Investigation of Forecast Accuracy and its Impact on the Efficiency of Data-Driven Forecast-Based Reservoir Operating Rules

Ehsan Mostaghimzadeh 1, Seyed Mohammad Ashrafi 1, Arash Adib 1 and Zong Woo Geem 2,*

1 Department of Civil Engineering, Faculty of Civil Engineering and Architecture, Shahid Chamran University of Ahvaz, Ahvaz 6135783151, Iran; e.mostaghimzadeh@gmail.com (E.M.); ashrafi@scu.ac.ir (S.M.A.); arashadib@scu.ac.ir (A.A.)
2 College of IT Convergence, Gachon University, Seongnam 13120, Korea
* Correspondence: zwgeem@gmail.com

Abstract: Today, variable flow pattern, which uses static rule curves, is considered one of the challenges of reservoir operation. One way to overcome this problem is to develop forecast-based rule curves. However, managers must have an estimate of the influence of forecast accuracy on operation performance due to the intrinsic limitations of forecast models. This study attempts to develop a forecast model and investigate the effects of the corresponding accuracy on the operation performance of two conventional rule curves. To develop a forecast model, two methods according to autocorrelation and wrapper-based feature selection models are introduced to deal with the wavelet components of inflow. Finally, the operation performances of two polynomial and hedging rule curves are investigated using forecasted and actual inflows. The results of applying the model to the Dez reservoir in Iran visualized that a 4% improvement in the correlation coefficient of the coupled forecast model could reduce the relative deficit of the polynomial rule curve by 8.1%. Moreover, with 2% and 10% improvement in the Willmott and Nash—Sutcliffe indices, the same 8.1% reduction in the relative deficit can be expected. Similar results are observed for hedging rules where increasing forecast accuracy decreased the relative deficit by 15.5%. In general, it was concluded that hedging rule curves are more sensitive to forecast accuracy than polynomial rule curves are.

Keywords: reservoir operation; hedging policy; forecast accuracy; stacking algorithm; ensemble learning; support vector regression

1. Introduction

Recently, water resource models have been used increasingly to study the vulnerability of human systems under probable water scarcity [1]. The forecast accuracy of the future surface and sub-surface inflows, which is always associated with different uncertainties, influences the efficiency of water resource planning and management [2,3]. Accurate prediction is crucial due to climate change and human activities that affect hydrological processes, leading to major changes in water distribution. According to the Fifth Report of the IPCC (Inter-Governmental Panel on Climate Change, 2014), the earth’s temperature is rising due to the frequent use of greenhouse gases, and these temperature increases will eventually alter natural flow patterns [4]. On the other hand, variable flow patterns ultimately impose serious challenges in reservoir operation. Generally, the release rate is determined using operating rule curves. However, these curves are fitted based on historical time series, and the release rate somehow depends on past events [5]. As such, possible inflow changes could decrease operating performance. To overcome this problem, a wide range of studies has been investigated to combine future conditions into operating policy. Liu et al. [6] suggested a Bayesian-based deep learning model to assess the impact of streamflow uncertainty on the efficiency of operation rule curves.
Uncertainty estimation results show that the effect of streamflow uncertainty is greater than corresponding values in model parameters. Kangrang et al. [7] evaluated reservoir operation policies against future flow patterns using climate change scenarios and land use patterns. The results showed fewer reservoir fluctuations compared to models without forecast terms. Zhang et al. [8] investigated the effects of forecast uncertainty on adaptive policies. This model was based on the simultaneous use of historical time series and the output of general circulation models (GCM). The results of applying this model to the Dongwushi reservoir in China showed the superiority of the proposed method as an effective approach in variable conditions caused by climate change. In order to derive the optimal strategy for a hydropower reservoir, Feng et al. [9] defined a novel approach according to the class-based evolutionary extreme learning machine (CEELM). The results show that the suggested method could properly drive operation rules for hydropower reservoirs. Anghileri et al. [10] introduced a consistent forecast-based operation model to supply the demands and to evaluate the influence of long-term inflow prediction on operation policies. The results showed that ensemble forecast data were approximately 35% less effective than meteorological data was to improve operation performances. It was also stated that the yearly component of the forecast is more influenced by seasonal factors. Ahmad and Hossain proposed a generic and scalable scheme for forecasting reservoir inflow to optimize reservoir operation using weather forecasts. The results of applying this model to over 23 dams in the U.S. with varying hydrological characteristics indicated more energy production without compromising flood risk [11]. Ashbolt and Perera [12] defined a series of historical inflow scenarios to investigate the effects of each scenario on a multi-reservoir system. The results showed that the short-term multi-objective simulation–optimization model had better performance compared to a long-term policy. Peng et al. [13] proposed a rainfall–runoff model and a stochastic dynamic programming model to effectively merge the ensemble precipitation forecast into an operation policy. The results show that hydropower energy generation increases compared to conventional rule curves. Jin and Lee [14] introduced a discrete hedging rule consisting of several hedging phases to determine reservoir release during drought periods. The results showed that the beginning storage-based rules were placed higher than the curves based on available water and induced earlier hedging. Ficchi et al. [15] suggested a combinatorial model benefiting from ensemble forecast data to improve the performance of a four-reservoir system in France. The results showed the superiority of the proposed model to manage the maximum and minimum boundary conditions. Ashrafi et al. [16] presented a wavelet-based forecast model according to the decomposed components of historical inflows in the Dez reservoir in Iran. Then, the performances of two predefined rule curves, including polynomial and hedging rules, were evaluated with and without the forecast term. The results showed that the forecast term could improve the evaluation criteria against standard operation policy.

In all of the above studies, a combinatorial policy corresponding to the forecast model was presented, while the effects of these forecast accuracies have not yet been thoroughly investigated. In other words, it is not possible to make an accurate estimate of the commensurate operation performance with the forecast accuracy, which may reduce performance in case of inaccurate forecast terms. As such, the purpose of this study is to first develop a monthly forecast model using data-driven models. To do so, two methods are introduced according to the wavelet components of inflows. Secondly, this forecast term is combined with two predefined rule curves to investigate the accuracy of the interactions with reservoir operation performance.

