water

Article

Enhancing Runoff Simulation Using BTOP-LSTM Hybrid Model in the Shinano River Basin

Silang Nimai [1], Yufeng Ren [2,*], Tianqi Ao [1,*], Li Zhou [1,3], Hanxu Liang [1] and Yanmin Cui [1]

[1] State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource & Hydropower, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, China
[2] China Yangtze Power Co., Ltd., Yichang 443133, China
[3] Institute for Disaster Management and Reconstruction, Sichuan University, Chengdu 610065, China
* Correspondence: ren_yufeng@ctg.com.cn (Y.R.); aotianqi@scu.edu.cn (T.A.)

Abstract: Runoff simulation is an ongoing challenge in the field of hydrology. Process-based (PB) hydrological models often gain unsatisfactory simulation accuracy due to incomplete physical process representations. While the deep learning (DL) models demonstrate their capacity to grasp intricate hydrological response processes, they still face constraints pertaining to the representative training data and comprehensive hydrological observations. In order to provide unobservable hydrological variables from the PB model to the DL model, this study constructed hybrid models by feeding the output variables of the PB model (BTOP) into the DL model (LSTM) as additional input features. These variables underwent feature dimensionality reduction using the feature selection method (Pearson Correlation Coefficient, PCC) and the feature extraction method (Principal Component Analysis, PCA) before input into LSTM. The results showed that the standalone LSTM performed well across the basin, with *NSE* values all exceeding 0.70. The hybrid models enhanced the simulation performance of the standalone LSTM. The *NSE* values increased from 0.75 to nearly 0.80 in a sub-basin. Lastly, if the BTOP output is directly fed into LSTM without feature dimensionality reduction, the model's accuracy significantly decreases due to noise interference. The *NSE* value decreased by 0.09 compared to the standalone LSTM in a sub-basin. The results demonstrated the effectiveness of PCC and PCA in removing redundant information within hydrological variables. These findings provide new insights into incorporating physical information into LSTM and constructing hybrid models.

Keywords: runoff simulation; process-based model; deep learning; lstm; feature dimensionality reduction; Hybridization

check for updates

Citation: Nimai, S.; Ren, Y.; Ao, T.; Zhou, L.; Liang, H.; Cui, Y. Enhancing Runoff Simulation Using BTOP-LSTM Hybrid Model in the Shinano River Basin. *Water* **2023**, *15*, 3758. https://doi.org/10.3390/w15213758

Academic Editor: Athanasios Loukas

Received: 10 September 2023
Revised: 9 October 2023
Accepted: 24 October 2023
Published: 27 October 2023

1. Introduction

Runoff simulation is an essential way of exploring hydrological processes and the laws of the water cycle. An accurate runoff simulation can provide a vital scientific basis for predicting flow velocity, mitigating short-term flood risks and managing long-term water resource systems [1–6]. However, due to the complex dynamic process of rainfall–runoff, which has high spatiotemporal variability and is influenced by various factors such as precipitation [7], terrain [8], and climate characteristics [9], accurately predicting runoff is a long-lasting challenge and remains one of the important topics in the study of surface hydrological processes [10–12].

Process-based (PB) hydrological models are widely utilized for simulating runoff and have applications in water resource development, flood control, disaster reduction, and urban planning [13–15]. These models incorporate well-defined physical concepts and employ precise mathematical and physical equations to represent the diverse processes involved in water circulation, with model parameters having physical meanings [16,17]. When PB models accurately capture the fundamental runoff response process during the calibration period, indicating correct parameterization, they can reliably simulate runoff even outside the range of historical observations using the model's internal equations [18,19]. Nevertheless, inadequate representation of physical processes and limitations in input data can result in unreasonable parameter calibration, leading to significant discrepancies between the simulated and observed discharge [20,21]. This drawback greatly limits the development of PB models and is also an urgent pain point to be addressed in the application of PB models in hydrological simulation.

Due to the rapid growth in hydrological observations and advancements in computing power, deep learning (DL) models, particularly the Long Short-Term Memory network (LSTM), are gaining significant attention for hydrological simulations [22–26]. In contrast to traditional recurrent neural networks (RNN), LSTM addresses the challenges of long-term information retention and short-term input loss and can be used to process and predict important events with very long intervals and delays in time series [27]. Specifically, LSTM excels at capturing the lag effects of hydrological events such as infiltration [28], evapotranspiration [29], and snow melt [30]. At present, LSTM has demonstrated unparalleled accuracy in runoff modeling. Kratzert et al. applied LSTM to simulate runoff in 241 catchments of the CAMELS (Catchment Attributes and Meteorology for Large-sample Studies) dataset, and the results were superior to the Sacramento Soil Moisture Accounting Model (SAC-SMA) coupled with the Snow-17 snow routine [30]. Kratzert et al. further constructed an EA–LSTM model capable of learning catchment attributes, which significantly outperformed two regionally calibrated hydrological models (VIC and mHM) across 531 catchments in the CAMELS dataset [31]. Additionally, Tian et al. applied four DL models, the echo state network (ESN), the nonlinear autoregressive exogenous inputs neural network (NARX), Elman recurrent neural network (ERNN), and LSTM to simulate runoff in Xiangjiang and Qujiang River basins, and the findings revealed that LSTM outperformed the other models in capturing the dynamics of time series [32]. In addition, Xiang et al. proposed the LSTM–seq2seq model for estimating hourly runoff based on rainfall observations, rainfall forecasts, runoff observations, and empirical monthly evapotranspiration data from multiple stations in the Clear Creek and Upper Wapsipinicon River in Iowa [33]. The LSTM–seq2seq model demonstrated favorable performance in two watersheds, suggesting its potential for short-term flood forecasting applications. However, as a black-box model, the application of LSTM is somewhat limited due to its weak representation of physical processes despite its remarkable computational efficiency and simulation accuracy [12].

As a pure data-driven DL model, LSTM requires a certain amount of data to capture the dynamic process of hydrological series, and the simulation performance largely depends on whether the training period data can represent the hydrological response process of the basin. Unlike LSTM, which constructs a mapping between meteorological observations (precipitation, minimum and maximum temperature, solar radiation and vapor pressure, etc.) and runoff, PB models are calculated through parameterization based on the principle of mass conservation. In addition to simulating runoff, these models can generate variables in the form of time series from internal hydrological processes like evapotranspiration, vegetation interception, and soil moisture transport [34]. As these variables are computed using mathematical physical equations and encompass comprehensive information, an appealing strategy has emerged: incorporating output variables of PB models into LSTM by constructing a hybrid model. This strategy has been proven to be an effective means of enhancing simulation performance. Liu et al. utilized simulated runoff of PB model (PRMS) as input features other than meteorological observations to construct the PRMS–LSTM hybrid model. The results indicated that the hybrid model performs well and can improve the out-of-distribution prediction with acceptable generalization accuracy [35]. Konapala et al. constructed a hybrid model by feeding the simulated runoff of the SAC model with five meteorological indicators into LSTM and applied it to 531 catchments of the CAMELS dataset [36]. The results showed that the performance of the hybrid model improved compared to the standalone SAC and LSTM, and the improvement was most pronounced in the catchment where the PB model completely failed (i.e., $NSE < 0$).

Additionally, Wi et al. tested the ability of hybrid models to simulate runoff under warming scenarios and found that a hybrid LSTM, when provided with evapotranspiration estimates from the SAC–SMA model as an additional input feature, produced more realistic runoff predictions [37]. In the field of DL, excessive input features may provide duplicate and irrelevant information, thereby reducing model accuracy and computational efficiency. For this consideration, the hybrid model construction approach mentioned above only uses discharge or evapotranspiration estimates as additional inputs to LSTM, ignoring other internal variables with physical information. However, data dimensionality reduction is an effective way to eliminate redundant and irrelevant information, and it has been widely used in DL. Feature selection is a kind of dimensionality reduction technique that has been applied in multiple fields [38,39]. The purpose of feature selection is to identify the most characterizing features of the target variables so as to improve the learning progress of DL models [40]. Due to the ability to measure the dependency between the input features and the target variables, the Pearson Correlation Coefficient (PCC) method has become the most commonly used feature selection method [41,42]. Chen et al. used the PCC method to remove features that were not relevant to the predicted target (photovoltaic power) and input the remaining features into the LSTM for prediction [43]. The results indicated that this method can achieve short-term prediction of photovoltaic power and reduce the impact of noise. Xie et al. proposed a multivariate Long Short-Term Memory network model (MV-LSTM) based on PCC feature selection to predict wind speed [44]. The results show that the simulation performance of MV–LSTM is better than the ARMA model and the singe-variable LSTM. Unlike feature selection, feature extraction, as a data dimensionality reduction technique, is achieved by altering the original representation of features. Taking Principal Component Analysis (PCA) as an example, this method maps high-dimensional data into low-dimensional space through a certain linear projection, and expects the maximum variance of the data on the projected dimension [45,46]. Xu et al. used PCA to extract the main components and remove elusive components when applying LSTM to wind speed prediction, thus improving the simulation accuracy [47]. Zhang et al. used PCA to reduce the dimension of input features before applying LSTM and GRU for runoff prediction [48]. The results showed that the PCA method can retain the core information while reducing the data dimension and effectively improve the accuracy of the deep RNN models. In summary, both the feature selection method, PCC, and the feature extraction method, PCA, have great potential in information preservation and noise removal of input features of deep learning models.

Therefore, to make the best use of the physical information output by the PB model, an appealing strategy is to perform dimensionality reduction on all output variables of the PB model before feeding them into LSTM. In this study, we selected the distributed hydrological model, BTOP, as the PB model and considered all output variables of BTOP. We utilized two common data dimensionality reduction methods to eliminate duplicate and irrelevant information: the PCC method for feature selection and the PCA method for feature extraction. We incorporated the dimensionality-reduced feature from both methods as additional input features into LSTM to develop hybrid models for runoff simulation. We then evaluated their performance against the BTOP, standalone LSTM, and hybrid model without feature dimensionality reduction preprocessing. Therefore, the study aimed to achieve the following main objectives: (1) Compare the performance of the process-based hydrological model BTOP with the data-driven LSTM in runoff simulation; (2) Evaluate the improvement of hybrid models constructed using BTOP model estimation as additional input features in runoff simulation compared to the standalone LSTM; (3) Verify the ability of feature dimensionality reduction methods (PCC and PCA) to eliminate redundant information within hydrological sequences.

2. Materials

2.1. Study Area

The Shinano River, originating from the base of Mount Kobushi in Honshu, is recognized as Japan's longest river at 367 km and boasts the third largest catchment area, spanning 11,900 km^2 (Figure 1). This river primarily traverses through Nagano and Niigata prefectures before discharging into the Sea of Japan in Niigata. The upper part of the Shinano River Basin (SRB), positioned centrally on the Japanese mainland, is known as the Chikuma River, with a length of 214 km and an area of 7163 km^2, accounting for 58% and 60% of SRB, respectively. Characterized by its alpine terrain, the Chikuma River basin predominantly features a stratum composed of andesite. The climate is the most typical inland climate. The southern region displays predominant climatic influences from the East China Sea, while complex weather conditions govern the northern areas due to its proximity to the Northern Land Region. Varied annual precipitation intensities are observed: roughly 1000–1400 mm in the upper course, around 1000 mm in the middle, and approximately 1400–1800 mm in the lower course. The mid-lower sections of the SRB's precipitation patterns interrelate with the climatic characteristics of the Sea of Japan, with 40–50% of their annual rainfall largely snowfall, which tends to occur between November and February. The subsequent period of high precipitation transpires amidst the June to July "plum rain" season. The region's ample water supply and fertile soil enhance its recognition as one of Japan's premier paddy-producing regions. The Shinano River is a flood-prone river, with 23 recorded floods between 1931 and 1960. At the same time, it is also the primary water source for industry and agriculture in the basin. Therefore, combined with the needs of flood control, irrigation, water supply, and power generation, simulation runoff in SRB is of great interest.

