chosen for forecasting. Access to the dataset can be requested by contacting the authors.

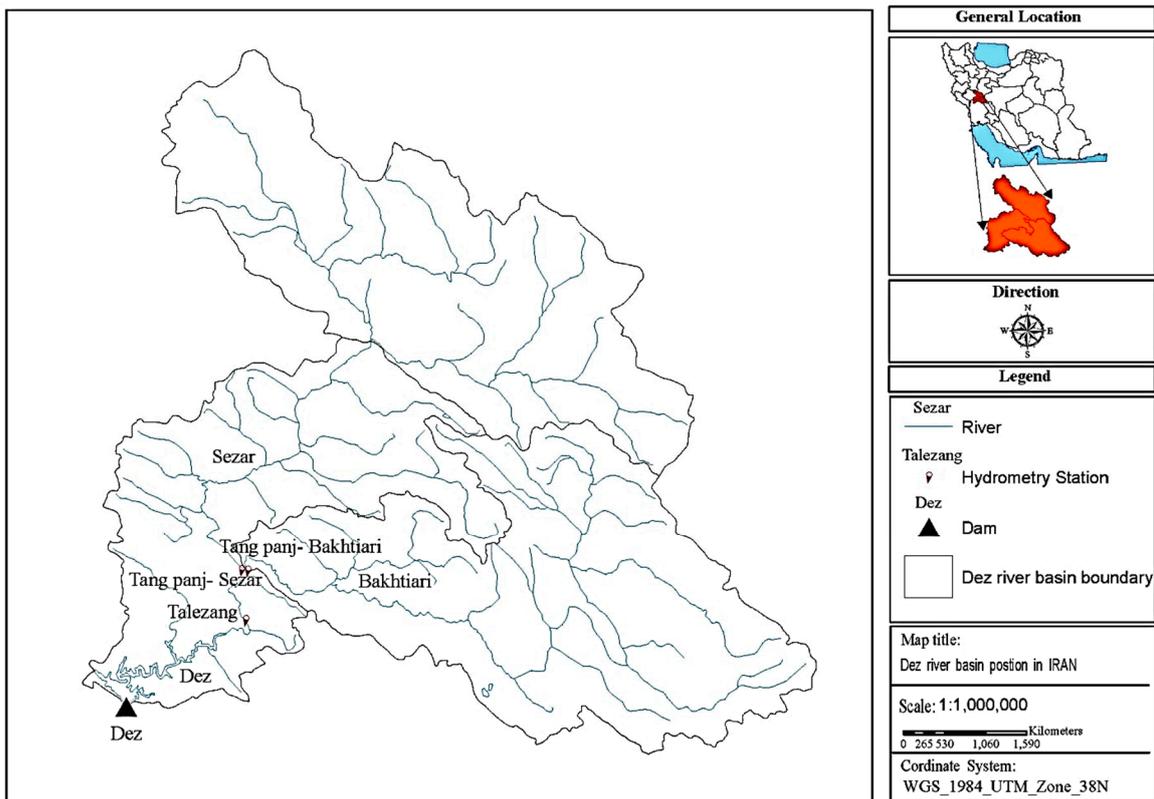


Figure 1. The location map of Dez Dam and its basin [9].

2.2. Overview of the Research Method and Performance Evaluation

The proposed hybrid autoregressive model consists two sub-models: Fourier Series Filtered ARIMA (FSF-ARIMA) for the linear part and RANN for the nonlinear part. Both sub-models have two processing phases: training (calibration), in which the models use historical observed inflow data for learning and then for forecasting (in which the models forecast the daily reservoir inflow for one year). First, FSF-ARIMA models were calibrated and used for forecasting the linear part of the reservoir inflow data, then the RANN model was trained and used for forecasting the non-linear part of the data.

In this study, reservoir inflow forecasting was carried out by using a conventional model (FSF-ARIMA model) and the proposed model (autoregressive hybrid model), and the results were compared with evaluation indices. Through comparing the models with observed data, the performance of each model was assessed in daily reservoir inflow forecasting. Error Index (EI) and coefficient of determination (R^2) were employed as the indices for performance evaluation of the autoregressive hybrid and FSF-ARIMA models in training and forecasting phases as below [2,6,9]:

$$EI = \frac{RMSE}{\bar{Q}_{obs}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{ft} - Q_{ot})^2}{n}} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Q_{ft} - Q_{ot})^2}{\sum_{i=1}^n (Q_{ot} - \bar{Q}_{obs})^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |Q_{ft} - Q_{ot}| \quad (4)$$

where Q_{ft} and Q_{ot} are the forecasted and observed daily reservoir inflow in the t th day of the forecasting horizon, and n is the number of data points. \bar{Q}_{obs} is the average of observed reservoir inflow. The training phase contains 12,775 data points and \bar{Q}_{obs} is $262 \text{ m}^3 / \text{s}$. The forecasting phase comprises 365 data sets where \bar{Q}_{obs} equals to $141 \text{ m}^3 / \text{s}$.

In the forecasting phase, to compare the one-year observed inflow hydrograph with the forecasted hydrographs of the models, EI , R^2 , and Average Cumulative Relative Error ($ACRE$) are applied for performance evaluation of autoregressive hybrid and FSF-ARIMA models. $ACRE$ is defined for determining the best forecasting duration as follows [2,6,9]:

$$ACRE = \frac{(\sum_{i=1}^m |Q_{fi} - Q_{oi}| / Q_{oi})}{m} \quad (5)$$

where $ACRE$ is the average cumulative relative error until a certain month, and m is the cumulative number of days until that certain month. $ACRE$ is used to evaluate the time-tendency of error of the models in the forecasting phase.