2. Materials and Methods

The conceptual structure of this study, which is shown in Figure 1, consists of three separate models: (1) the forecast model, (2) the reservoir operation model, and (3) the evaluation model. First, to develop a forecast model, two data-driven methods using discrete wavelet transform (DWT), an artificial neural network, a wrapper-based feature selection model, and a stacking learning structure are suggested. The output of this step is
the best forecast model in terms of evaluation indices. In the second step, three conventional
operation policies, including a standard operation policy (SOP), a polynomial rule curve,
and a hedging rule curve, are developed using a simulation–optimization model [17,18].
These rule curves are developed according to the essence of the inflow term, which may
be the output of the forecast model, the actual observed data, or eliminated from the rule
curve equation. Finally, the evaluation model investigates the results of the simulation–
optimization model in terms of operation indices and reveals the effects of the forecast
term and its corresponding accuracy on the operation performance.

2. Materials and Methods

In this study, two wavelet-based methods are used to develop a monthly forecast
model: (1) learning based on the autocorrelation between wavelet components and (2)
finding effective wavelet components using a feature selection model and training them
using a stacking structure. In the first approach, each decomposed component is exclusively
predicted using the autocorrelation between previous lags in the form of ACF and PACF
functions, which represent the temporal dynamics of a time series and measure the average
correlation between data points [19]. These functions are determined for different lags. It is
more common to have fewer ACF and PACF functions for longer lags. For each wavelet
component, these functions are estimated independently to identify the antecedent values
with more significant correlations. Then, the lags with greater function values are selected
as the predictors to estimate the future values of the wavelet components. Finally, the total
inflow is determined according to the aggregation of the predicted wavelet components. In
this approach, a three-layer perceptron network is used as a central learner. The sigmoid
function is selected as the transfer function for the hidden layer. Furthermore, the Haar
wavelet is used as the mother wavelet [16].

In the second approach, inflow is directly forecasted using wavelet components, unlike
the first approach, where the inflow was predicted indirectly as a result of the aggregation

Figure 1. Flowchart of the proposed models.

2.1. Forecast Model

In this study, two wavelet-based methods are used to develop a monthly forecast
model: (1) learning based on the autocorrelation between wavelet components and (2)
finding effective wavelet components using a feature selection model and training them
using a stacking structure. In the first approach, each decomposed component is exclusively
predicted using the autocorrelation between previous lags in the form of ACF and PACF
functions, which represent the temporal dynamics of a time series and measure the average
correlation between data points [19]. These functions are determined for different lags. It is
more common to have fewer ACF and PACF functions for longer lags. For each wavelet
component, these functions are estimated independently to identify the antecedent values
with more significant correlations. Then, the lags with greater function values are selected
as the predictors to estimate the future values of the wavelet components. Finally, the total
inflow is determined according to the aggregation of the predicted wavelet components. In
this approach, a three-layer perceptron network is used as a central learner. The sigmoid
function is selected as the transfer function for the hidden layer. Furthermore, the Haar
wavelet is used as the mother wavelet [16].

In the second approach, inflow is directly forecasted using wavelet components, unlike
the first approach, where the inflow was predicted indirectly as a result of the aggregation
of the predicted wavelet components. In this way, a wrapper-based feature selection model is first introduced to identify the practical components against primary inflow time series. This model is based on a single objective optimization using a multi-layer perceptron (MLP) and genetic algorithm (GA). Once the effective wavelet components are selected, a central learner consisting of a stacking structure is used to develop a direct wavelet-based forecast model. In the central core of the stacking structure, three well-known learners, such as support vector regression (SVR), MLP, and k nearest neighborhood (KNN) algorithms, are used to create an initial forecast for the inflow. Then, the outputs of the learners are considered as the inputs of the secondary learner to determine the effective weights that are related to each of them. This secondary learning section is based on the single-layer perceptron (SLP), which only uses a summation transfer function. Figure 2 visualizes the schematic of the stacking models used in this study [20].

As we can see, in the second approach, practical wavelet components are trained using triple models of KNN, kernel SVR, and MLP independently to directly estimate forecast time series. Then, the corresponding outputs are entered into the second layer, including SLP, to determine the weights of their effects. This model consists of a neuron with one transfer function, which only sums the weighted inputs. As such, the final weights indicate the effects of each sub-model in this forecasting problem. The critical point here is the differences in the training dataset for the first and second layers, which are applied to prevent the higher weights of the over-fitted models as potential results of the first layer. Figure 3 indicates the division of the data groups to obtain optimal results from the stacking algorithm, where the training data are used for the first layer. Then, the validation data train the second layer using the fitted values from the first layer, and finally, the test data are used to estimate the accuracy of the entire network.
The data division for the training and testing of the stacking algorithm.

As mentioned before, the discrete wavelet transform is used to extract the hidden variation frequencies in the inflow time series. This preprocessor is used in both of the suggested methods. DWT adopts the form of Equation (1) [21]:

$$\psi(a,b) \left( \frac{t - \gamma}{s} \right) = \frac{1}{s_0^{a/2}} \psi \left( \frac{t - b s_0^a}{s_0^a} \right)$$  \hspace{1cm} (1)

where $\psi$ is the mother wavelet, $a$ and $b$ are integers representing the dilation and translation values of the wavelet, respectively, $s_0$ denotes a dilation step whose value is unchanged and is greater than 1, and $\gamma_0$ symbolizes the location variable whose value is greater than zero. Generally, for practical reasons, the values for $s_0$ and $\gamma_0$ are chosen to be 2 and 1, respectively, which is the DWT dyadic grid arrangement (i.e., integer powers of two; logarithmic scaling of the translations and dilations) [22]. If a time series exhibits discrete properties with a value of $x(t)$ occurring at a discrete-time $t$, the wavelet coefficient becomes [23]:

$$W_{\psi}(a,b) = \frac{1}{2^{a/2}} \sum_{t=0}^{N-1} x(t) \cdot \psi \left( \frac{t - b}{2^{a}} \right)$$  \hspace{1cm} (2)

The wavelet coefficient for the DWT is calculated at scale $s = 2^a$ and location $\gamma = 2^a b$, revealing the variety of signals at different scales and locations [24]. Since most of the inflow data are sampled in discrete intervals, it makes sense to use DWT. Kamruzzaman et al. [25] concluded that the Haar wavelet makes outputs more compliant with the natural oscillations based on quadratic waves. The decomposition level is another fundamental parameter, as each decomposition level indicates a frequency band, but the maximum number of levels ($J_{\text{max}}$) with which a signal can be decomposed is determined using Equation (3) [25]:

$$J_{\text{max}} = \log_2 \left[ \frac{N}{N_w} \right] - 1$$  \hspace{1cm} (3)

where $N$ is the signal length and $N_w$ is an indicator for the length of the decomposition filter associated with the mother wavelet. Further investigations into the sensitivity analysis and checking the upper limits convinced us to decompose the time series to the fourth level. As clarified in the discussion of the second approach, the wrapper-based feature selection model uses a single-purpose optimization algorithm to identify effective wavelet components. This optimization problem is expressed as Equation (4), which is simplified in the form of Equation (5):

$$z = \left( \frac{1}{N} \sum_{i=1}^{N} (e_i)^2 \right) \times 100 + w \times n_f$$  \hspace{1cm} (4)
\[ z = E \times (1 + \beta n_f) \]  

