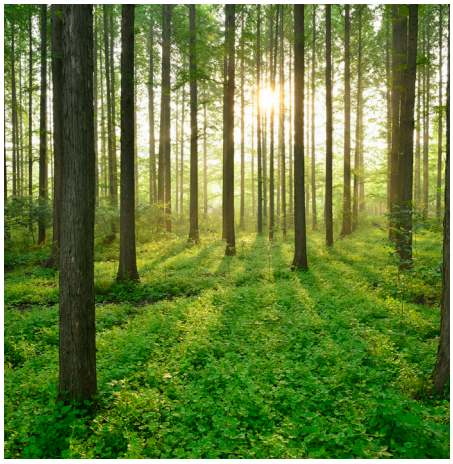
DEEP LEARNING-BASED RESERVOIR INFLOW FORECASTING

REPORT 2026:1228

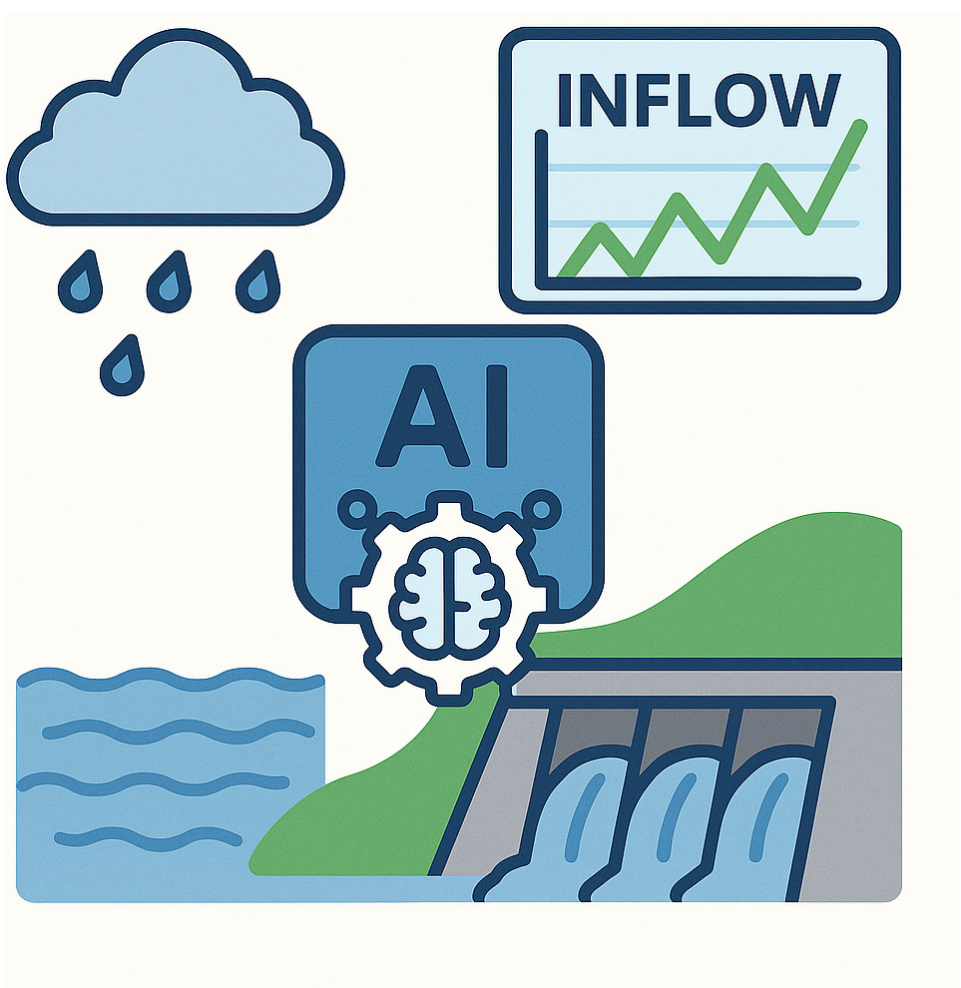




HUVA







Energiforsk

Deep learning-based reservoir inflow forecasting

A case study of the Överuman and Sädvajaure Reservoirs

SHICHENG LI (KTH), LUIGIA BRANDIMARTE (KTH)
JAMES YANG (VATTENFALL & KTH)

ISBN 978-91-89917-71-2| © Energiforsk November 2026

Energiforsk AB | Phone: 08-677 25 30 | E-mail: kontakt@energiforsk.se | www.energiforsk.se

Foreword

This project investigates the application of deep learning to predict reservoir inflows in Swedish hydropower systems, offering a data-driven tool to enhance decision-making for water resources management. It focuses on two regulated reservoirs in northern Sweden: Överuman in the Ume River basin and Sädvajaure in the Skellefte River basin, both of which are representative of cold-climate hydrological regimes influenced by snow accumulation and melt.

The project has been a part of Energiforsks research and development programme HUVA, which focuses on hydrological and hydraulical challenges of the hydropower industry. The financiers of this research are Vattenfall Vattenkraft AB, Fortum Sverige AB, Sydkraft Hydropower AB, Statkraft Sverige AB, Skellefteå Kraft AB, Holmen Energi AB, Tekniska Verken i Linköping AB, Mälarenergi Vattenkraft AB och Sollefteåforsen AB.

The authors of the report are Shicheng Li from KTH, Luigia Brandimarte from KTH and James Yang from KTH and Vattenfall R&D. The reference group for the project was comprised of Beatriz Quesada Montano (Sydkraft Hydropower AB), Björn Norell (Vattenregleringsföretagen AB), Imenne Ahlen (Fortum AB), Magnus Rylander (Holmen Energi AB), Mikael Sundby (Vattenfall AB), and Susanne Nyström (Vattenfall AB). At a later stage, Alexandra Hamm (Vattenfall AB) also joined.

The image on the front cover was generated by the authors using Chat GPT. The authors are responsible for the contents and conclusions of this report.

Summary

Accurate hydrological forecasting is crucial for the sustainable and efficient operation of hydropower dams. This project investigates the application of deep learning to predict reservoir inflows in Swedish hydropower systems, offering a data-driven tool to enhance decision-making for water resources management.

The study is conducted within the HUVA research program, coordinated by Energiforsk. It focuses on two regulated reservoirs in northern Sweden: Överuman in the Ume River basin and Sädvajaure in the Skellefte River basin, both of which are representative of cold-climate hydrological regimes influenced by snow accumulation and melt.

Using long-term hydro-meteorological datasets from 2000 to 2023, the project develops and tests deep learning-based models for daily inflow forecasting. The models are trained on key hydrological and meteorological variables, including precipitation, air temperature, water level, and historical inflow. The results demonstrate that deep learning approaches can reproduce the temporal dynamics of inflow with high accuracy (NSE > 0.94, RMSE < 9.0 m^3/s) and robustness across different hydrological regimes. The models perform well in short- and medium-term forecasting horizons, effectively capturing seasonal variability and extreme inflow events. Furthermore, compared with the HBV model, the deep learning approach model achieves higher predictive accuracy for both reservoirs under the tested conditions. Monte Carlo-based uncertainty analysis also shows predictive performance remains stable even under moderate perturbations.