In this study, we used Windows OS; programming language: MATLAB for RANN and R 2.13.0 for FSF-ARIMA. We stopped training when Mean Squared Error (MSE) was at the minimum. Training algorithm: Levenberg–Marquardt algorithm; number of hidden layers: 1; transfer functions of hidden layer: tangent-sigmoid (tansig) and log-sigmoid (logsig); transfer function in the output

layer: pure line. We also tested 1760 nonlinear autoregressive network with exogenous inputs (NARX)-recurrent neural network (RNN) (NARX-RNN) model structures differing in transfer functions, numbers of inputs (2–5), and neurons per hidden layer (1–22); input delays and output delays ranged from 1 to 10. We tested two training algorithms: Levenberg–Marquardt (LM) algorithm and traingdx, but LM was applied as the learning function, finally, because it has generally high accuracy and it is fast learning.

2.3. Previous Linear and Nonlinear Models

To examine the capability of the proposed hybrid model, the previous linear and nonlinear models (FSF-ARIMA RANN models) proposed by Banihabib et al. [9] were applied to the case study, and their results were compared to the result of the developed hybrid model.

The procedure for applying FSF-ARIMA to the seasonally variable reservoir inflow data for a one-year horizon is summarized as follows [9]. The FSF-ARIMA procedure requires normally distributed stationary reservoir inflow time series. First, the time series are normalized using a logarithm transformation [28]. This method has been successfully employed for inflow forecasting based on past investigations [29–31]. Then the mean and standard deviation of the logarithm-transformed data are computed. In the next step, a Fourier Series was used to remove the seasonal tendency in the logarithm transformed time series [1,9]. The FSF-ARIMA(p,d,q) model is used (Equation (5)) to forecast daily stream flow data [9]. Multiple FSF-ARIMA models were tried, and the most suitable model was selected using the Akaike Information Criterion (AIC) [32] as a well-known criterion for evaluating time series models. The Fourier-transformed data is defined as below:

$$Q_t' = \phi_1 Q_{t-1}' + \dots + \phi_p Q_{t-p}' + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (6)$$

where Q_t' and ε_t are the Fourier-transformed data and random error, respectively, at time step t . ϕ_i ($i = 1, 2, \dots, p$) and θ_j ($j = 1, 2, \dots, q$) are model parameters, p is the autoregressive model order, and q is the moving average model order [33]. The FSF-ARIMA model with the best results based on AIC and the number of parameters was used for forecasting and determining the linear part of the steam flow data. Then Q_t' is forecasted using Equation (5) for one year ahead.

Several FSF-ARIMA(p,1,q) models are developed and values of p and q are determined based on minimization of AIC. For each candidate model, we use Equation (6) to compute the AIC as below:

$$AIC = -2\log(\max \text{likelihood}) + 2(\text{number of parameters}) \quad (7)$$

Building RANN consists of selecting a learning function, inputs, and activation and training functions [9]. The Levenberg–Marquardt (LM) algorithm was applied as the learning function, which has generally high accuracy and is fast learning [34]. The neural networks involve bias, one hidden layer, and tangent-sigmoid or log-sigmoid as activation functions. A pure linear activation function was applied in the output layer. Besides, the appropriate number of output delays is determined by trial and error process.

2.4. The Proposed Autoregressive Hybrid Model

The proposed hybrid model benefits from the unique advantages of both autoregressive models, FSF-ARIMA, and RANN models, to recognize linear and nonlinear patterns of reservoir inflow time series. Both FSF-ARIMA and RANN models are autoregressive. Therefore, the hybrid model is also an autoregressive model. Developing the proposed hybrid model generally includes two steps. In the first step, an FSF-ARIMA model is applied to forecast linear elements of daily inflow time series. In the second step, a RANN model is employed to forecast FSF-ARIMA model residuals. Since FSF-ARIMA models cannot calculate nonlinear structures of the datasets, the residuals of the FSF-ARIMA model are the nonlinear part of the stream flow time series and can be forecasted by the nonlinear part of the proposed hybrid model (RANN). The Residuals of Linear Forecasting (RLF_t) can be computed as follows:

$$RLF_t = Q_{ot} - Q_{ARIMA} \tag{8}$$

where Q_{ot} is output discharge. To forecast the nonlinear part of reservoir inflow time series (RLF_t), the proposed autoregressive hybrid model employs the RANN model. The RANN model is trained with RLF_t time series to find the following autoregressive nonlinear relation:

$$RLF_t = F(RLF_{t-1}, RLF_{t-2}, RLF_{t-3}, \dots, RLF_{t-rn}, DOY) \tag{9}$$

where F is a nonlinear function estimated using the RANN; rn is the number of output delays and is tested for 1, 2, 3, 4, 5 in multiple RANN (RANN1 to RANN5); and DOY is the day of year of the forecasted day; RLF_t is the predicted flow at time step t . RLF_{t-y} is y -day delayed flow data (observed data in training and forecasted values in forecasting phase).