(5)

where \( z \) is the objective function of the feature selection problem, \( e_i \) is the forecast error of sample \( i \), \( t_i \) is the target vector, and \( n_f \) is the number of selected features. Equation (4) is obtained from the general equation for the feature selection problem, which is the weighted summation of the forecast error and the feature numbers [26]. The parentheses in this equation are the relative root mean square error (RRMSE), which represents the forecast error. Assuming (RRMSE = \( E \)) and (\( w = \beta \times E \)), the general form of Equation (4) is simplified as Equation (5). These assumptions are expressed from the nature of the trade-off between the two purposes of the minimum forecast error and the minimum features selected. This problem is solved using a binary version of GA [27]. According to methodology, a forecast model is developed either by an indirect forecast of wavelet components using ACF and PACF functions or by a direct inflow forecast using useful features. In the first approach, MLP is used to perform the learning process, while in the direct method, MLP, SVR, and KNN are used in the first layer, which is explained in detail. Since McCulloch and Pitts [28] proposed the idea of the neural network, this machine learning model has been frequently used in forecasting problems. Equation (6) can be used to introduce an artificial neural network in a general form:

\[ Y(x) = F \left( \sum_{i=1}^{L} W_i(t) \cdot X_i(t) + b \right) \]  

(6)

where \( x \) is the vector of the input variables, \( Y \) indicates the forecasted value, \( L \) is the number of hidden neurons, \( w \) is the vector of the connection weights, \( b \) is the bias value, and \( F \) represents the activation function. In this paper, we use the tangent sigmoid transfer function in the hidden layers and the linear transfer function in the output layer. The superiority of these functions in the forecast problem has been investigated in previous research [29]. SVR is based on the basic concepts of the support vector machine, which is usually adopted for solving regression problems [30]. The statistical learning theory coupled with an algorithm from optimization theory is adopted to train the SVR [31]. Applying a nonlinear mapping function, SVR maps the input vector to a high-dimensional feature space [32]. The purpose of SVR is to find a regression function that estimates the functional dependence between a set of sampled points \( x = \{x_1, x_2, \ldots, x_n\} \) (the input vector) and desired values \( y = \{y_1, y_2, \ldots, y_n\} \) [33]. The general form of the SVR regression function can be expressed as Equation (7) [30]:

\[ f(x) = \sum_{n=1}^{n} (\alpha_n - \alpha_n^*) K(x_i, x_j) + b \]  

(7)

where \( \alpha_n, \alpha_n^* \geq 0 \) are the Lagrangian multipliers that satisfy the equality \( \alpha_i \times \alpha_i^* = 0 \), \( K(x_i, x_j) \) is the kernel function, \( n \) is the total number of points, and \( b \) is a constant parameter. To achieve a more confident regression, the dimensionality of the input space is changed by kernel functions. Therefore, applying suitable kernel functions among the various predefined types (e.g., linear, polynomial, radial basis function, multi-layer perception, sigmoid, etc.) has a crucial impact on the performance of the SVR. The most commonly used function is the radial basis function (RBF) [34], which is defined as follows:

\[ K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \]  

(8)

here, \( \gamma > 0 \) is the kernel parameter [35]. In this study, the SVR parameters are optimized using a two-step grid search method that optimizes the model parameters more effectively [35]. We implemented the SVR models based on quadratic programming in MATLAB. KNN regression is a nonparametric method, where the information derived from the observed data is applied to forecast the number of predicted variables without defining a predetermined parametric relation between the predictor and targets. The basis
of this method is to calculate the similarity between the current \( x_r = \{x_{1r}, x_{2r}, \ldots, x_{mr}\} \) and historical \( x_t = \{x_{1t}, x_{2t}, \ldots, x_{mt}\} \) values of neighbors, as per Equation (9) [36]:

\[
D_{rt} = \sqrt{\sum_{i=1}^{m} w_i(x_{ir} - x_{it})^2} \quad t = 1, 2, \ldots, n \tag{9}
\]

where \( w_i (i = 1, 2, \ldots, m) \) are the weights of the predictors whose summation is equal to one, and \( m \) is the total number of points. The forecasted inflow \( (Y_r) \) is determined using the following probabilistic function of the observed inflow \( (T_j) \):

\[
Y_r = \sum_{j=1}^{k} F(D_{rj}) \times T_j \tag{10}
\]

where \( F(D_{rj}) \) is the kernel function of the \( k \) nearest neighbors (\( k \) observed data with the lowest distance from the current predictor), calculated as Equation (11):

\[
F(D_{rj}) = \frac{1/D_{rj}}{\sum_{j=1}^{k} 1/D_{rj}} \tag{11}
\]

In the KNN algorithm, the weights of the predictors and the number of neighbors \( (k) \) affect the final results; therefore, their optimum values should be calculated to achieve the most relevant results. The performance criteria to assess forecast models are the correlation coefficient \( (R) \), Willmott \( (WI) \), and Nash–Sutcliffe [37–39]. The \( R \) index presents valuable knowledge regarding the direction and level of the linear relationship between the output and the target. Nash takes values between \(-\infty, 1\) as the worst and best model performance, respectively. Values between \((0, 1)\) indicate the desired level of accuracy, while the values smaller than zero reveal that the average of the targets is a better estimation than the output values, which reflects undesirable performance [40]. The \( WI \) index usually overcomes the non-sensitivity features, as it determines the difference between the target and the output values that only have a power of 1 and is superior to other indices; however, values greater than 0.65 are possible, even for weak models [38].

