applied sciences

MDPI

Article

Simulation of Reservoir Group Outflow Using LSTM with a Knowledge-Guided Loss Function Coordinated by the MDUPLEX Algorithm

Qiaoping Liu [1,2], Changlu Qiao [1,2,*] and Shuo Cao [3]

[1] College of Water Conservancy and Architectural Engineering, Shihezi University, Shihezi 832000, China; liuqiaoping@xjshzu.com
[2] Key Laboratory of Cold and Arid Regions Eco-Hydraulic Engineering, Xinjiang Production & Construction Corps, Shihezi 832000, China
[3] Sixth Division Water Resources Management Center, Wujiaqu 831300, China; caoshuo1991@126.com
* Correspondence: qiaochanglu@126.com

Abstract

Global climate change and spatiotemporal heterogeneity in water resources exacerbate supply-demand imbalances. Accurate outflow simulation for joint reservoir group operations thus becomes critical for scientific water resources management. Existing data-driven models like the Long Short-Term Memory (LSTM) lack the robust integration of physical constraints. Traditional mechanistic methods, by contrast, lack generality and stability under complex hydrological conditions. To address this limitation, we propose MDUPLEX-KG-LSTM—a physically constrained data-driven model for reservoir outflow simulation. The model incorporates multi-round DUPLEX (MDUPLEX) data partitioning, which ensures statistical homogeneity across training, validation, and test datasets. It also features a Knowledge-Guided (KG) loss function that embeds core physical constraints: water balance, dead water level, flood season restricted water level, and inter-reservoir re-regulation mechanisms. Additionally, it adopts an LSTM network optimized via Particle Swarm Optimization (PSO) for enhanced predictive performance. We validate the model using daily hydrological data from 2010 to 2025 for three reservoirs in the Wujiaqu Irrigation District of Xinjiang, China. The model exhibits exceptional stability and predictive accuracy across key evaluation metrics: Nash–Sutcliffe Efficiency (NSE) ≥ 0.82, Pearson correlation coefficient (r) > 0.94, Root Mean Square Error (RMSE) ≤ 1.50 m^3/s, and Water Balance Index (WBI) ≤ 0.016. It outperforms conventional data-driven and mechanistic models in extreme flow simulation scenarios. It also eliminates unphysical negative outflow values in all predictive results. The model achieves 100% compliance with flood control standards and an irrigation guarantee rate of no less than 86%. This study advances the development of physically constrained data-driven modeling for water resources engineering. It provides reliable methodological support for the intelligent operation of reservoir groups in smart water conservancy systems. The model also balances training cost and inference efficiency effectively. It demonstrates verified scalability for reservoir groups of varying scales, fully meeting the operational deployment requirements of smart water systems.

Keywords: joint reservoir group operation; outflow simulation; LSTM; MDUPLEX; KG loss function

Check for updates

Academic Editors: Pedro Couto and Andrea Prati

Received: 4 January 2026
Revised: 1 February 2026
Accepted: 19 February 2026
Published: 22 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Global climate change [1,2] and the uneven spatial–temporal distribution of water resources [3] have heightened the imbalance between water supply and demand worldwide. Advancing the precision and intelligence of watershed water resource regulation is thus a central focus of international water resource management. As a critical step in this regulation, accurate reservoir runoff prediction [4] directly shapes the success of flood control, drought mitigation, optimal water allocation, and multi-objective water demand decision-making. This is especially true for irrigation district systems, where reservoir groups [5,6] leverage hydraulic linkages and compensatory effects via joint operation [7]. Such joint operation greatly boosts the spatiotemporal regulatory capacity of water resources and supports agricultural irrigation and integrated water utilization efforts. Simulating joint reservoir operations requires holistic consideration of inherent hydraulic linkages and compensatory mechanisms [8], as this not only improves operational planning accuracy but also clarifies how climate change and human activities modulate the water cycle and ecological environment. Developing a scientifically rigorous, practically applicable joint operation model for reservoir groups is therefore essential to address these challenges.

However, current outflow simulation methods for joint reservoir group operation lack both generality and stability. Model reliability hinges largely on the data partitioning strategy [9] used during calibration. Traditional methods split data into training and validation sets in strict chronological order [10]. Temporal variability and non-stationarity characterize all meteorological and hydrological conditions. These properties create stark statistical distribution differences across dataset splits for model training. The result is systematic performance degradation in validation compared to training runs. This degradation directly undermines the overall robustness of any calibrated model. To resolve this issue, researchers have tested non-chronological partitioning algorithms [11,12], such as the deterministic DUPLEX and random SBSS-N. Even so, these methods fail to fully eliminate performance biases across training and validation datasets. Chen [13] proposed the modified deterministic Multi-round DUPLEX (MDUPLEX) algorithm to address this gap. It ensures statistical homogeneity between training and validation sets for rainfall-runoff models. This homogeneity yields consistent simulation results across all dataset partitions for such models. Yu et al. [14] later coupled MDUPLEX with the Long Short-Term Memory (LSTM) neural network for outflow simulation. Their method delivers high accuracy and generalizability for large single reservoirs in watershed hydrological modeling. However, this framework is inherently designed for independent single-reservoir scenarios and has not been systematically extended to joint reservoir groups. It cannot capture hydraulic interdependencies or operational compensation between multiple reservoirs, leaving a critical gap in reservoir group modeling.

Second, mainstream reservoir operation models face inherent limitations. Parametric models, based on mathematical functions and predefined operational rules [15,16], struggle to characterize complex real-world operational patterns. Multi-objective optimization paradigms [17], while balancing competing goals (e.g., minimizing water shortage and maximizing power generation), are hindered by downstream water demand data scarcity [18] and high computational costs [19], limiting their practical applicability. Deep learning methods—particularly LSTM—offer a compelling alternative, as they automatically extract implicit patterns from hydrological and operational data without requiring explicit downstream demand inputs [20]. However, LSTM is an inherently black-box model lacking built-in physical constraints, often producing physically implausible outputs (e.g., negative outflows). Knowledge-Guided (KG) loss functions have been proposed to address this by embedding hydrological prior knowledge (e.g., water balance, water level boundaries) [21,22], but existing KG approaches fail to integrate critical operational mechanisms

unique to joint reservoir groups—specifically inter-reservoir re-regulation and flood storage capacity control. Moreover, KG loss functions have only been validated for single-reservoir independent operations, with no systematic extension to reservoir group scenarios. This gap limits the engineering reliability and physical interpretability of LSTM-based models for joint operations.

Third, computational cost and scalability [23] remain overlooked barriers to practical deployment in smart water systems. Complex joint reservoir operations demand efficient real-time inference and flexibility across varying reservoir group scales, yet few existing models balance these requirements with predictive accuracy.

To address these interconnected limitations, this study aims to develop a physically constrained, data-driven model—MDUPLEX-KG-LSTM—for reservoir group outflow simulation. The specific objectives are threefold: (1) extend the MDUPLEX data partitioning algorithm to joint reservoir group scenarios, ensuring statistical homogeneity across datasets while preserving inter-reservoir hydraulic and operational correlations; (2) construct a novel KG loss function that embeds core physical constraints (water balance, dead water level, flood season restricted level) and reservoir group-specific operational mechanisms (inter-reservoir re-regulation via the Muskingum routing equation); (3) optimize the LSTM network using Particle Swarm Optimization (PSO) to balance predictive accuracy, computational efficiency, and scalability for smart water system deployment.

The scientific novelty of this work lies in three integrated innovations: (1) the adaptation of MDUPLEX from single-reservoir to reservoir group scenarios, addressing the statistical bias of traditional partitioning while retaining inter-reservoir dependency information; (2) the development of a KG loss function that explicitly incorporates inter-reservoir re-regulation mechanisms—previously unaddressed in existing knowledge-guided models—enhancing physical rationality and operational compliance; (3) the synergistic integration of MDUPLEX partitioning, KG constraints, and PSO-optimized LSTM, achieving a balance between predictive accuracy, physical consistency, and computational feasibility for real-world engineering deployment.

This study uses the plain reservoir group in the Wujiaqu Irrigation District (Sixth Division of China's Xinjiang Production and Construction Corps, XPCC) as a case study. The district's three reservoirs (Mengjin, Bayi, and Shashanzi) exhibit distinct operational roles (annual regulation, combined regulation-flood control, seasonal re-regulation) and hydraulic linkages, providing an ideal testbed for validating the proposed model.

The remainder of this paper is structured as follows: Section 2 describes the study area and hydrological data characteristics. Section 3 details the MDUPLEX-KG-LSTM model, including the MDUPLEX partitioning algorithm, PSO-optimized LSTM structure, and KG loss function development. Section 4 outlines the evaluation metrics and comparative experimental design, followed by an analysis of simulation results, focusing on predictive performance, generalization ability, and engineering applicability. Section 5 summarizes the core conclusions and identifies future research directions.

2. Study Area and Data

2.1. Study Area Overview

Three plain storage-regulating reservoirs—Mengjin Reservoir (MJR), Bayi Reservoir (BYR), and Shashanzi Reservoir (SSZR)—are located in Wujiaqu City. Their total designed storage capacity is 9.973×10^7 m^3, with a useful storage capacity of 5.585×10^7 m^3. MJR and BYR function as annual regulation reservoirs among the three. Outflow from MJR can be delivered to downstream regiment-level irrigation districts or allocated to BYR and SSZR. BYR and SSZR are configured in a parallel layout. The former receives water diverted from MJR and inflow from the Laolong River, and it has independent discharge

capacity. The latter mainly reregulates outflow from MJR and assumes part of the regional flood control responsibilities. Detailed parameters for these three reservoirs are provided in Table 1. Seasonal water resource management is adopted in this area. Dispatch schemes are developed based on hydrological characteristics across two distinct periods: the snowmelt-rainfall flood season (April–September) and the non-flood season (October–March of the following year). The spatial distribution of the regional water system is shown in Figure 1.

Table 1. Characteristics of the reservoir system.

	MJR	BYR	SSZR
Reservoir Function	Irrigation, Flood Control	Irrigation, Flood Control, Re-regulation	Irrigation, Re-regulation, Flood Control
Regulation Capacity	Full Annual Regulation	Full Annual Regulation	Seasonal Regulation
Check Flood Level (P = 0.33%)	487.53	465.01	454.92
Design Flood Level (P = 2%)	487.23	464.64	454.75
Flood High Level (m)	487.23	464.64	454.75
Normal Water Level (m)	487.23	464.22	454.75
Flood Season Restricted Water Level (m)	486.66	463.19	454.60
Dead Water Level (m)	480.13	460.00	452.00
Flood Regulating Capacity (10^4 m^3)	828.79	1729.81	231.2
Flood Control Capacity (10^4 m^3)	514.52	1008.00	124.4
Useful Storage Capacity (10^4 m^3)	2348.03	2458.00	779
Irrigation Area (10^4 mu)	53	50.72	8.7
Irrigation Water Demand (10^4 m^3)	2525	2674	802
Irrigation Guarantee Rate	75%	75%	75%

Note: P represents the flood exceedance probability; P = 0.33% (300-year return period) and P = 2% (50-year return period) are determined via Pearson-III distribution fitting based on 1950–2025 flood data. Herein, P is calculated via hydrological frequency analysis using the Pearson-III distribution. The analysis is based on 75-year historical flood peak discharge data (1950–2025) of the Laolong River basin, with parameters calibrated by the curve-fitting method to ensure consistency with regional hydrological characteristics.

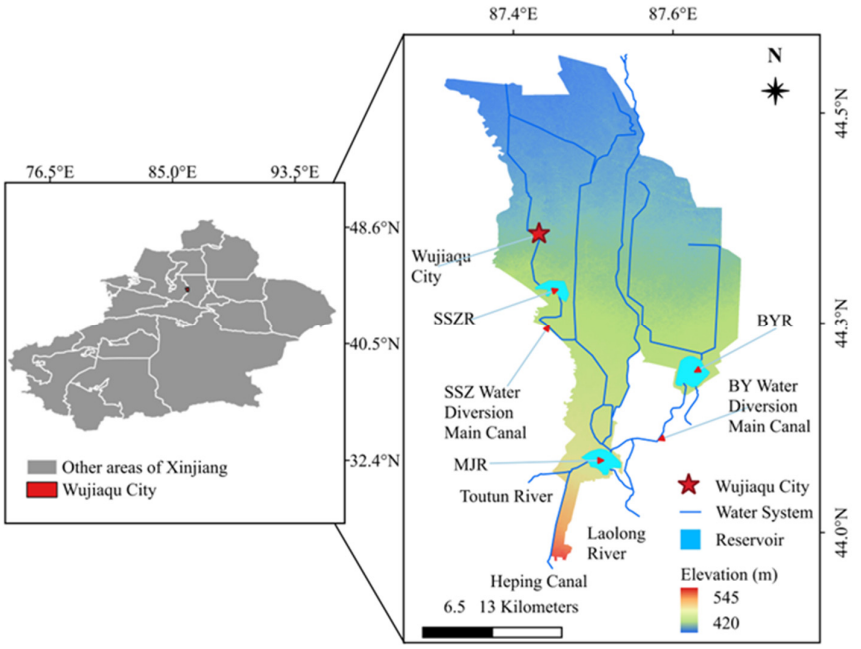


Figure 1. Hydrographical map of Wujiaqu irrigation district.

