

Extracting operation behaviors of cascade reservoirs using physics-guided long-short term memory networks

Yalian Zheng^{a,b}, Pan Liu^{a,b,*}, Lei Cheng^{a,b}, Kang Xie^{a,b}, Wei Lou^c, Xiao Li^{a,b}, Xinran Luo^{a,b}, Qian Cheng^{a,b}, Dongyang Han^{a,b}, Wei Zhang^d

^a State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

^b Hubei Provincial Key Lab of Water System Science for Sponge City Construction, Wuhan University, Wuhan 430072, China

^c Power China Huadong Engineering Corporation Limited, Hangzhou 311122, Zhejiang, China

^d Science and Technology Research Institute, China Three Gorges Corporation, Beijing 100038, China



ARTICLE INFO

Keywords:

LSTM
Data-driven
Extracting human behaviors
Physics-guided

ABSTRACT

Study region: Qingjiang cascade reservoir, China.

Study focus: Reservoirs regulate the natural streamflow to utilize water resources comprehensively. How to mine the existing massive reservoir operation data to describe human operation behaviors is a challenge. To address this issue, a data-driven method, Long short-term memory (LSTM), was used to simulate the reservoir outflow by inputting historical information. The physics-guided LSTM model, shortly named PG-LSTM, was formulated by using synthetic flood samples and physical constraints of water balance, boundary, and monotonicity.

New hydrological insights: (1) PG-LSTM can reproduce historical outflow with seasonal variations, or predict outflow without lags, (2) knowledge of reservoir operations can guide LSTM with the reduction of negative flow occurrence and the accurate identification of operation behaviors under extreme hydrological conditions, (3) specifically, compared with conventional LSTM, gradient boosting regression tree and conventional reservoir operation, PG-LSTM can improve the Nash-Sutcliffe efficiency of cascade reservoir during the test period from 0.50, 0.20, and 0.17 to 0.54 in the reproduction scenario, and from 0.84, 0.26, and 0.17 to 0.85 in the prediction scenario with five-fold cross-validation method. The PG-LSTM is helpful to describe human operation behaviors of reservoirs.

1. Introduction

Reservoirs are water resource projects for power generation, flood control, water supply, and comprehensive utilization, where efficient operations are essential for the implementation of reservoir functions (Gong et al., 2022; Lehner, 2011; Li, 2020b; Yang et al., 2018). With the development of science and technology, reservoir managers have recorded a large amount of historical hydrological information and operation data, causing an explosive increase in reservoir operation data (Li et al., 2020a; Yang et al., 2017). In reservoir operation, operators integrate multiple sources of information to form decisions based on their own experience and professional judgment (Zhang et al., 2021). Currently, it is a challenge to mine the existing massive data to describe human operation

* Corresponding author at: State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China.

E-mail address: liupan@whu.edu.cn (P. Liu).

behavior (Corani et al., 2009; Hossain and El-shafie, 2013; Yue et al., 2018).

Although operators' experience is difficult to quantify (Yang et al., 2019), historical operating data implies operation behaviors facing multiple inflow scenarios (Yang et al., 2020; Zhang et al., 2019b). Therefore, the decision-making process of the reservoir managers can be extracted by mining the historical operating data. If the human operation behavior can be extracted, the reservoir operation decision can be reproduced, and the outflow from upstream reservoirs can be estimated by the downstream ones, which is beneficial to manage cascade reservoirs.

Generally, reservoir operation can be implemented by models centered on the water balance principle (Cheng et al., 2021; Gong et al., 2021a; Wang et al., 2021), such as the optimization reservoir operation model (Ahmad et al., 2014; Huang et al., 2021; Ming et al., 2021) and the conventional one (Gong et al., 2021b; Li et al., 2022; Ming et al., 2017). However, due to the heterogeneity of the actual operation process on temporal and spatial scales (Bayoumi et al., 2019; Liu et al., 2019), these models are inconsistent with the actual reservoir operation because the cognition and assumptions are unsuitable. In contrast, data mining models make fewer assumptions about the operation process and can extract reservoir operation behaviors from actual operation data, which is a beneficial supplement to the traditional reservoir operation model (Moussa et al., 2016; Wang et al., 2011).

Data mining models comprise shallow learning and deep learning models. Shallow learning models (Mohammadifar et al., 2021), such as artificial neural network (ANN) (Gori and Tesi, 1992), support vector machine and regression (SVM and SVR) (Zhang et al., 2018a), decision trees (DT), and radial basis functions (RBFs), have been applied to derive reservoir operation rules in recent years. For instance, the Classification and Regression-Tree (CART) Algorithm, a kind of DT, was used to predict the outflow of nine reservoirs (Yang et al., 2016), and input variable selection (CIS) and Gaussian radial basis functions (RBFs) methods were combined to derive reservoir operation rules (Yang et al., 2017). Although shallow learning models are effective for nonlinear simulation, the computational difficulty of shallow learning models increases with the expansion of input dimension, leading to difficulties in extracting reservoir operation behaviors in various decision scenarios.

Deep learning methods have shown a broad application prospect in hydrology (Shen, 2017). Long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1996), a sequential deep learning model considered more efficient than other data mining models, has attracted widespread interest in time series predicting (Kratzert et al., 2019; Xu et al., 2021). Recently, there has been an increasing interest in the simulation of reservoir outflow by LSTM. For example, Zhang et al. (2018a) investigated the effect of different LSTM network parameter setting on the model performance, and the ability of LSTM to simulate reservoir outflow of different time scales and flow regimes. They further compared the performance of recurrent neural network (RNN), LSTM, and gated recurrent unit (GRU) in predicting the outflow of the Xiluodu reservoir (Zhang et al., 2019a). Yang et al. (2019) explored the ability of nonlinear autoregressive models with exogenous input (NARX), LSTM, and genetic algorithm-based NARX to simulate the outflow of a single reservoir under extreme floods and droughts. However, reservoir operating patterns can only be extracted by LSTM if there is sufficient support from

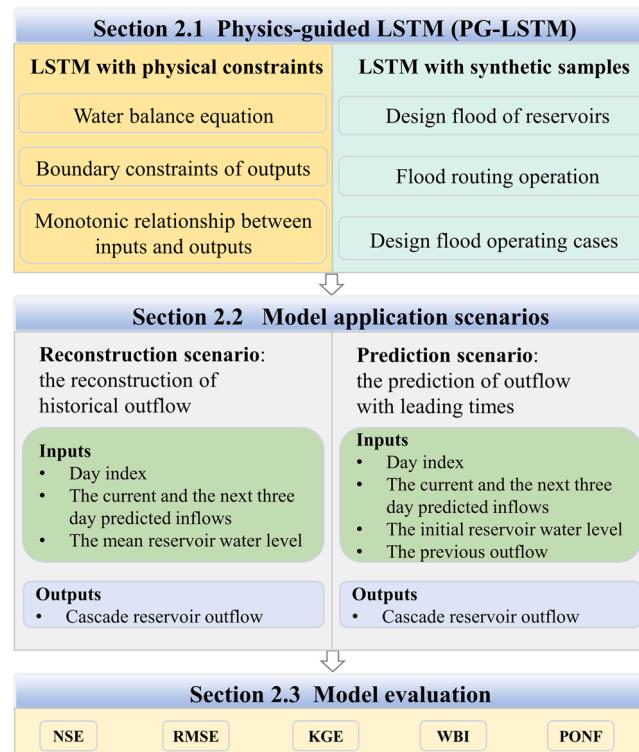


Fig. 1. The schematic diagram of the methodology.

historical operating data. As a black-box model, LSTM lacks physical mechanisms (Xie et al., 2021) causing abnormal values (such as negative flow), becoming a major limitation in extracting reservoir operation behaviors.

To achieve physical consistency in deep learning models, some researchers investigated hybrid strategies (Jiang et al., 2020; Markus et al., 2019). Read et al. (2019) tested components responsible for hybrid models, including process constraints derived from the energy conservation equation and the data obtained from process-based models. Wang et al. (2019) proposed a theory-guided neural network framework, where the loss function was guided by the governing equations, other physical constraints, engineering controls, and expert knowledge. Similarly, such hybrid strategies can be used to develop an improved model for extracting reservoir operation behaviors. Unfortunately, there is little attention paid to physical-based data-driven models in reservoir operation.

This study, therefore, aims to implant physical knowledge into LSTM to develop an improved model, named physics-guided LSTM (PG-LSTM), for extracting reservoir operation behaviors. Specifically, this study endeavors to answer the following questions: (1) What applications can deep learning models achieve? (2) To achieve the above applications, what physical mechanism can be added to the deep learning model? (3) What are the benefits of deep learning models combined with the physical mechanism compared with conventional deep learning models or shallow learning models?

The remainder of this paper is organized as follows: Section 2 presents the methodology for the PG-LSTM and benchmark models. Details on the chosen cascade reservoirs are described in Section 3, followed by results in Section 4 and discussion in Section 5. Finally, conclusions are drawn in Section 6.

2. Methodology

Fig. 1 provides an overview of the methodology comprising three steps: (1) Two physical mechanisms of reservoir operation were implanted into LSTM to build the PG-LSTM model (Section 2.1). (2) Model inputs and outputs were identified for different scenarios (Section 2.2). (3) Varied statistical criteria are used to assess the accuracy and physical consistency (Section 2.3).

2.1. Physics-guided LSTM (PG-LSTM)

2.1.1. Long-short term memory (LSTM)

RNNs comprise continuous units, and the unit calculation results are continuously recursive on the time scale. Different categories of RNNs differ in connections within the unit (Zhang et al., 2018b). However, if the units manifest a long-term dependence on the past information, it is difficult to obtain desirable results (LeCun et al., 2015). To overcome the pitfalls of traditional RNN, Hochreiter and Schmidhuber (1997) developed the LSTM neural network model to handle the long-term mapping relationship of time series data. A conventional LSTM unit contains forget gate, input gate, output gate, and a memory cell. The forget gate determines whether to discard information using the sigmoid function. The input gate and the tanh function save information to obtain candidate cell state values. The memory cell is utilized to store the information of the long-term memory, and the output gate decides the memory cell output. The conventional LSTM structure and the calculation procedure can be found in the study of Xu et al. (2021).