2.2. Water Resources Model

To investigate the effects of forecast accuracy on the performance of operation policies, two types of operation rule curves, including the polynomial rule curve (PRC) and the hedging rule curve (HRC), are provided. The reservoir operation model is presented as Equations (12)–(17) in the form of a constrained problem. As such, the penalty method is used to replace this problem with an unconstrained problem whose solution ideally converges to the solution of the original problem [41].

\[
\text{Min } Z = \sum_{t} \left( \frac{TD_t - R_t}{TD_t} \right)^2 + \sum_{j} C_j \times P_j \tag{12}
\]

\[
S_{t+1} = S_t + Q_t + P_t - R_t - E_t - \text{Spill}_t \tag{13}
\]

\[
\text{Spill}_t = \text{Max}\{0, (S_{t+1} - S_{\text{max}})\} \tag{14}
\]

\[
S_{\text{min}} \leq S_t \leq S_{\text{max}} \quad \forall t = 1, \ldots, T \tag{15}
\]

\[
P_j = f_j(\sum \text{Viol}_{ij}) \tag{16}
\]

\[
\text{Viol}_{ij} = \alpha(S_i - S_{\text{min}}) + \beta(R_{\text{Env}, t} - \text{ENV}_i) \tag{17}
\]

where \( Z \) is the operation objective function, \( TD_t \) stands for system demand, and \( S_t \) and \( S_{t+1} \) indicate storage at the beginning and end of month \( t \). \( Q_t, P_t, R_t, E_t, \) and \( \text{Spill}_t \) are inflow, precipitation, release, evaporation, and spill rate, respectively. \( \text{ENV}_i \) and \( R_{\text{Env}, t} \) are the environmental demand and release, respectively. Additionally, \( C_j \) is the weight of the penalty function estimated by sensitivity analysis, and \( P_j \) is the total violation value related
to the minimum environmental flow and minimum water level. The two coefficients $\alpha$ and $\beta$ magnify penalty values leading to faster convergence. In the case of feasible space, the penalty function ensures that the storage level is higher than the minimum level and also ensures that meeting the environmental flow requirements is the highest priority demand. The values for evaporation and precipitation are determined using the iterative method. To have a better overview, Equation (12) represents the objective function according to the minimization of the relative deficit, which is penalized in the opposite direction in case of violation constraints. Equations (13)–(15) represent the physical constraints, including mass balance and reservoir storage, which may lead to spillage or storage below the minimum level. Equation (16) determines the total violation of the constraints over the course of the entire period and are estimated independently in Equation (17). To operate the water resource system, two operating policies formed by polynomial and hedging rules are defined. The real-time state of the polynomial rule curve (RPRC) is as seen in Equation (18) [42]:

$$R_t = a_i \times S_t + b_i \times Q_t + c_i \times D_t + \varepsilon_i$$  \hspace{1cm} (18)

where, $a_i$, $b_i$, and $c_i$ are the rule curve coefficients, and $D_t$ and $\varepsilon_i$ are the total demand and an auxiliary parameter to help conversion. Other parameters were introduced earlier. According to Equation (18), release rates are achieved as a factor of current storage, demand, and the coming inflow rate. The coefficients of $a_i$, $b_i$, and $c_i$ are estimated during optimization. If the inflow term is removed from Equation (18), the static state of the polynomial rule curve (SPRC) is obtained. Similarly, using the observed inflow term instead of the forecast term leads to a complete polynomial rule curve (CPRC). To investigate the hedging policy, two hedging levels ($S_{\text{target}}$, $S_{\text{firm}}$) are defined, so the reservoir storage is divided into three zones according to Equation (19). For each zone, an allocation coefficient ($\alpha_t$) is defined, which determines the release rate [5]. As a result, there are five unknown variables for each month, including two hedging levels and three allocation coefficients that are estimated similarly during the optimization process. Equation (19) indicates the real-time state of the hedging policy (RHRC):

$$R_t = \begin{cases} 
\alpha_{t.1} \times \sum_i D_{i,t} & \text{if } S_{\text{target}} \leq S_t + Q_t \\
\alpha_{t.2} \times \sum_i D_{i,t} & \text{if } S_{\text{firm}} \leq S_t + Q_t \leq S_{\text{target}} \\
\alpha_{t.3} \times \sum_i D_{i,t} & \text{if } S_{\text{min}} \leq S_t + Q_t \leq S_{\text{firm}} 
\end{cases} \hspace{1cm} \text{Zone 1} \hspace{1cm} \text{Zone 2} \hspace{1cm} \text{Zone 3}$$  \hspace{1cm} (19)

where $\alpha_{t,j}$ represents the allocation coefficient at time $t$ for zone $j$. The other parameters were introduced earlier. In Equations (18) and (19), the priority of all of the demands is considered the same. Similarly, eliminating and replacing the inflow terms with observed values leads to a static and completed state of the hedging rule curve, respectively (SHRC and CHRC). To optimize the rule curve parameters, a real efficient version of GA is used in the form of a simulation–optimization model [27]. The random pairing and modified alpha blending operators are employed in the pairing and crossover steps [43]. The algorithm parameters are estimated by performing sensitivity analysis, and the optimal values of these parameters are summarized in Table 1.

**Table 1.** Applied parameters value for optimization models.

| Operation Strategy | Decision Variables | Population Size | Mating Rate | Mutation Rate |
|--------------------|--------------------|-----------------|-------------|---------------|
| SPRC               | 36                 | 140             | 0.5         | 0.16          |
| RPRC               | 48                 | 180             | 0.5         | 0.16          |
| CPRC               | 48                 | 180             | 0.5         | 0.16          |
| SHRC               | 60                 | 220             | 0.5         | 0.16          |
| RHRC               | 60                 | 220             | 0.5         | 0.16          |
| CHRC               | 60                 | 220             | 0.5         | 0.16          |
2.3. Evaluation Model

To investigate the impact of the forecast accuracy on the performance of the predefined rules, the optimization process is categorized according to the role of the inflow parameters: (1) without a forecast term or static state, (2) with a forecast term as a result of a coupled forecast–simulation–optimization model or real-time state, and (3) benefited from the real observed inflow or completed state. Finally, comparing the operation indices, the influence of the forecast term and its corresponding accuracy will be determined.