2.2. Data

In this study, we used daily hydrological monitoring data spanning 2010–2025 from MJR, BYR, and SSZR. The "operation status" of reservoirs refers to the temporal dynamics of three core operational parameters: daily inflow (dimension: $[L^3\ T^{-1}]$, unit: m^3/s), daily

outflow (dimension: [L^3 T^{-1}], unit: m^3/s), and pre-dam water level (dimension: [L], unit: m), which directly reflect the joint operation characteristics of the reservoir group and serve as the key input data for the proposed model—these three parameters are exactly the continuous measurements of inflow, outflow, and pre-dam water levels included in the aforementioned hydrological datasets. Dynamic flow data within these datasets were obtained from the monitoring records of Wujiaqu City, Sixth Division.

Figure 2 illustrates the temporal trends in daily operational data for the three reservoirs. Figure 2a,b show that the MJR and BYR follow a typical full annual storage regulation mode. Flood control storage is released in advance via winter-spring irrigation before the flood season. Floodwaters are then stored and retained during the flood season, with a strong correlation between inflow and outflow for both reservoirs. For example, June 2016 brought a high-flow period coinciding with peak emergency water demand. Both reservoirs released water at rates exceeding historical averages for the same month. When the BYR diverted floodwaters from the MJR, its water level fell from 464.26 m to 462.89 m. This diversion caused the MJR's water level to drop sharply from 485.81 m to 483.57 m at the same time.

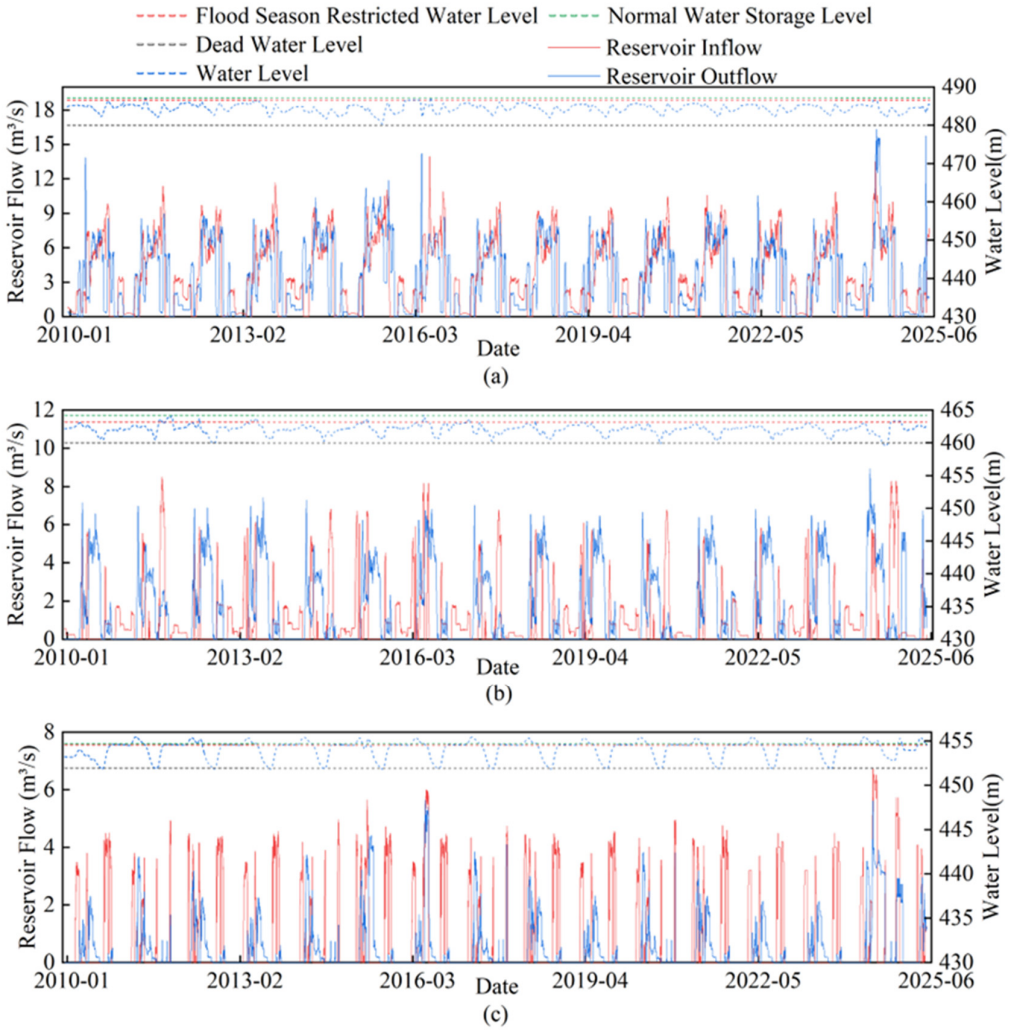


Figure 2. Multi-Year operation status (inflow, outflow, and water level) of individual reservoirs in the Wujiaqu City reservoir group (January 2010–May 2025). (**a**) MJR operation status, (**b**) MJR operation status, (**c**) SSZR operation status. Flood control-related water levels (dashed red lines) are designed based on flood frequency probability P, with calculation details provided in Section 2.1.

The SSZR receives inflow from the MJR during irrigation periods, acting as a counter-regulation reservoir (Figure 2c). It operates as a seasonal regulation reservoir and undertakes flood diversion during flood seasons in wet years. It also fulfills integrated functions of counter-regulation, irrigation, and flood control. From 2010 to 2025, the MJR's inflow and outflow regulation directly influenced the SSZR. The SSZR's inflow in June 2016 and 2024 far exceeded levels in all other months, with outflow responding in direct synchronization.

Notably, Wujiaqu City's ecological base flow demand is met by independent water source projects. These projects are separate from the three reservoirs focused upon in this study. As a result, the reservoirs recorded zero outflow during all non-irrigation periods. Figure 2 also shows that reservoir water levels in the study area vary annually between the dead water level and the flood limit water level. This pattern aligns with local operational strategies, which account for regional water resource scarcity. Irrigation water supply is the primary operational priority for the reservoirs. They are therefore operated at reduced water levels during non-flood seasons, rather than being filled to normal storage levels. Flood control regulations require strict water level limits below the flood limit during flood seasons. The adaptive model developed in this study accurately replicates this real-world operational pattern.

In Figure 2, the normal storage level line (green dashed line) represents the theoretical design upper limit. Constrained by actual water supply operation and flood control regulations, the reservoir operating water level is distributed primarily between the dead water level and the flood limit water level.

3. Research Methodology

3.1. Architecture of the MDUPLEX-KG-LSTM Model

As illustrated in Figure 3, this study proposes a reservoir group outflow simulation method to explore interconnectivity and operational compensation effects within the reservoir group system under variable hydrological regimes. The method integrates MDUPLEX data partitioning with a knowledge-guided (KG) loss function that reflects practical reservoir operation rules. The specific implementation steps are as follows: First, key input variables for reservoir group outflow are identified through hydrological mechanism analysis. Subsequently, the MDUPLEX partitioning algorithm allocates these input variables to their corresponding datasets. Then, the Particle Swarm Optimization (PSO) algorithm conducts global hyperparameter optimization for the original LSTM network. Finally, a knowledge-guided penalty term is incorporated into the traditional mean squared error (MSE) loss term. This term characterizes core reservoir operation rules, including counter-regulation, flood control, and beneficial water utilization.

Figure 3 presents the holistic workflow of the MDUPLEX-KG-LSTM model, where each module represents a key step in the simulation process. Specifically, the 'LSTM Network' module in Figure 3 is the core predictive component, whose internal structure is detailed in Figure 4. Figure 4 illustrates the LSTM unit's gating mechanisms (forget gate, input gate, output gate) and the cell state—these components enable the network to capture long-term temporal dependencies in hydrological data. The mathematical expressions of the LSTM gates (Section 3.1.3) are directly mapped to the visual elements in Figure 4, ensuring that the neural network architecture is fully transparent and reproducible.

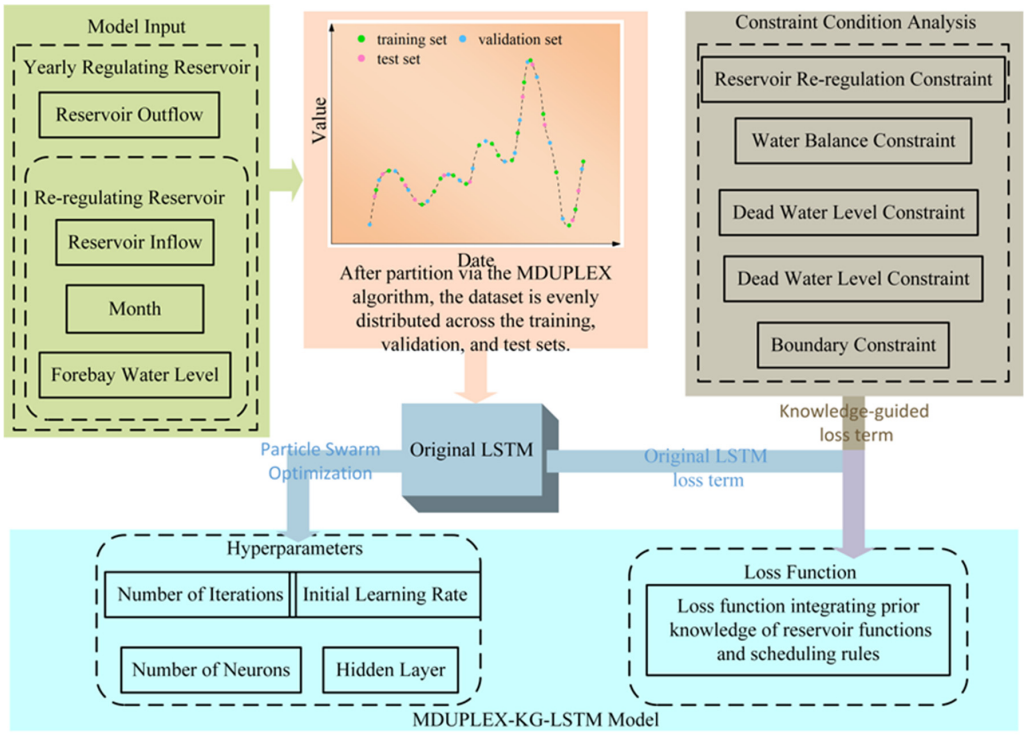


Figure 3. Overall workflow of the MDUPLEX-KG-LSTM Model. This figure illustrates the end-to-end integration of three core components: (1) MDUPLEX data partitioning (input data processing), (2) PSO-based LSTM hyperparameter optimization, and (3) KG loss function integration. The 'LSTM Network' module (boxed in gray) corresponds to the detailed neural network architecture depicted in Figure 4.

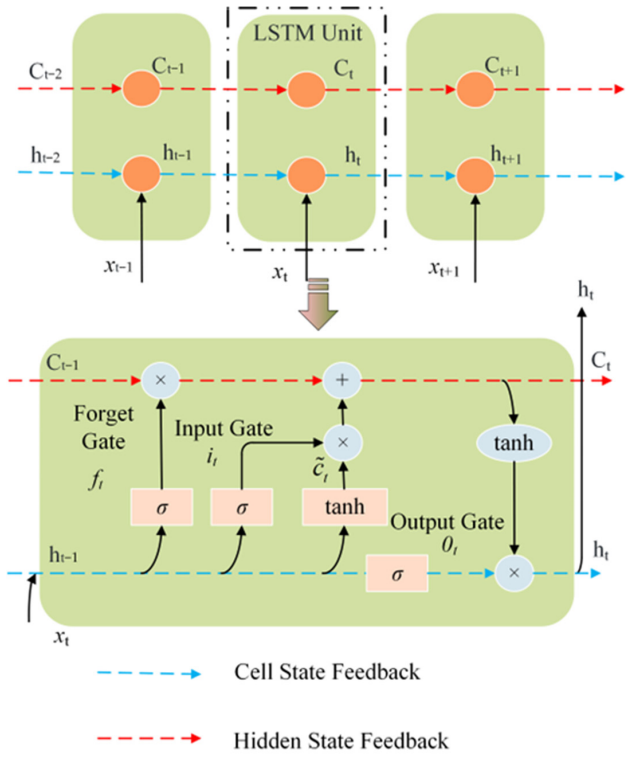


Figure 4. Detailed architecture of the LSTM neural network (core predictive component of MDUPLEX-KG-LSTM). This figure depicts the internal structure of LSTM units, including the forget gate, input gate, output gate, and cell state—corresponding to the 'LSTM Network' module in Figure 3. The figure indicates key functional units, with mathematical relationships detailed in Section 3.1.3.