2.1.2. Physical mechanisms of reservoir operation implanted into LSTM

Two physical mechanisms are integrated into LSTM as follows: (1) Penalty terms are added into the loss function of the deep learning model to encourage physical consistency, for violating knowledge of the water balance principle, outflow boundary, monotonic relationship between inflow and outflow, and (2) the training dataset of LSTM is supplemented by synthetic samples based on the design flood using the flood routing.

2.1.2.1. LSTM with physical constraints.

(1) Water balance principle of cascade reservoir

a. Water balance equation of single reservoir

$$V_{i,t+1} = V_{i,t} + (I_{i,t} - Q_{i,t})\Delta t \quad i = 1, 2, \dots, N \quad (1)$$

b. The hydraulic connection between the upstream reservoir i and the downstream reservoir $i + 1$.

$$I_{i+1,t} = I'_{i,t} + Q_{i,t} \quad i = 1, 2, \dots, N \quad (2)$$

where $V_{i,t+1}$ and $V_{i,t}$ are the initial and final water storage of reservoir i at period t respectively; $I_{i,t}$ and $Q_{i,t}$ are the inflow and release of reservoir i at period t respectively; Δt is the time interval for routing; N is the number of reservoirs in the studied cascade reservoir; $I'_{i,t}$ denotes the local inflow from the zone between the upstream reservoir i and the downstream reservoir $i + 1$ at period t . The water balance during the test period was checked by calculating the difference in water volume between inflow, outflow, and reservoir storage.

(2) Monotonic relationship between inflow and outflow

The internal relationship between reservoir inflow and outflow depends on the dynamic interaction between inflow intensity,

reservoir storage, reservoir outflow. The monotonic relationship between inflow and outflow is that the reservoir outflow should be not smaller than the original outflow if a small increase appears in reservoir inflow. This constraint is considered to ensure the robustness of deep learning models encountering perturbations.

The monotonic relationship between inflow and outflow is as follows:

$$\hat{Q}_{i,t} = f(I_{i,t}, \theta) \quad (3)$$

$$f(I_{i,t} + \Delta I_i, \theta) - f(I_{i,t}, \theta) \geq 0 \quad (4)$$

where $\hat{Q}_{i,t}$ is the simulated outflow of reservoir i at period t ; f denotes the simulation function of the deep learning network; θ is the parameters of the network; ΔI_i is a small increase of inflow.

(3) Boundary constraints of outflow

The boundary constraints of outflow are as follows:

$$\begin{aligned} Q_{l,i} &= \min\{\hat{Q}_{i,t}\} \\ Q_{u,i} &= \max\{\hat{Q}_{i,t}\} \\ Q_{l,i} &\leq \hat{Q}_{i,t} \leq Q_{u,i} \end{aligned} \quad (5)$$

where $Q_{l,i}$ is the lower boundary of outflow; $Q_{u,i}$ is the upper boundary of outflow.

(4) The loss function with physical constraints

The physical constraints, inspired by the physical understandings described above, are added to the conventional objective of data fitting. Thus, the PG-LSTM model can be built with the equation of loss function as follows:

$$\left\{ \begin{array}{l} Loss = \lambda_{data} loss_{data} + \lambda_{balance} loss_{balance} + \lambda_{monotonicity} loss_{monotonicity} + \lambda_{boundary} loss_{boundary} \\ loss_{data} = \sum_{i=1}^N \sqrt{\frac{1}{n_t} \sum_{t=1}^{n_t} (\hat{Q}_{i,t} - Q_{i,t})^2} \\ loss_{balance} = \sum_{i=1}^N \sqrt{\frac{1}{n_t} \left(\Delta W_i - \left(\sum_{t=1}^{n_t} (I_{i,t} - \hat{Q}_{i,t}) \times \Delta t \right) \right)^2} \\ loss_{monotonicity} = \frac{1}{n_t} \sum_{i=1}^N \sum_{t=1}^{n_t} \max[0, f(I_{i,t} + \Delta I_i, \theta) - f(I_{i,t}, \theta)] \\ g(\hat{Q}_{i,t}) = \begin{cases} |\hat{Q}_{l,i} - \hat{Q}_{i,t}| & \text{if } \hat{Q}_{l,i} > \hat{Q}_{i,t} \\ 0 & \text{if } \hat{Q}_{l,i} \leq \hat{Q}_{i,t} \leq \hat{Q}_{u,i} \\ |\hat{Q}_{u,i} - \hat{Q}_{i,t}| & \text{if } \hat{Q}_{u,i} < \hat{Q}_{i,t} \end{cases} \\ loss_{boundary} = \frac{1}{n_t} \sum_{i=1}^N \sum_{t=1}^{n_t} g(\hat{Q}_{i,t}) \end{array} \right. \quad (6)$$

where $Loss$ is the loss function of PG-LSTM; $loss_{data}$ is the root mean square error of simulated outflow; $loss_{balance}$ is the penalty term for violating long-term water balance constraint for simulated outflow; $loss_{monotonicity}$ is the penalty term for violating the monotonicity of simulated outflow; $loss_{boundary}$ is the penalty term for violating the boundary constraints of simulated outflow; λ is the weight coefficient for different loss items determined by the trial and error method; n_t is the length of time series in the test period; $Q_{i,t}$ is the observed outflow of reservoir i at period t ; and $I_{i,t}$ ($i = 1$) is the observed inflow of the first upstream reservoir, and $I_{i,t} = \hat{I}_{i,t} = I'_{i,t} + \hat{Q}_{i-1,t}$ ($i = 2, 3, \dots, N$) is calculated by the simulated outflow of the upstream reservoir and the local inflow according to the hydraulic connection of cascade reservoir; ΔW_i is the storage change of reservoir i during the test period.

2.1.2.2. LSTM with synthetic samples. An important goal of this study is the effective prediction of outflow facing extreme hydrological events. However, there are few high flow samples in the historical data under normal reservoir operation conditions, leading to the poor simulation of high flow. To solve this problem, the synthetic samples, based on the design flood operating cases, are used as the additional input to train the model. The synthetic samples can provide more information about the high flow events determined by the flood routing.

The steps of constructing LSTM with synthetic samples are shown in Fig. 2: (1) Collection of operational parameters and curves of

the studied cascade reservoir and acquisition of design flood information. (2) Establishing a flood routing model based on reservoir operating rules. (3) Analysis and extraction of the synthetic samples including information of cascade reservoir inflow, outflow, and water level.

2.1.3. Benchmark model

To quantitatively show the improvement of deep learning models, the gradient boosting regression tree (GBRT) model and the conventional reservoir operation (CRO) model are used as benchmark models.

2.1.3.1. Gradient boosting regression tree (GBRT). GBRT, as one of the most important machine learning methods, has been employed in engineering science, hydrology, and water science such as water demand (Carvalho and Filho, 2021), runoff (He et al., 2020), short-term load (Mayrink and Hippert, 2016) and solar power forecasting (Persson et al., 2017), addressing the problem of non-linearity simulating. GBRT consists of a sequential ensemble of weak learners (Friedman, 2001), usually, decision trees. After the creation of decision trees, the tree is first used to predict the outflow of cascade reservoirs with a set of explanatory variables. Then, the trees in the sequence focus on the prediction residuals of the previous trees. The above procedure repeated with the optimization of loss functions and several hyperparameters needs to be tuned to make the simulation better, such as the number of trees, the maximum tree depth, and the learning rate. In this study, the GBRT model was constructed by the python programming language and Scikit-learn package (Kramer, 2016; Rossum and Drake, 2012). Different combinations of the number of trees (10–100), the maximum tree depth (2–10), and the learning rate (0.01–0.1) were tested by a mesh search algorithm and tuned using five-fold cross-validation. The chosen hyperparameters can provide the best performance of the simulation. As a benchmark model, inputs and outputs of GBRT are the same as DL models.

2.1.3.2. Conventional reservoir operation (CRO). CRO (Gong et al., 2021a) infers release using the reservoir operating rule curves, which are used by reservoir operators in the face of normal hydrological conditions.

2.2. Model application scenarios

The PG-LSTM model includes constraints of the water balance, outflow boundary, monotonic relationship between inflow and outflow in the loss function, and the synthetic samples of the design flood in the training dataset. The application scenarios of PG-LSTM are further discussed in Sections 2.2.1 and 2.2.2.

2.2.1. Reproduction scenario: the reproduction of cascade reservoir historical outflow

The first application scenario is the reproduction of the historical outflow decision. The long series of the historical operating decision in periods without reservoirs can provide benchmark comparison solutions for the evaluation of optimized reservoir operation

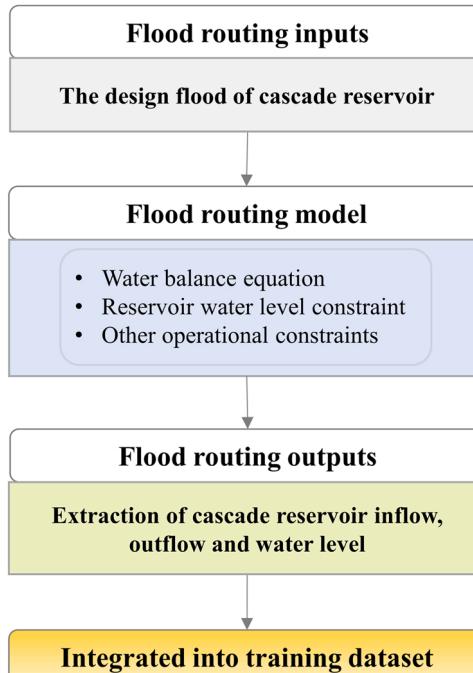


Fig. 2. The steps of synthesizing flood routing samples.

model, solving the problem that conventional reservoir operation method is harder to assess the benefits of optimized reservoir operation model. In the historical period without reservoirs, only the day index and cascade reservoir inflow information are available. However, different reservoir water level states should be considered for operating decision-making. Therefore, the mean of the operational water level on a daily scale, namely the mean reservoir water level, can be used to describe the seasonality of reservoir release. The inputs and outputs of models in the reproduction scenario are as shown in Fig. 3. The mapping relationship between model inputs and outputs can be determined as follows:

$$Q_{i,t}, Q_{i+1,t}, \dots, Q_{N,t} = f(X, \theta) \quad (7)$$

where $Q_{i,t}$ ($i = 1, 2, \dots, N$) is the release of reservoir i at period t ; f denotes the simulation function of the deep learning network in the reproduction scenario; X is model inputs in the reproduction scenario, including the day index t , the inflow of reservoir i ($i = 1$), namely $I_{1,t}, I_{1,t+1}, I_{1,t+2}, I_{1,t+3}$, local inflow between reservoir i and reservoir $i + 1$, namely $I'_{i,t}, I'_{i,t+1}, I'_{i,t+2}, I'_{i,t+3}$ ($i = 1, 2, \dots, N - 1$), and the static reservoir water level \bar{Z}_i ; θ is the parameters of the network in the reproduction scenario.