The study also assesses model transferability between the two reservoirs through transfer learning, showing that models trained on one site can be adapted to another using limited local data. This finding highlights the potential for scalable forecasting frameworks applicable to data-scarce catchments. However, local adaptation cannot fully compensate for fundamental differences in catchment and regulatory dynamics.

Overall, the project demonstrates that deep learning offers a promising complement to established hydrological methods, supporting HUVA's long-term goal of improving the precision, adaptability, and resilience of inflow forecasts for hydropower operation. The integration of such data-driven techniques can enhance operational planning, optimize water resource management, and strengthen preparedness for changing climatic and hydrological conditions.

Keywords

Reservoir inflow forecasting, hydropower, deep learning, hydrological modeling, uncertainty quantification, transfer learning

Sammanfattning

Noggranna hydrologiska prognoser är avgörande för en hållbar och effektiv drift av vattenkraftsdammar. I detta projekt undersöks hur djupinlärning kan användas för att förutsäga tillrinningen till vattenmagasin i svenska vattenkraftssystem. Syftet är att utveckla ett datadrivet verktyg som kan förbättra beslutsfattande inom vattenresurshantering.

Studien har genomförts inom Energiforsks forskningsprogramm HUVA. Fokus ligger på två reglerade magasin i norra Sverige: Överuman i Umeälvens avrinningsområde och Sädvajaure i Skellefteälvens avrinningsområde. Båda magasinen är representativa för hydrologiska förhållanden i kallt klimat, där snöackumulering och snösmältning har stor betydelse för vattenflödena.

Med hjälp av långsiktiga hydro-meteorologiska data från perioden 2000–2023 utvecklas och utvärderas djupinlärningsbaserade modeller för prognoser av daglig tillrinning. Modellerna tränas på centrala hydrologiska och meteorologiska variabler, såsom nederbörd, lufttemperatur, vattennivå och historisk tillrinning. Resultaten visar att djupinlärningsmetoderna med hög noggrannhet kan återskapa tillrinningens tidsmässiga variationer (NSE > 0,94, RMSE < 9,0 m^3/s) och uppvisar god robusthet under olika hydrologiska förhållanden. Modellerna presterar väl för både kort- och medellånga prognoshorisonter och fångar effektivt såväl säsongsvariationer som extrema tillrinningstillfällen. Jämfört med HBV-modellen uppnår djupinlärningsmodellen dessutom högre prognosnoggrannhet för båda magasinen under de testade förhållandena. En osäkerhetsanalys baserad på Monte Carlo-simuleringar visar också att modellernas prognosförmåga förblir stabil även vid måttliga störningar.

Studien utvärderar även modellernas överförbarhet mellan de två magasinen med hjälp av transfer learning. Resultaten visar att modeller som tränats för ett magasin kan anpassas till ett annat med hjälp av en begränsad mängd lokala data. Detta belyser potentialen för skalbara prognosramverk som även kan tillämpas i avrinningsområden där datatillgången är begränsad. Samtidigt visar resultaten att lokal anpassning inte fullt ut kan kompensera för grundläggande skillnader mellan avrinningsområdenas hydrologiska och regleringstekniska förutsättningar.

Sammanfattningsvis visar projektet att djupinlärning utgör ett lovande komplement till etablerade hydrologiska metoder och kan bidra till HUVA-programmets långsiktiga mål att förbättra precisionen, anpassningsförmågan och robustheten i tillrinningsprognoser för vattenkraftens drift. Genom att integrera sådana datadrivna metoder kan den operativa planeringen förbättras, vattenresurserna utnyttjas mer effektivt och beredskapen stärkas inför förändrade metereologiska och hydrologiska förhållanden.

List of content

1	**Background**		**7**
2	**Study area and data source**		**9**
	2.1	Study area	9
	2.2	Data source	10
3	**Methodology**		**13**
	3.1	LSTM model	13
	3.2	CNN-LSTM model	14
	3.3	Model setup and configuration	15
	3.4	Evaluation metrics	16
4	**Results and analysis**		**18**
	4.1	Performance evaluation	18
	4.2	Uncertainty quantification	21
	4.3	Model transferability	22
	4.4	Effect of lead times	25
	4.5	Comparison with HBV	29
	4.6	Effect of input features	34
5	**Concluding remarks**		**38**
6	**Acknowledgement**		**39**
7	**References**		**40**

1 Background

Hydropower plays a vital role in Sweden's energy system, providing both renewable electricity and crucial flexibility to balance the increasing shares of intermittent energy sources. Efficient hydropower operation depends fundamentally on accurate runoff and inflow forecasts, which support daily production planning, long-term reservoir management, and dam safety. Therefore, reliable hydrological predictions are essential for sustainable energy production, environmental adaptation, and climate-resilient water management.

Within the Swedish hydropower sector, the Program for Hydrological Development in the Hydropower Industry (HUVA) serves as the central platform for advancing hydrological forecasting and modeling. Established in 1981, HUVA brings together energy companies, SMHI, and academic partners to improve hydrological methods, enhance forecasting tools, and promote knowledge exchange across industry and research. The program focuses on strengthening runoff prediction capability through both model development and improved data availability, with ongoing efforts targeting higher spatial and temporal resolution, flexible model structures, and the integration of modern data-driven approaches.

Traditionally, the HBV model has been the operational backbone of hydrological forecasting in the Swedish hydropower industry. It is used daily for runoff calculation, short-term forecasting, and long-term hydrological assessments. However, as hydropower operations evolve under changing climatic and energy system conditions, there is an increasing demand for forecasts with finer temporal resolution and enhanced representation of nonlinear hydrological processes. Additionally, there is growing interest in complementing traditional models with data-driven approaches that leverage the increasing availability of high-resolution hydrological and meteorological data.

Recent developments in artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), offer promising alternatives or complementary tools for hydrological forecasting. DL models can learn patterns directly from large datasets, enabling the development of predictive models that support real-time forecasting and adaptive water management. By integrating such techniques with established hydrological practices, the hydropower industry can strengthen its forecasting capacity and enhance decision support for both operational and strategic planning.

Against this background, the present project, Deep Learning-Based Reservoir Inflow Forecasting, is initiated within the HUVA research framework to explore the applicability and benefits of DL models for reservoir inflow prediction. The study focuses on two regulated reservoirs in northern Sweden, Överuman and Sädvajaure, representing typical cold-climate hydrological regimes characterized by snow accumulation, melt-driven flow variability, and seasonal patterns. By developing different DL architectures for inflow forecasting, the project aims to:

- Evaluate the predictive performance of DL models for operational inflow forecasting,

- Quantify model uncertainty and robustness,
- Investigate transfer learning as a means to extend model applicability to data-scarce reservoirs, and
- Compare DL-based forecasts against HBV modeling.

This work contributes to HUVA's long-term goals of advancing hydrological forecasting methods, promoting digitalization in hydropower operations, and supporting the transition toward a smarter, climate-adaptive energy system.