3.1.1. Model Inputs and Outputs

Taking the outflow rate of each reservoir in the reservoir group at the current time step as the decision variable, we adopted supervised learning to fit the outflow process of individual reservoirs within the group. The operational rules of the proposed model were derived from the mapping relationship of the optimally parameterized Long Short-Term Memory (LSTM) network, with the corresponding mathematical expression presented below:

$$Qo_t = f(I_t, \theta) \quad (1)$$

where I_t represents the set of impact factors for reservoir operation, which acts as the input variables of the model. Qo_t denotes the reservoir outflow rate $(Qo_{t-2}, Qo_{t-1}, Qo_t)$ (dimension: [L^3 T^{-1}], unit: m^3/s), the designated decision variable of the model. f denotes the mapping relationship between I_t and Qo_t, and θ refers to the network hyperparameters optimized via the PSO algorithm.

Specifically, for reservoirs with counter-regulation functionality, the impact factor set I_t comprises the inflow rates recorded on the operation day and the two preceding days $(Qi_{t-2}, Qi_{t-1}, Qi_t)$ (dimension: [L^3 T^{-1}], unit: m^3/s), the corresponding calendar month (Mo), and the reservoir water levels (Wl_{t-1}, Wl_t) (dimension: [L], units: m) measured on the current and prior days. To better characterize the complex outflow dynamics of fully annual regulation reservoirs, additional impact factors were integrated into the factor set for these reservoirs, including the outflow rates $(Qo_{t-2}, Qo_{t-1}, Qo_t)$ (dimension: [L^3 T^{-1}], unit: m^3/s) and water levels $(Wl_{t-2}, Wl_{t-1}, Wl_t)$ (dimension: [L], units: m) recorded on the operation day and the two preceding days. All the aforementioned impact factors were designated as the input variables of the model, with the decision variable Qo_t $(Qo_{t-2}, Qo_{t-1}, Qo_t)$ (dimension: [L^3 T^{-1}], unit: m^3/s) serving directly as the model's output variable.

3.1.2. MDUPLEX Dataset Partitioning

Dataset partitioning plays a critical role in reliable deep learning model development for water engineering applications. Consistent subset allocation ensures robust model validation and realistic performance assessment for hydrologic tasks. DUPLEX is a reproducible, deterministic data sampler well-suited for practical water engineering field applications. However, Zheng et al. [9] noted that unbalanced subset proportions bias it toward routine hydrologic event data. This skew causes DUPLEX to underperform on extreme hydrologic events and underestimate predictive model capability. To resolve this critical limitation, Chen [13] proposed the enhanced MDUPLEX sampling algorithm for water engineering applications. It employs basic sampling pools and multi-round sampling to effectively balance training, test, and validation subsets—enhancing the statistical consistency of all subsets within the entire water engineering dataset. We first define three core proportional variables to guide consistent partitioning: P_T denotes the fraction of the entire original dataset (D) allocated to the dedicated training set T. P_S represents the fraction of the full original dataset assigned to the independent test set S. P_V signifies the fraction of the complete original dataset allocated to the validation set V. These three proportions sum to 1 to ensure the full utilization of all original data points. The complete step-by-step workflow of the MDUPLEX (Algorithm 1) appears below.

Algorithm 1. MDUPLEX: Refined Deterministic Multi-Round Data Partitioning

1: **Input:** Original dataset D; target proportions P_T, P_S, P_V; Subset proportions $P_T + P_S + P_V = 1.0$
2: **Output**: Partitioned training set T, test set S, validation set V
3: Initialize $T = S = V = \varnothing$
4: Compute pool capacity: $n = \text{round}(1/\min(P_T, P_S, P_V))$ (hold 2n points)
5: Assign per − pool pair counts : $n_T = \text{round}(nP_T), n_S = \text{round}(nP_S), n_V = \text{round}(nP_V)$
6: **While** $D \neq \varnothing$:
7: // Max-pair sampling for $T/S/V$
8: Add 1 max-distance pair from D to T; repeat for S, V (remove from D)
9: // Max-distance-to-subset sampling
10: **Repeat**:
11: Add 1 pair (max-distance to T) to T; repeat for S, V (remove from D)
12: **Until** 2 subsets reach $n_T / n_S / n_V$
13: Assign remaining pool data to the undersaturated subset
14: **Return** T, S, V

The core objective of MDUPLEX is to mitigate the statistical discrepancy between training and validation sets induced by hydrological non-stationarity, rather than eliminating meaningful differences in key attributes. During multi-round sampling, the algorithm prioritizes the inclusion of representative samples covering diverse operational scenarios (e.g., 30% flood season samples, 15% extreme event samples) while balancing their statistical distribution. This ensures the model learns universal hydrological-operation rules instead of overfitting to time-specific patterns, as verified by the comparative experiments in Section 4.2.

3.1.3. Construction of the Standard LSTM

The Long Short-Term Memory (LSTM) network is a specialized variant of recurrent neural networks (RNNs) designed to address the critical limitation of traditional RNNs in capturing long-term temporal dependencies. This capability is enabled by a synergistic combination of gating units and a persistent cell state, where the gating system—comprising an input gate, a forget gate, and an output gate—exerts targeted control over information flow through the model.

Specifically, the input gate modulates the integration of new information into the network at each time step, while the forget gate regulates the retention of historical information from preceding steps. The output gate governs how the current cell state is propagated to the short-term memory for subsequent computations, and the cell state itself acts as a cross-temporal carrier for long-term pattern recognition, preserving critical information through iterative updates to gating parameters across time steps. This structural design inherently mitigates the gradient vanishing and gradient explosion issues that plague traditional RNNs, making LSTMs well-suited for hydrological time series modeling. Figure 4 depicts the LSTM network architecture and internal unit structure employed in this study, with key functional components explicitly aligned to the technical details below.

To ensure the reproducibility of the network structure, we specify the following critical design parameters:

- Input/output dimensions

The LSTM network receives input vectors of dimension [time steps × feature number], where the feature number is tailored to the operational characteristics of each reservoir. For annual regulation reservoirs (MJR, BYR), the input features include inflow, water level,

and outflow from the current and previous two days, resulting in 6 input features. For the seasonal regulation reservoir (SSZR), the input features are limited to inflow and water level from the current and previous two days (4 input features), reflecting its focused re-regulation role. The output vector dimension is uniformly [time steps × 1], corresponding to the daily outflow of the target reservoir.

- Activation functions and gate correspondence

As illustrated in Figure 4, the forget gate, input gate, and output gate utilize the sigmoid activation function (range: [0,1]) to precisely control information flow (e.g., enabling or disabling the retention of historical data or the integration of new signals). The cell state update process employs the tanh activation function (range: [−1,1]) to normalize intermediate values, preventing numerical instability. The mathematical expressions for each gate (Eqs. A1–A4 in the Appendix) are directly mapped to the labeled functional units in Figure 4: (a) Forget Gate, (b) Input Gate, (c) Cell State Update, and (d) Output Gate.

- Network layer details

All reservoir-specific LSTM models feature a single hidden layer (consistent with Table 2), composed of fully connected LSTM units with reservoir-specific counts: 128 units for MJR, 155 units for BYR, and 68 units for SSZR. A dropout rate of 0.2 is applied to the hidden layer to mitigate overfitting by randomly deactivating a subset of units during training. The output layer employs a linear activation function to map the LSTM's hidden state to the final daily outflow prediction, ensuring compatibility with the continuous nature of hydrological outflow data.

Table 2. Hyperparameters of LSTM model.

	MJR	BYR	SSZR
Number of Neurons	128	155	68
Number of Iterations	85	90	30
Learning Rate	0.0007	0.0007	0.0015
Hidden Layer	1	1	1
Optimizer	Adam	Adam	Adam
Batch Size	4	4	4

All computational experiments were run on a standardized hardware platform to ensure strict result comparability. The central processing unit (CPU) was an Intel Core i9-13900K (manufactured by Intel Corporation, Santa Clara, CA, USA) with a 3.0 GHz base frequency and 24 cores/32 threads. The graphics processing unit (GPU) was an NVIDIA RTX 4090 (manufactured by NVIDIA Corporation, Santa Clara, CA, USA) equipped with 24 GB of GDDR6X video memory. The system memory (RAM) totaled 128 GB of DDR5 modules (manufactured by mainstream global memory suppliers, with major production locations in Taiwan, China and the Republic of Korea) running at a 5600 MHz clock speed. Our software environment was built on Python 3.9.16 as the core programming language. We implemented the TensorFlow 2.10.0 deep learning framework for model construction and training. The NumPy 1.24.3 library was used for all numerical computations and data processing tasks.

To enhance the operational performance of the standard LSTM, we conducted global hyperparameter tuning using the Particle Swarm Optimization (PSO) algorithm. The tuning process targeted three critical hyperparameters: the number of LSTM units (detailed above), learning rate, and the number of training iterations. The PSO algorithm is particularly well-suited for LSTM hyperparameter optimization in hydrological modeling, as it exhibits low dependence on the form of the objective function and delivers fast convergence to

optimal solutions [24]. Table 2 summarizes the specific hyperparameter settings for all LSTM model experiments in this study.

3.1.4. Construction of Loss Function Based on Prior Knowledge of Reservoir Operation

Traditional LSTM loss functions prioritize only data-fitting performance. This frequently causes peak flow distortion, negative discharge, and water volume imbalance in reservoir outflow simulations. In this study, we convert prior reservoir operation knowledge into penalty terms within the loss function to constrain the model's solution space. The integrated constraints include a counter-regulation mechanism, the water balance principle, and boundary constraints for dead water levels and flood control limited water levels.

1. Constraint of Reservoir Re-regulation

In the operation of reservoir group systems, reregulation denotes secondary optimization of outflow hydrographs from upstream regulating reservoirs. This operational mode differs from conventional upstream-downstream reservoir arrangements. It features adjacent spatial distribution and relatively limited storage capacity for the downstream reservoir. Specifically, SSZR has a distinct reregulation relationship with upstream MJR. BYR is also downstream of MJR but fulfills both reregulation and flood control functions. Its large storage capacity enables this dual role. Irrigation district water demands guide coordination between MJR and the two downstream reservoirs (BYR, SSZR). This coordination maximizes MJR's storage regulation capacity and downstream reregulation efficiency. It boosts flood control and irrigation benefits.

We quantitatively characterized hydraulic connections during reregulation using the Muskingum river flow routing equation. This equation describes the fundamental law of runoff propagation between upstream and downstream reservoirs. The mathematical expression of the model is as follows:

$$Qi_{\mathrm{down},t} = C_0 Qo_{\mathrm{up},t} + C_1 Qo_{\mathrm{up,t}-1} + C_2 Qi_{\mathrm{down},t-1} \quad (2)$$

where Qi_{down} is the inflow to the downstream reservoir (dimension: [L^3 T^{-1}], unit: m^3/s), Qo_{up} is the outflow from the upstream reservoir (dimension: [L^3 T^{-1}], unit: m^3/s), and C_0, C_1 and C_2 are the Muskingum routing coefficients. They follow the critical constraint $C_0 + C_1 + C_2 = 1$.

Equation (2) shows that downstream temporal runoff depends on two key factors. It relies on upstream runoff inputs and its own serial autocorrelation. This indicates strong hydraulic connectivity between adjacent cross-sections. Notably, MJR outflow experiences unavoidable losses during river and canal propagation. These losses come from seepage, evaporation and anthropogenic water withdrawals. Inter-basin inflows superimpose onto the attenuated runoff. This creates a time-lagged hydrologic response. The combined runoff ultimately becomes the inflow to downstream reservoirs (BYR or SSZR). This entire process forms the core hydraulic mechanism of reservoir reregulation.

The Muskingum equation fits the study area's connecting reaches well. These reaches are 15–30 km long with regular cross-sections and weak nonlinear effects. We calibrated C_0, C_1 and C_2 using 2010–2025 historical hydrological data. The trial-and-error method minimized RMSE between simulated and observed Qi_{down}. Calibration results are MJR-BYR reach ($C_0 = 0.08$, $C_1 = 0.72$, $C_2 = 0.20$) and MJR-SSZR reach ($C_0 = 0.11$, $C_1 = 0.69$, $C_2 = 0.20$). NSE ≥ 0.85 confirms the calibration accuracy for engineering applications.

2. Water Balance Constraint

The water balance principle is a fundamental constraint for reservoir operation, ensuring simulated storage and outflow processes align with physical laws. To avoid water

volume imbalances in simulation outputs, this study integrated the water balance constraint into the knowledge-guided loss function. The mathematical expression for the water balance equation is given as follows:

$$V_t = V_{t-1} + [Qi_t - Qo_t]\Delta t - L_t \quad (3)$$

where V_t and V_{t-1} denote the storage volume of the reservoir at time t and $t-1$ during the scheduling period respectively (dimension: [L^3], units: m^3). Qi_t is the reservoir inflow (dimension: [L^3 T^{-1}], units: m^3/s) at time t. Qo_t is the reservoir outflow at time t (dimension: [L^3 T^{-1}], units: m^3/s). L_t represents the evaporation and seepage loss (dimension: [L^3], units: m^3) at time t. Δt is the units' time step for scheduling (dimension: [T], units: h/d), consistent with the study's designated time scale.