2.2.2. Prediction scenario: the prediction of cascade reservoir outflow with leading times

The second application scenario is the prediction of cascade reservoir outflow with several leading times. For the prediction scenario, the currently available data regarding inflow, initial reservoir water level, and previous outflow can be used as model inputs. The inputs and outputs in the prediction scenario are as shown in Fig. 4. The mapping relationship between model inputs and outputs can be determined as follows:

$$Q_{i,t}, Q_{i+1,t}, \dots, Q_{N,t} = f'_t(X', \theta') \quad (t = 0, 1, 2, \dots, m) \quad (8)$$

where $Q_{i,t}, Q_{i+1,t}, \dots, Q_{N,t}$ ($t = 0, 1, 2, \dots, m$) is the release of reservoir i ($i = 1, 2, \dots, N$) with leading time t ; f'_t denotes the simulation function of the deep learning network with leading time t in the prediction scenario; X' is model inputs in the prediction scenario, including the day index t , the inflow of reservoir i ($i = 1$), namely $I_{1,t}, I_{1,t+1}, I_{1,t+2}, I_{1,t+3}$, local inflow between reservoir i and reservoir $i + 1$, namely $I'_{i,t}, I'_{i,t+1}, I'_{i,t+2}, I'_{i,t+3}$ ($i = 1, 2, \dots, N - 1$), the initial reservoir water level $Z_{i,t-1}$, and the outflow of reservoir i ($i = 1, 2, \dots, N$) in the previous day; θ' is the parameters of the network in the prediction scenario.

2.3. Model evaluation

2.3.1. Simulation accuracy

In this study, three statistical criteria are employed to quantify the model simulation accuracy.

(1) Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum_{t=1}^n (\hat{Q}_{i,t} - \bar{Q}_{i,t})^2}{\sum_{t=1}^n (Q_{i,t} - \bar{Q}_{i,t})^2} \quad (9)$$

(2) The root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{Q}_{i,t} - \bar{Q}_{i,t})^2} \quad (10)$$

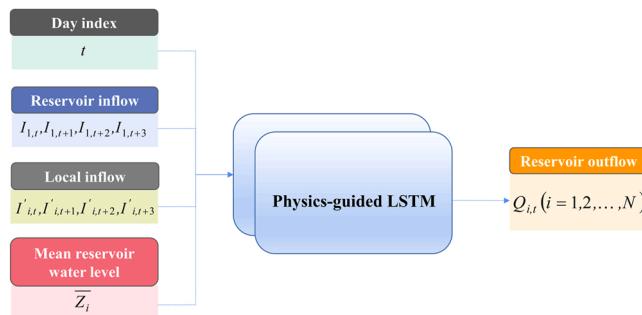


Fig. 3. The schematic diagram of inputs and outputs in the reproduction scenario.

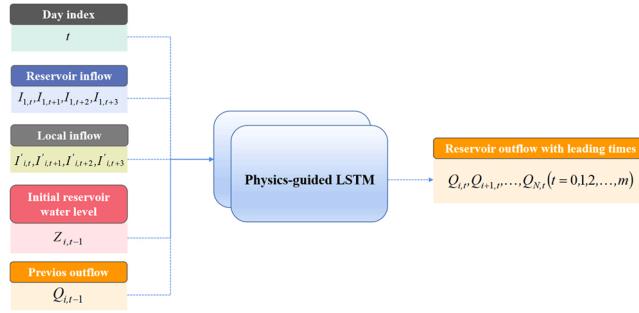


Fig. 4. The schematic diagram of inputs and outputs in the prediction scenario.

(3) The coefficient of Kling-Gupta efficiency (KGE) (Knoben et al., 2019):

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (11)$$

where n is the number of data samples; $Q_{i,t}$ is the observed outflow; $\bar{Q}_{i,t}$ is the mean of outflow observations; $\hat{Q}_{i,t}$ is the simulated outflow; $\alpha = \frac{\sigma_s}{\sigma_o}$ and $\beta = \frac{\mu_s}{\mu_o}$, σ_s and σ_o are the standard deviation of the simulated outflow and the observed outflow, respectively; μ_s and μ_o are the mean value of the simulated outflow and the observed outflow, respectively; r is the correlation coefficient of the simulated outflow and the observed outflow. A KGE value of 1 indicates the best physical consistency fit which can compensate for the NSE sensible to high flow and the slow response to low flow rate (Liu, 2020; Zhou et al., 2022)

2.3.2. Physical consistency

The water balance index (WBI) (Deng et al., 2019) and the percentage of negative flow (PONF) are used to evaluate the physical consistency of the model. WBI is a measure of bias between total simulated outflow and observed outflow. The more WBI is close to 1, the better the water balance of the model results. The WBI bias can be calculated as absolute values of WBI and 1.

(1) The water balance index (WBI) and the WBI bias:

$$WBI = \frac{\sum_{t=1}^n \hat{Q}_{i,t}}{\sum_{t=1}^n Q_{i,t}}, \quad WBI \text{ bias} = |1 - WBI| \quad (12)$$

(2) The percentage of negative flow (PONF):

Table 1

Hyperparameters for data-driven models.

Hyperparameter	Reproduction scenario			Prediction scenario		
	PG-LSTM	LSTM	GBRT	PG-LSTM	LSTM	GBRT
Epoch	200	200	\	200	200	\
Hidden size	32	32		32	32	
Hidden layer	2	2		2	2	
Dropout rate	0.2	0.2		0.2	0.2	
Optimizer	ADAM	ADAM		ADAM	ADAM	
Batch size	365	365		365	365	
Lag time (days)	20	20		20	20	
λ_{data}	1	1		1	1	
$\lambda_{balance}$	0.0001	0.0001		0.0001	0.0001	
$\lambda_{monotonicity}$	0.0001	0.0001		0.0001	0.0001	
$\lambda_{upper.boundary}$	3	3		10	10	
$\lambda_{lower.boundary}$	0.3	0.3		0.3	0.3	
Learning rate	0.005	0.005	0.1	0.005	0.005	0.1
Maximum tree depth	\		2	\		2
Number of trees			30			30
Normalization	Min-max					

$$PONF = \frac{\sum_{t=1}^n c_t}{n} \times 100\%, \quad c_t = \begin{cases} 1, & \text{if } \hat{Q}_{i,t} < 0 \\ 0, & \text{if } \hat{Q}_{i,t} \geq 0 \end{cases} \quad (13)$$

where c_i is the number of the simulated negative outflow.

2.3.3. Model settings

Different model settings are trained and tested to identify the optimal hyperparameter. The optimization of hyperparameters can refer to the previous study in Zhang et al. (2018a). The final hyperparameters are chosen by obtaining the optimal NSE of the cascade reservoir, as shown in Table 1. The LSTM, PG-LSTM, and GBRT are developed based on the PyTorch framework of Python. The CRO is implemented based on the C programming language.

3. Case study

3.1. Study area

The cascade reservoir used in this study is the Qingjiang (QJ) cascade reservoir, located in the middle and lower reaches of the QJ river basin (as shown in Fig. 5). The QJ river basin covers an area of 17,600 km² in the subtropical region between East Longitude 108.35°~111.35° and North Latitude 29.33°~30.50°. Shuibuya, Geheyan, and Gaobazhou hydropower stations (hereafter, SBY, GHY, and GBZ, respectively) are constructed in the middle and lower reaches of the QJ river basin. SBY has a multi-year regulation ability, which has a drainage area of 10,860 km² with a total storage capacity of 4.58 billion m³, and the surface area of SBY is 64.35 km² at the normal water level. GHY has an annual regulation ability, which has a drainage area of 14,430 km² with a total storage capacity of 3.34 billion m³, and the surface area of GHY is 67.94 km² at the normal water level. GBZ is the reverse regulating hydropower station of GHY with daily regulation ability, which has a drainage area of 15,650 km² with a total storage capacity of 489 million m³, and the surface area of GBZ is 31 km² at the normal water level. Since the time scale of this study is daily, only SBY and GHY reservoirs are considered.

3.2. Data description

The basic information and operation data of SBY and GHY reservoirs are obtained from Hubei Energy Group Co., Ltd. The day index, daily inflow, reservoir water level, and outflow of SBY and GHY are shown in Fig. 6.

The models are trained by information comprising day index, cascade reservoir predicted inflow of the current and the next three days, cascade reservoir outflow in the previous day, cascade reservoir water level of the previous day, and the mean water level of cascade reservoir. The day index ranges from 1 to 366 with only one of the 366 inputs corresponding to the current day. For example, if the date of the day is 10 February, the day index is 41. The mean water level is the mean of the operational water level on a daily scale, which is an attribute of reservoirs used as static inputs at each time step.