2 Study area and data source

2.1 STUDY AREA

Two lakes/reservoirs are selected as representative study sites for developing and validating DL-based inflow forecasting models: Överuman and Sädvajaure. Both reservoirs are located in northern Sweden and form integral components of the hydropower cascade systems operated within the Ume and Skellefte river basins, respectively. These basins are characterized by cold-climate hydrology, with distinct snow accumulation and melt periods that strongly influence seasonal flow regimes.

The Överuman Reservoir (Figure 1a) is a lake and regulating reservoir located in the upper reaches of the Ume River, spanning Storuman Municipality in Västerbotten County, Sweden, and Rana Municipality in Nordland County, Norway. Located northwest of Hemavan, it forms part of the main drainage area of the Ume River. The lake lies at an elevation of approximately 523.8 m above sea level, with a maximum depth of 78.8 m and a surface area of about 84 km² (SMHI, 2026).

As a reservoir, Överuman extends roughly 30 km in length and covers an effective regulated area of 75 km², of which about 3 km² lies within Norway. It has a storage volume of approximately 357 million m³ and a regulation amplitude of around 5 m. Water from the reservoir is conveyed through a 10 km-long tunnel to the Klippen hydropower station, enabling efficient flow regulation and energy production within the Ume River cascade system.

The Sädvajaure Reservoir (Figure 1b) is a natural lake and regulating reservoir located in Arjeplog Municipality, Lapland, within the upper reaches of the Skellefte River. It lies at an elevation of approximately 466.7 m above sea level, with a maximum depth of 120 m and a surface area of 32.1 km². The lake drains into the Skellefte River and forms part of its main catchment area (SMHI, 2026).

Hydrologically, Sädvajaure functions as a major storage reservoir for long-term regulation of hydropower production within the nearly fully developed Skellefte hydropower system. The reservoir underwent successive regulation phases in 1942, 1953, and 1985, resulting in a storage volume of approximately 605 million m³ and a regulation range of 460.7–477 m above sea level (Norrman, 1989).

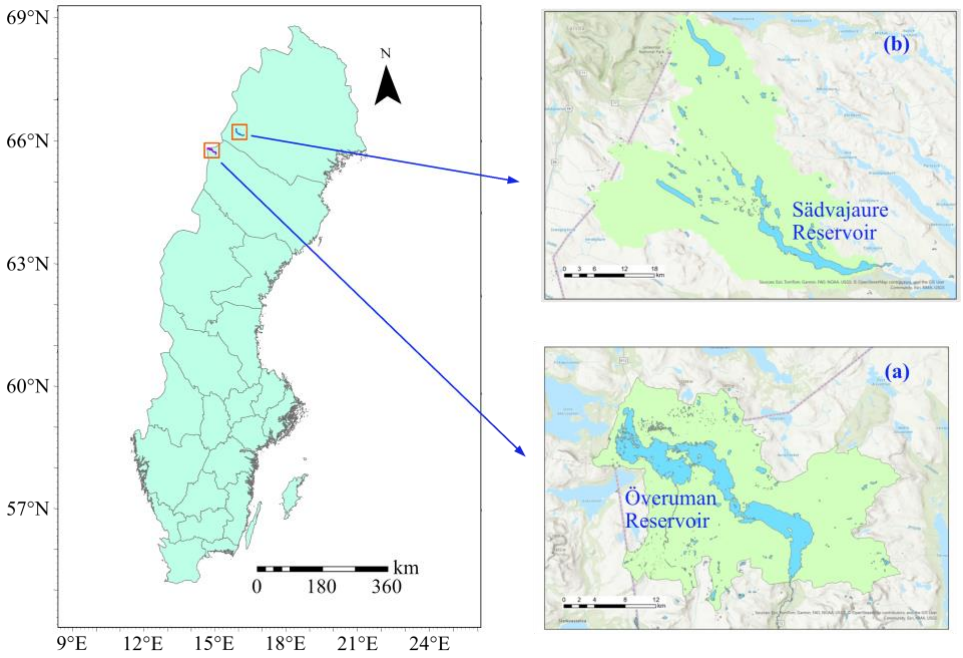


Figure 1. Study area: (a) Överuman and (b) Sädvajaure.

2.2 DATA SOURCE

Long-term hydro-meteorological datasets are collected for both study reservoirs. The data are provided by Vattenregleringsföretagen and Skellefteälvens Vattenregleringsföretag, and are supplemented with meteorological data from the Swedish Meteorological and Hydrological Institute (SMHI). For both reservoirs, data covering the period from January 1, 2000, to December 31, 2023, are obtained. The datasets include daily records of water level (*WL*), precipitation (*P*), air temperature (*T*), and inflow (Q_{in}). These datasets provide consistent temporal coverage spanning over two decades, enabling DL models to capture long-term hydrological variability and climatic influences on inflow dynamics.

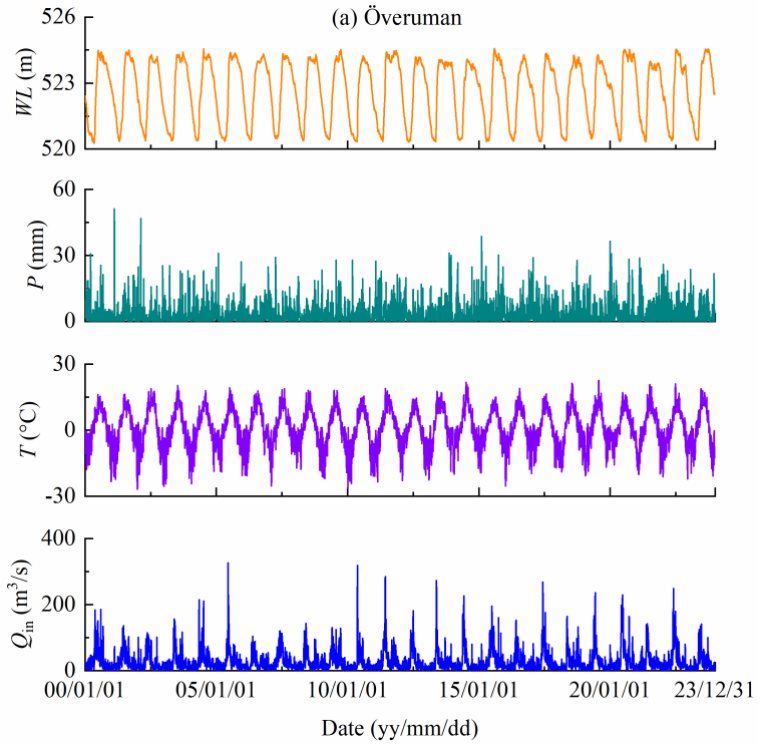
It is noted that the inflow records occasionally contain negative values, which are not physically meaningful. The negative inflows may originate from wind-induced surface water movement, water level oscillations, measurement uncertainty in stage-discharge relationships, and timing mismatches between observations. Anomalous negative inflows are replaced with linear interpolation to generate a continuous inflow time series, ensuring smooth temporal variation without introducing artificial trends. To reduce short-term variability and measurement noise, all variables are also smoothed using a two-day moving average. This small window size is selected to balance noise suppression and temporal resolution, ensuring that short-lived fluctuations are smoothed without distorting the underlying hydrological response.