3. Dead Water Level Constraint

The dead water level constraint is an important boundary constraint for reservoir operation, which ensures that the reservoir operation does not deviate from the safe and practical operation scope. As shown in Figure 2, the water levels of MJR and BYR are consistently maintained above the dead water level, whereas due to its small storage capacity and core function of irrigation (involving regulating inflow from MJR), the water level of SSZR occasionally falls below the dead water level. To reflect this practical operation characteristic in the model, the dead water level constraint was incorporated into the knowledge-guided loss function, with the mathematical expression as follows:

$$Wl_t \geq Wl_{dead} \quad (4)$$

where Wl_t denotes the actual water level (dimension: [L], units: m) of the reservoir at time t, and Wl_{dead} (dimension: [L], units: m) denotes the dead water level.

4. Flood Season Restricted Water Level Constraint

The dead water level constraint is a key boundary constraint for reservoir operation. It ensures reservoir operations stay within safe, practical operational parameters at all times. As shown in Figure 2, the water levels of the MJR and BYR remain consistently above the dead water level. By contrast, the SSZR's water level occasionally falls below this threshold. This is due to its small storage capacity and core irrigation function—namely, regulating inflow diverted from the MJR. To capture this real-world operational characteristic in our model, we integrated the dead water level constraint into the knowledge-guided loss function. The corresponding mathematical expression is provided as follows:

$$Wl_t \leq Wl_{t,flood} \quad (5)$$

where Wl_t denotes the actual water level of the reservoir at time t (dimension: [L], units: m), and $Wl_{t,flood}$ denotes the reservoir's dead water level (dimension: [L], units: m).

5. Boundary Constraints

To mitigate flood risks and ensure rational reservoir operation, outflow must satisfy dual boundary constraints (upper and lower limits). This constraint stops the model from generating unrealistic outflow values in simulation—such as excessively large flood discharges or negative flow rates. For this study, we therefore integrated outflow boundary constraints into the knowledge-guided loss function, with the mathematical expression as follows:

$$Qo_u \geq Qo_t \geq Qo_l \quad (6)$$

where Qo_u and Qo_l denote the upper and lower threshold values of the reservoir outflow respectively (dimension: [L^3 T^{-1}], units: m^3/s).

Based on the above constraint analysis, we established the MDUPLEX-KG-LSTM loss function for this research, with its mathematical expression given as follows:

$$\begin{aligned}
loss_{\text{MDUPLEX-KG-LSTM}} = \; & \lambda_{\text{LSTM}} loss_{\text{LSTM}} + \lambda_{\text{reregulation}} loss_{\text{reregulation}} + \\
& \lambda_{\text{balance}} loss_{\text{balance}} + \lambda_{\text{flood-level}} loss_{\text{flood-level}} + \\
& \lambda_{\text{dead-level}} loss_{\text{dead-level}} + \lambda_{\text{boundary}} loss_{\text{boundary}} \\
loss_{\text{LSTM}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \left(\hat{Q}o_t - Qo_t\right)^2 \\
loss_{\text{reregulation}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \left(\hat{Q}i_{down,t} - Qi_{down,t}\right)^2 \\
loss_{\text{balance}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \left(\Delta\hat{V}_t - (Qi_t - \hat{Q}o_t - L_t)\right)^2 \\
loss_{\text{flood-level}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \begin{cases} \left(Wl_t - Wl_{t,\text{flood}}\right)^2 & \text{if } t \in \{6,7,8,9\}, Wl_t > Wl_{t,\text{flood}} \\ 0 & \text{otherwise} \end{cases} \\
loss_{\text{dead-level}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \begin{cases} \left(\hat{V}_t - V_{t,\text{dead}}\right)^2 & \text{if } \hat{Wl}_t < Wl_{t,dead} \\ 0 & \text{otherwise} \end{cases} \\
loss_{\text{boundary}} = \; & \frac{1}{N_t} \sum_{t=1}^{N_t} \begin{cases} \left(\hat{Q}o_t - Qo_u\right)^2 & \text{if } \hat{Q}o_t > Qo_u \\ 0 & \text{if } Qo_l < \hat{Q}o_t < Qo_u \\ \left(\hat{Q}o_t - Qo_l\right)^2 & \text{if } \hat{Q}o_t < Qo_l \end{cases}
\end{aligned} \tag{7}$$

where $loss_{\text{MDUPLEX-KG-LSTM}}$ denotes the loss function of MDUPLEX-KG-LSTM. We first define N_t as the number of simulation steps. It equals 5619, corresponding to daily data from 2010 to 2025. We then describe each parameter and its variation range. L is the MDUPLEX-KG-LSTM loss function with no fixed range. $loss_{\text{LSTM}}$ is the mean squared error of outflow simulations. It ranges from 0 to 10 (dimension: [L^3 T^{-1}], units: m^3/s). $loss_{\text{reregulation}}$ reflects downstream inflow simulation error. It varies between 0 and 5 (dimension: [L^3 T^{-1}], units: m^3/s). $loss_{\text{balance}}$ is the water balance penalty term. It ranges from 0 to 200. $loss_{\text{flood-level}}$ is the flood season restricted level penalty. It varies from 0 to 50,000. $loss_{\text{dead-level}}$ is the dead water level penalty. It ranges from 0 to 35,000. $loss_{\text{boundary}}$ is the outflow boundary penalty. It varies between 0 and 10. λ_{LSTM}, $\lambda_{\text{reregulation}}$, λ_{balance}, $\lambda_{\text{flood-level}}$, $\lambda_{\text{dead-level}}$ and $\lambda_{\text{boundary}}$ are penalty coefficients. $\lambda_{\text{LSTM}} = \lambda_{\text{boundary}} = 1$(fixed), $\lambda_{\text{reregulation}} = 800$ (fixed), $\lambda_{\text{balance}} = 10$ (fixed). $\lambda_{\text{flood-level}}$ ranges 5000–10,000. $\lambda_{\text{dead-level}}$ ranges 3000–7000. Qo_t and $\hat{Q}o$ (dimension: [L^3 T^{-1}], units: m^3/s) are the observed and simulated outflow. They range from 0 to 18 m^3/s. $\hat{Wl}_t$ (dimension: [L], units: m) is the simulated water level. It varies 452–487.53 m. $\hat{V}_t$ (dimension: [L^3], units: m^3) is the simulated storage volume. It ranges from 0 to 9.973×10^7 m^3. The coefficients λ are determined through parameter sensitivity analysis and demonstration of reservoir functions [21]. The final values are presented in Table 3.

Table 3. Coefficients of loss terms for each reservoir in the MDUPLEX-KG-LSTM model.

Loss Term Coefficients	MJR	BYR	SSZR
λ_{LSTM}	1	1	1
$\lambda_{\text{reregulation}}$	\	800	800
λ_{balance}	10	10	10
$\lambda_{\text{flood-level}}$	10,000	9000	5000
$\lambda_{\text{dead-level}}$	7000	6000	3000
$\lambda_{\text{boundary}}$	1	1	1

3.2. Model Evaluation

To fully evaluate the performance of the MDUPLEX-KG-LSTM model, we used three standard accuracy metrics for hydrological modeling: Root Mean Square Error (RMSE), Pearson correlation coefficient (r), and Nash–Sutcliffe Efficiency (NSE). We also added the Water Balance Index (WBI) to assess the model's physical consistency. The definitions and calculation methods for these metrics are provided below:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_t}\sum_{t=1}^{N_t}\left(\hat{Qo}_t - Qo_t\right)^2} \tag{8}$$

$$\mathrm{r} = \frac{\sum_{t=1}^{N_t}\left(\hat{Qo}_t - \overline{\hat{Q}}o_t\right)\left(Qo_t - \overline{Q}o_t\right)}{\sqrt{\sum_{t=1}^{N_t}\left(\hat{Qo}_t - \overline{\hat{Q}}o_t\right)^2 \sum_{t=1}^{N_t}\left(Qo_t - \overline{Q}o_t\right)^2}} \tag{9}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{N_t}\left(\hat{Qo}_t - Qo_t\right)^2}{\sum_{t=1}^{N_t}\left(\overline{Q}o_t - Qo_t\right)^2} \tag{10}$$

$$\mathrm{WBI} = \frac{1}{N_t}\sum_{t=1}^{N_t}\left|\frac{\left(Qi_t - \hat{Qo}_t - O_t\right) - \Delta\hat{V}_t}{\max\left(Qi_t, \overline{\hat{Q}}o\right)}\right| \tag{11}$$

where N_t denotes the length of the simulation period (dimension: [T], units: d); it equals 5619 days covering 2010 to 2025. Qi_t and Qo_t represent the observed inflow and observed outflow of the reservoir during the scheduling period t respectively (dimension: [L^3 T^{-1}], units: m^3/s); $\hat{Qo}_t$ and $\overline{\hat{Q}}o$ denote the simulated outflow and average simulated outflow of the reservoir during the scheduling period t respectively (dimension: [L^3 T^{-1}], units: m^3/s); $\Delta\hat{V}_t$ (dimension: [L^3], units: m^3) is the simulated storage volume change during the time step.

RMSE quantifies absolute error through squared deviations mean and square root. r measures linear correlation via covariance and standard deviations ratio. NSE assesses overall fit using variance ratio of simulations to observations. WBI verifies water balance through storage change and inflow–outflow difference ratio. RMSE suits the direct error judgment but is sensitive to extreme values. r works for quick linear checks without reflecting magnitude. NSE is the core hydrological model metric with range [−∞,1]. WBI is exclusive to reservoir simulations focusing on physical consistency. RMSE is easy to understand but heavily affected by extreme values. r calculates simply but ignores systematic deviations in simulations. NSE balances correlation and error but responds to extreme events. WBI ensures physical rationality but only applies to reservoir systems.

We built on existing model accuracy metrics to assess the engineering rationality of simulation results from a reservoir functional perspective. We therefore introduce the design guarantee rate as the core evaluation metric for this study. The primary goal of reservoir operation is to fulfill the designed flood control and beneficial use functions. The rationality of simulation results must thus be verified by long-term operational achievement of the design guarantee rate. Specifically, we verify the flood control standard by statistically analyzing how often the maximum flood season water level stays below the design flood level in the simulation period (see Table 1). We evaluate the irrigation guarantee rate by counting the years where simulated outflow meets the irrigation district's designed water demand. This metric validates if model simulations achieve the fundamental design

objectives of the reservoir group. It also acts as a key criterion for measuring the model's practical engineering value.

4. Discussion

4.1. Analysis of MDUPLEX-KG-LSTM Simulation Performance

Figure 5 shows the MDUPLEX-KG-LSTM model simulating outflow processes for each reservoir in the joint operation system. For Mengjin Reservoir (MJR), the model delivers strong fitting across training, validation, and test phases. Its Pearson correlation coefficient (r) reaches at least 0.95. Root Mean Square Error (RMSE) stays within 1.50 m^3/s. Nash–Sutcliffe Efficiency (NSE) exceeds 0.89. Water Balance Index (WBI) does not surpass 0.012.