Data division is performed by a five-fold cross-validation method. Data normalization is accomplished using min-max normali-

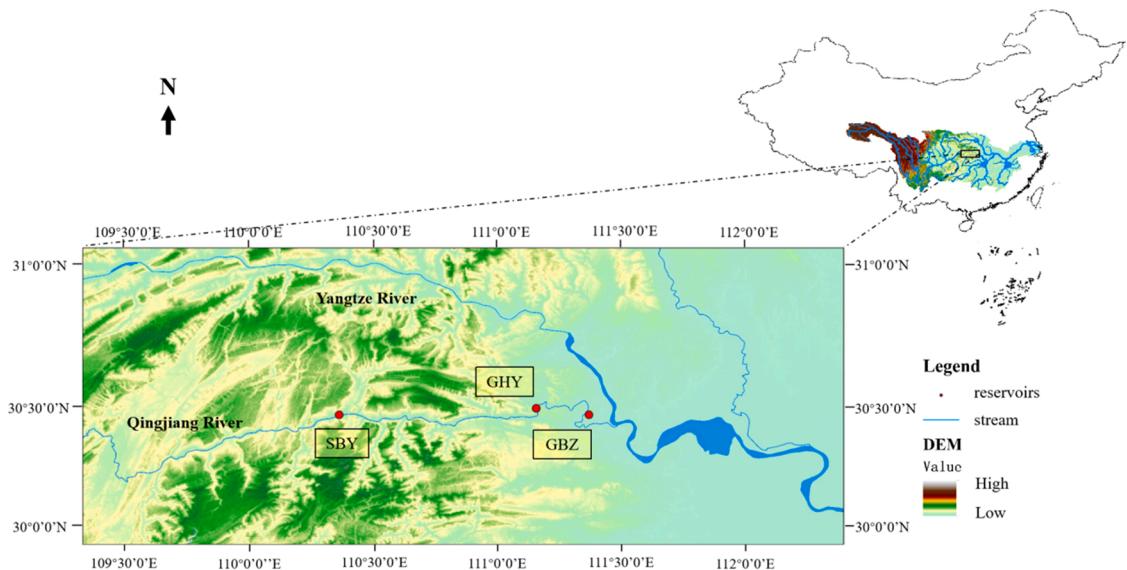


Fig. 5. The location of Qingjiang (QJ) cascade reservoir.

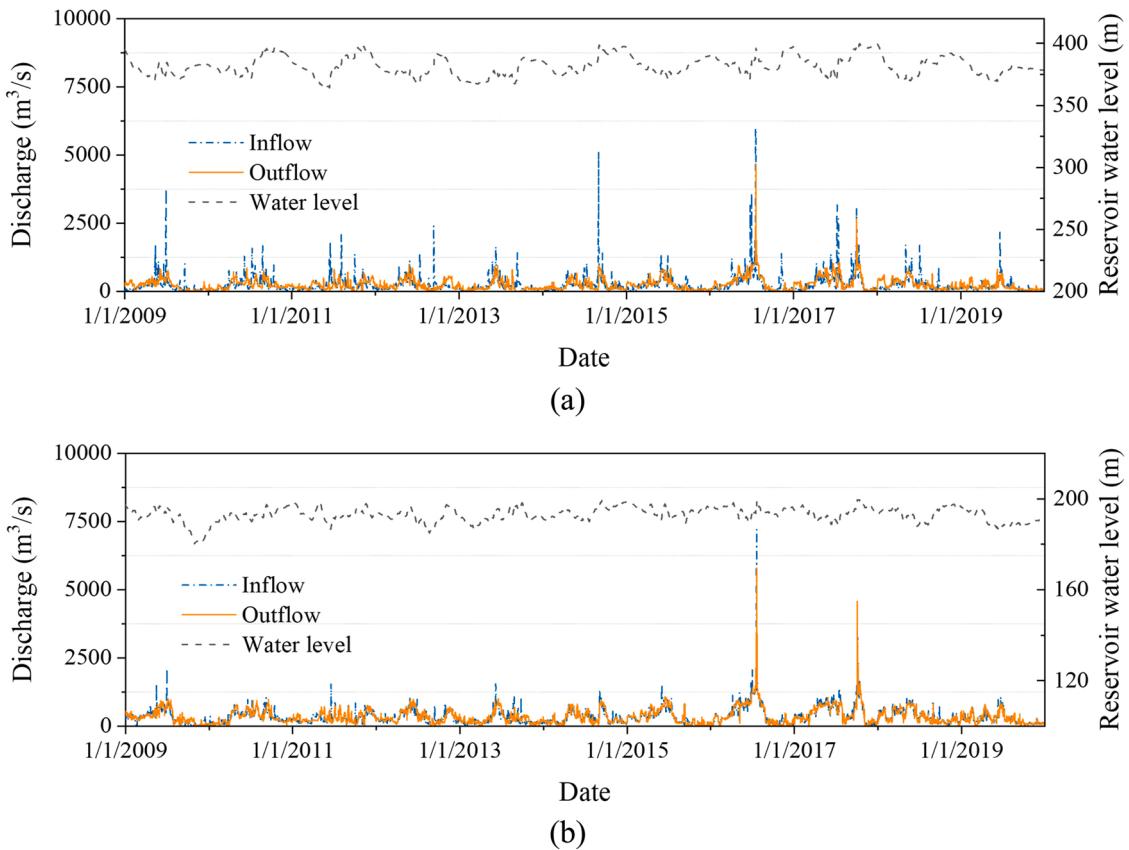


Fig. 6. The daily reservoir operation data of SBY(a), GHY(b), including the water level, inflow, and outflow.

zation. The normalized equation is as follows:

$$x_i = \frac{x_{i,\text{ori}} - x_{i,\text{min}}}{x_{i,\text{max}} - x_{i,\text{min}}} \quad (14)$$

where x_i are the standardized values for variables; $x_{i,\text{ori}}$ are the original values for variables; $x_{i,\text{min}}$, $x_{i,\text{max}}$ are the minimum and the

Table 2

The simulation accuracy of LSTM, PG-LSTM, GBRT, and CRO in the reproduction and prediction scenarios.

Scenario	Reservoir	Evaluation metrics	LSTM Test period	PG-LSTM Test period	GBRT Test period	CRO
Reproduction scenario: The reproduction of historical outflow	SBY	NSE	0.47	0.50	0.15	0.19
		RMSE (m^3/s)	148.15	135.50	194.88	187.61
		KGE	0.54	0.60	0.32	0.40
	GHY	NSE	0.52	0.58	0.24	0.15
		RMSE (m^3/s)	182.36	157.86	236.46	247.63
		KGE	0.63	0.66	0.43	0.38
	QJ cascade reservoir	NSE	0.50	0.54	0.20	0.17
		RMSE (m^3/s)	165.26	146.68	215.67	217.62
		KGE	0.59	0.63	0.38	0.39
Prediction scenario: The prediction of outflow with leading times	SBY	NSE	0.84	0.84	0.20	0.19
		RMSE (m^3/s)	84.4	82.02	189.26	187.61
		KGE	0.89	0.85	0.31	0.40
	GHY	NSE	0.84	0.85	0.31	0.15
		RMSE (m^3/s)	113.32	100.72	228.62	247.63
		KGE	0.86	0.88	0.42	0.38
	QJ cascade reservoir	NSE	0.84	0.85	0.26	0.17
		RMSE (m^3/s)	98.86	91.37	208.94	217.62
		KGE	0.88	0.87	0.37	0.39

maximum values of variables, respectively.

4. Results

The proposed method was applied to reproduction and prediction scenarios, respectively. First, the simulation results and the evaluation metrics were compared by LSTM, PG-LSTM, CRO, and GBRT (Section 4.1). Second, the simulation of high, intermediate, and low outflow was presented to explain how the synthetic samples affected the model performance (Section 4.2). After that, the reduction of the negative flow and the improvement of water balance by physical constraints was clarified (Section 4.3). Finally, the Applicability of reservoir operation by LSTM, PG-LSTM, CRO, and GBRT in the reproduction and prediction scenarios was analyzed (Section 4.4).

4.1. Comparison of simulation results of models

This section compares the simulation outflow of the cascade reservoir by LSTM, PG-LSTM, GBRT, and CRO. Table 2 summarizes the simulation accuracy of PG-LSTM compared with the results of LSTM, GBRT, and CRO. LSTM, GBRT, and CRO were trained as the benchmark model. As seen from Table 1, PG-LSTM had a higher simulation accuracy in the reproduction scenario in the test period: The average NSE of the cascade reservoir was 0.54, which was 0.04, 0.34, and 0.37 higher than the conventional LSTM, GBRT, and CRO models. In the prediction scenario, the average NSE of the cascade reservoir was 0.85, which was 0.01, 0.59, and 0.68 higher than the conventional LSTM, GBRT, CRO, respectively. Compared with deep learning models, the conventional learning models failed to absorb the data of reservoir operation. Compared with LSTM, the improvement of PG-LSTM lied in the physical mechanisms added into the network when the model inputs less information in the reproduction scenario. As a novel method, the physical constraints and the synthetic samples were added to the LSTM network to improve the model performance. It is indicated that the physical mechanisms can guide the model when less information is input. When the input information increased, the knowledge of reservoir operation guided the model less than the scenario with fewer data and information input.

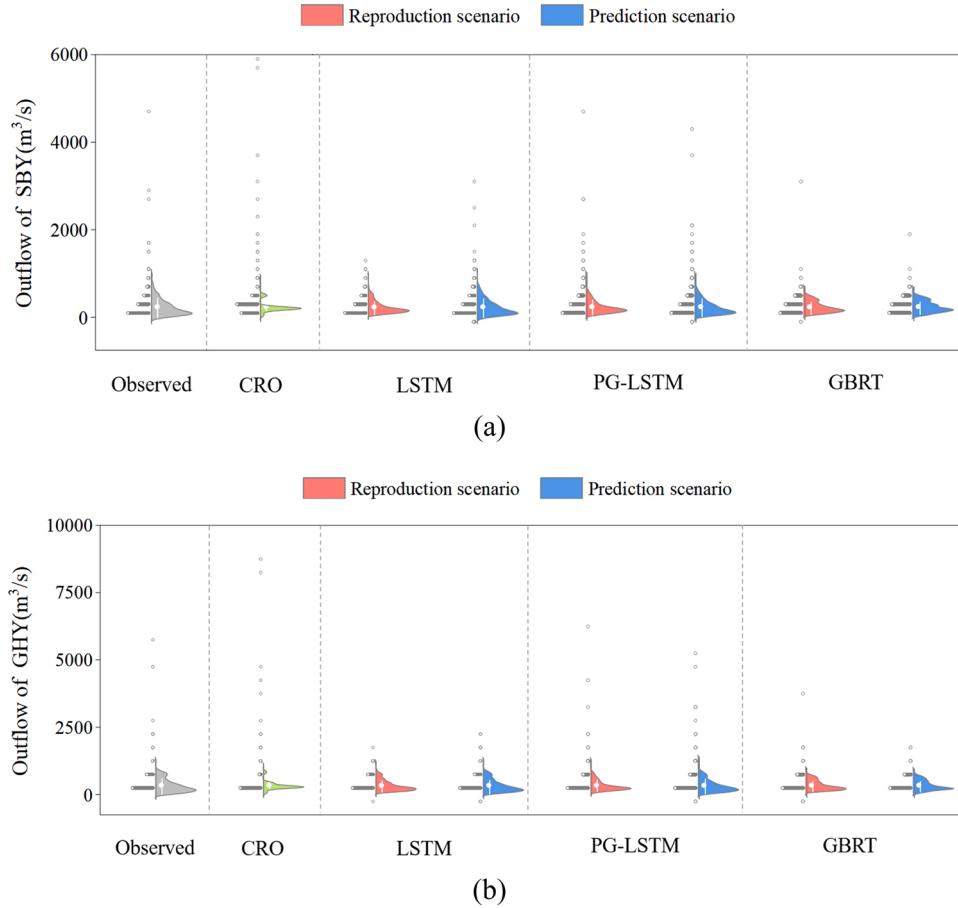


Fig. 7. The kernel density plot for analyzing the distribution of SBY (a) and GHY (b) outflow simulated by CRO, LSTM, PG-LSTM, and GBRT in the test period.