The RANN is inspired from our understanding of the human brain’s neural networks system. First, information processing is accomplished in elements which are identified as neurons; second, information is conveyed between neurons by using their connections; third, each connection has a specific weight that is a multiplier for information conveyed from one neuron to another; fourth, each neuron regularly uses a nonlinear activation function to compute its output. A RANN is described based upon network structure, training method, and activation function. Figure 2 shows a schematic diagram of a RANN model developed in this study; RANNs are dynamic recurrent nonlinear ANNs. In RANN networks, output delays act as dynamic memory in the reservoir inflow forecasting phase. After forecasting RLF_t using RANN, the forecasted reservoir inflows by autoregressive hybrid model (Q_{AH}) are calculated by the following equation:

$$Q_{AH} = Q_{ARIMA} + RLF_t \tag{10}$$

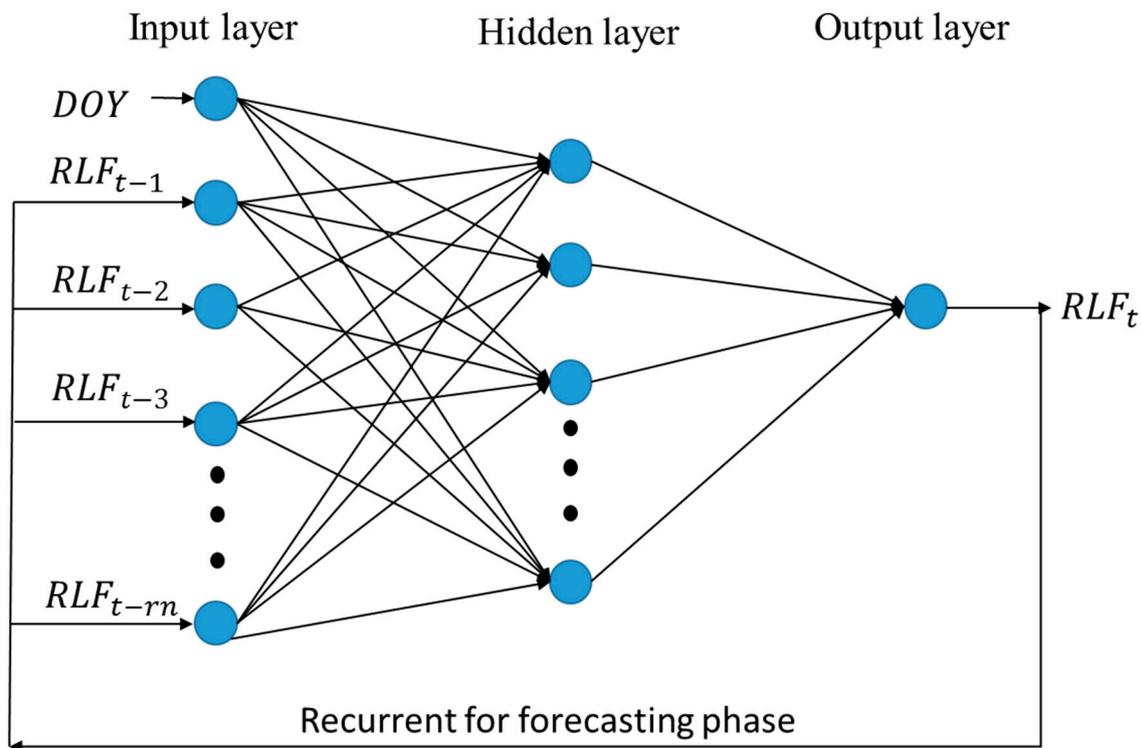


Figure 2. Recurrent Artificial Neural Network (RANN) structure.

In this research, the numbers of output delays from 1 to 5 in RANN1 to RANN5 models are examined. In addition, the number of neurons in the hidden layer is tested from 1 to 30. In each test, the output of RANN (RLF_t computed by Equation (8)) and target value of the outputs (RLF_t computed by Equation (7)) are compared, and the changing of weights and bias are repeated to

minimize the model error ($RMSE$). Then, the best RANN model is selected based on minimization of EI .

3. Results and Discussion

The best FSF-ARIMA model was chosen based on AIC and the number of parameters among multiple possible FSF-ARIMA models (Table 1). In addition, among the examined structures, FSF-ARIMA (5, 1, 5), FSF-ARIMA (5, 1, 6), and FSF-ARIMA (1, 1, 2) had the lowest AIC (about 13,760). Since AIC values are similar in these three structures, the best model is determined based on the minimum number of parameters. The FSF-ARIMA (1, 1, 2) with the lowest AIC and the minimum number of parameters among FSF-ARIMA models was chosen as the best model.

Table 1. The best Fourier-Series Filtered Autoregressive Integrated Moving Average (FSF-ARIMA) models based on Akaike Information Criterion (AIC) and number of parameters.

| Best Models | AIC | Number of Parameters |
|---------------------|-----------|----------------------|
| FSF-ARIMA (5, 1, 5) | 13,758.86 | 10 |
| FSF-ARIMA (5, 1, 6) | 13,759.23 | 11 |
| FSF-ARIMA (1, 1, 2) | 13,760.99 | 3 |
| FSF-ARIMA (1, 1, 1) | 13,762.64 | 2 |
| FSF-ARIMA (2, 1, 2) | 13,763.20 | 4 |
| FSF-ARIMA (4, 1, 1) | 13,763.83 | 5 |
| FSF-ARIMA (3, 1, 2) | 13,764.73 | 5 |
| FSF-ARIMA (4, 1, 2) | 13,765.43 | 6 |
| FSF-ARIMA (5, 1, 1) | 13,765.83 | 6 |
| FSF-ARIMA (4, 1, 3) | 13,767.53 | 7 |
| FSF-ARIMA (3, 1, 3) | 13,768.79 | 6 |
| FSF-ARIMA (3, 1, 5) | 13,769.35 | 8 |
| FSF-ARIMA (4, 1, 4) | 13,769.78 | 8 |
| FSF-ARIMA (0, 1, 1) | 14,657.45 | 1 |
| FSF-ARIMA (1, 2, 1) | 14,847.10 | 2 |

The best structure among the various structures of the RANN model to forecast the daily stream flow to the Dez reservoir was selected based on minimizing EI . The minimum EIs of training and forecasting phases for autoregressive hybrid models were 0.58 and 0.41, respectively. The selected autoregressive hybrid model uses RANN5 as its nonlinear part. RANN5 has a log sigmoid activation function, RLF_{t-1} , RLF_{t-2} , RLF_{t-3} , RLF_{t-4} , RLF_{t-5} , and DOY as the RANN model's inputs, 22 neurons in the hidden layer, and one neuron in the output layer (Table 2). The result of the best autoregressive hybrid model is selected as the proposed model and is compared with the previous FSF-ARIMA and RANN models.