Table 1 summarizes the main statistics of the collected data (January 1, 2000 – December 31, 2023). The Överuman Reservoir exhibits relatively small water-level fluctuations (520.3–524.6 m), while inflow ranges from 0.1 m^3/s to 326 m^3/s. The mean temperature (1.1 °C) and precipitation (3.1 mm) indicate a cold-climate

catchment with frequent low-temperature conditions. The Sädvajaure Reservoir shows a wider water-level variation (460.7–477.0 m) and higher inflow range (0–471.3 m³/s). Both reservoirs exhibit near-zero skewness for water level and temperature, suggesting approximately symmetric distributions. In contrast, precipitation and inflow exhibit positive skewness, indicating occasional high-magnitude events. The temporal variations of these variables are illustrated in Figure 2.

Table 1. Summary of hydro-meteorological data statistics (January 1, 2000 – December 31, 2023)

Reservoir	Statistics	*WL* (m)	*P* (mm)	*T* (°C)	Q_{in} (m³/s)
Överuman	Minimum	520.3	0	−26.8	0.1
	Maximum	524.6	51.1	22.7	326
	Mean	522.7	3.1	1.1	27.9
	Standard deviation	1.4	4.4	8.7	31.9
	Skewness	−0.4	2.7	−0.3	2.9
Sädvajaure	Minimum	460.7	0.0	−33.7	0.0
	Maximum	477.0	48.4	23.6	471.3
	Mean	469.8	1.5	0.7	39.9
	Standard deviation	5.4	3.3	10.0	57.4
	Skewness	−0.4	5.0	−0.5	2.7



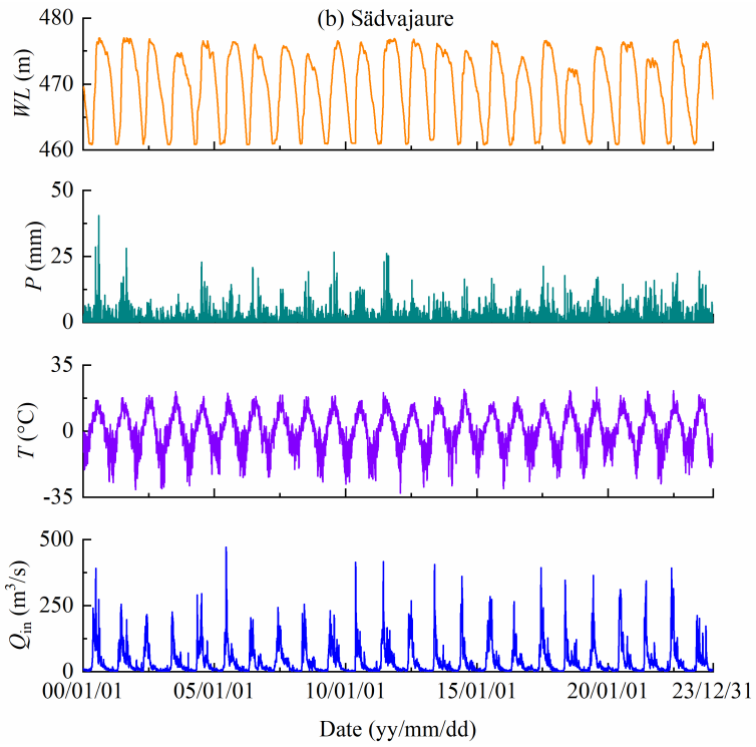


Figure 2. Long-term daily hydro-meteorological observations at the (a) Överuman and (b) Sädvajaure Reservoirs from 2000 to 2023.

3 Methodology

Two DL models are developed and evaluated for inflow forecasting: a Long Short-Term Memory (LSTM) model and a Convolutional Long Short-Term Memory (CNN-LSTM) model.

3.1 LSTM MODEL

The LSTM model is a type of recurrent neural network (RNN) designed to capture temporal dependencies in sequential data. It overcomes the limitations of traditional RNNs by incorporating memory cells and gated mechanisms (input, forget, and output gates), which enable it to retain relevant information over long time intervals while preventing gradient vanishing or explosion during training (DiPietro and Hager, 2020; Li and Yang, 2022).

Figure 3 shows a typical LSTM architecture with gate structures. The LSTM cell consists of three gates: the forget gate (f), the input gate (i), and the output gate (o). The forget gate determines which information from the previous cell state to discard. It takes the last hidden state, h_{t-1}, and the current input, x_t, to produce a value between 0 and 1 for each element in the cell state (DiPietro and Hager, 2020).

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

where ω_f = weight for forget gate, b_f = bias for forget gate, and σ = sigmoid activation function.

The input gate determines which new information to store in the cell state. It consists of a sigmoid layer that decides which values to update and a tanh (hyperbolic tangent) layer that generates new candidate values. The output of the input gate is expressed as:

$$i_t = \sigma\left(\omega_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

where ω_i = weight for input gate, and b_i = bias for input gate. The new candidate value C_n is computed by

$$C_n = \tanh\left(\omega_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{3}$$

where C_{ut} = updated cell state. Lastly, the output gate determines the next hidden state based on the updated cell state. The output of the output gate o_t is obtained by

$$o_t = \sigma\left(\omega_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

where ω_o = weight for output gate, and b_o = bias for output gate. The new hidden state becomes

$$h_t = o_t \times \tanh\left(C_{ut}\right) \tag{5}$$

The output (hidden state or cell state) at the current time step is used for predictions or passed to the next time step for iteration.

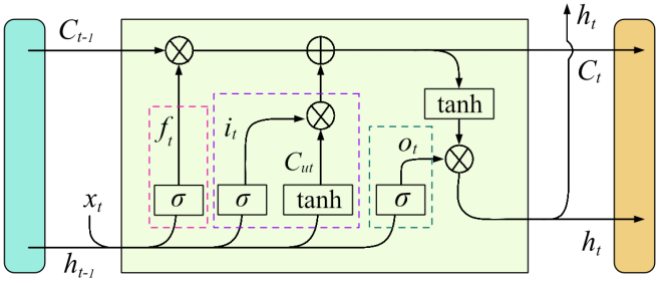


Figure 3. Typical architecture of an LSTM cell, adapted from Li and Yang (2022). x = sequential data, h = output vector, C = cell state, σ = sigmoid function, i = input gate, o = output gate, f = forget gate, and C_{ut} = updated cell state.

3.2 CNN-LSTM MODEL

The CNN-LSTM model combines convolutional and recurrent layers to leverage both spatial and temporal feature extraction (Alhussein et al., 2020). The convolutional layers act as automatic feature extractors, identifying short-term patterns or localized dependencies among the input variables (e.g., rapid changes in inflow or temperature); the LSTM layers, on the other hand, capture longer-term temporal dependencies.

A CNN operates on the principle of convolution, where a small filter (or kernel) slides across the input data to compute feature maps. Each convolution operation emphasizes specific local relationships among variables. Figure 4(a) illustrates a three-kernel convolutional structure (shown in different colors) designed for processing sequential data. Within the convolutional layer, each neuron is connected only to a local subset of neighboring neurons. This enables the network to capture localized patterns while reducing the number of trainable parameters and overall model complexity (LeCun et al., 2010).