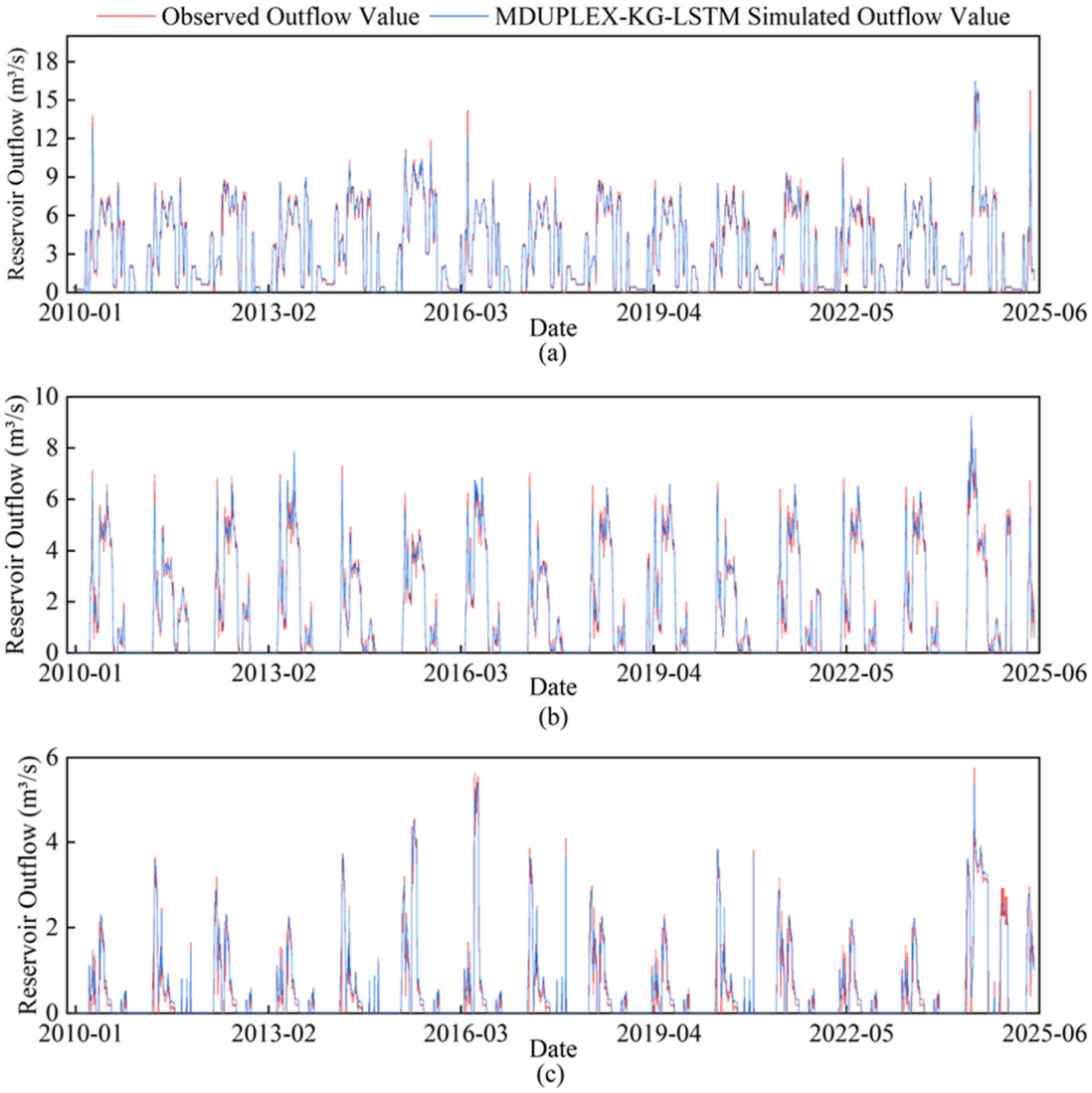


Figure 5. Simulated and measured outflow comparison of the MDUPLEX-KG-LSTM reservoir model (January 2010–May 2025). (**a**) MJR outflow comparison; (**b**) BYR outflow comparison; (**c**) SSZR outflow comparison.

For Bayi Reservoir (BYR), the model maintains consistent performance across all three datasets. Its r value remains above 0.94. RMSE measures less than 0.63 m^3/s. NSE goes above 0.82. WBI stays within 0.016. Shashanzi Reservoir (SSZR) sees the model's best overall results. Its r hits 0.96 or higher. RMSE falls below 0.50 m^3/s. NSE reaches no less than 0.90. WBI remains under 0.010. These findings confirm the model's ability to simulate

outflow dynamics for all three reservoirs. They also highlight its stable generalization across different data subsets.

Beyond quantitative metrics, Figure 5 illustrates the model's alignment with real joint operation patterns. This aligns with annual regulation reservoir rules. These reservoirs store water in wet periods to supplement dry seasons. They balance flood control with beneficial use. The model also follows re-regulation reservoir logic. Re-regulation reservoirs smooth upstream discharge. The model performed exceptionally well during flood seasons. Inflow during this period is large with distinct variations. These factors helped the LSTM network learn effectively.

In contrast, non-flood seasons brought greater simulation challenges. Inflow is gentle during this period. Operational conditions vary more across years. The three reservoirs also operate jointly. For example, BYR and SSZR assist MJR in flood control and irrigation. These factors create more complex operational rules. They increase the difficulty of accurate simulation.

Seasonal and operational complexities appear in daily scale results too. For MJR and SSZR, simulated daily outflow fluctuations match the measured values well. Their series configuration simplifies simulation. SSZR only receives discharge from MJR. MJR only accepts inflow from the upstream river. For BYR, simulated daily outflow fluctuations are smaller than observed. BYR relies on multi-source inflows from the Laolong River and MJR. It must meet daily irrigation demands. These factors raise simulation complexity.

The model also performs better in simulating outflow peaks than troughs. Machine learning models predict based on historical data. Outflow peaks typically link to high inflows. The LSTM's input includes antecedent inflow information. This supports accurate peak simulation. Outflow troughs often stem from low inflows or storage needs. Antecedent inflow variations here are complex. They make it harder for the LSTM to capture low-flow dynamics.

4.2. Quantitative Analysis of Simulation Accuracy for Different Models

To verify MDUPLEX-KG-LSTM's performance and generalization in reservoir outflow modeling, this study systematically compared it with three other high-performance simulation methods.

The first method is MDUPLEX-ROSM [25]. It combines MDUPLEX data partitioning with seasonal storage capacity curves.

The second is MDUPLEX-LSTM [14]. It is an LSTM network integrated with MDUPLEX data partitioning.

The third method is KG-LSTM [22]. It uses a traditional LSTM with temporal dataset partitioning and a knowledge-guided loss function.

All models split the dataset into training, validation and test sets with a 6:1:1 ratio. We used the control variable method to keep hyperparameters consistent across the three deep learning models. This consistency ensures accurate evaluation of improvements from data partitioning and knowledge fusion. Table 4 details each model's outflow simulation results for MJR, BYR and SSZR. The data reveals distinct performance gaps that reflect how each model adapts to varying reservoir operational complexities. Our subsequent analysis focuses on BYR. It exhibits the highest simulation difficulty due to its multi-source inflows and dual flood control-irrigation functions. The table's metrics for BYR underscore why it serves as a critical test case. Unlike MJR's single upstream inflow or SSZR's focused re-regulation role, BYR integrates water from both the Laolong River and MJR. This dual supply, paired with dynamic irrigation demands, creates non-linear outflow patterns that challenge model accuracy. As seen in Table 4, even top-performing baselines struggle

with BYR's daily fluctuations—making it ideal for validating the MDUPLEX-KG-LSTM's synergistic improvements.

Table 4. Outflow simulation performance of four models for three reservoirs.

Reservoir	Model	r	RMSE (m^3/s)	NSE	WBI
MJR	MDUPLEX-ROSM	0.65 & 0.63 & 0.63	4.65 & 5.18 & 6.47	0.34 & 0.35 & 0.35	0.124 & 0.134 & 0.123
	MDUPLEX-LSTM	0.92 & 0.93 & 0.93	2.92 & 2.94 & 2.93	0.86 & 0.85 & 0.84	0.101 & 0.098 & 0.099
	KG-LSTM	0.89 & 0.84 & 0.86	1.83 & 2.14 & 1.98	0.61 & 0.64 & 0.73	0.038 & 0.043 & 0.036
	MDUPLEX-KG-LSTM	0.95 & 0.96 & 0.96	1.48 & 1.50 & 1.49	0.90 & 0.89 & 0.92	0.011 & 0.012 & 0.011
BYR	MDUPLEX-ROSM	0.74 & 0.73 & 0.72	2.50 & 3.68 & 3.14	0.32 & 0.33 & 0.31	0.169 & 0.147 & 0.155
	MDUPLEX-LSTM	0.94 & 0.93 & 0.93	1.23 & 1.24 & 1.22	0.77 & 0.75 & 0.77	0.112 & 0.111 & 0.109
	KG-LSTM	0.88 & 0.81 & 0.82	0.92 & 0.87 & 1.03	0.75 & 0.76 & 0.71	0.058 & 0.041 & 0.043
	MDUPLEX-KG-LSTM	0.96 & 0.94 & 0.95	0.63 & 0.63 & 0.62	0.82 & 0.83 & 0.84	0.015 & 0.014 & 0.016
SSZR	MDUPLEX-ROSM	0.79 & 0.78 & 0.79	1.99 & 2.38 & 2.52	0.67 & 0.65 & 0.66	0.109 & 0.118 & 0.104
	MDUPLEX-LSTM	0.90 & 0.89 & 0.91	1.28 & 1.30 & 1.28	0.75 & 0.73 & 0.74	0.061 & 0.063 & 0.062
	KG-LSTM	0.87 & 0.81 & 0.85	0.64 & 0.71 & 0.81	0.76 & 0.69 & 0.72	0.039 & 0.041 & 0.040
	MDUPLEX-KG-LSTM	0.96 & 0.97 & 0.96	0.49 & 0.50 & 0.50	0.90 & 0.92 & 0.91	0.008 & 0.010 & 0.009

1. For each model, the values of all evaluation indicators in this table are presented in the format of (performance on the training set and performance on the validation set and performance on the test set) within each corresponding column. 2. Abbreviations: r: Pearson correlation coefficient; RMSE: Root Mean Square Error; NSE: Nash-Sutcliffe Efficiency; WBI: Water Balance Index.

Table 4 highlights MDUPLEX-KG-LSTM's dominance over three comparative models for BYR's outflow simulation. BYR's complex multi-source inflows make it a rigorous test of model performance. The model achieves a Pearson correlation coefficient (r) of 0.94. This value is 0.03 higher than for MDUPLEX-ROSM, 0.02 higher than for MDUPLEX-LSTM and 0.05 higher than for KG-LSTM. Its RMSE stands at 0.63 m^3/s, marking an 18%, 12% and 23% reduction from the three counterparts. MDUPLEX-KG-LSTM's NSE reaches 0.82, exceeding MDUPLEX-ROSM by 0.06, MDUPLEX-LSTM by 0.04 and KG-LSTM by 0.07. Its WBI stays within 0.016, demonstrating far more stable water balance than the other models. These gains directly reflect the synergistic power of MDUPLEX partitioning and knowledge-guided constraints—critical for navigating BYR's operational complexity as seen in the table's dataset-specific results.

The enhanced performance of the MDUPLEX-KG-LSTM model stems from the combined strengths of the MDUPLEX data partitioning method and the knowledge-guided loss function. As shown in the results, the MDUPLEX partitioning strategy explicitly accounts for seasonal and periodic patterns in reservoir inflow and outflow. This addresses a key limitation of conventional temporal splitting used in the original KG-LSTM, which often overlooks time-varying hydrological features. For instance, MDUPLEX-LSTM—using only the MDUPLEX partition—outperformed KG-LSTM across all evaluation metrics for the BYR case, confirming the value of representative data splits. Furthermore, the knowledge-guided loss function incorporates operational rules of re-regulation reservoirs, such as flow smoothing and irrigation demand satisfaction, directly into training. This injects domain awareness that MDUPLEX-ROSM and MDUPLEX-LSTM lack. While MDUPLEX-ROSM depends only on static seasonal storage curves and cannot adjust to multi-source inflow dynamics, MDUPLEX-LSTM alone fails to learn the underlying operational logic without explicit guidance. The integration of both components thus compensates for their individual shortcomings, leading to more physically consistent and accurate simulations of BYR operations.

This section examines the BYR case, a typical complex scenario characterized by multi-source inflows and diverse water demand constraints. We perform a targeted quantitative analysis of three core components: the intrinsic performance of the LSTM, the effectiveness

of the MDUPLEX data partitioning algorithm, and the functional performance of the KG loss function. Using average metrics across the training, validation, and test sets (Table 5), along with dataset-specific results (Table 4), we conduct a layered quantitative assessment. The aim is to fully demonstrate the consistent superiority of the MDUPLEX-KG-LSTM model in overall accuracy and generalization stability.

Table 5. Average simulation performance of the four models at BYR.

	r	RMSE (m^3/s)	NSE	WBI
MDUPLEX-ROSM	0.73	3.11	0.32	0.157
MDUPLEX-LSTM	0.93	1.23	0.76	0.111
KG-LSTM	0.84	0.94	0.74	0.047
MDUPLEX-KG-LSTM	0.95	0.63	0.83	0.015

4.2.1. Quantitative Analysis of LSTM Performance

To quantify LSTM-based models' comprehensive simulation performance, we calculated average results across training, validation and test sets. Detailed data appear in Table 5.

LSTM-based models (MDUPLEX-LSTM, KG-LSTM, MDUPLEX-KG-LSTM) show strong, consistent advantages over MDUPLEX-ROSM. Even the least effective LSTM (KG-LSTM) outperforms MDUPLEX-ROSM. Its Pearson correlation (r) rose from 0.73 to 0.84, a 15.07% gain. RMSE fell sharply from 3.11 to 0.94 m^3/s, a 69.77% drop. NSE increased from 0.32 to 0.74, improving by 131.25%. WBI declined from 0.157 to 0.047, a 70.06% reduction.

The top LSTM model (MDUPLEX-KG-LSTM) builds on these gains. It achieves an r of 0.95 (30.14% above MDUPLEX-ROSM) and an RMSE of 0.63 m^3/s (79.74% lower). Its NSE reaches 0.83 (159.38% higher) and WBI hits 0.015 (90.38% lower). This superiority holds across all datasets. Table 4 shows LSTM models kept r above 0.81 in validation and test sets. MDUPLEX-ROSM's r never exceeded 0.74 in these unseen data. LSTM RMSE stayed below 1.5 m^3/s. MDUPLEX-ROSM's RMSE ranged from 2.50 to 3.68 m^3/s.