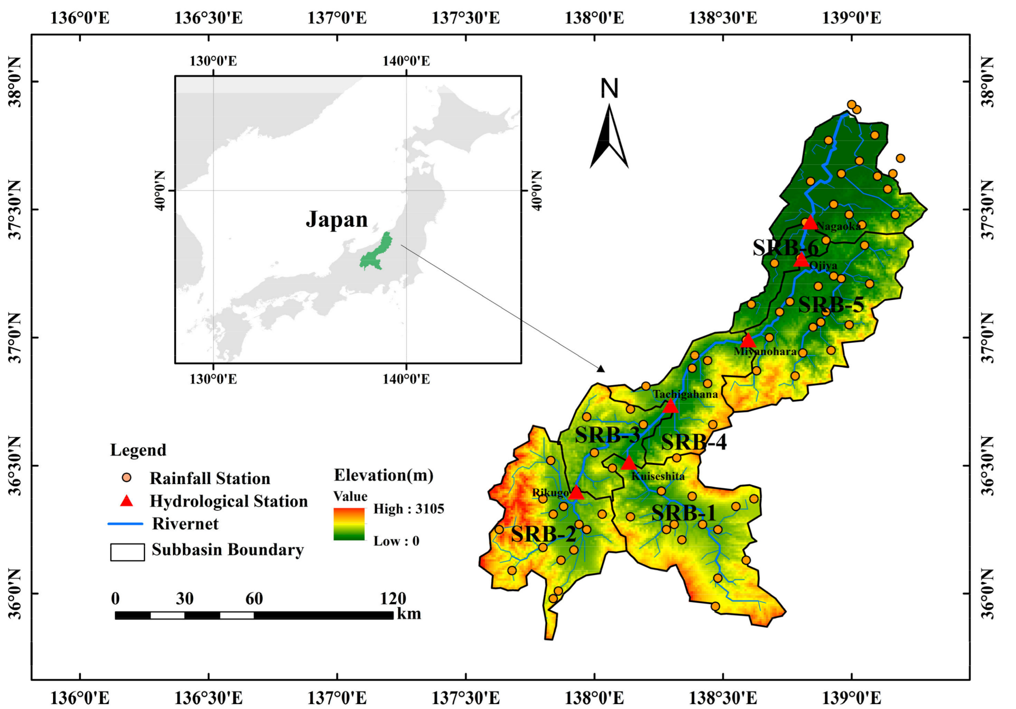


Figure 1. The location and topography of Shinano River Basin (SRB).

According to the location of hydrological stations and natural conditions, the BTOP model divides the entire basin into six sub-basins, named from SRB-1 to SRB-6 (Figure 1).

2.2. Data

The daily precipitation of seventy-seven rainfall stations and the daily discharge of six hydrological stations from 2002 to 2011 were collected from the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan. The daily actual evaporation data were obtained from the Global Land Evaporation Amsterdam Model (GLEAM) [49]. The above data could be directly used for runoff simulation by the standalone LSTM, while the following additional data are needed to drive the BTOP model: DEM, interception loss (PET_0), potential evapotranspiration (EP), leaf area vegetation index (LAI), soil and land use data. The DEM was sourced from MERIT-DEM (Multi-Error-Removed Improved Terrain DEM) data product developed by Yamazaki, D. et al. in Japan [50]. The PET with a 0.25° resolution was obtained from the GLEAM [49]. The EP came from the Climate Research Unit (CRU) with a resolution of 0.5° [51]. The LAI was downloaded from the National Environmental Information Center (NCEI) with monthly and 0.05° resolutions [52]. The Food and Agriculture Organization of the United Nations (FAO) provided the soil data [53]. The land use data obtained at a 500 m grid scale was acquired from the USGS Land Cover Research Institute [54]. Considering that such a large basin will cause huge calculation costs, the spatial resolution of the BTOP model was set to 1 km, and all input data were resampled to the same resolution using the nearest neighbor method.

3. Methodology

3.1. Evaluation Criteria

In order to quantitatively evaluate the overall performance of various models, five evaluation metrics, including mean absolute error (*MAE*), root means square error (*RMSE*), Pearson Correlation Coefficient (PCC), the Nash–Sutcliffe Efficiency (*NSE*) (Equations (1)–(4)) were selected in this study.

$$MAE = \frac{\sum_{i=1}^{n}\left|X_i^S - X_i^O\right|}{n} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i^S - X_i^O\right)^2}{n}} \tag{2}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(X_i^S - X_i^O\right)^2}{\sum_{i=1}^{n}\left(X_i^O - \overline{X^O}\right)^2} \tag{3}$$

$$PCC = \frac{\sum_{i=1}^{n}\left(X_i^O - \overline{X^O}\right)\left(X_i^S - \overline{X^S}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i^O - \overline{X^O}\right)^2}\sqrt{\sum_{i=1}^{n}\left(X_i^S - \overline{X^S}\right)^2}} \tag{4}$$

where X^O is the observed discharge data, X^S is the simulated discharge data by single BTOP, standalone LSTM and BTOP–LSTM hybrid models, i is the runoff time series, n is the time steps in the evaluation.

3.2. Hydrological Model: The Block-Wise Use of TOPMODEL (BTOP Model)

This study employed the BTOP model as the hydrological model due to its transparent physical mechanisms and ability to generate simulated internal variables. In order to extend the application scope of TOPMODEL from hillslope to macro-scale basin, the development of the distributed hydrological model BTOP started in 1999 at the University of Yamanashi, Japan [55,56]. The researchers preserved the basic equations of the model but altered the definition of the topographic index and introduced a new concept of mean groundwater travel distance [57]. The structure of the BTOP model mainly includes topographic preprocessing analysis (depression filling, topographic index calculation, subwatershed division, etc.), evapotranspiration calculation based on Shuttleworth–Wallace (S-W) equations [58], runoff generation based on TOPMODEL [59], and flow routing calculation based on Muskingum–Cunge equations [60]. The BTOP model has achieved extensive global application, such as the Malwathoya River Basin in Sri Lanka [61], the Mekong River Basin [62,63], the Fuji River Basin in Japan [64–66], and the Min River Basin in China [67].

The BTOP model, as a distributed model, introduces the concept of block-wise based on consideration of computational costs. The model consists of five parameters: block-average Manning coefficient (n_0c), drying function parameter (α), decay factor of lateral transmissivity (m), block-average saturation deficit (SD_{bar}), and groundwater dischargeability (D_0). These parameters are automatically calibrated using a shuffled complex evolution optimization algorithm (SCE-UA) developed at the University of Arizona [68]. Except for D_0, all other parameters are provided per block. In terms of output variables, beyond simulating essential values such as discharge, flow generation flux, and evapotranspiration for specific grids, BTOP can output the average values of internal variables upstream of a specified grid. Taking the calculation of net precipitation on a grid as an example, the BTOP model divides each grid into four zones: vegetation, root, unsaturated, and subsurface. For each BTOP grid, the net precipitation calculation occurs in the vegetation zone and is reduced by the canopy interception water storage and actual interception evaporation of the canopy (ET_0). The net precipitation of the basin is the average of the calculated values of all grids within the basin. A detailed listing of these internal variables is presented in Table 1.

Table 1. The summary of internal variables output by BTOP simulation.

Variable	Unit	Description
Effec.P	mm	Net precipitation
ET	mm	Actual evaporation
ET_0	mm	Actual interception evapotranspiration
SD	mm	The soil moisture saturation deficit
S_{rz}	mm	The root zone water storage
S_{uz}	mm	The unsaturated zone water storage
Q_{oft}	mm	Flow generation flux, the sum of simulated Hortonian overland flow flux, saturation excess runoff flux and groundwater flux
Q_v	mm	Groundwater recharge flux

3.3. Long Short-Term Memory Network (LSTM)

For a long time, latent variable models have faced problems of long-term information preservation and short-term input loss (owing to vanishing and exploding gradients). Inspired by the logic gate of computer systems, an early solution to mitigate these problems was introduced through Long Short-Term Memory (LSTM) [25,69]. LSTM is analog to standard recurrent neural networks, but each ordinary recurrent node is replaced by a memory cell [70]. Each memory cell is equipped with an internal state and multiple gates (the forget gate f_t, the input gate i_t and the output gate O_t) to determine when a hidden state should be updated or reset through a dedicated mechanism.

The data feeding into the LSTM cell is the input of the current time step (x_t) and the hidden state of the previous time step (h_{t-1}), as shown in Figure 2. Firstly, LSTM introduces the concept of candidate memory cell ($\widetilde{C}_t$), and uses *tanh* as the activation function with a value range of $(-1, 1)$. This leads to the following equation at time step t, where W_c is the weight parameter; b_c is the bias parameter:

$$\widetilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$$

Next, LSTM introduces three gates to control the information flow. Specifically, the forget gate determines whether to keep the current value of the memory or flush it, the input gate determines how much of the candidate memory cell's value should be added to the current memory cell state (C_t), the output gate determines whether the memory cell should influence the output at the current time step [71]. Three fully connected layers with sigmoid activation functions (σ) compute the values of the input, forget, and output gates.

The equations at time step t as follows, where W_f, W_i, W_o are weight parameters, b_f, b_i, b_o are bias parameters:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{6}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

$$O_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \tag{8}$$

Then, combining forget gate, update gate, previous memory cell state (C_{t-1}) and candidate memory cell state, the current memory cell internal state at time step t could be updated by Equation (9), where $\odot$ is Hadamard (elementwise) product operator:

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \tag{9}$$

At last, LSTM applies output gates and the current memory cell state to compute the output of the memory cell, which is called the hidden state (h_t). The equation at time step t is as follows:

$$h_t = O_t \odot tanh(C_t) \tag{10}$$

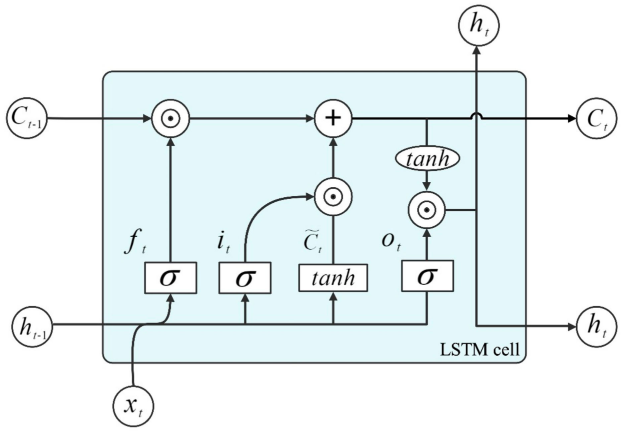


Figure 2. Basic structure of LSTM cell. The σ and *tanh* represent sigmoid and hyperbolic tangent activation functions, respectively. $\odot$ is Hadamard (elementwise) product operator. F_t, i_t, and O_t represent the forget gate, update gate, and output gate, respectively. x_t denotes the input data, C_t denotes the cell state, h_t denotes the hidden state, and $\widetilde{C}_t$ denotes the candidate memory cell state.

3.4. Feature Dimensionality Reduction Method

3.4.1. Pearson Correlation Coefficient (PCC)

Pearson Correlation Coefficient (PCC), a statistical measure introduced by Carl Pearson, quantifies the degree of linear relationship between research variables. Nowadays, in the field of machine learning, as a feature selection method of the feature dimension reduction branch, PCC is often applied to identify the most characterizing features of the target variable [40,41,72]. This method filters the features with the highest correlation and discards the rest to obtain the optimal subset of features. The mathematical representation for PCC is provided in Equation (4). The closer the absolute value of the correlation coefficient is to 1, the stronger the correlation is. The closer the correlation coefficient is to 0, the weaker the correlation is.

3.4.2. Principal Component Analysis (PCA)

Principal Component Analysis (PCA), initially introduced by Carl Pearson [73], is a prevalent feature extraction method utilized in machine learning. By employing a covariance matrix, PCA has the capacity to distill primary comprehensive variables from the original dataset, preserving integral information while discarding superfluous data [45,46].