To evaluate the effectiveness of the proposed approach, the simulated outflow distribution of CRO, LSTM, PG-LSTM, and GBRT was compared during the test period, as shown in Fig. 7. The distributions of the model results showed concentrated characteristics, but the characteristics were not identical. The results of the observed outflow, LSTM, PG-LSTM, and GBRT models showed a single-peaked distribution shape while the results of CRO showed a double-peaked distribution shape. For outflow less than $2500 \text{ m}^3/\text{s}$, the distributions of LSTM and PG-LSTM results were closer to that of the observed outflow. However, LSTM was difficult to realize the simulation of high flow while PG-LSTM can enable high flow simulation. Thus, the distribution of PG-LSTM results was closest to that of the observed outflow. Meanwhile, Fig. 7a and Fig. 7b show the consistent model performance in SBY and GHY reservoirs. Therefore, only the simulation results of SBY were analyzed in detail in the following sections.

4.2. Effect of synthetic samples based on the design flood

In the QJ cascade reservoir system, the outflow of SBY and GHY was in the range of 0– $1500 \text{ m}^3/\text{s}$, as can be seen from Fig. 6. Due to the inadequate high inflow of actual operation, the existing operating data mostly included reservoir operation behaviors with normal operating conditions. It is insufficient to handle the rare and intense flood events. One approach to tackle the above problem of imbalanced data is to synthesize the high flow operating cases. By calculating the design flood from 20% to 0.01%, some high flow samples can be generated.

As shown in Fig. 8, GBRT presented poor flow simulation in reproduction and prediction scenarios. LSTM showed poor flow simulation in the reproduction scenario while the high flow simulation in the prediction scenario was more effective than that of the reproduction scenario by considering the outflow and reservoir water level of the previous day with the simulated peak flow of $2050 \text{ m}^3/\text{s}$. PG-LSTM can effectively simulate the high flow with the simulated peak flow of $4750 \text{ m}^3/\text{s}$, $4385 \text{ m}^3/\text{s}$ in reproduction and prediction scenarios, respectively. Compared with other models, PG-LSTM was most consistent with the observed value.

The simulated peak flow of CRO was $5940 \text{ m}^3/\text{s}$, and the observed value of peak flow was $4660 \text{ m}^3/\text{s}$. The relative error of CRO peak was 0.27 while the relative error of PG-LSTM peak in reproduction and prediction scenarios was 0.02 and 0.06, respectively. Thus, PG-LSTM outperformed CRO.

PG-LSTM outperformed LSTM in extracting high flow since the relative error of peak of LSTM was larger and the peak flow timing of LSTM lagged. Specifically, the relative error of peak was reduced using PG-LSTM compared with the conventional LSTM model: the relative error of peak decreased from 0.80 to 0.02 in the reproduction scenario, from 0.55 to 0.06 in the prediction scenario. The performance of LSTM in high flow simulation was poor due to the rarity of high flow samples in the training dataset. Thus, the synthetic high flow samples assist in providing more information for deep learning models.

However, PG-LSTM was less effective in simulating intermediate-low flows, as shown in Fig. 9. After the high flow samples were added to the model training dataset, the model paid less attention to the intermediate-low flow, resulting in the overestimation performance. In actual reservoir operations, outflow decisions for extreme hydrological conditions are critical for reservoir operation. Although the addition of design flood operating samples would make the model focus less on the intermediate-low flow, it is still recommended to add the operating samples of the design flood for the model training.

4.3. Effect of physical constraints derived from reservoir operation

To improve the physical consistency of deep learning networks, some physical constraints were added to the loss function. As shown in Table 3, the mean WBI bias of the QJ cascade reservoir for LSTM, PG-LSTM, GBRT, and CRO in the reproduction scenario were 0.03, 0.01, 0.11, and 0.04, which indicated that the water balance of the PG-LSTM model was more desirable. Similarly, this viewpoint can be drawn from the prediction scenario.

The conventional LSTM for outflow simulation appeared negative flow in the prediction scenario. The mean PONF of the QJ cascade reservoir for LSTM and PG-LSTM in the prediction scenario were 1.03 and 0.14, respectively. The results of PONF suggested

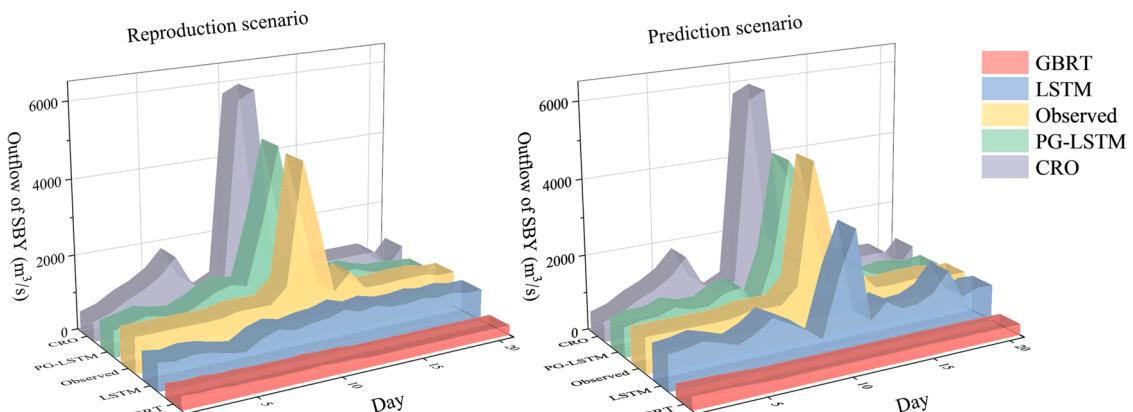


Fig. 8. Improved outflow peak simulation by synthetic samples based on the design flood in SBY reservoir from 7/10/2016 to 7/30/2016.

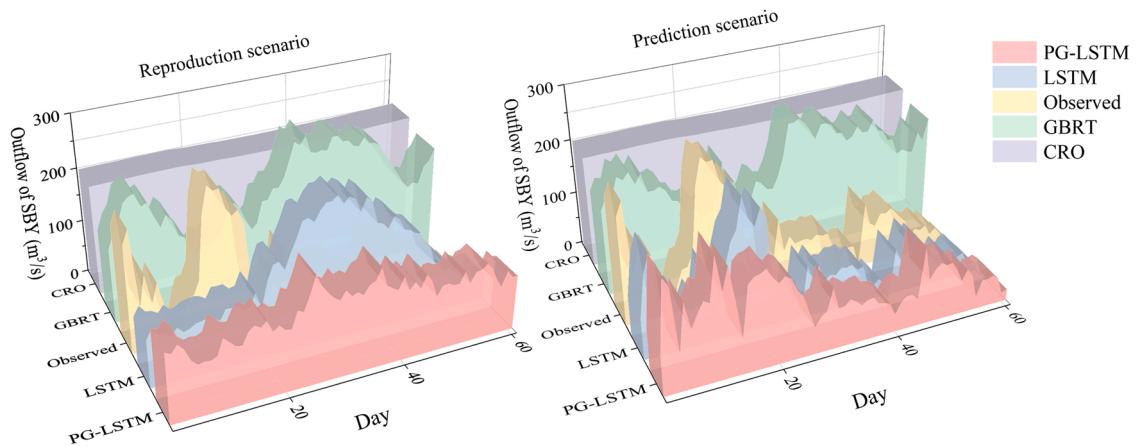


Fig. 9. Comparison of simulated intermediate-low flow in SBY reservoir from 9/1/2015 to 10/31/2015.

Table 3

The physical consistency of LSTM, PG-LSTM, GBRT, and CRO in the reproduction and prediction scenarios.

Scenario	Reservoir	Evaluation metrics	LSTM Test period	PG-LSTM Test period	GBRT Test period	CRO
Reproduction scenario: The reproduction of historical outflow	SBY	WBI bias	0.03	0.00	0.11	0.02
		PONF (%)	0.00	0.00	0.00	0.00
	GHY	WBI bias	0.03	0.01	0.11	0.05
		PONF (%)	0.03	0.00	0.00	0.00
	QJ cascade reservoir	WBI bias	0.03	0.01	0.11	0.04
		PONF (%)	0.02	0.00	0.00	0.00
Prediction scenario: The prediction of outflow with leading times	SBY	WBI bias	0.02	0.00	0.12	0.02
		PONF (%)	1.98	0.20	0.00	0.00
	GHY	WBI bias	0.02	0.01	0.11	0.05
		PONF (%)	0.08	0.08	0.00	0.00
	QJ cascade reservoir	WBI bias	0.02	0.01	0.12	0.04
		PONF (%)	1.03	0.14	0.00	0.00

that PG-LSTM can effectively reduce the occurrence of negative flow.

The results showed that the addition of physical constraints can improve model physical consistency. The proposed PG-LSTM network can decrease the water balance bias and reduce the occurrence of negative flow.