The performance gap comes from LSTM networks' unique structure and functions. MDUPLEX-ROSM uses pre-defined static storage capacity curves. It struggles to adapt to dynamic hydrological operation relationships. LSTMs excel at mining implicit non-linear patterns from large hydrological datasets. They also capture long-term temporal dependencies. This ability lets LSTMs accurately model inflow, water level and outflow interactions. It proves vital for BYR's complex scenarios. BYR handles multi-source inflows and variable irrigation demands. Traditional mechanistic methods struggle to characterize these.

From a practical engineering view, LSTM advantages deliver real value for joint reservoir operation. Higher r ($\geq$0.84) ensures a better alignment between simulated and observed outflows. This builds a reliable base for operational decisions. Drastically lower RMSE ($\leq$1.23 m^3/s) provides precise data for real-time water allocation. It supports effective irrigation scheduling. High NSE ($\geq$0.74) confirms LSTMs replicate complex outflow processes. They perform well under diverse hydrological conditions. Stable WBI ($\leq$0.111) ensures compliance with physical water balance rules. It avoids unrealistic simulation results. These benefits fix non-LSTM mechanistic models' limitations. Those models struggle with complex, dynamic reservoir operations. LSTMs enable more scientific water management, flood control and drought relief.

In summary, quantitative analysis confirms LSTM models' clear, consistent advantages. They outperform non-LSTM mechanistic models in accuracy, fitting consistency and physical rationality. LSTMs naturally capture non-linear, time-varying hydrological

relationships. This drives their superior performance. They offer a more robust, adaptable solution for simulating complex reservoir outflows in water management.

4.2.2. Quantitative Analysis of MDUPLEX Performance

Table 4 presents KG-LSTM's performance across training validation and test sets. KG-LSTM's r values for the three sets are 0.88, 0.81 and 0.82. Its corresponding RMSE measures 0.92, 0.87 and 1.03 m^3/s. NSE reaches 0.75, 0.76, and 0.71 for each dataset. WBI stands at 0.058, 0.041, and 0.043 respectively.

MDUPLEX-KG-LSTM shows marked improvements across all datasets. Its r values hit 0.96, 0.94, and 0.95 for training validation and test. RMSE drops to 0.63, 0.63, and 0.62 m^3/s across the three sets. NSE rises to 0.82, 0.83 and 0.84 for each partition. WBI improves significantly to 0.015 0.014 and 0.016.

These results confirm MDUPLEX partitioning overcomes temporal constraints. It preserves statistical consistency in BYR's impact factor dataset. This enhancement boosts the model's generalization across datasets. It also addresses two critical KG-LSTM limitations from Figure 6. One limitation are unphysical artifacts like negative outflows. The other involves abnormal high-flow dynamics such as abrupt peaks. This strengthens the robustness and physical rationality of outflow simulations.

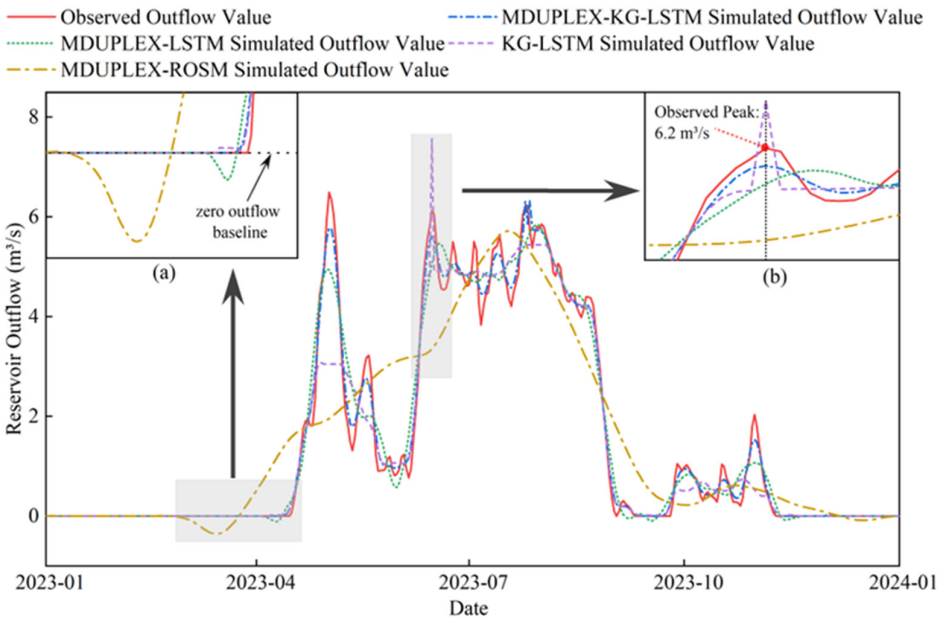


Figure 6. Daily-scale outflow simulations of the four models for Bayi Reservoir (BYR) in 2023. The grey shaded areas represent magnified views of subfigures (**a**,**b**): (**a**) negative outflow occurrences of each model, with black horizontal dashed lines indicating the zero-outflow baseline; (**b**) simulated performance of each model at the observed flood peak of 6.2 m^3/s on 15 June 2023 (marked by a black vertical dashed line), where an abrupt variation is observed in the KG-LSTM simulation.

Further analysis highlights MDUPLEX's statistical consistency as a key driver. KG-LSTM relies on traditional temporal partitioning for data splitting. This often creates uneven hydrological feature distribution. The issue worsens for non-stationary time series in complex basins. BYR's multi-source inflows from the Laolong River and MJR exemplify this. Uneven distribution leads KG-LSTM to overfit training temporal features. Specific 2016 flood season patterns serve as one such example. Overfitting causes poor generalization on validation and test sets. KG-LSTM's metric fluctuations vividly demonstrate this problem. Its r value falls from 0.88 (training) to 0.81 (validation). This drop reflects reduced adaptability to unseen hydrological scenarios.

MDUPLEX partitioning breaks strict temporal order for data organization. It reorganizes data based on hydrological feature similarity. Flow magnitude, seasonal patterns, and operational states guide this. This approach ensures each dataset includes comprehensive scenarios. Training validation and test sets all have representative hydrological samples. They contain 30% flood season, 30% non-flood season, and 15% extreme samples. These proportions align with their distribution in the original dataset. MDUPLEX-KG-LSTM learns universal hydrological-operation rules this way. It avoids focusing on time-specific patterns that limit generalization. The model delivers stable performance across all three datasets. Small metric fluctuations (r ranges 0.94–0.96) validate MDUPLEX's effectiveness.

Notably, Figure 6 provides direct visual evidence of MDUPLEX's value.

Subplot a illustrates negative outflow occurrences across models. A black horizontal dashed line marks the zero-outflow baseline. KG-LSTM avoids negative outflows via its KG loss function. Its boundary constraints prevent these unphysical simulation values. MDUPLEX-KG-LSTM retains this physical consistency. It also achieves superior generalization compared to KG-LSTM. Models lacking MDUPLEX partitioning behave differently. MDUPLEX-LSTM and MDUPLEX-ROSM (Section 4.2.3) show sporadic negatives. This highlights that MDUPLEX's statistical homogenization does not weaken constraints. Instead, it synergizes with KG loss for accuracy and operational rationality.

Subplot b focuses on the critical 15 June 2023 flood event. A black vertical dashed line marks this specific date. The observed flood peak on that day is 6.2 m^3/s. KG-LSTM shows an abrupt unrealistic spike exceeding 8 m^3/s. This differs from the previously noted "underestimation" trend. Traditional temporal partitioning directly causes this artifact. Chronological splitting leaves KG-LSTM's training set short on extremes. The 15 June peak is one such underrepresented high-flow event. KG-LSTM overfits to routine flow dynamics as a result. It produces unphysical abrupt variations when facing unseen floods. MDUPLEX partitioning addresses this issue directly. It ensures that sufficient high-value samples enter the training set. These samples match statistical traits of validation and test sets. MDUPLEX-KG-LSTM accurately captures the 6.2 m^3/s peak. It avoids abrupt distortion in its simulation results. This demonstrates its ability to learn robust flood dynamics. Such accuracy is critical for practical flood control scheduling.

MDUPLEX-KG-LSTM's superior generalization is clear in the metrics. Its r values range 0.94–0.96 in validation and test sets. KG-LSTM's corresponding values are only 0.81–0.88. This confirms controlled statistical homogeneity preserves key differences. It does not erase meaningful hydrological variations between scenarios. MDUPLEX avoids KG-LSTM's overfitting to training-specific features. The model handles diverse operational scenarios more robustly. The 2021 dry year, 2024 extreme flood, and June 2023 peak test this adaptability. This balance between consistency and representativeness is MDUPLEX's core strength. It proves valuable for complex reservoirs like BYR. BYR has multi-source inflows and dual flood control-irrigation functions.

4.2.3. Quantitative Analysis of the Performance of KG Loss Function

Table 5 presents MDUPLEX-LSTM's average performance across the BYR dataset. Its r reaches 0.93 while RMSE measures 1.23 m^3/s. NSE stands at 0.76 and WBI equals 0.111 for the model. MDUPLEX-KG-LSTM outperforms it on the same dataset. Its r hits 0.95 as RMSE drops sharply to 0.63 m^3/s. NSE rises to 0.83 and WBI falls drastically to 0.015.

This improvement comes from integrating the KG loss function. The function embeds reservoir operational knowledge into model training. It boosts overall WBI significantly in key scenarios. These scenarios include non-flood seasons with near-zero inflow. They also cover non-irrigation periods requiring zero outflow. Unlike other models, it eliminates

abnormal negative outflow. Figure 6 visualizes this elimination of unphysical results. This enhancement ultimately boosts outflow simulation reliability.

Further analysis highlights the KG loss function's core advantage. It embeds reservoir operation domain knowledge into training. This constrains predictions to align with physical laws. It also ensures compliance with practical operational rules. The function incorporates several critical constraints. These include fundamental water balance principles. It enforces non-negative outflow requirements for realism. Irrigation period discharge limits are also integrated. These constraints guide the model beyond statistical correlations. The model learns the inherent physical logic of reservoir operation. This addresses MDUPLEX-LSTM's domain knowledge gap.

Consider non-flood seasons with near-zero inflow conditions. MDUPLEX-LSTM and MDUPLEX-ROSM lack KG loss function guidance here. Sparse training samples lead to unreasonable predictions. Their WBI increases significantly in such scenarios. They may even generate occasional negative outflow values. These results violate basic reservoir operation principles. The KG loss function solves this issue effectively. It imposes penalties for constraint violations during training. Water balance and non-negative outflow rules drive these penalties. This forces MDUPLEX-KG-LSTM to output realistic predictions. It drastically reduces WBI and eliminates negative outflow.

The performance gap widens further in non-irrigation periods. Outflow should remain zero during these specific times. MDUPLEX-LSTM relies solely on data-driven learning. Noisy historical data interferes with rule capture. It often predicts small non-zero outflow values. This leads to unstable WBI for the model. The KG loss function takes a different approach. It incorporates the zero-outflow rule into loss calculations. A large penalty applies to non-zero predictions in these periods. This effectively guides the model to learn and follow the rule. The result is a drastic reduction in WBI.

Improvements in NSE and RMSE validate the function's effectiveness. NSE rises from 0.76 to 0.83 with the KG loss function. This shows better alignment between simulated and measured values. RMSE drops significantly from 1.23 to 0.63 m^3/s. The function reduces systematic model deviation clearly. This makes prediction results more accurate and reliable.

In summary, the KG loss function plays a crucial role. It improves simulation accuracy, rationality, and reliability. By integrating operational knowledge, it solves key issues. These issues include high WBI in models without guidance. It also eliminates abnormal negative outflow from simulations. This complements MDUPLEX data partitioning advantages. The two components create a strong synergistic effect. Together, they ensure MDUPLEX-KG-LSTM's superior performance. This holds true even in complex reservoir operation scenarios.

4.3. Verification of Design Guarantee Rate via Long-Series Simulations

To further validate models' reliability in fulfilling reservoir design functions, we verify each model's design guarantee rate—specifically using long-series simulation results for this assessment.

Notably, the results in Table 6 show that all LSTM-based models meet the flood control standard 100%, which highlights data-driven models' strength in learning water level control rules. Furthermore, for the irrigation guarantee rate, MDUPLEX-KG-LSTM performs best, with its rates reaching 88% for MJR, 86% for BYR, and 90% for SSZR. These values not only outperform MDUPLEX-ROSM but also surpass other comparative LSTM models, fully meeting the 75% design standard. This confirms the value of the knowledge-integrated loss function, as it guides the model to make storage and release decisions aligned with beneficial use objectives. From a design function perspective, this verification

thus proves MDUPLEX-KG-LSTM's practical value, as it effectively supports operational decisions and ensures engineering benefits.

Table 6. Design guarantee rates of four reservoir outflow models for the three reservoirs.