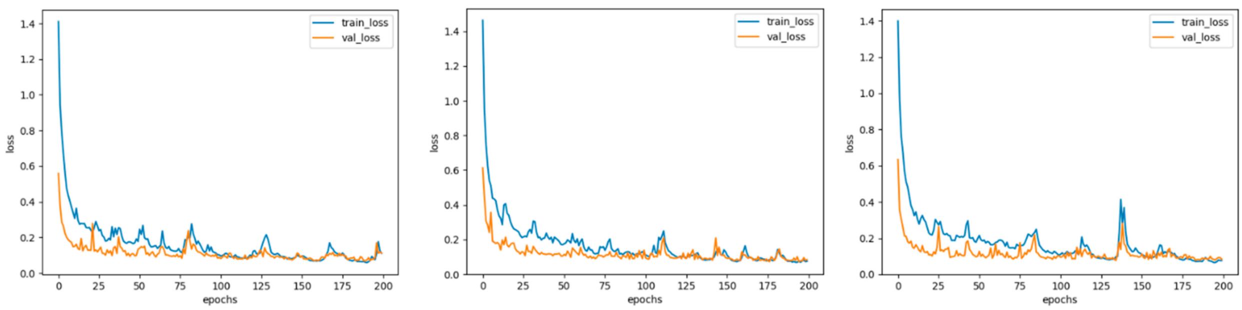
The objective of PCA is to map high-dimensional data into low-dimensional space through a certain linear projection, with the expectation of maximizing the amount of information (variance) in the projected dimension, thereby using fewer data dimensions to retain the characteristics of original data points. In principal component analysis, feature vectors corresponding to the largest eigenvalues are selected, and the data are mapped in the reference frame composed of these eigenvectors to achieve the purpose of dimensionality reduction. More detailed explanations are referred to Hotelling, H [74].

3.5. Model Design

In the simulation of the BTOP model, the calibration and validation periods were designated as 2002–2008 and 2009–2011, respectively. In the simulation of LSTM, 2002–2007 was chosen as the training period, 2008 was chosen as the validation period (to verify the rationality of model hyperparameters), and 2009–2011 was chosen as the testing period. All models employed equal lengths of training periods, and their performance comparison occurred during the same testing period.

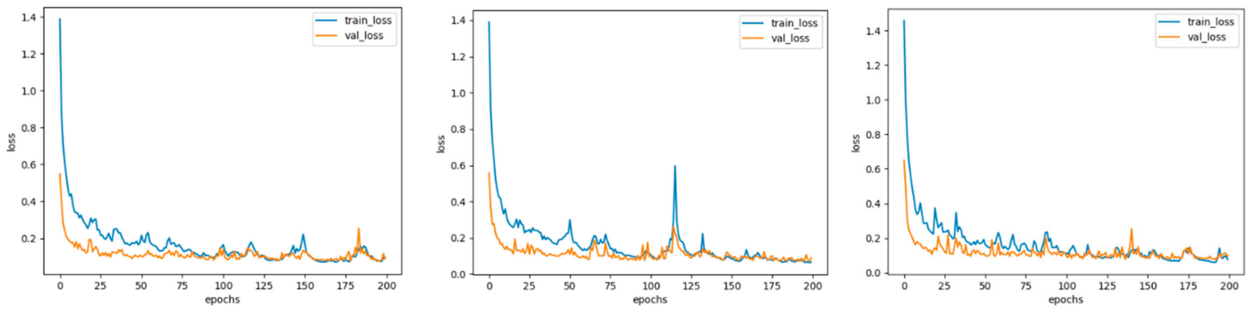
3.5.1. Standalone LSTM Settings

The standalone LSTM was built using the Keras library in Python 3.8 based on a two-layer structure of 40 hidden nodes per layer. To avoid overfitting and underfitting, we performed several preliminary simulations for optimal tuning of hyperparameters. To capture the hydrological dynamic response of SRB, we set the input step to 30 days, and the input features were composed of 3 variables, including actual precipitation, actual evaporation, and historical observed discharge. The actual precipitation was the surface average precipitation calculated using the Thiessen Polygon method based on the gauged precipitation data, while the actual evaporation was the surface average value calculated based on the gridded satellite product. For efficient and more stable learning, all input features as well as the output data, were normalized by subtracting the mean and dividing by the standard deviation before input into the model [30]. The batch size and the number of iteration steps (epochs) were set to 128 and 200, respectively. To enhance the model's generalization capability, we applied the L2 regularization method (set to 0.01), thus preventing overfitting of the training period data via penalizing larger weights [75]. Additionally, Adam was selected as our optimization algorithm with a learning rate of 0.005. Finally, we used the mean-squared error (*MSE*) as an objective criterion in this study. The mean-squared error reduction curves of the standalone model during the training and validation period at 6 sub-basins are shown in Figure 3. The model exhibited low bias during the training periods and low variance during the validation periods, which manifested that the hyperparameter settings were reasonable. There were no overfitting or underfitting phenomena throughout the entire process, indicating that the standard LSTM model effectively captured the rainfall–runoff processes in the Shinano River Basin and had strong generalization ability. The model was executed ten times on each sub-basin, and the average value was taken as the result.



(**a**) (**b**) (**c**)

Figure 3. *Cont.*



(**d**) (**e**) (**f**)

Figure 3. Loss function process of the standalone LSTM at 6 sub-basins. Train_loss: mean-squared error during the training process; val_loss: mean-squared error during the validation process; epochs: iteration steps; loss: mean-squared error. (**a**) SRB-1; (**b**) SRB-2; (**c**) SRB-3; (**d**) SRB-4; (**e**) SRB-5; (**f**) SRB-6.

3.5.2. Construction of Hybrid Models

After a runoff simulation of the BTOP model, in addition to the simulated discharge, the model will output eight other information-rich internal variables, as depicted in Table 1. These variables, similar to the observed precipitation and evaporation, can provide useful information for the runoff simulation of LSTM. In the field of DL, providing more input features tends to improve model accuracy. However, when superabundant features are fed into the model, the noise and redundant information will diminish model accuracy and greatly reduce the computational efficiency. Therefore, using data dimensionality reduction operation to preprocess input features is the mainstream method. In this study, two fundamental strategies were employed to construct Hybrid-1 and Hybrid-2, while Hybrid-3 was established as a benchmark model:

1. Hybrid-1: Pearson correlation analysis, a common feature selection method, was employed to screen the feature variables demonstrating high correlation with the observed discharge as additional input to the standalone LSTM.
2. Hybrid-2: Employing principal component analysis as a feature dimensionality reduction method, converted eight input features into a few principal components and fed them as additional features into the standalone LSTM.
3. Hybrid-3: All output variables from the BTOP model were directly fed into the standalone LSTM as additional features without any dimensionality reduction.
4. The hybrid model adopted the same model structure as the standalone LSTM (avoiding overfitting and underfitting), with variations in input variables only. The hybrid models were executed ten times on each sub-basin, and the average value was taken.

The framework of this study is shown in Figure 4.

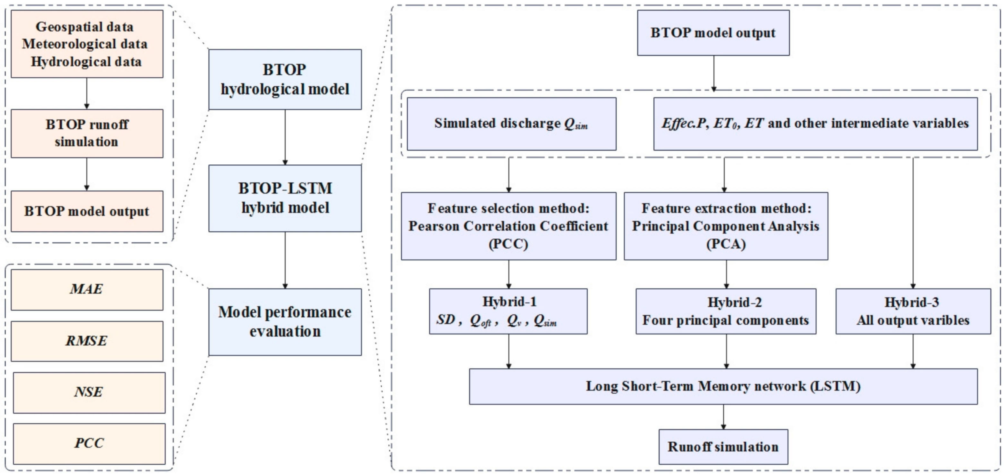


Figure 4. The framework of this study.

4. Results and Discussion

4.1. BTOP Model Simulation

Table 2 presents the evaluation criteria of the BTOP simulation at six sub-basins. The smaller the values of *MAE* and *RMSE*, and the closer the values of *NSE* and *PCC* are to 1, the better the simulation performance. The results indicated that the BTOP model performed well apart from two sub-basins (SRB-5 and SRB-6) close to the outlet of the SRB. For example, satisfying efficiency was gained at SRB-1 with *MAE* of 19.62 m^3/s, *RMSE* of 37.54 m^3/s, *NSE* of 0.77 and *PCC* of 0.89 for calibration, in addition to 19.01 m^3/s, 31.19 m^3/s, 0.63 and 0.82 for validation, respectively. The poor simulation performance of SRB-5 and SRB-6 could potentially be attributed to the unreasonable sub-basin division within the BTOP model. In addition, since the flood in the lower courses of SRB is largely affected by snowmelt from March to April, the BTOP model significantly underestimated the discharge during flood periods in SRB-5 and SRB-6 due to the lower observed rainfall. Moreover, the Shinano River is the main source of water for industry and agriculture in the basin, and more than half of the population in the basin depends on the Shinano River for water supply. In combination with the needs of flood control, irrigation, water supply and power generation, numerous large comprehensive utilization projects were built in the lower courses of the basin (Niigata has six flood control reservoirs with a total storage capacity of nearly 100 million cubic meters). Therefore, human activities such as reservoir storage and irrigation may also lead to the poor simulation performance of the BTOP model in SRB-5 and SRB-6.

Table 2. Evaluation criteria for BTOP runoff simulations.

Sub-Basin	Calibration Period				Validation Period			
	MAE (m^3/s)	*RMSE* (m^3/s)	*NSE*	*PCC*	*MAE* (m^3/s)	*RMSE* (m^3/s)	*NSE*	*PCC*
SRB-1	19.62	37.54	0.77	0.89	19.01	31.19	0.63	0.82
SRB-2	37.15	64.72	0.56	0.81	33.79	53.07	0.66	0.88
SRB-3	63.79	103.09	0.70	0.89	62.34	94.72	0.66	0.89
SRB-4	125.31	155.40	0.65	0.85	125.34	155.81	0.55	0.82
SRB-5	205.73	324.06	0.14	0.54	199.18	340.60	0.15	0.59
SRB-6	210.34	338.52	0.19	0.54	192.25	333.84	0.17	0.59

It is worth mentioning that regardless of the simulation performance, the internal variables of all sub-basins were calculated using the same mathematical equation under the basic structure of the model, following the basic law of mass conservation. Therefore, in order to investigate whether the output variables of the BTOP model follow basic physical laws and cover physical information, we chose the internal variable of SRB-1 for analysis, while the other five sub-basins have the same conclusion. Referring to SRB-1 as an example, after a runoff simulation, the BTOP model outputted variables in the form of a time series (Figure 5), including the variables listed in Table 1 and simulated discharge (Q_{sim}). In the natural world, affected by factors such as temperature, humidity, atmospheric pressure, and wind, both precipitation and evapotranspiration exhibit significant seasonal and inter-annual variation characteristics, which were also reflected in BTOP model calculations (Figure 5a–c). As a process-based hydrologic model, the parameters of the BTOP model have clear physical significance and can describe the hydrologic process more accurately by solving continuous and dynamic equations. For instance, after a rainfall event (refer to Figure 5a), there was a significant decrease in soil moisture saturation deficit (Figure 5d). Moreover, since the root zone is located above the unsaturated zone in the BTOP model structure, the decrease in the root zone water storage (Figure 5e) was usually accompanied by a surge in unsaturated zone water storage (Figure 5f). On the other hand, as shown in Figure 5g,i, the comparison between Q_{oft} and Q_{sim} confirms that the simulated discharge was primarily determined by flow generation flux. In conclusion, the output variables of the BTOP model, derived via mass and energy equations, potentially offered informative physical data.