4.4. Applicability of models in the reproduction and prediction scenarios

The trained models were applied to the reproduction and prediction scenarios and the performance of the model was evaluated.

4.4.1. Reproduction scenario: the reproduction of cascade reservoir historical outflow

Fig. 10 is the comparison of the reproduced outflow using LSTM, PG-LSTM, GBRT, and CRO. After the models extracted the experience of reservoir operators, the historical information was applied to the model to obtain outflow reproduced by LSTM, PG-LSTM, GBRT, and CRO. As can be seen from Fig. 9, the results of LSTM and GBRT cannot accommodate the seasonal variation of reservoir inflow and presented auto-correlation characteristics. The results of CRO were similar to that of PG-LSTM, which meant that

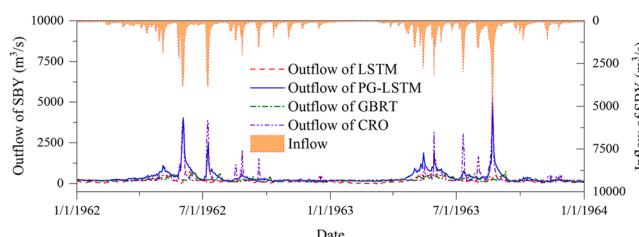


Fig. 10. Comparison of the reproduction historical outflow from 1/1/1962 to 1/1/1964 using LSTM, PG-LSTM, GBRT, and CRO.

PG-LSTM can infer the rule curves used by reservoir operators. The results of PG-LSTM were better than that of CRO. For example, when the inflow was less than 2500 m³/s, PG-LSTM tended to produce a smaller outflow, which was consistent with the operating experience of the QJ cascade reservoir operators. The reasons for the difference between simulation results of CRO and PG-LSTM can be explained as follows: CRO is usually used for the preparation of the annual reservoir operating plan, which cannot consider the forecast information of inflow, resulting in conservative operating decisions. According to Fig. 10, CRO tended to release more flow during the flood season to make a safe decision and preferred releasing less outflow to ensure the realization of reservoir functions in the non-flood season. Considering the forecast information of inflow, the results of PG-LSTM were closer to the actual operation than those of CRO.

Further, the daily reservoir water level and power generation can be obtained by using the simulated outflow of PG-LSTM. With the data of 1 May 1962 to 30 April 1963 as an example, the results showed that the total power generation of the operation period of CRO for the QJ cascade reservoir was 7.60 billion kWh while that of PG-LSTM was 9.26 billion kWh. In contrast, the total power generation of the cascade reservoir calculated by the deterministic optimization operation is 9.46 billion kWh, indicating that PG-LSTM can be used as a better benchmark compared with the CRO method for evaluation of the optimal operation.

4.4.2. Prediction scenario: the prediction of cascade reservoir outflow with several leading times

Table 4 shows the NSE of SBY and GHY outflow simulations. As can be seen from Table 4, the deep learning model outperformed the extraction of reservoir operation behaviors compared to GBRT and CRO. As for the deep learning models, PG-LSTM was better than LSTM in the three lead times with NSE of PG-LSTM higher than that of LSTM in the test period. In addition, the improvement of NSE of PG-LSTM in the test period was greater as the lead time increased, and the improvement percentage of PG-LSTM to LSTM was 4.66% with three days ahead.

Fig. 11 is the comparison of the outflow predicted by LSTM, PG-LSTM, and GBRT with several lead times. As can be seen from Fig. 11, the performance of PG-LSTM decreased more slowly than LSTM and GBRT with the increase of lead times. Fig. 12 shows the results of high flow prediction with several lead times using LSTM, PG-LSTM, and GBRT. From Fig. 12, it can be seen that there was a serious lag in the predicted outflow of LSTM as the lead time increased while PG-LSTM can still ensure accurate high-flow predictions. Therefore, PG-LSTM is more robust than LSTM.

5. Discussion

5.1. Verification of the physical mechanism effects

From Section 4, it was known that synthetic samples using the design flood operating case improved the peak outflow simulation and the loss function revised by physical constraints enhanced the model physical consistency. Two new experiments were conducted in Section 5.1, in which the above two physical mechanisms were added to LSTM separately to verify the effects of the physical mechanism.

Table 5 is the comparison of model performance in extracting cascade reservoir outflow behaviors using different modification ways with LSTM. Table 5 illustrates that adding the operating cases of the design flood can gain higher NSE than revising the loss function with physical constraints, which meant the former is more conducive to improving the NSE of the model. However, only considering the synthetic samples method cannot fully satisfy the physical constraints of reservoir operation. For example, the results of LSTM with synthetic samples had more negative flow and less water balance satisfaction than PG-LSTM. Thus, physical constraints are beneficial to limit the solution space.

Table 4
Comparison of overall model performance in extracting cascade reservoir outflow behaviors with different lead times.

Evaluation metric	Lead time	Reservoir	LSTM	PG-LSTM	GBRT	CRO	Improvement percentage of PG-LSTM to LSTM (%)
			Test period	Test period	Test period		
NSE	The current day	SBY	0.84	0.84	0.20	0.19	0.00
		GHY	0.84	0.85	0.31	0.15	1.19
		QJ cascade reservoir	0.84	0.85	0.26	0.17	0.60
	One day ahead	SBY	0.74	0.74	0.18	0.19	0.00
		GHY	0.76	0.78	0.30	0.15	2.63
		QJ cascade reservoir	0.75	0.76	0.24	0.17	1.32
	Two days ahead	SBY	0.66	0.69	0.17	0.19	4.55
		GHY	0.70	0.73	0.28	0.14	4.29
		QJ cascade reservoir	0.68	0.71	0.23	0.17	4.42
	Three days ahead	SBY	0.62	0.65	0.15	0.19	4.84
		GHY	0.67	0.70	0.26	0.14	4.48
		QJ cascade reservoir	0.65	0.68	0.21	0.17	4.66

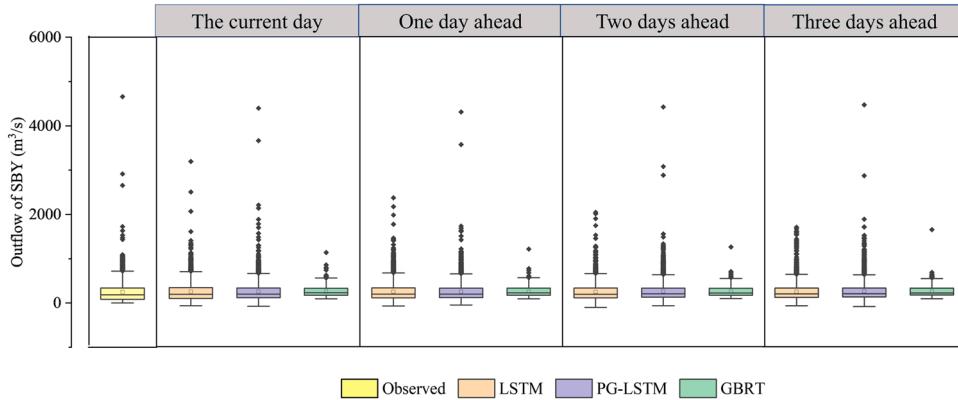


Fig. 11. Comparison of the predicted outflow and the observed outflow with several lead times.

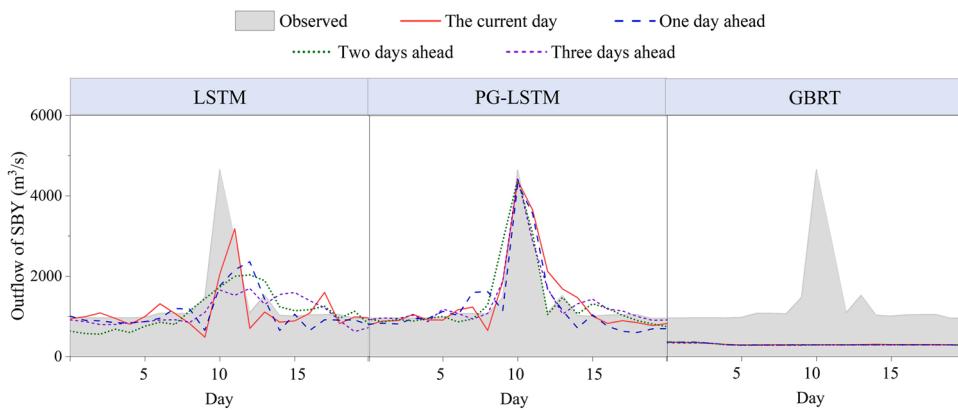


Fig. 12. Comparison of high flow prediction with several lead times from 7/10/2016 to 7/30/2016.

Consequently, two reasons can explain why PG-LSTM is more effective than LSTM in the cascade reservoir operation simulation as follows.

- (1) The design flood operating cases provide knowledge to LSTM when the extreme flood occurs.
- (2) The penalty terms constraint the range of outflow and promote the model to better reflect the reservoir operation behaviors.

The two physical mechanisms, one is from the perspective of the domain prior knowledge to guide the model, and the other is to realize the guidance of the physical mechanism to the deep learning model by increasing the sample richness. Such a physical mechanism-guided method can be applied to other research areas: building loss function or activation function containing domain control equations, variable constraints; obtaining more training samples from physical models to increase the richness of samples. The two physical mechanisms can be used to build physically guided data-driven models for different domains.

5.2. Representation of the hydraulic connection in cascade reservoirs for deep learning models

Hydraulic connection is a key element in the cascade reservoir operation. The proposed models in this study took the cascade reservoir as a whole, in which the model inputs were the inflow of the upstream reservoir, the local flow between cascade reservoir, etc., and outputs were the outflow of each reservoir in the cascade reservoir. In this way, the hydraulic connection has been considered in the loss function of PG-LSTM when the model simultaneously outputs the outflow of SBY and GHY. Thus, the hydraulic connection between the upstream and downstream reservoir can be explicitly included in the proposed model. A new experiment (Experiment 2) was added in Section 5.2 to compare the original experiment (Experiment 1). The new experiment separated the reservoirs of the cascade reservoir to obtain the outflow of the single reservoir, and the hydraulic connection was calculated step by step from upstream to downstream. The comparative analysis of simulation results of Experiment 1 and Experiment 2 was as follows.