In this study, the single reservoir system of Dez in Iran is investigated according to Figure 4. The Dez river consists of the Sezar and Bakhtiari branches and joins the Karun river at the BandeGhir and forms the Great Karin river. Table 2 presents the essential reservoir information. Several studies have been conducted to improve the performance of this system to define a systematic simulation. In this study, all of models were developed to cover a period of 60 years, from October 1957 to December 2017. To develop the forecast models, 70% of data were used for training, 15% were used for validation, and 15% were for the testing processes. However, to evaluate the operating policies, 70% and 30% of the data were used in the optimization and simulation stages, respectively. Additionally, all of the demands were considered to be lumped to facilitate the simulation process.

| Reservoir Specification                  | Value | Reservoir Specification                  | Value |
|-----------------------------------------|-------|-----------------------------------------|-------|
| Mean annual discharge (m³/s)            | 254   | Regulation ability                      | Multiyear |
| Normal water level (m)                  | 352   | Installed capacity (MW)                 | 520   |
| Minimum operation level (m)             | 310   | NWL storage (billion m³)                | 45    |
| NWL storage (billion m³)                | 3174  | Design head (m)                         | 165   |
| MOL storage (billion m³)                | 1126  | Number of units                         | 8     |

3. Results

In the first forecast method, five independent networks for wavelet components were developed, where some of the lags between 1 to 10 were considered as predictors according to the ACF and PACF functions. The final inflow was estimated as the summation of
these forecasted components. In the second method, useful features against one month ahead of inflow were identified among five wavelet components. Table 3 indicates that the minimum objective function of the wrapper method (Equation (5)) was gained using the approximation component of layer four and the detail components of layers one, three, and four with corresponding lags.

Table 3. Feature selection output—second method.

| Kind of Features | Objective Function ($z$) |
|------------------|--------------------------|
| App41-App42-App43-App47 | 32.36 |
| D11-D12-D15-D18 | |
| D31-D36-D39 | |
| D41-D45-D48 | |

The feature magnifier, $\beta$ (Equation (5)), was estimated using a trial-and-error approach during the optimization process. This parameter can alter the feature selection policy, where the smaller value of $\beta$ causes the model to be more flexible and leads to the use of more features. The higher value of $\beta$ magnifies the influence of $n_i$ and usually results in the use of a fewer number of features [44]. In this study, the $\beta$ coefficient was estimated as 0.21, which allowed the model to apply more features as long as the objective function was downward.

When the practical features of inflow were identified, the stacking algorithm was used to perform the learning process to complete the proposed forecast model. In the stacking algorithm, the central learners of the first layer have a significant impact on the results [45]. These models must estimate the outputs as accurately as possible and must simultaneously not overfit the input data [45]. In this study, the KNN, MLP, and SVR sub-models were used as the first layer learners, and the SLP was applied as the second layer, which is an aggregator. Then, all of the sub-model parameters were calibrated through a trial-and-error method. To increase the regression flexibility, an RBF kernel was used in the SVR sub-model, and all of the lagrangian coefficients ($\alpha_i^-$, $\alpha_i^+$) resulting in support vectors, weights, and bias values were calculated using a quadratic optimization. In this regard, the calibration of the SVR parameters, including the support range ($\varepsilon$), cost of violation ($C$), and the number of kernels ($\sigma$), was a serious challenge. It was observed that the fewer values of $\sigma$ induced more localization to optimization behavior, which led to sharper peaks, while higher $\sigma$ neglected the dynamics of the data, which significantly deviated the outputs from the initial behavior [45]. On the other hand, the higher values of $\varepsilon$ made the model more flexible so that it could try more sets of data, and the smaller values limited the optimization to the smaller ranges, which improved the accuracy [46]. The simultaneous calibration of these parameters was very challenging, which imposed more calculation costs on the model. In this study, these parameters were estimated as $\varepsilon = 0.1$, $C = 88$, and $\sigma = 1$. Additionally, in the KNN algorithm, the number of neighbors was set to $k = 15$. Another point raised here was about the weights of the nearest neighbors. In this study, whenever we talk about KNN, we are referring to the weighted version, which uses the weighted average of the nearest neighbors for each sample. The nearest neighbors are determined based on the Euclidean distance criterion according to Equation (9). In the MLP sub-model, the number of neurons was set to 10, and the sigmoid and linear transfer functions were also used in the hidden and output layers, respectively. Figure 5 illustrates that the minimum weight of the stacking sub-models was gained for SVR in the second layer due to the concept of the learning process [47]. It can be observed that SVR was overfitted on the training data and therefore could not perform correctly in the training phase of the second layer, which generally uses validation data. On the contrary, KNN performed moderately in the first and second training phases, which led to higher weight values. It can be seen that 54% of the final output is estimated using MLP, which could have originated from the accurate and straightforward nature of the learning process in this algorithm [28].
Weights of sub-models in stacking algorithm.

Table 4 shows the final performance of the forecast model for two proposed methods in the form of evaluation indices:

Table 4. Evaluation criteria for forecast models.

| Proposed Method     | R       | WI     | Nash  |
|---------------------|---------|--------|-------|
| Autocorrelation     | 0.922   | 0.873  | 0.839 |
| Ensemble learning   | 0.962   | 0.966  | 0.903 |

As shown, the ensemble learning method, which directly forecasts inflows, performed better than the other one, which forecasts the wavelet components independently. More evaluation shows that the correlation coefficient R in the second approach has the highest value, equal to 0.962, which proves desirable estimation in terms of linearity. Additionally, the WI index behaves very similar to R, where it records a number above 0.972, which is expected. Similarly, the Nash index is estimated as 0.903, which shows satisfied regression with better performance than the first method [39]. Figure 6 visualizes the simulation of the inflow forecast for the test data according to the proposed methods.

It can be seen that despite apparently desirable performances, neither method performed well in terms of peak estimation. However, the direct forecast model handled the initial data pattern better than the autocorrelation method, which estimated some sample points as a negative number. How much the preprocessing improved the evaluation indices should be noted [16]. Figure 7 indicates the percentage of improvement resulting from preprocessing divided by each index.

Figure 6. Comparison of observed and forecast data of proposed methods.
More evaluation revealed that the WI index showed the least amount of impact from preprocessing, as the greatest change in the second method was recorded, about 16%. Moreover, this process affected the indices of the second method, which could be due to the adverse effects of the detail component of the second layer. This component had a very low correlation with inflow, which caused it to be omitted during feature selection. Therefore, the absence of this parameter in the ensemble training improved learning performance. Figure 8 illustrates the peak values of the estimated inflow obtained by the proposed approaches compared to the observed data.

As it can be clearly seen, neither method estimated the peak samples accurately. This fact could be due to the small number of samples for peak data [48]. Figure 6 indicates that the autocorrelation method is more effective in estimating these points, which could be due to the independent forecasting of the components, where each component tends to be estimated separately according to its interior behavior. In summary, the second method, i.e., ensemble learning based on a stacking algorithm, is considered to be coupled with the proposed simulation–optimization model.