After the convolution computation, the resulting feature maps undergo a nonlinear transformation through activation functions Figure 4(b). The Rectified Linear Unit (ReLU) activation function is employed to process the convolution outputs. The kernel-processed results Y_j^k are expressed as:

$$Y_j^k = \mathrm{ReLU}\left(b_j^k + \sum_{i=1}^{N_{k-1}} \mathrm{conv}\left(x_i^{k-1}, \omega_{ij}^{k-1} \right) \right) \tag{6}$$

where where x_j^{k-1} = input value of layer $(k-1)$, ω_{ij}^{k-1} = convolution kernel, b_j^k = threshold valve of layer k, and N_k = number of input features. Subsequently, pooling layers are applied to reduce the temporal dimension of the feature maps, retaining the most informative characteristics of the input sequence. These pooling operations enhance computational efficiency, prevent overfitting, and provide a degree of translation invariance (Zha et al., 2022). After passing through multiple convolutional, activation, and pooling layers, the input signals are transformed into a set of high-level extracted features (Y), which serve as compact representations of the original data for subsequent processing by the LSTM layers.

$$Y = \mathrm{ReLU}\left(\mathrm{ReLU}\left(\cdots \mathrm{ReLU}\left(b_j^k + \sum_{i=1}^{N_{k-1}} \mathrm{conv}\left(x_i^{k-1}, \omega_{ij}^{k-1}\right)\right)\right)\right) \quad (7)$$

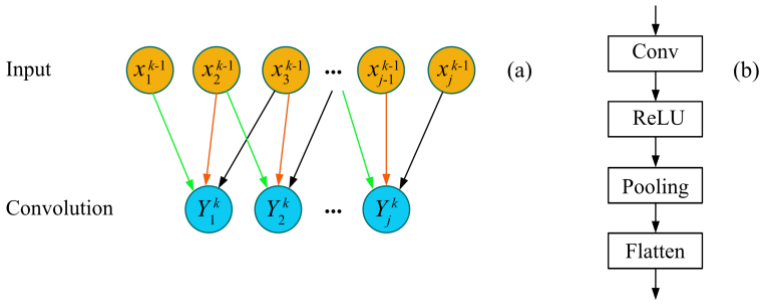


Figure 4. Illustration of a CNN: (a) convolution operation and (b) network structure, adapted from Li et al. (2024).

3.3 MODEL SETUP AND CONFIGURATION

The DL-based inflow forecasting framework consists of five main stages: data preparation, normalization, model configuration, training, and testing. 80% of the sequences are used for training (January 1, 2000 – February 10, 2019) and the remaining 20% for testing (February 11, 2019 – December 31, 2023), which is a typical split ratio in ML modeling (Betts et al., 2022; Katipoğlu, 2023). Daily hydro-meteorological variables (WL, P, T, Q_{in}) are used to construct input-target sequences via a sliding-window approach with a 10-day sequence length, enabling the model to learn temporal dependencies over short-term periods. Each input sequence corresponds to the previous 10 days of meteorological and hydrological conditions, and the model predicts the inflow of the following day (lead time = 1 day).

It is important to clarify the differences in forecasting definitions between the DL and HBV models. In HBV, forecasting relies on externally supplied meteorological forecasts to generate future inflows. In contrast, DL forecasting in this study refers to predicting future inflows from historical data, without using future meteorological forecasts.

To stabilize training and ensure balanced scaling across variables, both the inputs and target were normalized using the z-score standardization:

$$X' = \frac{X - \mu_X}{\sigma_X} \quad (8)$$

where μ_X, σ_X = mean and standard deviation computed from the training set, and X, X' = original and normalized variable. Normalization parameters are applied to both training and testing data.

A univariate LSTM network is developed with the following layers:

- Sequence input layer matching the number of input channels.

- Single LSTM layer with 128 hidden units to capture temporal dependencies.
- Fully connected layer with one output neuron predicting daily inflow.

The network is trained using the Adam optimizer with a maximum of 100 epochs, mean-squared-error (MSE) loss function, and sequence-padding direction set to 'left'.

The CNN-LSTM model consists of the following major components:

- Sequence input layer with four features per time step.
- Sequence folding and unfolding layers to convert sequential data into a form suitable for convolution.
- Convolutional block with a 2-D convolution layer (kernel size = 10, 32 filters), followed by batch normalization, ELU activation, and average-pooling.
- Flattening layer to transform feature maps into a 1-D sequence.
- Two stacked LSTM layers (128 and 40 units, respectively) for capturing long-term temporal dependencies.
- Dropout layer (rate = 0.2) to reduce overfitting.

The model is trained using the Adam optimizer with a piecewise learning-rate schedule. The initial learning rate is 0.01, reduced by a factor of 0.8 every 20 epochs. L2 regularization (0.01) is applied to prevent overfitting, and a gradient threshold of 1 is set to stabilize learning. Training is performed for 100 epochs with a mini-batch size of 24.

For both models, the predicted and observed inflow sequences are de-normalized using the stored normalization parameters. To ensure physical consistency, any negative predicted inflows are replaced with linearly interpolated values to produce a continuous and realistic inflow series.

3.4 EVALUATION METRICS

To quantitatively assess the predictive performance of the DL-based inflow forecasting models, four commonly used statistical indicators are adopted: the Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). These metrics collectively evaluate the accuracy, bias, and overall agreement between simulated and observed inflows.

The NSE measures the predictive power of a model by comparing the residual variance to the variance of the observed inflows. An NSE value of 1 indicates perfect agreement between model predictions and observations.

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{obs}} - Q_{\mathrm{in},i}^{\mathrm{sim}})^2}{\sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{obs}} - \bar{Q}_{\mathrm{in},i}^{\mathrm{obs}})^2} \tag{9}$$

where $Q_{\mathrm{in},i}^{\mathrm{obs}}$, $Q_{\mathrm{in},i}^{\mathrm{sim}}$ = observed and simulated inflows at time step i, $\bar{Q}_{\mathrm{in}}^{\mathrm{obs}}$ = mean observed inflow, and N = number of samples.

The KGE metric integrates three key components: correlation, bias, and variability. It provides a balanced assessment of model performance by quantifying both accuracy and hydrological realism.

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{10}$$

where r = Pearson correlation coefficient between $Q_{\mathrm{in},i}^{\mathrm{obs}}$ and $Q_{\mathrm{in},i}^{\mathrm{sim}}$, β = bias ratio, and γ = variability ratio. They expressed by

$$r = \frac{\sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{obs}} - \bar{Q}_{\mathrm{in}}^{\mathrm{obs}})(Q_{\mathrm{in},i}^{\mathrm{sim}} - \bar{Q}_{\mathrm{in}}^{\mathrm{sim}})}{\sqrt{\sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{obs}} - \bar{Q}_{\mathrm{in}}^{\mathrm{obs}})^2 \sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{sim}} - \bar{Q}_{\mathrm{in}}^{\mathrm{sim}})^2}}$$
$$\beta = \frac{\mu_{\mathrm{sim}}}{\mu_{\mathrm{obs}}}, \quad \gamma = \frac{C_v^{\mathrm{sim}}}{C_v^{\mathrm{obs}}}, \quad C_v = \frac{\sigma}{\mu} \tag{11}$$

where μ_{obs}, μ_{sim} = mean of observed and simulated inflows, σ_{obs}, σ_{sim} = standard deviation of observed and simulated inflows, μ, σ = mean and standard deviation of inflow (either observed or simulated), and C_v = coefficient of variation.