Table 2. Result of the best autoregressive hybrid models.

| Model Name | Hybrid Model Input | Number of Hidden Layer Neurons | Number of Output Delays | Transfer Function of Hidden Layer | EI | R ² | MAE |
|-----------------|--------------------|--------------------------------|-------------------------|-----------------------------------|------|----------------|-----|
| Hybrid (1,14,1) | 1 | 14 | 1 | logsig | 0.47 | 0.88 | 27 |
| Hybrid (5,22,1) | 5 | 22 | 1 | logsig | 0.41 | 0.86 | 21 |
| Hybrid (1,17,4) | 1 | 17 | 4 | logsig | 0.55 | 0.74 | 29 |
| Hybrid (5,5,2) | 5 | 5 | 2 | tansig | 0.65 | 0.63 | 34 |

Highlighted row represents the best model.

Comparing the models based on evaluation indices showed that the proposed autoregressive hybrid model decreases the error of fitting in the training phase. Coefficient of determination (R^2) and error index (EI) were used for examining the error of fitting in the training phase. The higher value of R^2 for the proposed model indicates that the proposed model enhanced the goodness of

fitting in training phase (Table 3). In addition, the results show that the proposed autoregressive hybrid model had a smaller EI than the previous FSF-ARIMA and RANN models (Table 3). In the training phase, the EI decreased by 0.44 and 0.27, and also R^2 increased by 0.67 and 0.03 by using the proposed model compared to the FSF-ARIMA and RANN models, respectively. The comparison indicates the improvement of fitting by the proposed model in the training phase. Therefore, the results indicate significant improvement via capturing the flow pattern in the training phase by the proposed hybrid autoregressive model compared to the previous models.

Table 3. Comparison of the models based on evaluation indices.

| Models | EI (Training) | EI (Forecasting) | R^2 (Training) | R^2 (Forecasting) | MAE (Training) | MAE (Forecasting) |
|-----------|---------------|------------------|------------------|---------------------|----------------|-------------------|
| Hybrid | 0.58 | 0.41 | 0.77 | 0.86 | 70 | 21 |
| FSF-ARIMA | 0.87 | 0.85 | 0.10 | 0.50 | 105 | 70 |
| RANN | 0.62 | 0.68 | 0.74 | 0.60 | 77 | 30 |

The regression evaluation metrics help to determine how close the predicted values are to the actual ones. However, they do not evaluate whether the model properly fits the data while the residuals are usually dedicated to evaluating this. Therefore, we evaluated the forecasting reliability of the proposed models by examining for auto-correlation in the errors [35,36]. Figures 3–5 illustrate the autocorrelation function (ACF) diagram for Dez reservoir inflow forecasting by FSF-ARIMA model, RANN model, and Hybrid model, respectively. As we can see, most of the ACF values fall into the 95% confidence bounds, and they show a decreasing trend for increasing lag times.

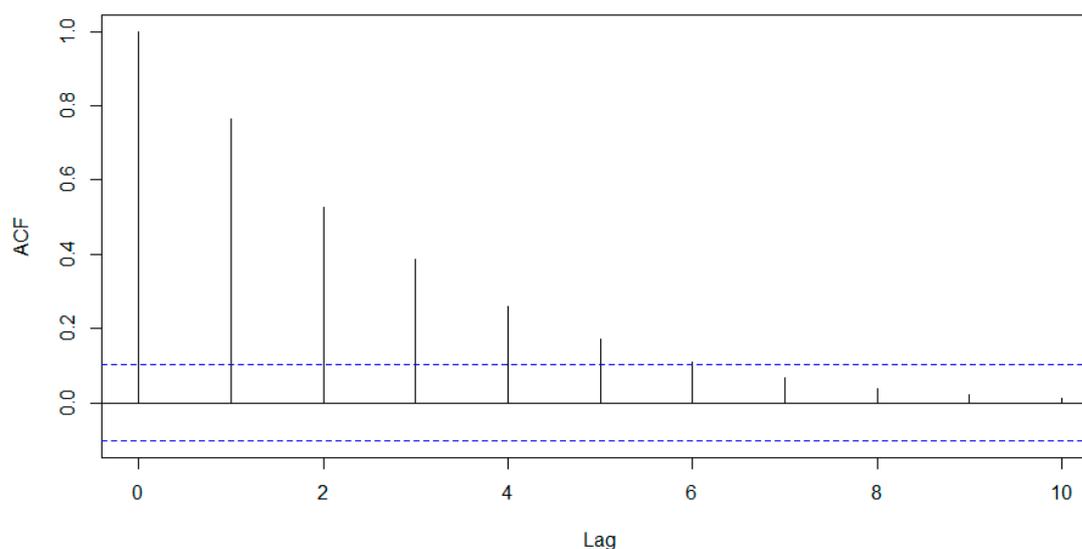


Figure 3. autocorrelation function (ACF) diagram for Dez reservoir inflow forecasting by FSF-ARIMA model; the blue dotted lines represent 95% confidence bounds.