To investigate the impact of forecast accuracy on the performance of operation rules, the results were evaluated according to a forecast–simulation–optimization model. As shown in Table 5, the addition of the inflow forecast term led to an improvement for the RPRC that was 7.5% higher than the static state of the SPRC in terms of the relative deficit. The results also showed the better performance of the RPRC in other indices such as quantitative reliability and maximum vulnerability as 7% and 4%, respectively. On the other hand, comparing the RPRC and CPRC, which benefited from complete forecast, i.e., R = 1, revealed that a 4% improvement in the correlation coefficient decreased the relative deficit by 8.1%. Moreover, with 2% and 10% improvement in WI and Nash, respectively, the same 8.1% reduction could be expected. Accordingly, it can be concluded that the
performance of a real-time polynomial rule curve was more sensitive to the WI index of the coupled forecast model. Similarly, the influence of forecast accuracy against the operation indices was evaluated for the hedging policy, where a 4% increase in $R$ decreased the relative deficit by 15.5%. The same improvements were gained for other indices, including 6.25%, 5.28%, and 1.36% for reliability, quantitative reliability, and maximum vulnerability, respectively. In general, it can be concluded that the hedging rule curves were more sensitive to the forecast accuracy than the polynomial rule curves.

Table 5. Effects of forecast accuracy on operation criteria.

| Evaluation Criteria       | Polynomial Rule Curves | Hedging Rule Curve | SOP  |
|---------------------------|------------------------|--------------------|------|
|                           | SPRC | RPRC | CPRC | SHRRC | RHRC | CHRC |     |
| Relative Deficit          | 0.53 | 0.49 | 0.45 | 0.49  | 0.45 | 0.38 | 0.55|
| Reliability               | 0.6  | 0.67 | 0.725| 0.63  | 0.71 | 0.77 | 0.79|
| Vulnerability (Max)       | 0.77 | 0.74 | 0.72 | 0.74  | 0.73 | 0.7  | 1   |
| Quantitative Reliability  | 0.66 | 0.70 | 0.76 | 0.68  | 0.72 | 0.79 | 0.81|
| Reliability               |      |      |      |       |      |      |     |

More evaluation revealed that the decision variables for all of the optimized rule curves were on feasible space. In other words, considering all of the proposed rule curves, the constraints of the minimum storage and minimum environmental demands were completely satisfied. Another point is about the range of variations in the objective function, where every little change in the squared relative deficit reflected a significant improvement in the operation performance, originating from the nature of Equation (12). Therefore, any slight changes in the objective function resulting from the forecast accuracy must be achieved as much as possible. Table 5 also indicates the superiority of real-time and completed policies over standard operation policy in the case of volume-based indices such as vulnerability [14].

The simulation models illustrated the magnified influence of forecast accuracy on the operation policies during late summer and early autumn for the polynomial rules. For example, increasing the forecast accuracy could improve the allocation reliability of autumn by 15%, while this number declined to 6% for spring. Similarly, the improvement of the forecast accuracy decreased the vulnerability by 31% and 27% for the autumn and spring, respectively, which indicates the variable influence in different seasons [49]. The same results were obtained for the hedging policy, where the forecast term accuracy could fluctuate during different seasons.

Reflecting on the effects of forecast accuracy, Figure 9 visualizes the variations of the operational indices for the complete operating rules against the real-time rule curves. As it can be inferred, the relative deficit was generally more sensitive to the forecast accuracy. In contrast, the maximum vulnerability was the least sensitive index in both policies, which could be related to the nature of these indices. As this index is the maximum value of the deficits that occurred, which are directly related to the volume of the allocated water, the improvement of this factor at any value means a significant increase in demand allocation. Therefore, the vulnerability-based indices are generally less sensitive to forecast terms.
Figure 9. Variation of operational indices for completed rule.

Figure 10 represents how the completed rule curves, including CPRC and CHRC, reduced the water deficit over the last ten years of the simulation period. In these figures, the non-white cells indicate the improvement in the total deficit. In other words, the intensity of the non-white cells can be determined as the difference of the total deficit between the completed and real-time rule curves. As such, the darker the cells, the more the percentage of the total deficit reduced due to the improvement in the forecast accuracy. Additionally, the total deficit variations are represented using a color spectrum. According to these figures, the total deficit was reduced more in the case of intensity reduction (darker cells) and the number of periods (number of non-white cells) following the hedging policy. This fact is related to the nature of the hedging policy, which dictates that keeping the water level high makes the rule more dependent on the forecasting terms [48]. Hence, this dependency is reflected in more variations related to higher forecast accuracy.

Figure 10. Visualization of deficit variation over the last ten years of the simulation period against real-time rule curves. (a) CPRC, (b) CHRC.

4. Discussion

Further discussion about the reservoir storage fluctuation reflected the effects of forecast accuracy on the range of usable storage space. In the months leading to the wet season, the higher forecast accuracy led to more flexible operation and resulted in more lower reservoir water levels [49]. A comparison of the results from RPRC and CPRC illustrates that the completed forecast model caused the storage to drop toward dead storage levels by 6% more than the inaccurate forecast term. The same results were obtained for hedging rules where CHRC or using more accurate forecast terms utilized more than 7% of the potential storage space than RHRC. Subsequently, spillage rates where accurate awareness about the future flows reduced the risk of reservoir filling were affected [16]. Using
actual inflows in the case of CPRC, and CHRC reduced the spillage rate by 7% and 10% compared to the rules using semi-accurate forecast terms in RPRC and RHRC, respectively.

Similarly, this improvement was observed in the months leading to the dry season, where more accurate awareness about future inflows kept the water level as close to the maximum level as possible, and more potential reservoir storage was reserved for the coming dry season [49]. The superiority of the forecast-based rule curves confirmed the significant role of the future inflow term as a part of the mathematical rule equation. On the other hand, it was observed that the improvement of the forecast accuracy could have twofold effects on the performance of coupled policies. Hence, improving forecast accuracy as much as possible is crucial in order to have better insight on unstable future conditions.

As we concluded, forecast accuracy has a huge impact on the performance of forecast-based operation policies. The question now is how we can be optimistic about these improvements. Today, climate change has intensified as a result of greenhouse gas production, which is on a dramatic upward trend. This phenomenon can alter rainfall patterns. On the other hand, rainfall–runoff regimes have been significantly triggered by human activities and land use alterations. Therefore, all of these factors have an exponential effect on the performance of forecast models and also deviate the estimation accuracy against observational records. This issue can be considered to be one of the most obvious drawbacks of forecast-based policies, where the accuracy of the forecasted inflows plays a key role in gaining desirable performance during the application of such policies.