5.2.1. Experiment 1: a cascade reservoir mode

In Experiment 1, the hydraulic connection was presented in the loss function. The results of Experiment 1 have been shown in

Table 5

Comparison of model performance in extracting cascade reservoir outflow using different modification methods with LSTM.

Scenario	Reservoir	Evaluation metrics	LSTM with synthetic samples	LSTM with physical constraints	PG-LSTM
			Test period		Test period
			Test period		
Reproduction scenario: The reproduction of historical outflow	SBY	NSE	0.49	0.45	0.50
		RMSE(m^3/s)	138.40	151.07	135.50
		KGE	0.63	0.54	0.60
		WBI bias	0.03	0.02	0.00
		PONF (%)	0.00	0.00	0.00
		NSE	0.57	0.52	0.58
	GHY	RMSE(m^3/s)	160.22	184.55	157.86
		KGE	0.68	0.61	0.66
		WBI bias	0.03	0.01	0.01
		PONF (%)	0.00	0.00	0.00
		NSE	0.53	0.49	0.54
		RMSE(m^3/s)	149.31	167.81	146.68
Prediction scenario: The prediction of outflow with leading times	SBY	KGE	0.66	0.58	0.63
		WBI bias	0.03	0.02	0.01
		PONF (%)	0.00	0.00	0.00
		NSE	0.84	0.84	0.84
		RMSE(m^3/s)	81.82	86.16	82.02
		KGE	0.88	0.86	0.85
	GHY	WBI bias	0.02	0.01	0.00
		PONF (%)	1.43	0.53	0.20
		NSE	0.86	0.85	0.85
		RMSE(m^3/s)	100.49	108.79	100.72
		KGE	0.88	0.86	0.88
		WBI bias	0.04	0.00	0.01
	QJ cascade reservoir	PONF (%)	0.65	0.00	0.08
		NSE	0.85	0.85	0.85
		RMSE(m^3/s)	91.15	97.48	91.37
		KGE	0.88	0.86	0.87
		WBI bias	0.03	0.01	0.01
		PONF (%)	1.04	0.27	0.14

Table 2 and **Table 3**. The models took the information of cascade reservoir (for instance, the inflow of SBY, local flow between SBY and GHY, etc.) as the input, and output the outflow of SBY and GHY simultaneously.

5.2.2. Experiment 2: a single reservoir mode

In Experiment 2, models were applied independently to each reservoir, and the hydraulic connection was calculated step by step

Table 6

The simulation accuracy of Experiment 1 and Experiment 2 in the reproduction and prediction scenarios.

Scenario	Reservoir	Evaluation metrics	Experiment 1		Experiment 2	
			LSTM	PG-LSTM	LSTM	PG-LSTM
			Test period	Test period	Test period	Test period
Reproduction scenario: The reproduction of historical outflow	SBY	NSE	0.47	0.50	0.48	0.52
		RMSE (m^3/s)	148.15	135.50	144.73	133.87
		KGE	0.54	0.60	0.57	0.61
		NSE	0.52	0.58	0.54	0.59
		RMSE (m^3/s)	182.36	157.86	176.45	157.01
		KGE	0.63	0.66	0.62	0.68
	GHY	NSE	0.50	0.54	0.51	0.56
		RMSE (m^3/s)	165.26	146.68	160.59	145.44
		KGE	0.59	0.63	0.60	0.65
		NSE	0.84	0.84	0.85	0.84
		RMSE (m^3/s)	84.4	82.02	82.25	81.03
		KGE	0.89	0.85	0.88	0.87
Prediction scenario: The prediction of outflow with leading times	SBY	NSE	0.84	0.84	0.85	0.84
		RMSE (m^3/s)	84.4	82.02	82.25	81.03
		KGE	0.89	0.85	0.88	0.87
		NSE	0.84	0.85	0.86	0.86
		RMSE (m^3/s)	113.32	100.72	105.14	99.49
		KGE	0.86	0.88	0.89	0.89
	GHY	NSE	0.84	0.85	0.86	0.85
		RMSE (m^3/s)	98.86	91.37	93.70	90.26
		KGE	0.88	0.87	0.89	0.88
		NSE	0.84	0.85	0.86	0.85
		RMSE (m^3/s)	98.86	91.37	93.70	90.26
		KGE	0.88	0.87	0.89	0.88

from upstream to downstream, based on the simulated outflow of the upstream reservoir. The hydraulic connection can be formulated as follows:

$$I_{GHY,t} = I'_{SBY-GHY,t} + \hat{Q}_{SBY,t} \quad (15)$$

where $I_{GHY,t}$ is the inflow of GHY; $I'_{SBY-GHY,t}$ is the local flow between SBY and GHY; $\hat{Q}_{SBY,t}$ is the simulated outflow of SBY.

The models for SBY only input the operational information of SBY and output SBY outflow. Similarly, the models for GHY input the operational information of GHY. Specifically, the inflow of GHY was calculated according to Eq. (15), and then GHY models output the outflow of GHY.

From Table 6, it can be seen that Experiment 1 and Experiment 2 have similar NSE performance in the test period. As PG-LSTM that listed in Table 7, Experiment 1 obtains a better physical consistency with PONF of 0.14 and WBI bias of 0.01 in the prediction scenario while that of Experiment 2 is 0.21 and 0.01, respectively. Because Experiment 1 owns better physical consistency and can output the outflow of the cascade reservoir simultaneously, the model structure of Experiment 1 is recommended in this study.

6. Conclusion

This study investigated the effectiveness of PG-LSTM to extract cascade reservoir operation behaviors. The simulation accuracy and physical consistency of PG-LSTM were evaluated in reproduction and prediction scenarios. The main conclusions are as follows:

- (1) This study provides a reservoir operation behavior mining method, which can be used for two scenarios: one is to reproduce the historical cascade reservoir outflow with significant seasonal variations, and the other is to predict the outflow with several lead times without lags.
- (2) To achieve accurate simulation of reservoir operators' experience, physical mechanisms of reservoir operation are integrated into LSTM. The proposed physical mechanism of reservoir operation comprises two parts: one is to add the design flood operating samples into the training dataset of LSTM, the other is to utilize the water balance equation, the monotonic relationship between inflow and outflow, and the boundary constraint of outflow.
- (3) Considering the operating samples based on the design flood and the historical operation data as the model training dataset, PG-LSTM can enhance the understanding of LSTM for the operation behaviors under extreme hydrological conditions with more accurate high flow simulation. Moreover, using the water balance constraint, boundary constraints, and monotonic constraint can help PG-LSTM effectively reduce the occurrence of the negative flow and decrease the bias of the water balance index, making the mining operation behavior more consistent with the basic principles of reservoir operation. Overall, compared with LSTM, GBRT, and CRO, PG-LSTM can improve NSE of cascade reservoir during the test period from 0.50, 0.20, 0.17 to 0.54 in the reproduction scenario and from 0.84, 0.26, and 0.17 to 0.85 in the prediction scenario with five-fold cross-validation method.

The proposed PG-LSTM is helpful to describe human operation behaviors, which is capable of providing benefits for cascade reservoir operations and changing the situation that reservoir operation depends on manual experience. Further research should concentrate on expanding the employment of physically guided data-driven models in water resources utilization.

CRediT authorship contribution statement

Yalian Zheng: Conceptualization, Methodology, Software, Writing – original draft. **Pan Liu:** Conceptualization, Writing – review & editing. **Lei Cheng:** Supervision. **Kang Xie:** Methodology. **Wei Lou:** Visualization. **Xiao Li:** Software. **Xinran Luo:** Editing. **Qian**

Table 7

The physical consistency of Experiment 1 and Experiment 2 in the reproduction and prediction scenarios.

Scenario	Reservoir	Evaluation metrics	Experiment 1		Experiment 2	
			LSTM Test period	PG-LSTM Test period	LSTM Test period	PG-LSTM Test period
Reproduction scenario: The reproduction of historical outflow	SBY	WBI bias	0.03	0.00	0.03	0.02
		PONF (%)	0.00	0.00	0.00	0.00
	GHY	WBI bias	0.03	0.01	0.03	0.03
		PONF (%)	0.03	0.00	0.00	0.00
	QJ cascade reservoir	WBI bias	0.03	0.01	0.03	0.03
		PONF (%)	0.02	0.00	0.00	0.00
Prediction scenario: The prediction of outflow with leading times	SBY	WBI bias	0.02	0.00	0.00	0.00
		PONF (%)	1.98	0.20	1.03	0.33
	GHY	WBI bias	0.02	0.01	0.01	0.02
		PONF (%)	0.08	0.08	0.13	0.08
	QJ cascade reservoir	WBI bias	0.02	0.01	0.01	0.01
		PONF (%)	1.03	0.14	0.58	0.21

Cheng: Editing. **Dongyang Han:** Software, Validation. **Wei Zhang:** Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data and codes generated or used in the study are available from the corresponding author by request.

Acknowledgments

This study was supported by the National Natural Science Foundation of China, China (Grant No. 51861125102, and Grant No. 52109011), the Joint Funds of the National Natural Science Foundation of China, China (Grant No. U1865201), and China Three Gorges Corporation Research Project, China (Contract No: 202103511). The authors would like to thank the editor and the anonymous reviewers for their comments, which helped significantly in advancing the quality of the study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ejrh.2022.101034](https://doi.org/10.1016/j.ejrh.2022.101034).