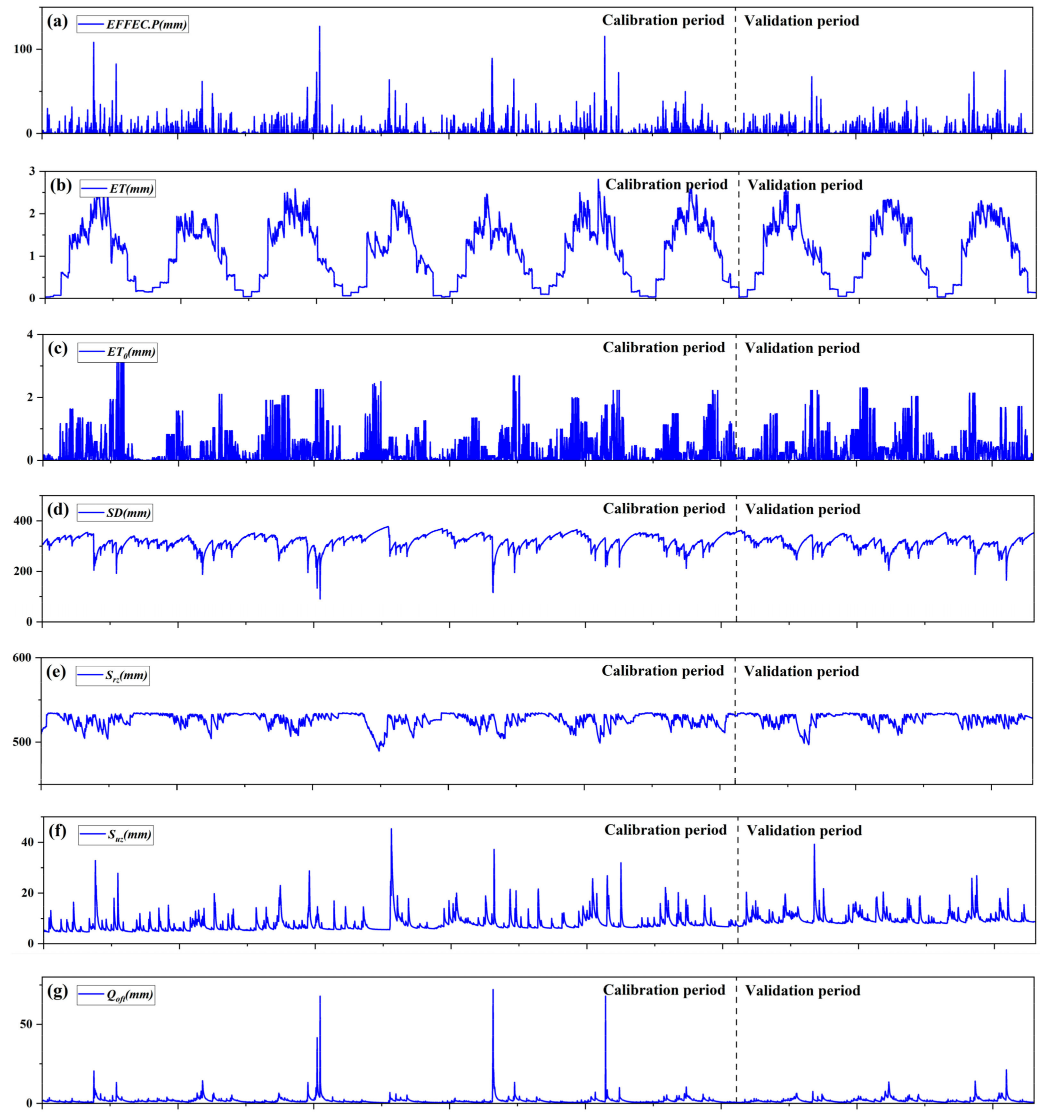


Figure 5. *Cont.*

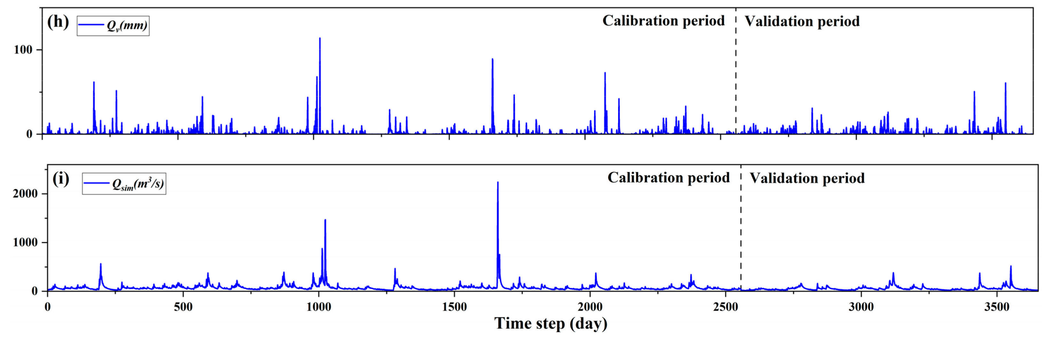


Figure 5. Variables output by BTOP at SRB-1. (**a**) Is the net precipitation (*Effec.P*), (**b**) is the actual evaporation (*ET*), (**c**) is the actual interception evapotranspiration (ET_0), (**d**) is the soil moisture saturation deficit (*SD*), (**e**) is the root zone water storage (S_{rz}), (**f**) is the unsaturated zone water storage (S_{uz}), (**g**) is the flow generation flux (Q_{oft}), (**h**) is the groundwater recharge flux (Q_v), (**i**) is the simulated discharge (Q_{sim}).

4.2. Feature Dimensionality Reduction Results

4.2.1. Feature Selection (PCC)

The rainfall–runoff process has a significant hysteresis effect due to factors such as basin size, soils, geology, slope, land use, precipitation amount and duration, the timing of peak rainfall intensity, and antecedent precipitation [76,77]. Therefore, we explored the correlation between the output variables and the observed discharge, considering various lag times, as shown in Table 3. Variables like net precipitation (*Effec.P*), soil moisture saturation deficit (*SD*), flow generation flux (Q_{oft}), and groundwater discharge flux (Q_v) output from the BTOP model demonstrated robust correlations with observed discharge (Q_{obs}) over three different lag times as they reached their peak when the lag time was one day. *Effec.P*, Q_{oft}, and Q_v displayed a significant positive correlation with Q_{obs}, while the *SD* showed a significant negative correlation with Q_{obs}. Remarkably, the relationship of actual evaporation (*ET*), actual interception evapotranspiration (ET_0), root zone water storage (S_{rz}), and unsaturated zone water storage (S_{uz}) to Q_{obs} was somewhat weak but typically peaked at a one-day lag time. Therefore, this study determined the lag time as one day in the Shinano River Basin.

In order to perform feature selection, we conducted a correlation analysis between various variables, including internal variables, simulated discharge, and observed discharge with a one-day lag. As shown in Figure 6, *Effec.P*, Q_{oft}, Q_v, and Q_{sim} variables present high correlations among themselves and a clear negative correlation with *SD*. Moreover, due to the model design, actual evapotranspiration mainly occurs in the root zone. Therefore, *ET* calculated by the model showed a significant negative correlation with S_{rz}. Feature selection not only removes features unrelated to the training target but also removes a redundant feature that provides the same information. In this study, Q_{obs} was mainly associated with five variables: *Effec.P*, *SD*, Q_{oft}, Q_v and Q_{sim}. However, the two variables may provide redundant information due to the strong correlation between *Effec.P* and Q_v (0.91 in all six sub-basins). Therefore, based on the consideration of a simplified model, the input features *SD*, Q_{oft}, Q_v, and Q_{sim} determined by the PCC feature selection method were selected to construct Hybrid-1.

4.2.2. Feature Extraction (PCA)

This study employed principal component analysis to eliminate the correlation between input variables and achieve feature dimensionality reduction. Table 4 shows the eigenvalues and cumulative variance contribution rates of the correlation coefficient matrix. Eight internal variables (indexed in Table 1) and simulated discharge output from the BTOP model could be expressed as nine uncorrelated principal components by linear expression. The fundamental of PCA is to extract principal components with eigenvalues greater than one and cumulative variance contribution rates of about 85%. This principle can ensure that the extracted principal components have a strong explanatory ability. In each sub-basin, the eigenvalues of the first four principal components were all greater than 1, and the cumulative variance contribution rates reached around 85%. Therefore, based on the fundamentals, in order to preserve the original data's main information and avoid data redundancy, the first four principal components were selected as extracted features to construct Hybrid-2 in this study. The four features extracted from nine out variables of SRB-1 are shown in Figure 7; the input features of other sub-basins were extracted using the same method.

Table 3. Correlation coefficient between output variables of BTOP model and observed discharge under different lag times (* indicates that the variable is related to observed discharge with a *p*-value of less than 5%, ** indicates that the variable is related to observed discharge with a *p*-value of less than 1%).

Sub-Basin	Lag Time (day)	*Effec.P*	*ET*	ET_0	*SD*	S_{rz}	S_{uz}	Q_{oft}	Q_v
SRB-1	0	0.43 **	0.19 **	0.19 **	−0.58 **	0.17 **	0.43 **	0.71 **	0.60 **
	1	0.69 **	0.18 **	0.28 **	−0.58 **	0.17 **	0.47 **	0.78 **	0.80 **
	2	0.37 **	0.16 **	0.20 **	−0.38 **	0.11 **	0.34 **	0.37 **	0.39 **
SRB-2	0	0.41 **	0.35 **	0.24 **	−0.68 **	0.05 *	−0.11 *	0.72 **	0.54 **
	1	0.55 **	0.35 **	0.30 **	−0.67 **	0.04 *	0.11 **	0.76 **	0.67 **
	2	0.35 **	0.22 **	0.33 **	−0.54 **	0.01	−0.12 **	0.51 **	0.42 **
SRB-3	0	0.36 **	0.28 **	0.20 **	−0.64 **	0.14 **	0.04	0.67 **	0.51 **
	1	0. 69 **	0.28 **	0.33 **	−0.67 **	0.14 **	0.04 *	0.86 **	0.82 **
	2	0.41 **	0.26 **	0.25 **	−0.49 **	0.09 **	0.02	0.45 **	0.46 **
SRB-4	0	0.24 **	0.22 **	0.17 **	−0.55 **	0.15 **	0.03	0.57 **	0.41 **
	1	0.66 **	0.22 **	0.31 **	−0.62 **	0.16 **	0.04	0.85 **	0.78 **
	2	0.43 **	0.20 **	0.26 **	−0.45 **	0.11 **	0.01	0.55 **	0.50 **
SRB-5	0	0.20 **	0.18 **	0.13 **	−0.23 **	0.16 **	−0.01	0.36 **	0.31 **
	1	0.51 **	0.17 **	0.25 **	−0.25 **	0.16 **	−0.01	0.56 **	0.56 **
	2	0.30 **	0.15 **	0.20 **	−0.15 **	0.12 **	−0.04	0.32 **	0.32 **
SRB-6	0	0.19 **	0.12 **	0.11 **	−0.25 **	0.19 **	0.02	0.36 **	0.30 **
	1	0.50 **	0.11 **	0.23 **	−0.27 **	0.19 **	0.03	0.56 **	0.56 *
	2	0.30 **	0.09 **	0.18 **	−0.18 **	0.15 **	−0.01	0.34 **	0.33 **