5. Conclusions

In order to evaluate the impact of forecast accuracy on the performance of the reservoir operation, some forecast-based operating rule curves were extracted by applying various methodologies and data-driven intelligent models, and the results were compared based on different performance criteria. The developed models were implemented for operation on the Dez reservoir in southwest Iran over a long period of time. Analyzing and discussing the achieved results, the following conclusions can be presented:

1. The wavelet-based decomposed components of inflow time series are positively correlated with the inflow term and can be used directly with longer time horizon estimations. This can be explained mathematically with the concepts of the discrete wavelet transform since each component represents a degree of similarity between the mother wavelet and the initial time series for different frequencies.

2. Each machine learning algorithm can only express some parts of the learning space between the input and output attributes, depending on their computational structure. Therefore, combinatorial approaches that benefit from the combined capabilities of various intelligent methods can significantly improve learning performance.

3. Implementing a forecast term into the operating rule curves remarkably improves the performance of the reservoir operation compared to conventional policies. This could be due to the ability of glancing into future conditions, which can make the rule curves more flexible and compatible with the flow patterns, while having no estimation of the future inflows generally leads to wrong decisions.

4. The performance criteria represent the average values over a long-term horizon, while the effects of forecasting on the performance of the operating rules are varied throughout different conditions. Specifically, while the accuracy of data-driven methods dramatically depends on the sample size, the extracted forecast-based rules may result in considerable errors for extreme events, such as peak flows or severe droughts.

5. The effect of forecast accuracy on operating rules is not constant and strongly depends on the conceptual structure and mathematical formulation of the operating rules. Among the different considered operating rules in the present study, the forecast accuracy affects the hedging rule the most.
Author Contributions: E.M. generated the analysis code under the supervision of S.M.A. and A.A. The text of the paper was written by E.M. and S.M.A. The figures were prepared by E.M. and S.M.A. Z.W.G. edited the paper and supervised the research direction. S.M.A. obtained the funds to perform the research, and Z.W.G. obtained the funds to publish the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Shahid Chamran University of Ahvaz (SCU) under grant number SCU.C99.31254. This research was also supported by the Energy Cloud R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT (2019M3F2A1073164).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable.