References

- Ahmad, A., El-Shafie, A., Razali, S.F.M., Mohamad, Z.S., 2014. Reservoir optimization in water resources: a review. *Water Resour. Manag.* 28 (11), 3391–3405. <https://doi.org/10.1007/s11269-014-0700-5>.
- Bayoumi, A., Gomaa, E., Hamdy, A., 2019. Heterogeneous reservoir characterization (Upper Bahariya Case study). *Acad. Res. Community Publ.* 2 (4), 514.
- Carvalho, T., Filho, F., 2021. Variational mode decomposition hybridized with gradient boost regression for seasonal forecast of residential water demand. *Water Resour. Manag.* 1–15.
- Cheng, Q., Ming, B., Liu, P., Huang, K., Gong, Y., Li, X., Zheng, Y., 2021. Solving hydro unit commitment problems with multiple hydraulic heads based on a two-layer nested optimization method. *Renew. Energy*. 172, 317–326. <https://doi.org/10.1016/j.renene.2021.02.126>.
- Corani, G., Rizzoli, A.E., Salvetti, A., Zaffalon, M., 2009. Reproducing human decisions in reservoir management: the case of lake Lugano. *Information Technologies in Environmental Engineering*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 252–263.
- Deng, C., Liu, P., Wang, W., Shao, Q., Wang, D., 2019. Modelling time-variant parameters of a two-parameter monthly water balance model. *J. Hydrol.* 573, 918–936.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 29 (5), 1189–1232.
- Gong, Y., Liu, P., Liu, Y., Huang, K., 2021a. Robust operation interval of a large-scale hydro-photovoltaic power system to cope with emergencies. *Appl. Energy* 290, 116612.
- Gong, Y., et al., 2021b. Deriving pack rules for hydro-photovoltaic hybrid power systems considering diminishing marginal benefit of energy. *Appl. Energy* 304, 117858. <https://doi.org/10.1016/j.apenergy.2021.117858>.
- Gong, Y., Liu, P., Ming, B., Feng, M., Huang, K., Wang, Y., 2022. Identifying the functional form of operating rules for hydro-photovoltaic hybrid power systems. *Energy*. 243, 123027. <https://doi.org/10.1016/j.energy.2021.123027>.
- Gori, M., Tesi, A., 1992. On the problem of local minima in backpropagation. *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (1), 76–86. <https://doi.org/10.1109/34.107014>.
- He, X., Luo, J., Li, P., Zuo, G., Xie, J., 2020. A hybrid model based on variational mode decomposition and gradient boosting regression tree for monthly runoff forecasting. *Water Resour. Manag.* 34 (2), 865–884. <https://doi.org/10.1007/s11269-020-02483-x>.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Zhou, L., Liu, P., Gui, Z., Zhang, X., Liu, W., Cheng, L., Xia, J., 2022. Diagnosing structural deficiencies of a hydrological model by time-varying parameters. *J. Hydrol.* 605, 127305. <https://doi.org/10.1016/j.jhydrol.2021.127305>.
- Hossain, M.S., El-shafie, A., 2013. Intelligent systems in optimizing reservoir operation policy: a review. *Water Resour. Manag.* 27 (9), 3387–3407. <https://doi.org/10.1007/s11269-013-0353-9>.
- Huang, K., Liu, P., Ming, B., Kim, J.-S., Gong, Y., 2021. Economic operation of a wind-solar-hydro complementary system considering risks of output shortage, power curtailment and spilled water. *Appl. Energy* 290, 116805. <https://doi.org/10.1016/j.apenergy.2021.116805>.
- Jiang, S., Zheng, Y., Solomatine, D., 2020. Improving AI system awareness of geoscience knowledge: symbiotic integration of physical approaches and deep learning. *Geophys. Res. Lett.*
- Knoben, W., Freer, J.E., Woods, R.A., 2019. Technical note: inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrol. Earth Syst. Sci.* 23 (10), 4323–4331.
- Kramer, O., 2016. Machine Learning for Evolution Strategies. Springer International Publishing.
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Nearing, G., 2019. Benchmarking a catchment-aware long short-term memory network (LSTM) for large-scale hydrological modeling. *Hydrol. Earth Syst. Sci. Discuss.* 1–32.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444. <https://doi.org/10.1038/nature14539>.
- Lehner, B., et al., 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Front. Ecol. Environ.* 9 (9), 494–502.
- Li, H., Liu, P., Guo, S., Cheng, L., Yin, J., 2020a. Climatic control of upper Yangtze River flood hazard diminished by reservoir groups. *Environ. Res. Lett.* 15 (12), 124013 <https://doi.org/10.1088/1748-9326/abc4fe>.
- Li, X., et al., 2020b. Reducing lake water-level decline by optimizing reservoir operating rule curves: a case study of the Three Gorges Reservoir and the Dongting Lake. *J. Clean. Prod.* 264, 15. <https://doi.org/10.1016/j.jclepro.2020.121676>.
- Li, X., et al., 2022. Derivation of operating rule curves for cascade hydropower reservoirs considering the spot market: a case study of the China's Qing River cascade-reservoir system. *Renew. Energy* 182, 1028–1038. <https://doi.org/10.1016/j.renene.2021.11.013>.
- Liu, D., 2020. A rational performance criterion for hydrological model. *J. Hydrol.* 590, 125488 <https://doi.org/10.1016/j.jhydrol.2020.125488>.

- Liu, D.D., et al., 2019. Variability of spatial patterns of autocorrelation and heterogeneity embedded in precipitation. *Hydrol. Res.* 50 (1), 215–230. <https://doi.org/10.2166/nh.2018.054>.
- Markus, et al., 2019. Deep learning and process understanding for data-driven Earth system science. *Nature*.
- Hochreiter, S., Schmidhuber, J., 1996. LSTM can solve hard long time lag problems. NIPS'96: Proceedings of the 9th International Conference on Neural Information Processing Systems. 473–479.
- Mayrink, V., Hippert, H.S., 2016. A hybrid method using Exponential Smoothing and Gradient Boosting for electrical short-term load forecasting. In: 2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI).
- Ming, B., Liu, P., Cheng, L., 2021. An integrated framework for optimizing large hydro-photovoltaic hybrid energy systems: capacity planning and operations management. *J. Clean. Prod.* 306, 127253 <https://doi.org/10.1016/j.jclepro.2021.127253>.
- Ming, B., et al., 2017. Optimizing utility-scale photovoltaic power generation for integration into a hydropower reservoir by incorporating long- and short-term operational decisions. *Appl. Energy* 204, 432–445. <https://doi.org/10.1016/j.apenergy.2017.07.046>.
- Mohammadifar, A., Gholami, H., Golzari, S., Collins, A.L., 2021. Spatial modelling of soil salinity: deep or shallow learning models? *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-021-13503-7>.
- Moussa, S., Fawzy, D., Badr, N., 2016. The evolution of data mining techniques to big data analytics: an extensive study with application to renewable energy data analytics. *Asian J. Appl. Sci.* 4 (3).
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles - ScienceDirect. *J. Hydrol.* 10 (3), 282–290.
- Persson, C., Bacher, P., Shiga, T., Madsen, H., 2017. Multi-site solar power forecasting using gradient boosted regression trees. *Sol. Energy* 150 (jul.), 423–436.
- Read, J.S., et al., 2019. Process-guided deep learning predictions of lake water temperature. *Water Resour. Res.* 55 (11), 9173–9190. <https://doi.org/10.1029/2019WR024922>.
- Rossum, G.V., Drake, F.L. (Eds.), 2012. *The Python Language Reference*. Manual Network Theory Ltd, p. 1.
- Shen, C., 2017. A trans-disciplinary review of deep learning research for water resources scientists. *Water Resour. Res.*
- Wang, N., Zhang, D., Chang, H., Li, H., 2019. Deep Learning of Subsurface Flow via Theory-guided Neural Network.
- Wang, Y., Liu, P., Dou, M., Li, H., Ming, B., Gong, Y., Yang, Z., 2021. Reservoir ecological operation considering outflow variations across different time scales. *Ecological Indicators* 125, 107582. <https://doi.org/10.1016/j.ecolind.2021.107582>.
- Wang, L., Tao, G., Wang, Z., 2011. Application research of data mining on reservoir characterization. International Conference on Electronic & Mechanical Engineering & Information Technology 1, 153–156. <https://doi.org/10.1109/EMEIT.2011.6022885>.
- Xie, K., et al., 2021. Physics-guided deep learning for rainfall-runoff modeling by considering extreme events and monotonic relationships. *J. Hydrol.* 603, 127043 <https://doi.org/10.1016/j.jhydrol.2021.127043>.
- Xu, W., Liu, P., Cheng, L., Zhou, Y., Liu, Y., 2021. Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renew. Energy* 163, 772–782.
- Yang, G., Guo, S., Liu, P., Li, L., Xu, C., 2017. Multiobjective reservoir operating rules based on cascade reservoir input variable selection method. *Water Resour. Res.* 53 (4), 3446–3463.
- Yang, G., Guo, S.L., Liu, P., Liu, X.F., Yin, J.B., 2020. Heuristic input variable selection in multi-objective reservoir operation. *Water Resour. Manag.* 34 (2), 617–636. <https://doi.org/10.1007/s11269-019-02456-9>.
- Yang, S., Yang, D., Chen, J., Zhao, B., 2019. Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. *J. Hydrol.* 579, 124229.
- Yang, T., Gao, X., Sorooshian, S., Li, X., 2016. Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme. *Water Resour. Res.* 52 (3) (n/a-n/a).
- Yang, Z., Liu, P., Cheng, L., Wang, H., Ming, B., Gong, W., 2018. Deriving operating rules for a large-scale hydro-photovoltaic power system using implicit stochastic optimization. *J. Clean. Prod.* 195 (SEP.10), 562–572.
- Yue, Q., Liu, F.S., Diao, Y.F., Liu, Y.M., 2018. Research and application of a big data-driven intelligent reservoir management system. *J. Coast. Res.* 270–279. <https://doi.org/10.2112/si82-039.1>.
- Zhang, D., et al., 2018a. Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. *J. Hydrol.* 565, 720–736. <https://doi.org/10.1016/j.jhydrol.2018.08.050>.
- Zhang, J., Cai, X., Lei, X., Liu, P., Wang, H., 2021Z. Real-time reservoir flood control operation enhanced by data assimilation. *J. Hydrol.* 598, 126426. <https://doi.org/10.1016/j.jhydrol.2021.126426>.
- Zhang, D., Lindholm, G., Ratnaweera, H., 2018b. Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. *J. Hydrol.* 556, 409–418. <https://doi.org/10.1016/j.jhydrol.2017.11.018>.
- Zhang, D., et al., 2019a. Simulating reservoir operation using a recurrent neural network algorithm. *Water* 11 (4). <https://doi.org/10.3390/w11040865>.
- Zhang, X., et al., 2019b. Real-time reservoir flood control operation for cascade reservoirs using a two-stage flood risk analysis method. *J. Hydrol.* 577, 123954 <https://doi.org/10.1016/j.jhydrol.2019.123954>.