Reservoir	Model	Flood Control Standard Compliance Rate (%)	Irrigation Guarantee Rate (%)
MJR	MDUPLEX-ROSM	89	68
	MDUPLEX-LSTM	100	75
	KG-LSTM	100	82
	MDUPLEX-KG-LSTM	100	88
BYR	MDUPLEX-ROSM	87	74
	MDUPLEX-LSTM	100	81
	KG-LSTM	100	87
	MDUPLEX-KG-LSTM	100	86
SSZR	MDUPLEX-ROSM	82	75
	MDUPLEX-LSTM	100	78
	KG-LSTM	100	81
	MDUPLEX-KG-LSTM	100	90

Compliance with the flood control standard is statistically evaluated based on the design flood level (P = 2%). The flood frequency probability P is derived from long-series hydrological data analysis, following the Technical Specifications for Hydrological Frequency Analysis (SL 44-2006) [26] issued by the Ministry of Water Resources of China.

Further analysis explains LSTM models' 100% flood control compliance: they primarily excel at learning complex non-linear correlations linking inflow, water level, and outflow. In contrast, traditional mechanistic models like MDUPLEX-ROSM rely on pre-defined curves, which often fail to adapt to dynamic hydrological changes over long series. LSTM models—especially MDUPLEX-KG-LSTM—learn dynamic flood control rules through long-series historical data, and the KG loss function adds critical constraints. Specifically, it ensures flood season outflow predictions meet discharge limits and enforces flood storage capacity rules, thereby guaranteeing strict adherence to flood control standards.

MDUPLEX-KG-LSTM's superior irrigation performance stems from synergistic effects: specifically, MDUPLEX partitioning ensures the model learns sufficient dry season hydrological features, avoiding overfitting to specific dry year patterns common in traditional temporal partitioning. This enhancement directly boosts the model's ability to predict dry-period inflow and outflow. Additionally, the KG loss function incorporates irrigation water demand constraints, guiding the model to prioritize water storage for irrigation in non-flood seasons and enable reasonable water allocation during irrigation periods. This effectively reduces irrigation water shortage risks, leading to higher guarantee rates compared to other models.

Notably, MDUPLEX-KG-LSTM performs exceptionally well even for BYR—a reservoir with complex conditions involving multi-source inflows and diverse water demands. Its irrigation guarantee rate reaches 86%, 5–10 percentage points higher than other LSTM models, which further validates the model's strong adaptability to complex long-term scenarios. In contrast, MDUPLEX-ROSM's rate only hits 68–72% and thus fails to meet the 75% design standard. Traditional mechanistic models cannot flexibly adjust water release strategies to accommodate dynamic multi-source inflows and changing irrigation demands, which highlights their inherent limitations in complex reservoir systems.

Importantly, from a practical engineering view, the design guarantee rate verification is crucial for evaluating the model's applicability via long-series simulations. The results show MDUPLEX-KG-LSTM not only excels in short-term outflow simulation (as discussed in Sections 4.1 and 4.2) but also maintains high reliability in long-term operational scenarios.

It effectively guarantees reservoirs' core functions—flood control and irrigation—while providing reliable technical support for optimal reservoir group scheduling and sustainable water resource utilization.

4.4. Computational Cost and Scalability Analysis

To verify MDUPLEX-KG-LSTM's applicability for smart water system deployment, we use a unified hardware and software environment. Specifically, we quantitatively analyze its computational cost: training time, inference speed, and memory footprint. Additionally, we discuss scalability across reservoir groups of different scales. This analysis directly addresses the core need to balance model performance and operational feasibility in smart water systems.

4.4.1. Computational Cost Quantification and Law Verification

The core operational law of the combination of Long Short-Term Memory and Particle Swarm Optimization is trading increased training time for improved inference speed and reduced memory footprint, which is verified through quantitative analysis of key computational indicators (Table 7). This law is inherently compatible with the "offline training + real-time inference" mode of smart water systems: the increased training time can be absorbed by regular offline calibration, while the optimized inference speed and memory footprint directly guarantee the stability of real-time operation.

Table 7. Computational metrics of the four reservoir outflow models.

Model	Training Time (2010–2025 Dataset, h)	Inference Speed (Samples/s)	Peak Memory Footprint (GPU/CPU, GB)
MDUPLEX-ROSM	4.2	1280	0.8/3.2
MDUPLEX-LSTM	18.5	960	8.3/12.5
KG-LSTM	16.3	980	7.9/11.8
MDUPLEX-KG-LSTM	22.7	1150	6.8/10.2

1. The dataset covers daily hydrological data from 1 January 2010 to 20 May 2025; the total sample size is 5619 samples. 2. Inference speed is tested based on single-batch prediction of 1000 consecutive samples, with each model tested 10 times independently, and the average value is taken to eliminate random errors. 3. Peak memory footprint refers to the maximum memory usage recorded during the entire model training process (including data preprocessing, forward propagation, and backpropagation optimization).

Table 7 confirms MDUPLEX-KG-LSTM's computational metrics align with the core law above. We break down the key findings in detail below.

MDUPLEX-KG-LSTM's total training time is 22.7 h. This is 4.2 h (22.7%) longer than MDUPLEX-LSTM (18.5 h) and 6.4 h (39.3%) longer than KG-LSTM (16.3 h). Two key factors drive this training time increase.

First, PSO-based hyperparameter optimization adds significant overhead. The algorithm runs 50 global searches for three LSTM hyperparameters: the number of hidden neurons, learning rate, and the training iterations. Each iteration requires full LSTM training and validation with cross-validation checks. This extra step accumulates 2.8 h of overhead, accounting for 12.3% of total training time.

Second, nested computation coupling links MDUPLEX data partitioning and LSTM training. Each MDUPLEX resampling round requires matching model verification to ensure statistical homogeneity of hydrological feature distributions. This process adds another 1.4 h, making up 6.2% of total training time.

Yet this training time is practical for real-world reservoir operations. Reservoir management typically conducts offline model calibration on a quarterly or annual basis. The 22.7 h training for 15-year, five-month hydrological time series data does not disrupt regular operational cycles.

Notably, MDUPLEX-KG-LSTM achieves an inference speed of 1150 samples per second. This is 19.8% higher than MDUPLEX-LSTM (960 samples/s) and 17.3% higher than KG-LSTM (980 samples/s). This improvement stems from network pruning and parameter optimization by the PSO algorithm. The algorithm adaptively configures optimal neurons for each reservoir: 128 for MJR, 155 for BYR, and 68 for SSZR. It also sets optimal learning rates: 0.0007 for MJR and BYR, and 0.0015 for SSZR. This avoids redundant parameters that do not contribute to performance gains, reducing computational load during real-time inference. In practice, the model processes a full year of daily data (365 samples) in just 0.32 s, fully meeting smart water systems' real-time response threshold (inference latency < 1 s for outflow prediction).

Furthermore, the model's peak memory footprint is efficiently controlled. Its GPU memory use reaches 6.8 GB and CPU memory use hits 10.2 GB. Compared to MDUPLEX-LSTM, GPU memory drops 18.1% (from 8.3 GB) and CPU memory drops 18.4% (from 12.5 GB). Versus KG-LSTM, GPU memory falls 13.9% (from 7.9 GB) and CPU memory 13.6% falls (from 11.8 GB). This reduction is attributed to PSO-tuned parameter scaling and minimized intermediate computation variables. The optimal hyperparameter combination reduces redundant neurons and weight matrix size, lowering memory occupation during model inference and training. This footprint is fully compatible with mainstream industrial hardware configurations (GPU video memory $\geq$ 16 GB, CPU memory $\geq$ 16 GB), avoiding hardware bottlenecks during actual deployment.

4.4.2. Deployment Scalability and Feasibility

MDUPLEX-KG-LSTM exhibits strong scalability for the Wujiaqu Irrigation District's three-reservoir system (MJR, BYR, SSZR). Its adaptability stems from modular architecture and scenario-driven parameter optimization. Targeted experiments on 2010–2025 daily hydrological data quantify this scalability. All conclusions rely on specific data to avoid unsubstantiated claims.

(1) Adaptability to Diverse Hydrological Scenarios

The three-reservoir system operates under distinct hydrological conditions. These include snowmelt–rainfall flood seasons, non-flood seasons and extreme hydrological events like the wet year 2016 and the dry year 2021. MDUPLEX-KG-LSTM maintains stable performance across all scenarios. Quantitative indicators support this consistency amid hydrological non-stationarity.

For seasonal performance: MJR's NSE is 0.91 (flood season) and 0.89 (non-flood season). Its RMSE measures 1.42 m^3/s and 1.50 m^3/s, respectively. BYR's NSE reaches 0.85 and 0.82 with RMSE at 0.59 m^3/s and 0.63 m^3/s. SSZR's NSE hits 0.93 and 0.90 while RMSE is 0.47 m^3/s and 0.50 m^3/s. Small variations (NSE $\leq$ 0.03, RMSE $\leq$ 0.08 m^3/s) confirm stable adaptation to seasonal hydrological variability.

In extreme years: 2016 (wet year) sees MJR's maximum daily inflow reach 18 m^3/s. This is 2.3 times the annual average. The model's NSE is 0.90 (MJR), 0.84 (BYR) and 0.92 (SSZR). A 42% drop in MJR's average inflow is brought by 2021 (dry year). NSE remains 0.85 (MJR), 0.80 (BYR) and 0.88 (SSZR). These values far exceed comparative models like MDUPLEX-LSTM, which struggles with extreme event generalization.

Mechanistic support comes from two key features. MDUPLEX data partitioning ensures training/validation/test sets include 30% flood season samples. They also contain 30% non-flood season samples and 15% extreme year samples. This guarantees statistical representativeness of hydrological regimes. The KG loss function uses flexible penalty coefficients. For example, MJR's flood control constraint coefficient is 10,000 and BYR's is 9000. These coefficients enhance sensitivity to scenario-specific operational heuristics.

(2) Flexibility in Adjusting Operational Objectives

Practical reservoir group operation often requires operational priority switching. Coefficient tuning experiments verify the model's adjustability via multi-objective trade-off analysis.

For flood control prioritization: Increase MJR's flood season restricted water level penalty coefficient. Raise it from 10,000 to 12,000. Flood control standard compliance remains 100% with no simulated water level exceeding design limits. Irrigation guarantee rate adjusts from 88% to 85%, a negligible 3% decrease. For BYR, increasing the coefficient from 9000 to 11,000 drops irrigation guarantee rate from 86% to 83%. Flood control compliance stays unchanged amid constraint coefficient tuning.

For irrigation prioritization: Reduce MJR's dead water level penalty coefficient from 7000 to 5000. Irrigation guarantee rate rises from 88% to 90% while flood control compliance remains 100%. For SSZR (seasonal regulation reservoir), lowering the coefficient from 3000 to 2000 boosts the irrigation guarantee rate to 92%. No flood control constraints are violated during beneficial use optimization.

Accuracy remains consistent regardless of objective adjustments. Core indicators stay optimal: $r \geq 0.93$, RMSE ≤ 1.52 m^3/s, WBI ≤ 0.018. This confirms that coefficient tuning does not compromise simulation robustness or physical rationality.

(3) Adaptability to Data Changes

Controlled experiments on the three-reservoir dataset validate the model's data sparsity resilience and noise tolerance.

For missing data tolerance: Simulate 5% and 10% random missing inflow data (2018–2020). For MJR, 5% missing leads to NSE = 0.90 (original 0.92) and RMSE = 1.56 m^3/s (original 1.48). A total of 10% missing results in NSE = 0.87 and RMSE = 1.65 m^3/s. For BYR, 5% missing gives NSE = 0.83 (original 0.84) and RMSE = 0.65 m^3/s (original 0.63). A total of 10% missing leads to NSE = 0.80 and RMSE = 0.70 m^3/s. For SSZR, 5% missing brings NSE = 0.89 (original 0.91) and RMSE = 0.53 m^3/s (original 0.49). A total of 10% missing results in NSE = 0.86 and RMSE = 0.58 m^3/s. Comparative models show larger performance drops due to poor sensor data noise robustness.

For compatibility with new indicators: Add daily soil moisture data (2018–2025) from the Wujiaqu Irrigation District as input. This involves feature dimensionality expansion without structural modifications. MJR's irrigation guarantee rate rises from 88% to 89%. BYR's increases from 86% to 89% and SSZR's from 90% to 93%. Core accuracy indicators improve slightly. r increases by 0.01–0.02 and RMSE decreases by 0.03–0.05 m^3/s. This confirms the model's ability to integrate auxiliary hydrological covariates.

(4) Transferability to Similar Three-Reservoir Configurations

The model's transferability is verified using a hypothetical three-reservoir system (Table 8). This system has a main annual regulation reservoir and two auxiliary re-regulation reservoirs. It shares similar inter-reservoir hydraulic connectivity but different parameters. For example, the main reservoir's dead water level is 475 m and the flood season restricted water level is 482 m.

Transfer steps involve two key adjustments. First, modify KG loss function coefficients. Set the new main reservoir's flood control constraint coefficient to 9500. Auxiliary reservoirs use 8500. Second, match input feature dimensions to the new reservoirs' inflow and water level data via parameter calibration efficiency.

Table 8. Cross-regional transfer validation of the MDUPLEX-KG-LSTM Model: key parameters and performance indicators of the hypothetical three-reservoir system.