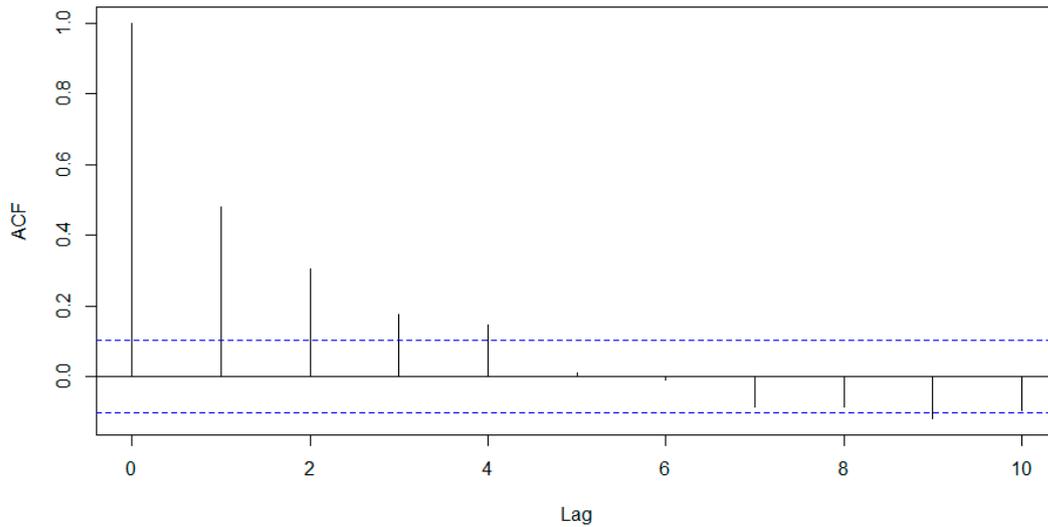


Figure 4. ACF diagram for Dez reservoir inflow forecasting by RANN model; the blue dotted lines represent 95% confidence bounds.

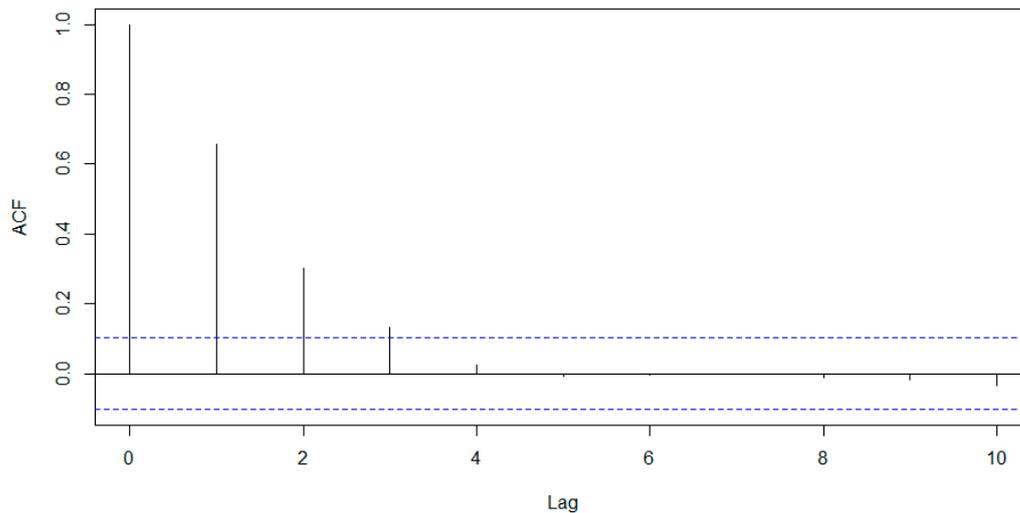


Figure 5. ACF diagram for Dez reservoir inflow forecasting by hybrid model; the blue dotted lines represent 95% confidence bounds.

The comparison of the observed and model-based hydrographs reveals that the proposed model follows the observed flow variation better than the previous models. Figure 6 shows the comparison of forecasted hydrographs in Cubic Meter per Second (CMS) by the models versus the observed data. However, the inflow hydrograph proposed by the RANN model follows the observed hydrograph better than that of the FSF-ARIMA model but with considerable error in forecasting peak flows (Figure 6). Moreover, Figure 6 displays the capability of the proposed model in following the peaks and low points of observed hydrograph. The hybrid model forecasts peak flows better than the RANN model. The proposed hybrid model forecasts the inflow values of the hydrograph precisely, whereas the FSF-ARIMA model overestimates the inflow except for the maximum peak flow (Figure 6). In addition, the proposed autoregressive hybrid model forecasts the maximum peak flow considerably better than the RANN model by reducing the relative error from 361% to 57% compared to the previous study [9].