The RMSE measures the overall magnitude of the model's prediction errors, giving greater weight to large deviations. Smaller RMSE values indicate higher predictive accuracy.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Q_{\mathrm{in},i}^{\mathrm{sim}} - Q_{\mathrm{in},i}^{\mathrm{obs}})^2} \tag{12}$$

The MAE quantifies the average magnitude of absolute prediction errors and is less sensitive to outliers compared with RMSE. It provides an intuitive measure of mean deviation between observed and simulated inflows.

$$\mathrm{MAE} = \frac{1}{N} \sum_{i=1}^{N} |Q_{\mathrm{in},i}^{\mathrm{sim}} - Q_{\mathrm{in},i}^{\mathrm{obs}}| \tag{13}$$

4 Results and analysis

4.1 PERFORMANCE EVALUATION

The predictive performance of the developed DL models is evaluated using the statistical metrics described in Section 3.4. Table 2 summarizes the performance for inflow forecasting at the Överuman and Sädvajaure Reservoirs. Overall, both models demonstrate high predictive skill, with NSE and KGE values exceeding 0.8 in all cases. For Överuman, the LSTM model achieves slightly higher efficiency (NSE = 0.94, KGE = 0.94) and lower errors (RMSE = 8.97 m^3/s, MAE = 5.96 m^3/s) than the CNN-LSTM, suggesting that the added convolutional component provides no performance enhancement for this reservoir. At Sädvajaure, both models perform exceptionally well, with the LSTM model yielding the highest accuracy (NSE = 0.99, KGE = 0.98) and the lowest RMSE and MAE values (6.34 m^3/s and 3.66 m^3/s, respectively). The results indicate that the proposed DL frameworks can effectively capture the nonlinear hydrological processes governing inflow dynamics, although model performance varies slightly with site characteristics and data complexity. The CNN-LSTM model integrates convolutional layers for feature extraction; however, its performance is somewhat lower than that of the LSTM model, particularly for the Överuman dataset. This may be attributed to the limited spatial variability in the input variables, which reduces the benefit of convolutional feature learning.

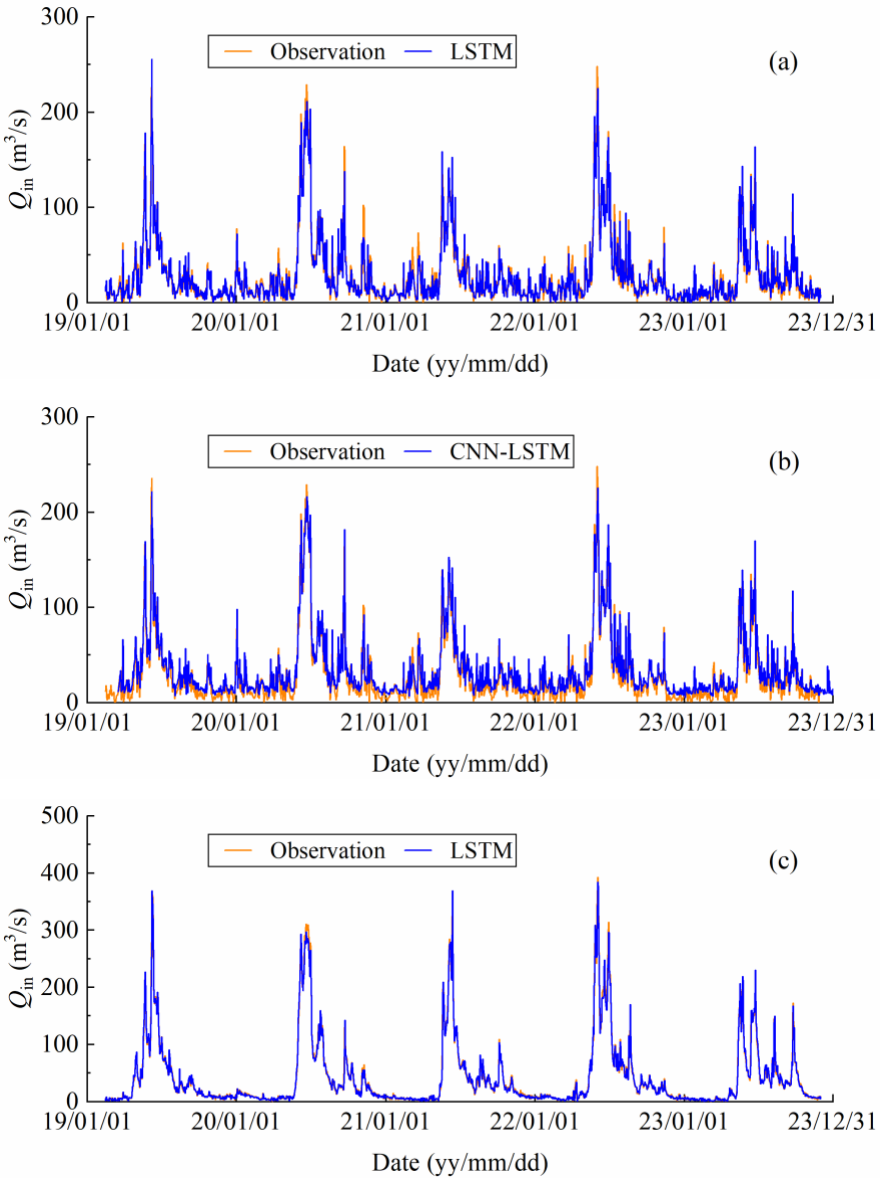
Table 2. Performance metrics of LSTM and CNN-LSTM models

Reservoir	Model	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
Överuman	LSTM	0.94	0.94	8.97	5.96
	CNN-LSTM	0.89	0.81	10.75	8.16
Sädvajaure	LSTM	0.99	0.98	6.34	3.66
	CNN-LSTM	0.98	0.90	8.73	5.51

The time-series comparisons between observed and simulated inflows are presented in Figure 5. At Överuman, both models accurately reproduce the temporal dynamics of the inflow series, capturing the timing and magnitude of major inflow peaks. The LSTM model (Figure 5a) demonstrates slightly higher accuracy, particularly during low-flow periods, reflecting its strong ability to learn sequential dependencies in hydrological time series. In contrast, the CNN-LSTM model (Figure 5b) tends to overestimate low-flow conditions and produces slightly larger deviations during peak events. This overestimation may result from the convolutional layers, which smooth short-term variability and amplify minor fluctuations. Despite these discrepancies, both models adequately capture the seasonal flow pattern and overall interannual variability, confirming their reliability for data-driven inflow forecasting under regulated reservoir conditions.