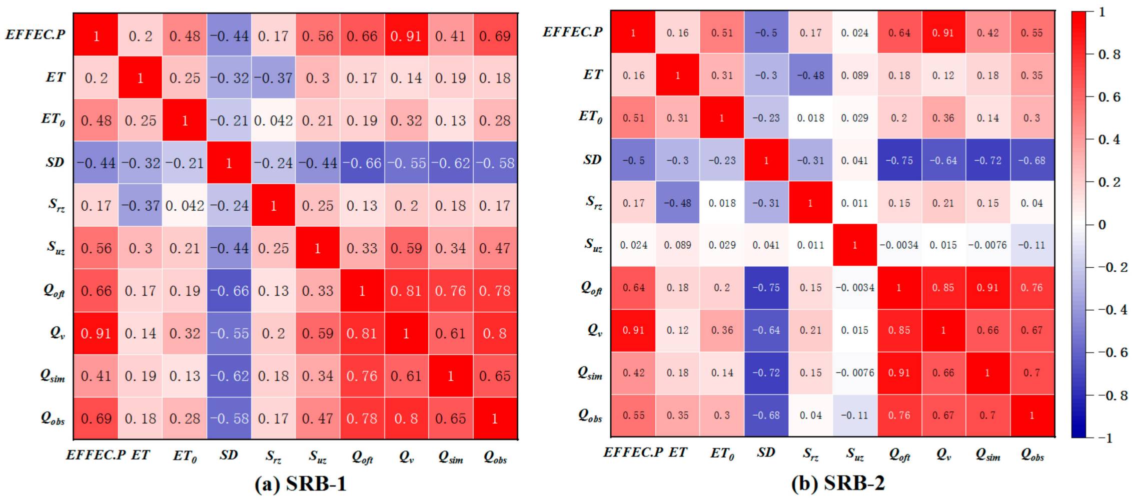


Figure 6. *Cont.*

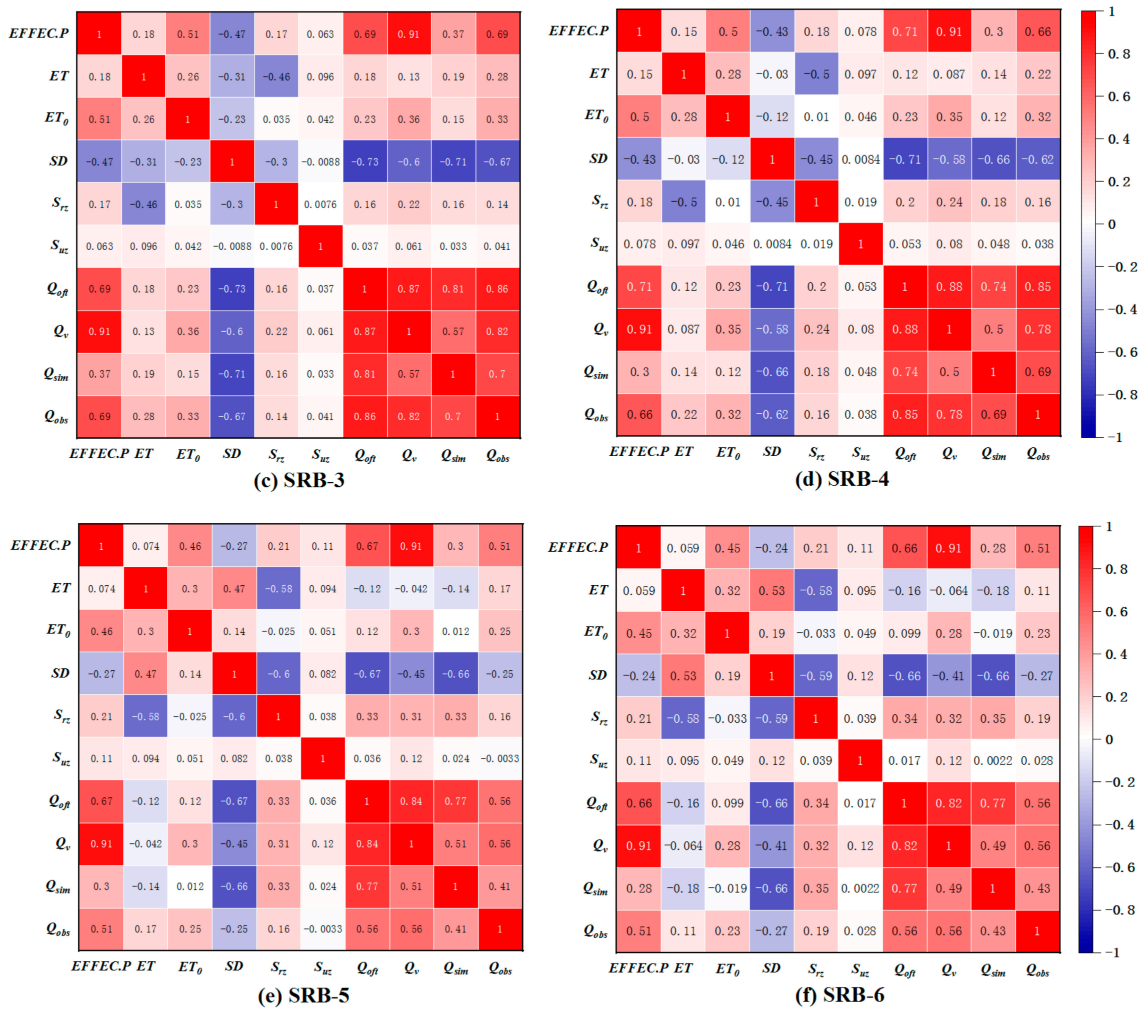


Figure 6. Correlation between variables with a lag time of one day. *EFFEC.P*: the net precipitation, *ET*: the actual evaporation, ET_0: the actual interception evapotranspiration, *SD*: the soil moisture saturation deficit, S_{rz}: the root zone water storage, S_{uz}: the unsaturated zone water storage, Q_{oft}: the flow generation flux, Q_v: the groundwater recharge flux, Q_{sim}: the simulated discharge, Q_{obs}: the observed discharge. (**a**) SRB-1, (**b**) SRB-2, (**c**) SRB-3, (**d**) SRB-4, (**e**) SRB-5, (**f**) SRB-6.

4.3. Comparison of Performance between Varied Models

After reducing the feature dimensionality, three hybrid models were constructed according to the method outlined in Section 3.5.2. The performance of these models was compared against that of the BTOP model and the standalone LSTM during the testing period. Figure 8 depicts the simulated and observed discharge from five models in six sub-basins. Table 5 summarizes the specific evaluation criteria.

The results indicated that, across all sub-basins, both the data-driven LSTM and hybrid models were superior to the process-based distributed hydrological model BTOP. In the two sub-basins where BTOP simulation failed (SRB-5 and SRB-6), LSTM and the three hybrid models still achieved significant accuracy, denoted by *NSE* values approximating 0.80. From this, it is evident that even with less data used (without inputting terrain attributes, hydrological and meteorological station locations, etc.), LSTM-based models outperform distributed hydrological model BTOP in terms of accuracy in runoff simulation.

Table 4. The eigenvalues and cumulative variance contribution rates of principal components of 6 sub-basins.

	SRB-1		SRB-2		SRB-3	
Ingredient	**Eigenvalues**	**Cumulative (%)**	**Eigenvalues**	**Cumulative (%)**	**Eigenvalues**	**Cumulative (%)**
1	4.207	46.740	4.146	46.063	4.059	45.102
2	1.395	62.240	1.534	63.107	1.485	61.598
3	1.138	74.882	1.102	75.353	1.122	74.068
4	1.002	83.873	1.001	86.445	1.003	85.143
5	0.560	91.172	0.596	93.102	0.621	92.048
6	0.297	94.773	0.378	97.303	0.407	96.635
7	0.227	98.069	0.169	99.178	0.190	98.743
8	0.130	99.516	0.049	99.720	0.077	99.604
9	0.044	100.000	0.025	100.000	0.036	100.000
	SRB-4		**SRB-5**		**SRB-6**	
Ingredient	**Eigenvalues**	**Cumulative (%)**	**Eigenvalues**	**Cumulative (%)**	**Eigenvalues**	**Cumulative (%)**
1	3.918	43.537	3.413	42.368	3.487	42.078
2	1.592	61.222	1.955	64.086	2.005	64.352
3	1.139	73.879	1.672	75.085	1.554	75.361
4	1.007	84.885	1.004	85.857	1.072	85.897
5	0.605	91.798	0.416	92.239	0.381	92.285
6	0.395	96.184	0.236	95.905	0.235	95.916
7	0.224	98.674	0.172	98.531	0.133	98.528
8	0.084	99.603	0.092	99.549	0.092	99.546
9	0.036	100.000	0.041	100.000	0.041	100.000

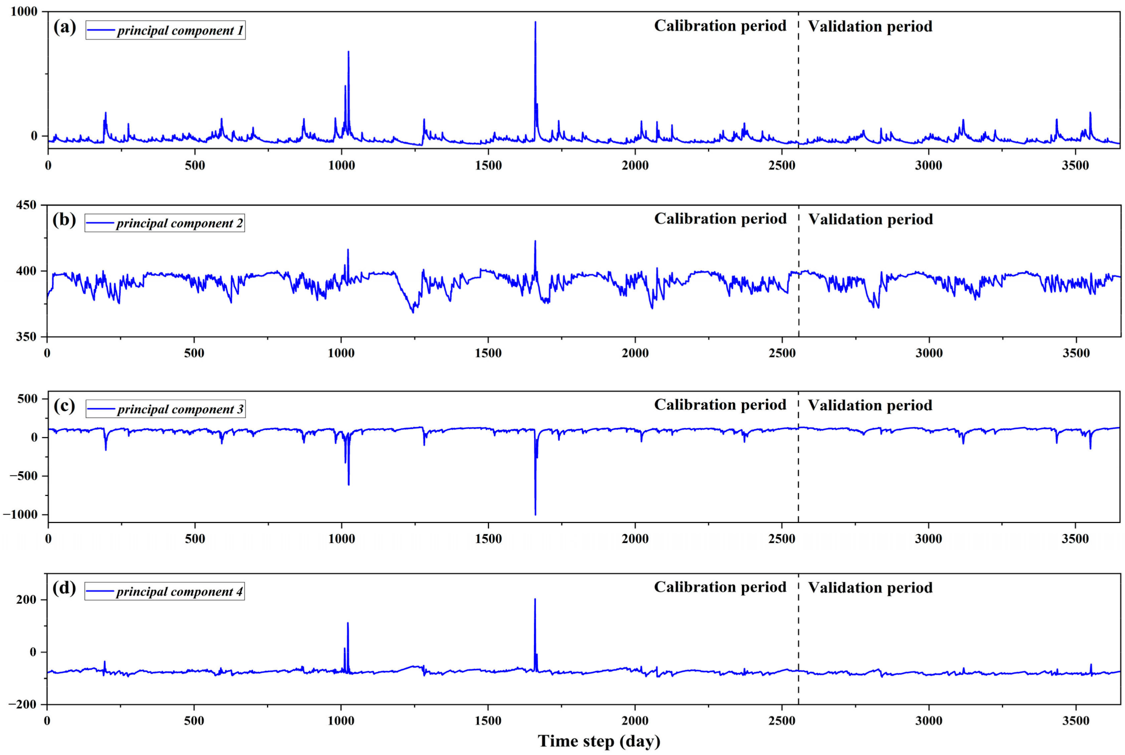


Figure 7. Four features extracted using the PCA method in SRB-1.

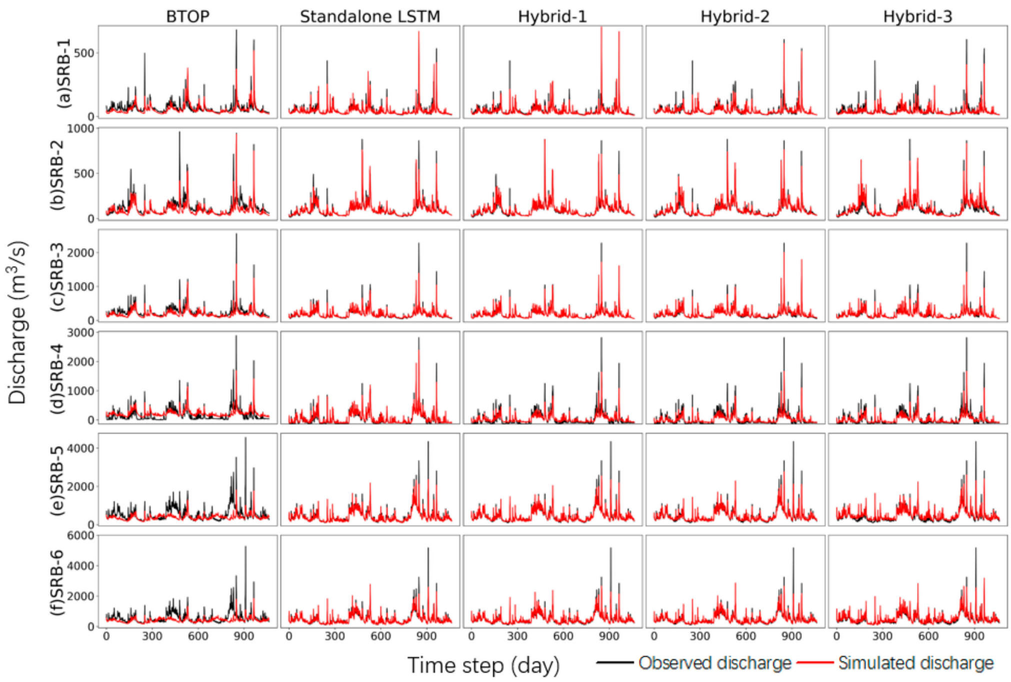


Figure 8. Comparison of runoff simulation at six sub-basins by BTOP, Standalone LSTM, Hybrid-1, Hybrid-2, and Hybrid-3.