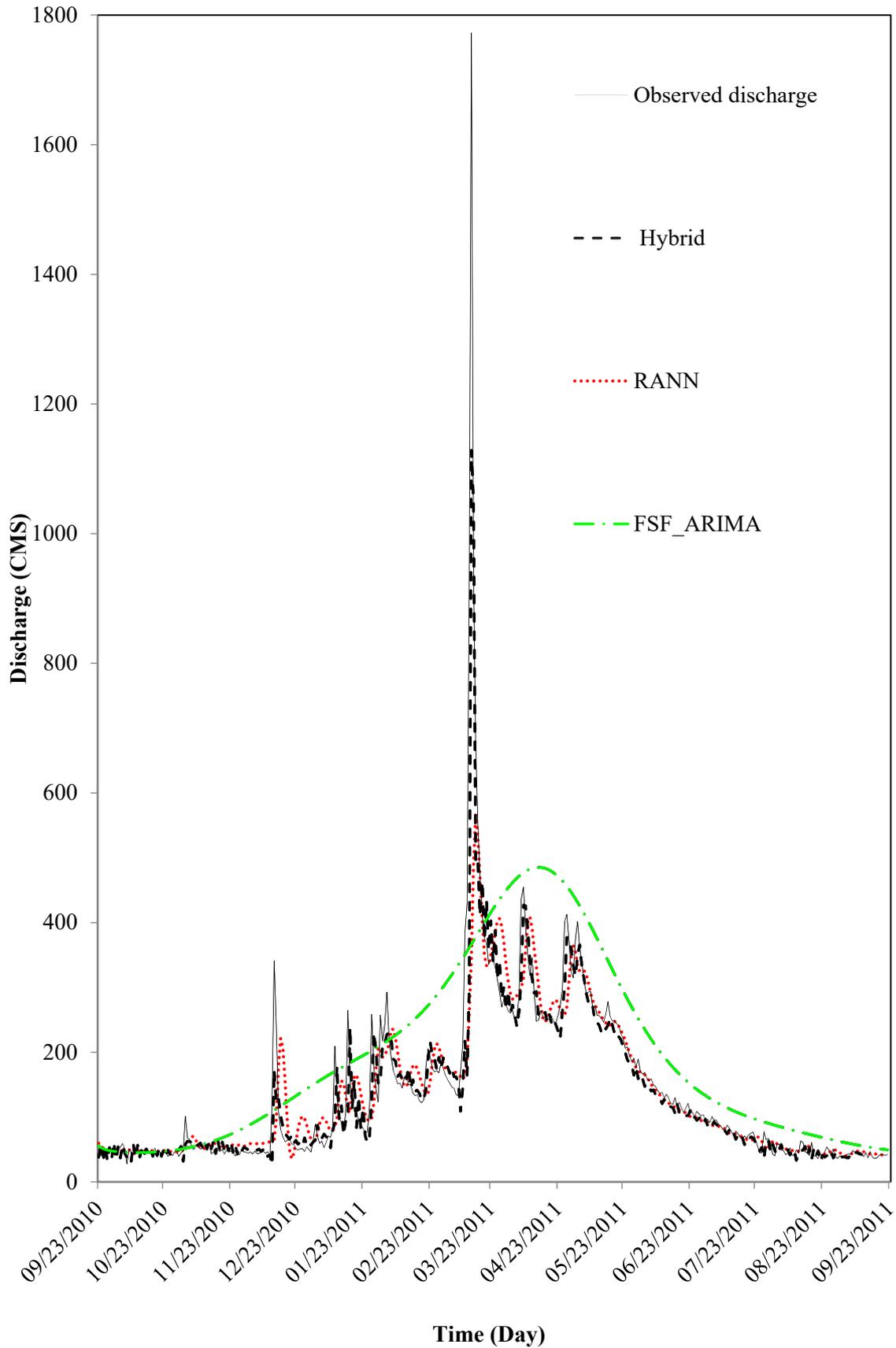


Figure 6. Comparison of observed and forecasted hydrograph by proposed and previous models.

The evaluation of the models based on the *ACRE* indicates that the monthly forecasting-performance of the proposed model is better than the previous models (Figure 7). In most of the

months, *ACRE* for the autoregressive hybrid model is less than for the previous models. Most of the months show *ACRE* higher than 0.4 for the FSF-ARIMA model. However, the proposed model decreases *ACRE* to less than 0.15. In addition, the maximum of *ACRE* for the proposed model is considerably less than the previous models. Furthermore, the maximum and average of *ACRE* for the FSF-ARIMA model are 1.2 and 0.7 for the peak season (from January to May), respectively. Those values are 0.225 and 0.15 for the proposed model. Indeed, the hybrid model has a potential to decrease maximum and average forecasting error by 81% and 80%, respectively.