At Sädvajaure, both models achieve exceptionally high accuracy, closely reproducing the magnitude and timing of inflow peaks. The LSTM model (Figure 5c) provides nearly perfect alignment with the observed hydrograph, capturing both high- and low-flow periods with minimal error. The CNN-LSTM model (Figure 5d) also produces excellent agreement. However, it exhibits a slight

tendency to smooth the sharp rising limbs of flood peaks, likely due to convolutional filtering. The strong performance at Sädvajaure can be attributed to the clearer hydrological signal (weaker fluctuations), which allows both models to learn seasonal and interannual patterns effectively. These results demonstrate that DL-based frameworks can accurately model inflow processes, even in highly regulated systems.



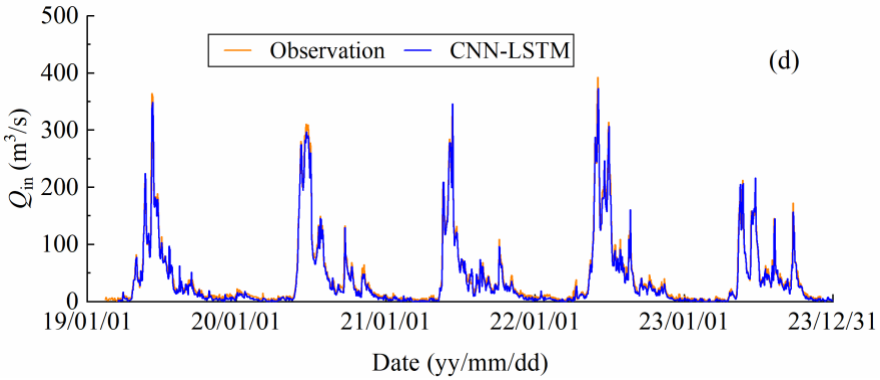


Figure 5. Comparison of observed and predicted inflow at the (a, b) Överuman and (c, d) Sädvajaure Reservoirs.

The scatter plots in Figure 6 provide a complementary assessment of model performance. Most data points are closely aligned along the 1:1 reference line, indicating strong predictive capability of both DL models. At Överuman (Figure 6a), the LSTM predictions show tighter clustering around the 1:1 line, particularly in the mid- to high-flow range, confirming its slightly better accuracy, as reflected in the statistical metrics (Table 2). The CNN-LSTM model, however, exhibits a mild positive bias for low-flow conditions, consistent with its tendency to overestimate, as observed in the time-series plots (Figure 5b).

For Sädvajaure (Figure 6b), both models demonstrate excellent linearity and minimal scatter across the entire data range, reflecting their robustness in capturing the more stable hydrological dynamics of this reservoir. The near-perfect alignment with the 1:1 line underscores the models' ability to accurately reproduce both peak and base flows, with the LSTM exhibiting marginally higher precision. Collectively, these results confirm that both models perform reliably across different hydrological regimes, with the LSTM model exhibiting slightly superior generalization performance.

Given its higher accuracy and lower structural complexity, the LSTM model is, therefore, selected for subsequent analysis and interpretation in the following sections.

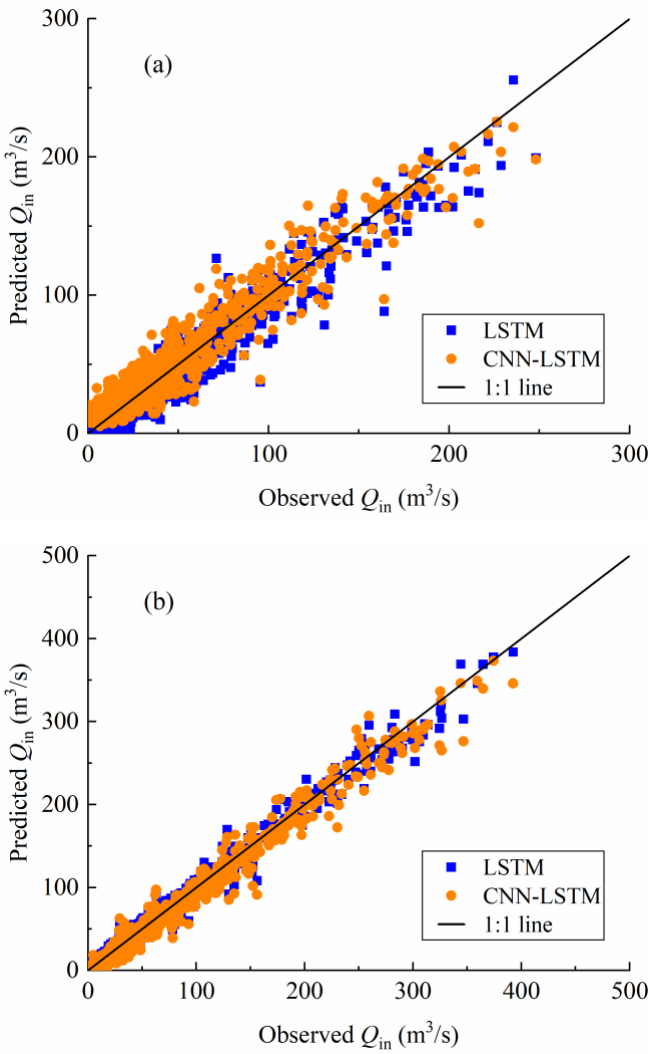


Figure 6. Scatter plots for the (a) Överuman and (b) Sädvajaure Reservoirs: observation vs. prediction.

4.2 UNCERTAINTY QUANTIFICATION

Although the model demonstrates high predictive accuracy, it is essential to evaluate the uncertainty associated with its forecasts under varying conditions. Since direct measurement-error information is unavailable, uncertainty is artificially introduced by perturbing the model inputs with controlled levels of Gaussian noise. Specifically, Monte Carlo (MC) simulations are conducted with noise amplitudes ranging from 1% to 10% of each variable's mean value, representing plausible levels of observational and modeling uncertainty. For each noise level, the LSTM model is repeatedly executed 1000 times to generate an ensemble of inflow predictions, from which the mean forecast and 95% confidence interval (CI) are computed. The uncertainty is quantified as the width of the CI bounds around the mean prediction.

The model uncertainty of the inflow forecasts under different noise scenarios is illustrated in Figure 7. In both reservoirs, the model uncertainty increases systematically with higher noise levels, reflecting the propagation of input variability through the model. At low noise levels (1–3%), the uncertainty bounds remain narrow (below 1 m^3/s for most cases); the model predictions are stable and only weakly sensitive to small perturbations in the input variables. As the noise amplitude increases beyond 5%, both the interquartile range (IQR) and whisker length expand considerably, suggesting a growing divergence among ensemble predictions. For the Överuman Reservoir, the median uncertainties corresponding to 1%, 3%, 5%, 7%, and 10% noise levels are 0.09 m^3/s, 0.26 m^3/s, 0.44 m^3/s, 0.60 m^3/s, and 0.84 m^3/s, respectively. For the Sädvajaure Reservoir, the respective values are 0.13 m^3/s, 0.39 m^3/s, 0.62 m^3/s, 0.83 m^3/s, and 1.15 m^3/s. These results reveal that the LSTM model is highly robust to input uncertainty, with only moderate increases in predictive variability even when input noise reaches 10%.