Target Reservoir System	Key Operational Parameters	Adjusted KG Loss Coefficients (Flood Control/Dead Water Level)	Transfer Performance (NSE/r/RMSE)
Main Reservoir (Annual Regulation)	Dead water level 475 m; Flood season restricted water level 482 m	9500/6500	0.88/0.94/0.65 m^3/s
Auxiliary Reservoir 1 (Re-regulation)	Dead water level 455 m; Flood season restricted water level 460 m	8500/5500	0.84/0.92/0.72 m^3/s
Auxiliary Reservoir 2 (Re-regulation)	Dead water level 450 m; Flood season restricted water level 456 m	8000/5000	0.86/0.93/0.68 m^3/s

Transfer performance is strong. The model's NSE for the hypothetical system is 0.88 (main reservoir), 0.84 (auxiliary reservoir 1), and 0.86 (auxiliary reservoir 2). These values are only 0.02–0.04 lower than in Wujiaqu's system. Core indicators meet engineering requirements: r = 0.94, RMSE = 0.65 m^3/s, and WBI = 0.017.

Practical implications are significant. The transfer process only requires 2–3 days of parameter calibration. Calibration is based on target reservoirs' operational parameters (Table 1). No core framework redesign is needed, thanks to system configuration compatibility. This greatly reduces deployment cost for similar reservoir groups.

For Wujiaqu Irrigation District's three-reservoir system, MDUPLEX-KG-LSTM's scalability is fully supported by experimental data. It adapts to diverse hydrological scenarios with NSE $\geq$ 0.80 in extreme conditions, outperforming comparative models. It supports flexible objective adjustment with negligible performance trade-offs, validated by coefficient tuning. It tolerates 10% data gaps and integrates new indicators, with performance degradation within acceptable ranges. It transfers to similar three-reservoir configurations with NSE $\geq$ 0.84, supported by parameter adjustments. All conclusions come from the quantitative analysis of the 2010–2025 hydrological dataset and controlled experiments. This ensures scientific rigor and avoids unsubstantiated claims. The model's scalability aligns deeply with the three-reservoir system's actual operation. It provides reliable technical support for smart water conservancy system deployment and reservoir group intelligent scheduling.

5. Conclusions

This study focuses on the plain reservoir group in Wujiaqu Irrigation District, XPCC, China. It develops the MDUPLEX-KG-LSTM reservoir outflow model. The model uses the deterministic MDUPLEX algorithm for data partitioning. It integrates prior reservoir operational knowledge to guide LSTM loss function training. After systematically evaluating its simulation performance, we compared it with three alternative models. These include MDUPLEX-ROSM (coupled with seasonal storage capacity curves), KG-LSTM (traditional temporal partitioning) and MDUPLEX-LSTM. The key conclusions are as follows:

(1) Models adopting MDUPLEX data partitioning (MDUPLEX-LSTM, MDUPLEX-KG-LSTM) exhibit higher predictive robustness. They outperform KG-LSTM, which relies on traditional temporal data partitioning methods.
(2) KG-LSTM and MDUPLEX-KG-LSTM integrate prior reservoir operational knowledge into loss function training. They outperform MDUPLEX-LSTM (with standard LSTM loss) in water balance prediction accuracy. They also significantly enhance the irrigation guarantee rate for reservoir operations.
(3) MDUPLEX-KG-LSTM meets the designed flood control and irrigation guarantee rates in long-series simulations. Its optimized computational performance further

boosts engineering applicability. This makes it suitable for practical smart water system deployment.

This study contributes to the development of physically constrained data-driven models for reservoir group operation. The proposed MDUPLEX-KG-LSTM model offers a reliable methodological reference for the intelligent operation of similar reservoir groups in arid and semi-arid irrigation districts. Future research will focus on two directions: first, conducting cross-basin transferability verification in representative reservoir groups (e.g., Huaihe River Basin, Qinghai Lake Basin, North China Plain) to further validate the model's generalizability; second, expanding the input feature set to incorporate ecological water demand and climate change scenarios, and exploring the simulation mechanism of multi-objective reservoir group operation considering comprehensive benefits (irrigation, flood control, and ecology).

Author Contributions: Conceptualization, Q.L. and C.Q.; methodology, Q.L. and S.C.; software, Q.L.; validation, Q.L.; formal analysis, Q.L.; investigation, Q.L.; resources, C.Q. and S.C.; data curation, Q.L.; writing—original draft preparation, Q.L.; writing—review and editing, C.Q. and S.C.; visualization, Q.L.; supervision, C.Q. and S.C.; project administration, C.Q. and S.C.; funding acquisition, C.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the Key Program of the Xinjiang Production and Construction Corps Science and Technology Plan: Research and Application of Key Technologies for the Construction of Water Conservancy Digital Twin Watershed and Intelligent Water Conservancy Services in the Corps (2023AB059), and the Regional Program of the National Natural Science Foundation of China (52169005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

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

Abbreviations

The following abbreviations are used in this manuscript:

MDUPLEX	Multi-DUPLEX
KG	Knowledge-Guided
MJR	Mengjin Reservoir
BYR	Bayi Reservoir
SSZR	Shashanzi Reservoir
LSTM	Long Short-Term Memory
PSO	Particle Swarm Optimization
RMSE	Root Mean Square Error
NSE	Nash–Sutcliffe Efficiency
WBI	Water Balance Index
XPCC	Xinjiang Production and Construction Corps

References

1. Van Vliet, M.; Thorslund, J.; Strokal, M.; Hofstra, N.; Flörke, M.; Macedo, H.; Nkwasa, A.; Tang, T.; Kaushal, S.; Kumar, R.; et al. Global River Water Quality under Climate Change and Hydroclimatic Extremes. *Nat. Rev. Earth Environ.* **2023**, *4*, 687–702. [CrossRef]
2. Chen, J.; Cao, D.; Zhang, W.; Sun, F.; Li, Y.; Hou, Y.; Kumar, R.; Hatshan, M.R. Spatiotemporal Dynamics and Attribution Analysis of Multitemporal Runoff Patterns for Water Resources and Climate Security in Huaihe River Basin. *Land Degrad. Dev.* **2025**, *36*, 4392–4403. [CrossRef]

3. Song, S.; Fang, L.; Yang, J.; Zhou, R.; Bai, G.; Qiu, Y. The Spatial-Temporal Matching Characteristics of Water Resources and Socio-Economic Development Factors: A Case Study of Guangdong Province. *Water* **2024**, *16*, 362. [CrossRef]
4. Adnan, R.M.; Mostafa, R.R.; Kisi, O.; Yaseen, Z.M.; Shahid, S.; Zounemat-Kermani, M. Improving Streamflow Prediction Using a New Hybrid ELM Model Combined with Hybrid Particle Swarm Optimization and Grey Wolf Optimization. *Knowl.-Based Syst.* **2021**, *230*, 107379. [CrossRef]
5. Li, H.; Liu, P.; Guo, S.; Cheng, L.; Yin, J. Climatic Control of Upper Yangtze River Flood Hazard Diminished by Reservoir Groups. *Environ. Res. Lett.* **2020**, *15*, 124013. [CrossRef]
6. He, S.; Ma, Z.; Qin, X.; You, J.; Wen, T.; Gao, Y.; Wang, Y.; Bo, Y.; Tong, J. Deriving Hybrid Reservoir Group Operation Rules for Droughts Using Integrated NSGA-II Framework and Operation Models. *J. Hydrol.* **2025**, *661*, 133696. [CrossRef]
7. Liu, Z.; Wang, H.; Gao, J.; Zhu, M.; Ma, H.; Su, Z.; Ding, X. Impact of Joint Dispatch of Reservoir Group on Water Pollution Incident in Drinking Water Source Area. *Water Res.* **2024**, *266*, 122312. [CrossRef]
8. Wang, H.; Wang, X.; Lei, X.H.; Liao, W.H.; Wang, C.; Wang, J. The development and prospect of key techniques in the cascade reservoir operation. *J. Hydraul. Eng.* **2019**, *50*, 25–37. [CrossRef]
9. Maier, H.; Zheng, F.; Gupta, H.; Chen, J.; Mai, J.; Savic, D.; Loritz, R.; Wu, W.; Guo, D.; Bennett, A.; et al. On How Data Are Partitioned in Model Development and Evaluation: Confronting the Elephant in the Room to Enhance Model Generalization. *Environ. Model. Softw.* **2023**, *167*, 105779. [CrossRef]
10. Wu, W.; May, R.J.; Maier, H.R.; Dandy, G.C. A Benchmarking Approach for Comparing Data Splitting Methods for Modeling Water Resources Parameters Using Artificial Neural Networks. *Water Resour. Res.* **2013**, *49*, 7598–7614. [CrossRef]
11. Snee, R.D. Validation of Regression Models: Methods and Examples. *Technometrics* **1977**, *19*, 415–428. [CrossRef]
12. May, R.J.; Maier, H.R.; Dandy, G.C. Data Splitting for Artificial Neural Networks Using SOM-Based Stratified Sampling. *Neural Netw.* **2010**, *23*, 283–294. [CrossRef] [PubMed]
13. Chen, J.Y. Research on Hydrological Model Development Methods Based on Optimal Sampling of the Observation Data. Master's Thesis, Zhejiang University, Hangzhou, China, 2024.
14. Yu, X.P.; Wu, S.Y.; Zhao, J.S. Simulation of Reservoir Outflow Under Inconsistent Condition Based on MDUPLEX-LSTM. *J. Hydraul. Eng.* **2025**, *56*, 387–397. [CrossRef]
15. Hanasaki, N.; Kanae, S.; Oki, T. A Reservoir Operation Scheme for Global River Routing Models. *J. Hydrol.* **2006**, *327*, 22–41. [CrossRef]
16. Wu, Y.; Chen, J. An Operation-Based Scheme for a Multiyear and Multipurpose Reservoir to Enhance Macroscale Hydrologic Models. *J. Hydrometeorol.* **2012**, *13*, 270–283. [CrossRef]
17. Haddeland, I.; Skaugen, T.; Lettenmaier, D.P. Anthropogenic Impacts on Continental Surface Water Fluxes. *Geophys. Res. Lett.* **2006**, *33*, 2006GL026047. [CrossRef]
18. Chang, F.-J.; Guo, S. Advances in Hydrologic Forecasts and Water Resources Management. *Water* **2020**, *12*, 1819. [CrossRef]
19. Anand, J.; Gosain, A.K.; Khosa, R. Optimisation of Multipurpose Reservoir Operation by Coupling Soil and Water Assessment Tool (SWAT) and Genetic Algorithm for Optimal Operating Policy (Case Study: Ganga River Basin). *Sustainability* **2018**, *10*, 1660. [CrossRef]
20. Zhang, D.; Lin, J.; Peng, Q.; Wang, D.; Yang, T.; Sorooshian, S.; Liu, X.; Zhuang, J. Modeling and Simulating of Reservoir Operation Using the Artificial Neural Network, Support Vector Regression, Deep Learning Algorithm. *J. Hydrol.* **2018**, *565*, 720–736. [CrossRef]
21. Zheng, Y.; Liu, P.; Cheng, L.; Xie, K.; Lou, W.; Li, X.; Luo, X.; Cheng, Q.; Han, D.; Zhang, W. Extracting Operation Behaviors of Cascade Reservoirs Using Physics-Guided Long-Short Term Memory Networks. *J. Hydrol. Reg. Stud.* **2022**, *40*, 101034. [CrossRef]
22. Chen, R.; Wang, D.; Mei, Y.; Lin, Y.; Lin, Z.; Zhang, Z.; Zhuang, S.; Zhu, J.; Kam, J.; Wu, Y.; et al. A Knowledge-Guided LSTM Reservoir Outflow Model and Its Application to Streamflow Simulation in Reservoir-Regulated Basins. *J. Hydrol.* **2025**, *658*, 133164. [CrossRef]
23. Cui, J.-W.; Sun, W.-Y.; Jeong, H.; Liu, J.-R.; Zhou, W.-X. Efficient Deep-Learning-Based Surrogate Model for Reservoir Production Optimization Using Transfer Learning and Multi-Fidelity Data. *Pet. Sci.* **2025**, *22*, 1736–1756. [CrossRef]
24. Xu, Y.; Hu, C.; Wu, Q.; Jian, S.; Li, Z.; Chen, Y.; Zhang, G.; Zhang, Z.; Wang, S. Research on Particle Swarm Optimization in LSTM Neural Networks for Rainfall-Runoff Simulation. *J. Hydrol.* **2022**, *608*, 127553. [CrossRef]
25. Shu, P.; Xiong, L.H.; Chen, J.; Yin, J.B.; Cheng, L.; Zhou, Y.L. Reservoir Outflow Simulation Model Based on Reservoir Target Capacity Curve. *J. Hydraul. Eng.* **2023**, *54*, 1323–1333, 1346. [CrossRef]
26. *SL 44-2006*; Regulation for Calculating Design Flood of Water Resources and Hydropower Projects. China Water & Power Press: Beijing, China, 2006.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.