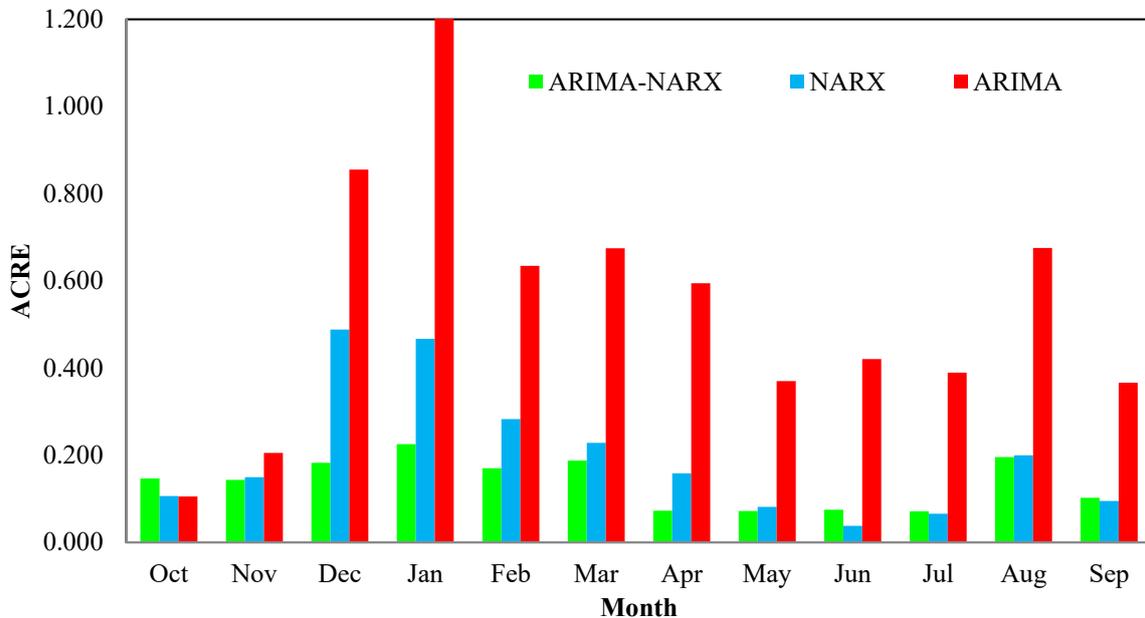


Figure 7. Monthly variations of *ACRE*.

4. Conclusions

In this study, an autoregressive hybrid model developed to forecast long-horizon daily inflow for optimal operation of multi-purpose reservoir with flood control, hydropower generation, domestic water supply, and irrigation goals. The proposed hybrid model comprises two parts: FSF-ARIMA (for linear part) and RANN model (nonlinear part). Since FSF-ARIMA models cannot calculate nonlinear structures of the datasets, the residuals of the FSF-ARIMA model are the nonlinear part of the stream flow time series and can be forecasted by the nonlinear part of the proposed hybrid model (RANN). The best RANN developed in this study had the log sigmoid activation function, five recurrent RLF_{t-1} , RLF_{t-2} , RLF_{t-3} , RLF_{t-4} , RLF_{t-5} , 22 neurons in the hidden layer, and one neuron in the output layer. The best autoregressive hybrid model is selected as the proposed model and compared with the previous linear and nonlinear auto-regressive models (FSF-ARIMA and RANN). Comparison of the proposed autoregressive hybrid model and the previous models showed that the forecasting of long-term daily flow was significantly enhanced by the proposed hybrid model as follows:

- The results demonstrated significant improvement in capturing the inflow pattern by the proposed autoregressive hybrid model from the previous models.
- The monthly variations of the forecasting accuracy were extensively improved by the proposed model throughout the year and during peak season.
- The proposed autoregressive hybrid model forecasted the peak flows more precisely than the previous models.

Finally, the achievement of this research compared to the current forecasting models [9] is proposing an autoregressive hybrid model which improves the ability for forecasting reservoir inflow especially for peak flows which occurred during a one-year horizon. In [9], NARX-RNN and

ARIMA models were employed. The results of ARIMA showed *EI* values equal to 0.87 and 0.85 for the training and forecasting period, respectively. Moreover, the results of NARX-RNN showed *EI* values equal to 0.62 and 0.68 for training and forecasting period, respectively. Therefore, compared to Table 3, the hybrid model developed in the current study outperforms both ARIMA and NARX-RNN which were presented by [9]. The findings of this study can be used for optimum allocation of water resources and releasing reservoir water for optimal operation of multi-purpose reservoirs, especially in operating dams for flood control and generating hydropower. Although the hybrid model outperformed the RANN and FSF-ARIMA models, the ACF plots revealed that all models were unable to make reliable forecasts. It is worth mentioning that employing a more advanced model such as long short-term memory (LSTM) and deep learning techniques like the Convolutional Neural Network (CNN), as well as their combination, CNN-LSTM [35,36], can improve the accuracy of forecasting. In addition, these new models have shown more reliable forecasts [35,36]. Therefore, there are open avenues to compare the results of this study with LSTM and CNN-LSTM in future investigations.

Author Contributions: M.E.B., principal investigator, supervised the project; R.B. ran the models and generated the results; M.V. revised the paper.

Funding: This research was supported by the Khuzestan Water and Power Authority (KWPA) through research project #89-02-02-017.

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

References

1. Valenzuela, O.; Rojas, I.; Rojas, F.; Pomares, H.; Herrera, L.J.; Guillén, A.; Marquez, L.; Pasadas, M. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets Syst.* **2008**, *159*, 821–845.
2. Banihabib, M.E.; Mousavi, F.S.; Jamali, F.S. Artificial neural network model to study the spatial and temporal correlation between stations in inflow forecasting. In Proceedings of the 3rd Iran Water Resources Management Conference, Tabriz, Iran, 14–16 October 2008.
3. Kisi, I.; Cigizoglu, K. Reservoir management using artificial neural networks. In Proceedings of 14th. Reg. Directorate of DSI (State Hydraulic Works), Istanbul, Turkey, 11–12 June 2005.
4. Noakes, D.J.; McLeod, A.I.; Hipel, K.W. Forecasting monthly river flow time series. *Int. J. Forecast.* **1985**, *1*, 179–190.
5. Tawfik, M. Linearity versus non-linearity in forecasting Nile River flows. *Adv. Eng. Softw.* **2003**, *34*, 515–524.
6. Valipour, M.; Banihabib, M.E.; Behbahani, S.M.R. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J. Hydrol.* **2013**, *476*, 433–441.
7. Yurekli, K.; Kurunc, A.; Ozturk, F. Application of linear stochastic models to monthly flow data of Kelkit Stream. *Ecol. Model.* **2005**, *183*, 67–75.
8. Aggarwal, S.; Goel, A.; Singh, V.P. Stage and discharge forecasting by SVM and ANN techniques. *Water Resour. Manag.* **2012**, *26*, 3705–3724.
9. Banihabib, M.E.; Bandari, R.; Peralta, R.C. Auto-regressive neural-network models for long lead-time forecasting of daily flow. *Water Resour. Manag.* **2019**, *33*, 159–172.
10. Partal, T.; Cigizoglu, H.K. Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *J. Hydrol.* **2008**, *358*, 317–331.
11. Sattari, M.T.; Pal, M.; Apaydin, H.; Ozturk, F. M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey. *Water Resour.* **2013**, *40*, 233–242.
12. Zhang, Z.; Zhang, Q.; Singh, V.P. Univariate streamflow forecasting using commonly used data-driven models: Literature review and case study. *Hydrol. Sci. J.* **2018**, *63*, 1091–1111.
13. Xie, M. Prediction of Daily Net Inflows for Management of Reservoir Systems. Master's thesis, McGill University, Montréal, QC, Canada, 2001.
14. Sattari, M.T.; Yurekli, K.; Pal, M. Performance evaluation of artificial neural network approaches in forecasting reservoir inflow. *Appl. Math. Model.* **2012**, *36*, 2649–2657.