In addition, the simulation performance of both Hybrid-1 and Hybrid-2 surpassed that of the standalone LSTM, as they incorporate additional input features from the dimensionality-reduced BTOP output variables. Especially in the case of SRB-6, the *MAE* value decreased from 81.87 m^3/s to 75.66 m^3/s and 81.41 m^3/s, the *RMSE* value reduced from 184.85 m^3/s to 167.10 m^3/s and 168.24 m^3/s, the *NSE* value increased from 0.75 to 0.80 and 0.79, and the PCC value rose from 0.88 to 0.90. These results underlined the superior simulation performance achieved by Hybrid-1 and Hybrid-2, indicating that the output variables of the BTOP model can provide more physical information to the standalone LSTM, thereby improving simulation accuracy.

More input features do not guarantee higher simulation accuracy of LSTM because redundant and irrelevant features will interfere with the LSTM learning process. When utilizing all variables provided by the BTOP model as additional input features, Hybrid-3 manifested a substantial decrease in simulation accuracy across nearly all sub-basins compared to the standalone LSTM. Illustratively, the *MAE* value increased by 4.56 m^3/s, the *RMSE* value increased by 4.86 m^3/s, the *NSE* value decreased by 0.09, and the PCC value decreased by 0.01 at SRB-1. The results suggested that employing feature selection methods such as PCC and feature extraction methods like PCA can significantly reduce redundant and irrelevant information within hydrological sequences. Filtering the output variables of the BTOP model through two feature dimensionality reduction methods and subsequently incorporating these variables into the input features of the standalone LSTM can effectively improve the performance of runoff simulation.

It is worth noting that in the sub-basins where the BTOP simulation was not satisfied (SRB-5 and SRB-6), the output variables of the model differ significantly from the true values but still provide useful physical information for LSTM. This conclusion is similar to that of Yu et al. and Konapala et al. [36,78]. This may be due to the fact that in the PB model simulation, even if the simulation performance is poor, the results of calibration and validation periods are calculated using the same set of parameters and belong to the same distribution. This is consistent with the basic rules of data during training, validation, and testing in deep learning. Instead of learning the measured data, using the output of the PB model as an additional input feature to the LSTM allows the LSTM to learn the data simulated by the PB model (the results of the PB model calculations). Therefore, even if the PB model simulation accuracy is low, as long as the simulated data follows the same distribution during the calibration and validation periods, it can provide physical information to the LSTM. The exceedance probability was utilized to examine the ability of varied models to capture the flow of different magnitudes. Dividing the flow duration curve (FDC) into three segments for analysis, including the high-flow segment (probability less than 5%), the mid-flow segment (probability between 20% and 70%), and the low-flow segment (probability greater than 70%) [79].

Table 5. Evaluation criteria for discharge simulations of varied models.

Sub-Basin Number	Metrics	BTOP	Standalone LSTM	Hybrid-1	Hybrid-2	Hybrid-3
SRB-1	*MAE* (m^3/s)	18.95	10.50	11.96	10.50	15.06
	RMSE (m^3/s)	31.43	26.92	25.95	25.01	31.78
	NSE	0.63	0.73	0.75	0.76	0.64
	PCC	0.83	0.86	0.88	0.89	0.85
SRB-2	*MAE* (m^3/s)	34.59	18.80	18.38	19.17	26.01
	RMSE (m^3/s)	53.82	46.02	40.80	41.62	45.52
	NSE	0.66	0.75	0.80	0.80	0.76
	PCC	0.88	0.87	0.90	0.90	0.89
SRB-3	*MAE* (m^3/s)	63.29	32.75	29.75	36.74	36.23
	RMSE (m^3/s)	95.95	68.19	61.84	61.07	74.19
	NSE	0.66	0.83	0.86	0.86	0.80
	PCC	0.90	0.92	0.93	0.94	0.90
SRB-4	*MAE* (m^3/s)	126.09	48.51	37.82	44.17	54.02
	RMSE (m^3/s)	157.14	77.96	69.60	72.58	84.39
	NSE	0.55	0.89	0.91	0.90	0.87
	PCC	0.81	0.95	0.96	0.95	0.94
SRB-5	*MAE* (m^3/s)	203.60	71.50	67.43	73.54	103.70
	RMSE (m^3/s)	345.39	157.43	152.89	151.74	165.59
	NSE	0.15	0.82	0.83	0.84	0.80
	PCC	0.62	0.92	0.92	0.92	0.91
SRB-6	*MAE* (m^3/s)	196.34	81.87	75.66	81.41	112.22
	RMSE (m^3/s)	338.47	184.85	167.10	168.24	196.35
	NSE	0.17	0.75	0.80	0.79	0.72
	PCC	0.62	0.88	0.90	0.90	0.88

As shown in Figure 9, the BTOP model significantly underestimated the flow during the whole period at all sub-basins except SRB-4. At SRB-4, the BTOP model tended to underestimate the high-flow and overestimate both the mid-flow and low-flow, as depicted in Figure 9d. Due to the impact of snowmelt, the BTOP model greatly underestimated the runoff during flood periods in SRB-5 and SRB-6 downstream of the basin. The standalone LSTM significantly underestimated the high flow of SRB-1, indicated in Figure 9a, mildly overestimated the mid-flow of SRB-4 (Figure 9d), and slightly underestimated the mid-flow of SRB-5 (Figure 9e). Except for the above, the FDC from the standalone LSTM simulated results matched well with the observed curve in the entire probability.

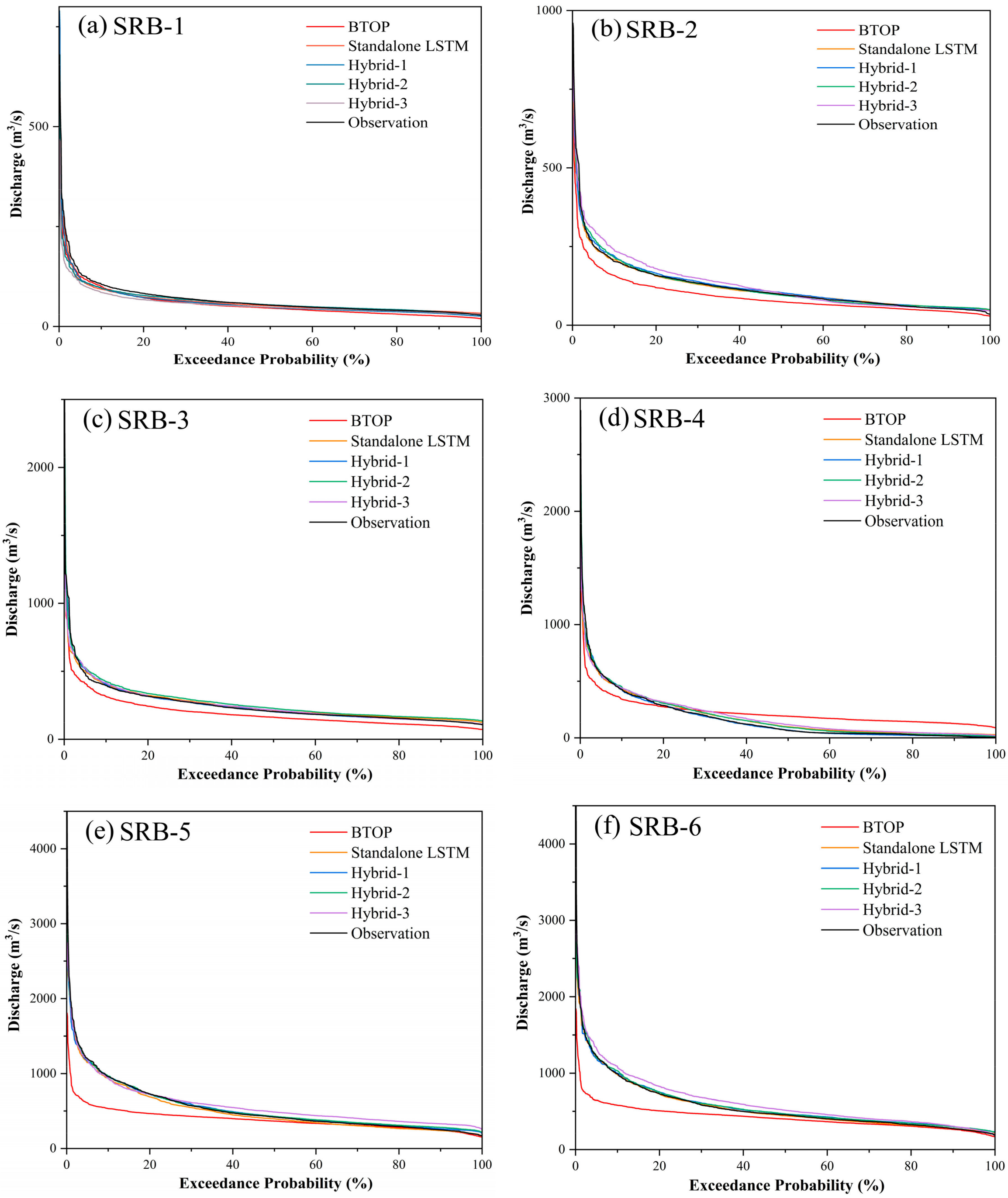


Figure 9. Comparison of the exceedance probabilities of the observed discharge and simulated discharge from varied models at six sub-basins. (**a**) SRB-1, (**b**) SRB-2, (**c**) SRB-3, (**d**) SRB-4, (**e**) SRB-5, (**f**) SRB-6.

Hybrid-1 and Hybrid-2 were in line with the observed discharge except for a slight underestimation of high flow in SRB-1 (Figure 9a) and a slight overestimation of mid-flow in SRB-4 (Figure 9d). Moreover, the Hybrid-3 showed a significant underestimation of high flow in SRB-1 and an overestimation of mid-flow in SRB-2, SRB-4, SRB-5, and SRB-6 (Figure 9b,d–f).

Compared to the BTOP model and the standard LSTM, Hybrid-1 and Hybrid-2 can better fit the observed runoff hydrograph and flow duration curve, indicating a better simulation performance. The primary source of error within the hybrid model stems from the simulation of high-flow periods, which is influenced by whether similar high-flow events appear during the training period.

5. Conclusions

This study applied a process-based BTOP model and data-driven LSTM to simulate runoff in six sub-basins of the Shinano River Basin in Japan. Subsequently, two hybrid models, referred to as Hybrid-1 and Hybrid-2, were constructed by integrating output variables of the BTOP model into the LSTM as input features after implementing two feature dimensionality reduction methods (PCC and PCA). Comprehensive evaluation indicators were employed to evaluate the performance of the BTOP model, the standalone LSTM, and hybrid models, and the exceedance probability was used to test the ability of varied models to capture the flow of different magnitudes. The main conclusions are as follows:

(1) The data-driven LSTM demonstrated superior and more stable simulation performance compared to the process-based BTOP model. The BTOP model failed to simulate two downstream basins with *NSE* of only 0.15 and 0.17. On the contrary, LSTM performed well across the basin, with *NSE* values all exceeding 0.70. Moreover, the BTOP model significantly underestimated the discharge of the six sub-basins, while the flow duration curve simulated by LSTM fitted well with that of the observed discharge.

(2) Feeding the output variables from the BTOP model into LSTM as input features could provide LSTM with more physical information for learning, thereby improving simulation accuracy. Hybrid-1 and Hybrid-2 displayed comparable performances throughout the basin, both outperforming the standalone LSTM. Moreover, the exceedance probability shows that the fit of the flow duration curve of Hybrid-1 and Hybrid-2 was better than that of the standalone LSTM, and the main error of the two hybrid models originates from the slight underestimation of high flow.