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

References
1. Turner, S.W.D.; Doering, K.; Voisin, N. Data-Driven Reservoir Simulation in a Large-Scale Hydrological and Water Resource Model. Water Resour. Res. 2020, 56, 1–16. [CrossRef]
2. Bayesteh, M.; Azari, A. Comparison of the performance of stochastic models in the generation of synthetic monthly flows data: A case study on Marun river. J. Appl. Res. Water Wastewater 2019, 6, 117–125.
3. Ashrafi, S.M.; Dariane, A.B. Coupled Operating Rules for Optimal Operation of Multi-Reservoir Systems. Water Resour. Manag. 2017, 31, 4505–4520. [CrossRef]
4. Stocker, T.; Plattner, G.K.; Dahe, Q. IPCC Climate Change 2013: The Physical Science Basis-Findings and Lessons Learned. In Proceedings of the EGU General Assembly Conference, Vienna, Austria, 27 April–2 May 2014.
5. Ashrafi, S.M. Two-Stage Metaheuristic Mixed-Integer Nonlinear Programming Approach to Extract Optimum Hedging Rules for Multireservoir Systems. J. Water Resour. Plan. Manag. 2021, 147, 04021070. [CrossRef]
6. Liu, Y.; Qin, H.; Zhang, Z.; Yao, L.; Wang, Y.; Li, J.; Liu, G.; Zhou, J. Deriving reservoir operation rule based on Bayesian deep learning method considering multiple uncertainties. J. Hydrol. 2019, 579, 124207. [CrossRef]
7. Kangrang, A.; Prasanchum, H.; Hornwichian, R. Active future rule curves for multi-purpose reservoir operation on the impact of climate and land-use changes. J. Hydro-Environ. Res. 2019, 24, 1–13. [CrossRef]
8. Zhang, W.; Liu, P.; Wang, H.; Chen, J.; Lei, X.; Feng, M. Reservoir adaptive operating rules based on both historical streamflow and future projections. J. Hydrol. 2017, 553, 691–707. [CrossRef]
9. Feng, Z.K.; Niu, W.J.; Cheng, C.T. A quadratic programming approach for fixed head hydropower system operation optimization considering power shortage aspect. J. Water Resour. Plan. Manag. 2017, 143, 06017005. [CrossRef]
10. Anghileri, D.; Voisin, N.; Castelletti, A.; Pianosi, F.; Nijsen, B.; Lettenmaier, D.P. Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. Water Resour. Res. 2016, 52, 4209–4225. [CrossRef]
11. Ahmad, S.K.; Hossain, F. A generic data-driven technique for forecasting of reservoir inflow: Application for hydropower maximization. Environ. Model. Softw. 2019, 119, 147–165. [CrossRef]
12. Ashbolt, S.C.; Perera, B.J.C. Multiobjective Optimization of Seasonal Operating Rules for Water Grids Using Streamflow Forecast Information. J. Water Resour. Plan. Manag. 2018, 144, 05018003. [CrossRef]
13. Peng, A.; Zhang, X.; Peng, Y.; Xu, W.; You, F. The application of ensemble precipitation forecasts to reservoir operation. Water Sci. Technol. Water Supply 2019, 19, 588–595. [CrossRef]
14. Jin, Y.; Lee, S. Comparative Effectiveness of Reservoir Operation Applying Hedging Rules Based on Available Water and Beginning Storage to Cope with Droughts. J. Water Resour. Plan. Manag. 2019, 133, 1897–1911. [CrossRef]
15. Ficchi, A.; Raso, L.; Dorchies, D.; Pianosi, F. Optimal Operation of the Multireservoir System in the Seine River Basin Using Deterministic and Ensemble Forecasts. J. Water Resour. Plan. Manag. 2015, 142, 05015005. [CrossRef]
16. Ashrafi, S.M.; Mostaghimzadeh, E.; Adib, A. Applying wavelet transformation and artificial neural networks to develop forecasting-based reservoir operating rule curves. Hydrol. Sci. J. 2020, 65, 2007–2021. [CrossRef]
17. Nohara, D.; Nishioka, Y.; Hori, T.; Sato, Y. Real-Time Reservoir Operation for Flood Management Considering Ensemble Streamflow Prediction and Its Uncertainty. In Advances in Hydroinformatics; Springer: Singapore, 2016; pp. 333–347. [CrossRef]
18. Ashrafi, S.M. Investigating Pareto front extreme policies using a semi-distributed simulation model for Great Karun River Basin. J. Hydraul. Struct. 2019, 5, 75–88.
19. Hagan, M.T.; Behr, S.M. The Time Series Approach to Short Term Load Forecasting. IEEE Trans. Power Syst. 1987, 2, 785–791. [CrossRef]
20. Wolpert, D.H.; Macready, W.G. Efficient method to estimate Bagging’s generalization error. Mach. Learn. 1999, 35, 41–55. [CrossRef]
21. Polikar, R. The story of wavelets. In Physics and Modern Topics in Mechanical and Electrical Engineering; Mastorakis, N., Ed.; World Scientific and Engineering Society Press: Athens, Greece, 1999; pp. 192–197.
22. Mallat, S.G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans. Pattern Anal. Mach. Intell. 1989, 11, 674–693. [CrossRef]
23. Partal, T.; Earth, M.K. Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey). Phys. Chem. Earth Parts 2006, 31, 1189–1200. [CrossRef]
24. Kamruzzaman, M.; Metcalfe, A.V.; Beecham, S. Wavelet-Based Rainfall–Stream Flow Models for the Southeast Murray Darling Basin. J. Hydrol. Eng. 2014, 19, 1283–1293. [CrossRef]
25. Artiğas, M.Z.; Elias, A.G.; de Campra, P.F. Discrete wavelet analysis to assess long-term trends in geomagnetic activity. Phys. Chem. Earth Parts A/B/C 2006, 31, 77–80. [CrossRef]
26. Dash, M.; Liu, H. Feature selection for classification. Intell. Data Anal. 1997, 1, 131–156. [CrossRef]
27. Holland, J.H. Genetic algorithms and the optimal allocation of trials. SIAM J. Comput. 1973, 2, 88–105. [CrossRef]
28. McCulloch, W.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biol. 1990, 52, 99–115. [CrossRef]
29. Moazami, S.; Abdollahipour, S.; Zakeri Niri, S.; Ashrafi, S.A. Hydrological Assessment of Daily Satellite Precipitation Products over a Basin in Iran. J. Hydraul. Struct. 2016, 2, 35–45. [CrossRef]
30. Cortes, C.; Vapnik, V. Support-Vector Networks. Mach. Learning 1995, 20, 273–297. [CrossRef]
31. Yu, P.; Chen, S.; Chang, I.F. Support vector regression for real-time flood stage forecasting. J. Hydrol. 2006, 328, 704–716. [CrossRef]
32. Wu, M.C.; Lin, G.F.; Lin, H.Y. Improving the forecasts of extreme streamflow by support vector regression with the data extracted by self-organizing map. Hydrol. Processes 2014, 28, 386–397. [CrossRef]
33. Kisi, O.; Cimen, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. J. Hydrol. 2011, 399, 132–140. [CrossRef]
34. Azamathulla, H.; FC, W. Support vector machine approach for longitudinal dispersion coefficients in natural streams. Appl. Soft Comput. 2011, 11, 2902–2905. [CrossRef]
35. Hua, X.G.; Ni, Y.Q.; Ko, J.M.; Wang, K.Y. Modeling of Temperature–Frequency Correlation Using Combined Principal Component Analysis and Support Vector Regression Technique. J. Comput. Civ. Eng. Am. Soc. Civ. Eng. (ASCE) 2007, 21, 122–135. [CrossRef]
36. Araghinejad, S. Data-Driven Modeling: Using MATLAB® in Water Resources and Environmental Engineering; Springer Science & Business Media: Dordrecht, The Netherlands, 2016; Volume 67.
37. Nash, J.E.; Sutcliffe, J.V. River Flow Forecasting Through Conceptual Models Part I-a Discussion of Principles. J. Hydrol. 1970, 10, 282–290. [CrossRef]
38. Willmott, C.J. On the validation of models. Phys. Geogr. 1981, 2, 184–194. [CrossRef]
39. Shamseldin, A.A.; Press, S.J. Bayesian parameter and reliability estimation for a bivariate exponential distribution parallel sampling. J. Econom. 1984, 24, 363–378. [CrossRef]
40. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmer, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 1983, 50, 885–900. [CrossRef]
41. Reid, D. Genetic Algorithms in Constrained Optimization. Math. Comput. Model. 1996, 23, 87–111. [CrossRef]
42. Bolouri, Y. Effects of Extracting Multi-Reservoir Systemes Operational Rule Curves for Different Reservoirs in System Efficiency. Master’s Thesis, Faculty of Agriculture and Natural Resource, Soil and Water Department Tehran University, Karaj, Iran, 2011.
43. Estelman, L.; Schaffer, J. Real-coded genetic algorithms and interval-schemata. Found. Genet. Algorithms 1993, 2, 187–202.
44. Forsati, R.; Moayedikia, A.; Keikha, A. A Novel Approach for Feature Selection based on the Bee Colony Optimization. Int. J. Comput. Appl. 2017, 43, 975–8887. [CrossRef]
45. Li, Y.; Liang, Z.; Hu, Y.; Li, B.; Xu, B.; Wang, D. A multi-model integration method for monthly streamflow prediction: Modified stacking ensemble strategy. J. Hydroinform. 2020, 22, 310–326. [CrossRef]
46. Liu, P.; Li, L.; Chen, G.; Rheinheimer, D.E. Parameter uncertainty analysis of reservoir operating rules based on implicit stochastic optimization. J. Hydrol. 2014, 514, 102–113. [CrossRef]
47. Kalteh, A.M. Wavelet Genetic Algorithm-Support Vector Regression (Wavelet GA-SVR) for Monthly Flow Forecasting. Water Resour. Manag. 2015, 29, 1283–1293. [CrossRef]
48. Li, H.; Liu, P.; Guo, S.; Ming, B. Hybrid two-stage stochastic methods using scenario-based forecasts for reservoir refill operations. J. Water Resour. Plan. Manag. 2018, 144, 04018080. [CrossRef]
49. Tan, Q.F.; Wang, X.; Wang, H.; Wang, C.; Lei, X.H.; Xiong, Y.S.; Zhang, W. Derivation of optimal joint operating rules for multi-purpose multi-reservoir water-supply system. J. Hydrol. 2017, 551, 253–264. [CrossRef]