15. Jeong, D.I.; Kim, Y.O. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrol. Process. Int. J.* **2005**, *19*, 3819–3835.
16. Lin, G.F.; Wu, M.C.; Chen, G.R.; Tsai, F.Y. An RBF-based model with an information processor for forecasting hourly reservoir inflow during typhoons. *Hydrol. Process. Int. J.* **2009**, *23*, 3598–3609.
17. Mehr, A.D.; Kahya, E.; Bagheri, F.; Deliktas, E. Successive-station monthly streamflow prediction using neuro-wavelet technique. *Earth Sci. Inform.* **2014**, *7*, 217–229.
18. Mohammadi, K.; Eslami, H.; Dardashti, S.D. Comparison of regression, ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj). *J. Agric. Sci. Technol.* **2005**, *7*, 17–30.
19. Pekarova, P.; Pekar, J. Long-term discharge prediction for the Turnu Severin station (the Danube) using a linear autoregressive model. *Hydrol. Process. Int. J.* **2006**, *20*, 1217–1228.
20. Shalamu, A. Monthly and seasonal streamflow forecasting in the Rio Grande Basin. Ph.D. Thesis, New Mexico State University, Las Cruces, NM, USA, 2009.
21. Coulibaly, P.; Anctil, F.; Bobee, B. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J. Hydrol.* **2000**, *230*, 244–257.
22. Hassan, M.; Shamim, M.A.; Hashmi, H.N.; Ashiq, S.Z.; Ahmed, I.; Pasha, G.A.; Naeem, U.A.; Ghumman, A.R.; Han, D. Predicting streamflows to a multipurpose reservoir using artificial neural networks and regression techniques. *Earth Sci. Inform.* **2015**, *8*, 337–352.
23. Karunasinghe, D.S.; Liong, S.Y. Chaotic time series prediction with a global model: Artificial neural network. *J. Hydrol.* **2006**, *323*, 92–105.
24. Pulido-Calvo, I.; Portela, M.M. Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. *J. Hydrol.* **2007**, *332*, 1–15.
25. Riad, S.; Mania, J.; Bouchaou, L.; Najjar, Y. Predicting catchment flow in a semi-arid region via an artificial neural network technique. *Hydrol. Process.* **2004**, *18*, 2387–2393.
26. Wang, W.; Van Gelder, P.H.; Vrijling, J.; Ma, J. Forecasting daily streamflow using hybrid ANN models. *J. Hydrol.* **2006**, *324*, 383–399.
27. Huang, W.; Xu, B.; Chan-Hilton, A. Forecasting flows in Apalachicola River using neural networks. *Hydrol. Process.* **2004**, *18*, 2545–2564.
28. Cryer, J.D.; Chan, K.S. *Time Series Analysis: With Applications in R*; Springer-Verlag: New York, NY, USA, 2008.
29. Kote, A.S.; Jothiprakash, V. Reservoir inflow prediction using time lagged recurrent neural networks. In Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology, IEEE, Nagpur, Maharashtra, India, 16–18 July 2008; pp. 618–623.
30. Jothiprakash, V.; Kote, A.S. Effect of pruning and smoothing while using M5 model tree technique for reservoir inflow prediction. *J. Hydrol. Eng.* **2010**, *16*, 563–574.
31. Xu, W.; Fu, X.; Li, X.; Wang, M. Data transformation models utilized in Bayesian probabilistic forecast considering inflow forecasts. *Hydrol. Res.* **2019**, *50*, 1267–1280.
32. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control.* **1974**, *19*, 716–723.
33. Salas, J.; Delleur, J.; Yevjevich, V.; Lane, W. Applied modeling of hydrological time series. Water Resources Publications: Littleton, CO, USA, 1980.
34. Banihabib, M.E.; Arabi, A.; Salha, A.A. A dynamic artificial neural network for assessment of land-use change impact on warning lead-time of flood. *Int. J. Hydrol. Sci. Technol.* **2015**, *5*, 163–178.
35. Livieris, I.E.; Pintelas, E.; Stavroyiannis, S.; Pintelas, P. Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms* **2020**, *13*, 121.
36. Livieris, I.E.; Pintelas, E.; Kiriakidou, N.; Stavroyiannis, S. An advanced deep learning model for short-term forecasting U.S. natural gas price and movement. In Proceedings of the 16th International Conference on Artificial Intelligence Applications and Innovations (AIAI), Halkidiki, Greece, 5–7 June, 2020.