(3) The feature selection method, PCC, and the feature extraction method, PCA, can effectively eliminate noise within the hydrological sequences. Feeding all the BTOP model estimates into LSTM (Hybrid-3) does not enhance simulation performance but instead leads to poor simulation accuracy due to redundant and irrelevant information. Notably, Hybrid-3 exhibited overestimation behavior in the mid-flow, resulting in model accuracy far lower than Hybrid-1 and Hybrid-2 and even lower than the standalone LSTM. It is confirmed that implementing the feature dimension reduction method before constructing the hybrid model is an effective strategy to improve simulation accuracy.

Overall, the results indicated that the output of the BTOP model can provide more physical information for the LSTM learning process, and reducing the dimensionality of variables before constructing a hybrid model can significantly enhance the simulation accuracy. This strategy provides new insights for constructing a hybrid model for runoff simulations in the hydrological field. However, due to the black-box nature of deep learning models, we are unable to quantitatively determine what the model learns from dimensionality-reduced features, whether in feature selection or feature extraction. That is to say, although the experimental results indicate that the dimensionality-reduced features can effectively improve the performance of LSTM simulations, it is difficult for us to directly determine the optimal feature dimensionality reduction scheme according to the learning process of the model. Therefore, future research will focus on applying multiple feature dimensionality reduction methods and conducting a comprehensive comparative analysis. In addition, we will also validate our experimental conclusions on large datasets such as CAMELS and apply transfer learning techniques to improve the performance of deep learning models on data-scarce basins.

Author Contributions: Conceptualization, S.N., L.Z. and H.L.; methodology, S.N., Y.R. and T.A.; software, S.N.; validation, S.N., Y.R. and T.A.; formal analysis, S.N., L.Z. and Y.C.; data curation, Y.R.; writing—original draft preparation, S.N.; writing—review and editing, L.Z., Y.R. and T.A.; funding acquisition, L.Z., Y.R. and T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Open Research Fund (202208) from Key Laboratory of Construction and Safety of Water Engineering of the Ministry of Water Resources, China Institute of Water Resources and Hydropower Research, the Joint Fund Project (U2240214) from National Natural Science Foundation of China, Ministry of Water Resources of the People's Republic of China and China Three Gorges Corporation-Yangtze River Water Science Research, Key R&D Project (XZ202101ZY0007G) from the Science and Technology Department of Tibet, and the Postdoctoral Interdisciplinary Innovation Fund (JCXK2239) from Sichuan University.

Data Availability Statement: Not applicable.

Acknowledgments: We gratefully acknowledge the anonymous reviewers for their comments and suggestions that improved this manuscript.

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

References

1. Hirpa, F.A.; Salamon, P.; Alfieri, L.; Thielen-del Pozo, J.; Zsoter, E.; Pappenberger, F. The Effect of Reference Climatology on Global Flood Forecasting. *J. Hydrometeorol.* **2016**, *17*, 1131–1145. [CrossRef]
2. Croley, T.E.; He, C.S. Distributed-Parameter Large Basin Runoff Model. I: Model Development. *J. Hydrol. Eng.* **2005**, *10*, 173–181. [CrossRef]
3. Pagano, T.C.; Wood, A.W.; Ramos, M.-H.; Cloke, H.L.; Pappenberger, F.; Clark, M.P.; Cranston, M.; Kavetski, D.; Mathevet, T.; Sorooshian, S.; et al. Challenges of Operational River Forecasting. *J. Hydrometeorol.* **2014**, *15*, 1692–1707. [CrossRef]
4. Liu, L.; Zhou, L.; Ao, T.; Liu, X.; Shu, X. Flood Hazard Analysis Based on Rainfall Fusion: A Case Study in Dazhou City, China. *Remote Sens.* **2022**, *14*, 4843. [CrossRef]
5. Shan, Y.; Yan, C.; Liu, J.; Liu, C. Predicting Velocity and Turbulent Kinetic Energy inside an Emergent Phragmites Australis Canopy with Real Morphology. *Environ. Fluid Mech.* **2023**, *23*, 943–963. [CrossRef]
6. Liu, C.; Shan, Y. Impact of an Emergent Model Vegetation Patch on Flow Adjustment and Velocity. *Proc. Inst. Civ. Eng.-Water Manag.* **2022**, *175*, 55–66. [CrossRef]
7. Duan, Z.; Ren, Y.; Liu, X.; Lei, H.; Hua, X.; Shu, X.; Zhou, L. A Comprehensive Comparison of Data Fusion Approaches to Multi-Source Precipitation Observations: A Case Study in Sichuan Province, China. *Environ. Monit. Assess.* **2022**, *194*, 422. [CrossRef]
8. Zhu, Y.; Liu, L.; Qin, F.; Zhou, L.; Zhang, X.; Chen, T.; Li, X.; Ao, T. Application of the Regression-Augmented Regionalization Approach for BTOP Model in Ungauged Basins. *Water* **2021**, *13*, 2294. [CrossRef]
9. Du, J.; Yu, X.; Zhou, L.; Ren, Y.; Ao, T. Precipitation Characteristics across the Three River Headwaters Region of the Tibetan Plateau: A Comparison between Multiple Datasets. *Remote Sens.* **2023**, *15*, 2352. [CrossRef]
10. Beven, K. Linking Parameters Across Scales—Subgrid Parameterizations and Scale-Dependent Hydrological Models. *Hydrol. Process.* **1995**, *9*, 507–525. [CrossRef]
11. Liu, W.; Pan, J.; Ren, Y.; Wu, Z.; Wang, J. Coupling Prediction Model for Long-Term Displacements of Arch Dams Based on Long Short-Term Memory Network. *Struct. Control Health Monit.* **2020**, *27*, e2548. [CrossRef]
12. Yaseen, Z.M.; El-Shafie, A.; Jaafar, O.; Afan, H.A.; Sayl, M.N. Artificial Intelligence Based Models for Stream-Flow Forecasting: 2000–2015. *J. Hydrol.* **2015**, *530*, 829–844. [CrossRef]
13. Tian, P.; Lu, H.; Feng, W.; Guan, Y.; Xue, Y. Large Decrease in Streamflow and Sediment Load of Qinghai-Tibetan Plateau Driven by Future Climate Change: A Case Study in Lhasa River Basin. *Catena* **2020**, *187*, 104340. [CrossRef]
14. Salvadore, E.; Bronders, J.; Batelaan, O. Hydrological Modelling of Urbanized Catchments: A Review and Future Directions. *J. Hydrol.* **2015**, *529*, 62–81. [CrossRef]
15. Arnone, E.; Zoratti, V.; Formetta, G.; Bosa, S.; Petti, M. Predicting Peakflows in Mountain River Basins and Data-Scarce Areas: A Case Study in Northeastern Italy. *Hydrol. Sci. J.* **2023**, *68*, 432–447. [CrossRef]
16. Gupta, H.V.; Nearing, G.S. Debates—The Future of Hydrological Sciences: A (Common) Path Forward? Using Models and Data to Learn: A Systems Theoretic Perspective on the Future of Hydrological Science. *Water Resour. Res.* **2014**, *50*, 5351–5359. [CrossRef]
17. Wagener, T.; Gupta, H.V. Model Identification for Hydrological Forecasting under Uncertainty. *Stoch. Environ. Res. Risk Assess.* **2005**, *19*, 378–387. [CrossRef]
18. Renard, B.; Kavetski, D.; Kuczera, G.; Thyer, M.; Franks, S.W. Understanding Predictive Uncertainty in Hydrologic Modeling: The Challenge of Identifying Input and Structural Errors. *Water Resour. Res.* **2010**, *46*, W05521. [CrossRef]

19. Vache, K.B.; McDonnell, J.J. A Process-Based Rejectionist Framework for Evaluating Catchment Runoff Model Structure. *Water Resour. Res.* **2006**, *42*, W02409. [CrossRef]
20. Clark, M.P.; Nijssen, B.; Lundquist, J.D.; Kavetski, D.; Rupp, D.E.; Woods, R.A.; Freer, J.E.; Gutmann, E.D.; Wood, A.W.; Brekke, L.D.; et al. A Unified Approach for Process-Based Hydrologic Modeling: 1. Modeling Concept. *Water Resour. Res.* **2015**, *51*, 2498–2514. [CrossRef]
21. Clark, M.P.; Bierkens, M.F.P.; Samaniego, L.; Woods, R.A.; Uijlenhoet, R.; Bennett, K.E.; Pauwels, V.R.N.; Cai, X.; Wood, A.W.; Peters-Lidard, C.D. The Evolution of Process-Based Hydrologic Models: Historical Challenges and the Collective Quest for Physical Realism. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3427–3440. [CrossRef] [PubMed]
22. Zhu, S.; Luo, X.; Yuan, X.; Xu, Z. An Improved Long Short-Term Memory Network for Streamflow Forecasting in the Upper Yangtze River. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1313–1329. [CrossRef]
23. Nearing, G.S.; Kratzert, F.; Sampson, A.K.; Pelissier, C.S.; Klotz, D.; Frame, J.M.; Prieto, C.; Gupta, H.V. What Role Does Hydrological Science Play in the Age of Machine Learning? *Water Resour. Res.* **2021**, *57*, e2020WR028091. [CrossRef]
24. Feng, D.; Liu, J.; Lawson, K.; Shen, C. Differentiable, Learnable, Regionalized Process-Based Models With Multiphysical Outputs Can Approach State-Of-The-Art Hydrologic Prediction Accuracy. *Water Resour. Res.* **2022**, *58*, e2022WR032404. [CrossRef]
25. Liu, L.; Zhou, L.; Gusyev, M.; Ren, Y. Unravelling and Improving the Potential of Global Discharge Reanalysis Dataset in Streamflow Estimation in Ungauged Basins. *J. Clean. Prod.* **2023**, *419*, 138282. [CrossRef]
26. Xiao, Q.; Zhou, L.; Xiang, X.; Liu, L.; Liu, X.; Li, X.; Ao, T. Integration of Hydrological Model and Time Series Model for Improving the Runoff Simulation: A Case Study on BTOP Model in Zhou River Basin, China. *Appl. Sci.* **2022**, *12*, 6883. [CrossRef]
27. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
28. Li, Q.; Wang, Z.; Shangguan, W.; Li, L.; Yao, Y.; Yu, F. Improved Daily SMAP Satellite Soil Moisture Prediction over China Using Deep Learning Model with Transfer Learning. *J. Hydrol.* **2021**, *600*, 126698. [CrossRef]
29. Ahmed, A.A.M.; Deo, R.C.; Feng, Q.; Ghahramani, A.; Raj, N.; Yin, Z.; Yang, L. Hybrid Deep Learning Method for a Week-Ahead Evapotranspiration Forecasting. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 831–849. [CrossRef]
30. Kratzert, F.; Klotz, D.; Brenner, C.; Schulz, K.; Herrnegger, M. Rainfall–Runoff Modelling Using Long Short-Term Memory (LSTM) Networks. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6005–6022. [CrossRef]
31. Kratzert, F.; Klotz, D.; Shalev, G.; Klambauer, G.; Hochreiter, S.; Nearing, G. Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine Learning Applied to Large-Sample Datasets. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 5089–5110. [CrossRef]
32. Tian, Y.; Xu, Y.-P.; Yang, Z.; Wang, G.; Zhu, Q. Integration of a Parsimonious Hydrological Model with Recurrent Neural Networks for Improved Streamflow Forecasting. *Water* **2018**, *10*, 1655. [CrossRef]
33. Xiang, Z.; Yan, J.; Demir, I. A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning. *Water Resour. Res.* **2020**, *56*, e2019WR025326. [CrossRef]
34. Tsai, W.-P.; Feng, D.; Pan, M.; Beck, H.; Lawson, K.; Yang, Y.; Liu, J.; Shen, C. From Calibration to Parameter Learning: Harnessing the Scaling Effects of Big Data in Geoscientific Modeling. *Nat. Commun.* **2021**, *12*, 5988. [CrossRef]
35. Lu, D.; Konapala, G.; Painter, S.L.; Kao, S.-C.; Gangrade, S. Streamflow Simulation in Data-Scarce Basins Using Bayesian and Physics-Informed Machine Learning Models. *J. Hydrometeorol.* **2021**, *22*, 1421–1438. [CrossRef]
36. Konapala, G.; Kao, S.-C.; Painter, S.L.; Lu, D. Machine Learning Assisted Hybrid Models Can Improve Streamflow Simulation in Diverse Catchments across the Conterminous US. *Environ. Res. Lett.* **2020**, *15*, 104022. [CrossRef]
37. Wi, S.; Steinschneider, S. Assessing the Physical Realism of Deep Learning Hydrologic Model Projections Under Climate Change. *Water Resour. Res.* **2022**, *58*, e2022WR032123. [CrossRef]
38. Saeys, Y.; Inza, I.; Larrañaga, P. A Review of Feature Selection Techniques in Bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef]
39. Zhang, D.; Chen, S.; Zhou, Z.-H. Constraint Score: A New Filter Method for Feature Selection with Pairwise Constraints. *Pattern Recognit.* **2008**, *41*, 1440–1451. [CrossRef]
40. Wang, J.; Wu, L.; Kong, J.; Li, Y.; Zhang, B. Maximum Weight and Minimum Redundancy: A Novel Framework for Feature Subset Selection. *Pattern Recognit.* **2013**, *46*, 1616–1627. [CrossRef]
41. Lin, S.-S.; Zhang, N.; Zhou, A.; Shen, S.-L. Time-Series Prediction of Shield Movement Performance during Tunneling Based on Hybrid Model. *Tunn. Undergr. Space Technol.* **2022**, *119*, 104245. [CrossRef]
42. Pathy, A.; Meher, S.; Balasubramanian, P. Predicting Algal Biochar Yield Using eXtreme Gradient Boosting (XGB) Algorithm of Machine Learning Methods. *Algal Res.* **2020**, *50*, 102006. [CrossRef]
43. Chen, H.; Chang, X. Photovoltaic Power Prediction of LSTM Model Based on Pearson Feature Selection. *Energy Rep.* **2021**, *7*, 1047–1054. [CrossRef]
44. Xie, A.; Yang, H.; Chen, J.; Sheng, L.; Zhang, Q. A Short-Term Wind Speed Forecasting Model Based on a Multi-Variable Long Short-Term Memory Network. *Atmosphere* **2021**, *12*, 651. [CrossRef]
45. Yang, J.; Zhang, D.; Frangi, A.F.; Yang, J. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 131–137. [CrossRef]
46. Zhang, Y.; Yan, B.; Aasma, M. A Novel Deep Learning Framework: Prediction and Analysis of Financial Time Series Using CEEMD and LSTM. *Expert Syst. Appl.* **2020**, *159*, 113609. [CrossRef]

47. Xu, W.; Liu, P.; Cheng, L.; Zhou, Y.; Xia, Q.; Gong, Y.; Liu, Y. Multi-Step Wind Speed Prediction by Combining a WRF Simulation and an Error Correction Strategy. *Renew. Energy* **2021**, *163*, 772–782. [CrossRef]
48. Zhang, J.; Chen, X.; Khan, A.; Zhang, Y.; Kuang, X.; Liang, X.; Taccari, M.L.; Nuttall, J. Daily Runoff Forecasting by Deep Recursive Neural Network. *J. Hydrol.* **2021**, *596*, 126067. [CrossRef]
49. Martens, B.; Miralles, D.G.; Lievens, H.; van der Schalie, R.; de Jeu, R.A.M.; Fernandez-Prieto, D.; Beck, H.E.; Dorigo, W.A.; Verhoest, N.E.C. GLEAM v3: Satellite-Based Land Evaporation and Root-Zone Soil Moisture. *Geosci. Model Dev.* **2017**, *10*, 1903–1925. [CrossRef]
50. Yamazaki, D.; Ikeshima, D.; Tawatari, R.; Yamaguchi, T.; O'Loughlin, F.; Neal, J.C.; Sampson, C.C.; Kanae, S.; Bates, P.D. A High-Accuracy Map of Global Terrain Elevations. *Geophys. Res. Lett.* **2017**, *44*, 5844–5853. [CrossRef]
51. The Potential Evapotranspiration (EP) from the Climatic Research Unit (CRU) of the School of Environmental Sciences (ENV) at the University of East Anglia (UEA). Available online: http://www.cru.uea.ac.uk/ (accessed on 25 March 2022).
52. The Leaf Area Index (LAI) of the National Centers for Environmental Information. Available online: https://www.ncei.noaa.gov/Data/Avhrr-Land-Leaf-Area-Index-and-Fapar/Access/ (accessed on 25 March 2022).
53. FAO Digital Soil Map of the World (DSMW). Available online: http://www.fao.org/Land-Water/Land/Land-Governance/Landresources-Planning-Toolbox/Category/Details/En/c/1026564/ (accessed on 25 March 2022).
54. LP DAAC-MCD12Q1. Available online: https://lpdaac.usgs.gov/Products/Mcd12q1v006/ (accessed on 25 March 2022).
55. Takeuchi, K.; Ao, T.Q.; Ishidaira, H. Introduction of Block-Wise Use of TOPMODEL and Muskingum-Cunge Method for the Hydro-Environmental Simulation of a Large Ungauged Basin. *Hydrol. Sci. J.-J. Sci. Hydrol.* **1999**, *44*, 633–646. [CrossRef]
56. Ao, T.; Ishidaira, H.; Takeuchi, K. Study of Distributed Runoff Simulation Model Based on Block Type Topmodel and Muskingum-Cunge Method. *Proc. Hydraul. Eng.* **1999**, *43*, 7–12. [CrossRef]
57. Takeuchi, K.; Hapuarachchi, P.; Zhou, M.; Ishidaira, H.; Magome, J. A BTOP Model to Extend TOPMODEL for Distributed Hydrological Simulation of Large Basins. *Hydrol. Process.* **2008**, *22*, 3236–3251. [CrossRef]
58. Zhou, M.C.; Ishidaira, H.; Hapuarachchi, H.P.; Magome, J.; Kiem, A.S.; Takeuchi, K. Estimating Potential Evapotranspiration Using Shuttleworth-Wallace Model and NOAA-AVHRR NDVI Data to Feed a Distributed Hydrological Model over the Mekong River Basin. *J. Hydrol.* **2006**, *327*, 151–173. [CrossRef]
59. Beven, K.J.; Kirkby, M.J.; Freer, J.E.; Lamb, R. A History of TOPMODEL. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 527–549. [CrossRef]
60. Barry, D.; Bajracharya, K. On the Muskingum-Cunge Flood Routing Method. *Environ. Int.* **1995**, *21*, 485–490. [CrossRef]
61. Navarathinam, K.; Gusyev, M.A.; Hasegawa, A.; Magome, J.; Takeuchi, K. Agricultural Flood and Drought Risk Reduction by a Proposed Multi-Purpose Dam: A Case Study of the Malwathoya River Basin, Sri Lanka. In Proceedings of the 21st International Congress on Modelling and Simulation (MODSIM 2015), Queensland, Australia, 29 November–4 December 2015; Weber, T., McPhee, M.J., Anderssen, R.S., Eds.; Modelling & Simulation Soc Australia & New Zealand Inc.: Christchurch, New Zealand, 2015; pp. 1600–1606.
62. Ishidaira, H.; Sawada, H.; Masumoto, T. Studies on the Mekong and River Basin-Modelling of Hydrology Water Resources. *Hydrol. Process.* **2008**, *22*, 1243–1245. [CrossRef]
63. Hapuarachchi, H.A.P.; Takeuchi, K.; Zhou, M.; Kiem, A.S.; Georgievski, M.; Magome, J.; Ishidaira, H. Investigation of the Mekong River Basin Hydrology for 1980-2000 Using the YHyM. *Hydrol. Process.* **2008**, *22*, 1246–1256. [CrossRef]
64. Liu, L.; Zhou, L.; Li, X.; Chen, T.; Ao, T. Screening and Optimizing the Sensitive Parameters of BTOPMC Model Based on UQ-PyL Software: Case Study of a Flood Event in the Fuji River Basin, Japan. *J. Hydrol. Eng.* **2020**, *25*, 05020030. [CrossRef]
65. Zhou, L.; Rasmy, M.; Takeuchi, K.; Koike, T.; Selvarajah, H.; Ao, T. Adequacy of Near Real-Time Satellite Precipitation Products in Driving Flood Discharge Simulation in the Fuji River Basin, Japan. *Appl. Sci.* **2021**, *11*, 1087. [CrossRef]
66. Zhou, L.; Koike, T.; Takeuchi, K.; Rasmy, M.; Onuma, K.; Ito, H.; Selvarajah, H.; Liu, L.; Li, X.; Ao, T. A Study on Availability of Ground Observations and Its Impacts on Bias Correction of Satellite Precipitation Products and Hydrologic Simulation Efficiency. *J. Hydrol.* **2022**, *610*, 127595. [CrossRef]
67. Liu, L.; Ao, T.; Zhou, L.; Takeuchi, K.; Gusyev, M.; Zhang, X.; Wang, W.; Ren, Y. Comprehensive Evaluation of Parameter Importance and Optimization Based on the Integrated Sensitivity Analysis System: A Case Study of the BTOP Model in the Upper Min River Basin, China. *J. Hydrol.* **2022**, *610*, 127819. [CrossRef]
68. Duan, Q.; Sorooshian, S.; Gupta, V. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. *Water Resour. Res.* **1992**, *28*, 1015–1031. [CrossRef]
69. Hochreiter, S.; Schmidhuber, J. LSTM Can Solve Hard Long Time Lag Problems. In *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference*; Mozer, M.C., Jordan, M.I., Petsche, T., Eds.; Advances in Neural Information Processing Systems; M I T Press: Cambridge, UK, 1997; Volume 9, pp. 473–479.
70. Yang, S.; Yu, X.; Zhou, Y. LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example. In *2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)*; IEEE: Shanghai, China, 2020; pp. 98–101. [CrossRef]
71. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef]
72. Zhang, Z.; Qin, H.; Liu, Y.; Wang, Y.; Yao, L.; Li, Q.; Li, J.; Pei, S. Long Short-Term Memory Network Based on Neighborhood Gates for Processing Complex Causality in Wind Speed Prediction. *Energy Convers. Manag.* **2019**, *192*, 37–51. [CrossRef]

73. Karl Pearson, F.R.S. LIII. On Lines and Planes of Closest Fit to Systems of Points in Space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **2010**, *2*, 559–572. Available online: https://www.tandfonline.com/doi/abs/10.1080/14786440109462720 (accessed on 26 July 2023). [CrossRef]
74. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24*, 417–441. [CrossRef]
75. Cui, W.; Zhang, Y.; Zhang, X.; Li, L.; Liou, F. Metal Additive Manufacturing Parts Inspection Using Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 545. [CrossRef]
76. Hood, M.J.; Clausen, J.C.; Warner, G.S. Comparison of Stormwater Lag Times for Low Impact and Traditional Residential Development. *J. Am. Water Resour. Assoc.* **2007**, *43*, 1036–1046. [CrossRef]
77. Mao, G.; Wang, M.; Liu, J.; Wang, Z.; Wang, K.; Meng, Y.; Zhong, R.; Wang, H.; Li, Y. Comprehensive Comparison of Artificial Neural Networks and Long Short-Term Memory Networks for Rainfall-Runoff Simulation. *Phys. Chem. Earth Parts ABC* **2021**, *123*, 103026. [CrossRef]
78. Yu, Q.; Jiang, L.; Wang, Y.; Liu, J. Enhancing Streamflow Simulation Using Hybridized Machine Learning Models in a Semi-Arid Basin of the Chinese Loess Plateau. *J. Hydrol.* **2023**, *617*, 129115. [CrossRef]
79. Lei, H.; Zhao, H.; Ao, T.; Hu, W. Quantifying the Reliability and Uncertainty of Satellite, Reanalysis, and Merged Precipitation Products in Hydrological Simulations over the Topographically Diverse Basin in Southwest China. *Remote Sens.* **2022**, *15*, 213. [CrossRef]