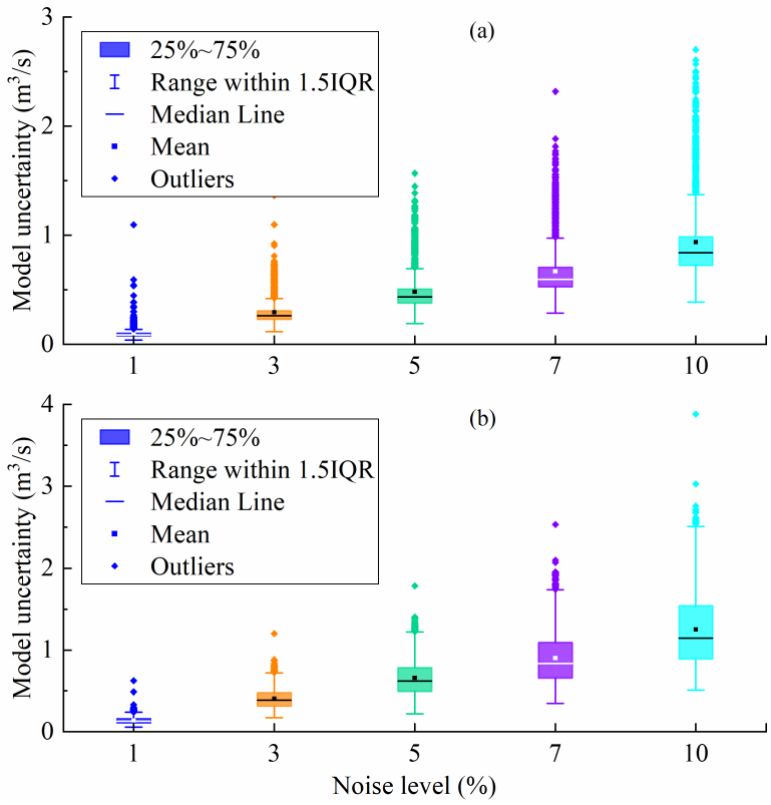


Figure 7. Model uncertainty at the 95% confidence level for the (a) Överuman and (b) Sädvajaure Reservoirs. IQR = interquartile range.

4.3 MODEL TRANSFERABILITY

To assess cross-reservoir generalization, the model trained on one reservoir is applied to another reservoir. Table 3 and Table 4 summarize the model transferability between the Överuman and Sädvajaure Reservoirs. The reference model in each case refers to the locally trained LSTM using long-term observational data (Table 2). In the direct-transfer case, the pretrained model

performs poorly when applied to a new reservoir, with NSE values dropping below zero. This indicates that inflow-climate relationships learned from the source reservoir cannot be directly generalized to another site without adaptation, likely due to differences in catchment characteristics, hydrological responses, and data distributions. This limitation motivates the use of transfer learning, which enables the pre-trained model to be fine-tuned with a small portion of local data, allowing it to adapt to site-specific hydrological and climatic characteristics while reducing the need for extensive local training data.

Transfer learning is an ML technique in which knowledge from one task or dataset (the source domain) is reused to improve performance on a related but different task or dataset (the target domain) (Hosna et al., 2022; Yao et al., 2023). Instead of training a new model from scratch, the pretrained model's parameters, particularly those capturing temporal and nonlinear relationships, are used as an initial foundation and then fine-tuned using a smaller amount of data from the target site. In this study, the pre-trained model developed for the Överuman (or Sädvajaure) Reservoir is fine-tuned (LSTM weights and biases) with limited Sädvajaure (or Överuman) data, allowing it to retain learned temporal patterns while adapting to the distinct hydrological and climatic conditions of the target reservoir.

As shown in Table 3 and Table 4, once transfer learning is applied, the predictive performance improves rapidly by fine-tuning the pre-trained model with a small fraction of local data. Even with as little as 2–4% of local data, NSE rises from negative to moderate values (>0.6), indicating successful correction of the systematic biases learned from the source domain. As the proportion of local data increases to 10–20%, the fine-tuned models nearly match, and sometimes surpass, the performance of the locally trained reference models. The improvement is particularly pronounced in the Överuman-to-Sädvajaure direction, where transfer learning achieves near-perfect recovery with limited fine-tuning. This suggests a more substantial hydrological similarity or more transferable inflow-forcing relationships between these two sites. Overall, the results show that transfer learning can significantly reduce data and computational requirements for new reservoirs while maintaining high prediction accuracy.

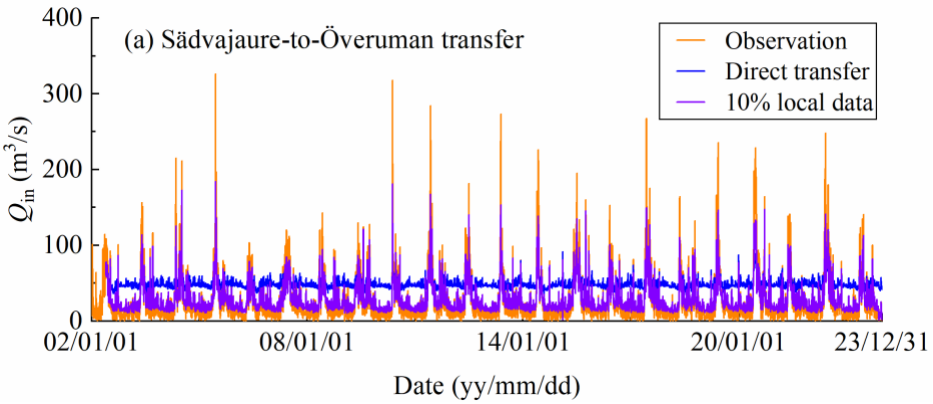
Figure 8 compares the inflow forecasts obtained by direct transfer and transfer learning with 10% local data. The direct transfer models exhibit poor predictive performance when applied across reservoirs, failing to reproduce both the magnitude and timing of inflow variations. This deficiency suggests that reservoir-specific hydrological and climatic conditions significantly impact model behavior, thereby limiting the generalizability of features learned solely from the source domain. The fine-tuned models effectively recover seasonal flow dynamics and capture the general magnitude of inflows. Nonetheless, high (underestimated) and low flows (overestimated) remain challenging to reproduce precisely, particularly at Överuman. This suggests that local adaptation cannot fully compensate for fundamental differences in catchment and regulatory dynamics. However, transfer learning remains useful for data-scarce reservoirs, requiring minimal retraining effort and offering a practical balance between model generalization and site-specific calibration.

Table 3. Model transferability using transfer learning: From Sädvajaure to Överuman

Local data	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
– (Reference model)	0.94	0.94	8.97	5.96
0% (Direct transfer)	−0.08	−0.10	33.25	29.92
2%	0.54	0.31	21.60	17.65
4%	0.67	0.42	18.32	14.06
6%	0.75	0.51	16.03	11.04
8%	0.80	0.56	14.34	9.85
10%	0.85	0.62	12.63	8.13
20%	0.95	0.81	6.88	3.77

Table 4. Model transferability using transfer learning: From Överuman to Sädvajaure

Local data	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
– (Reference model)	0.99	0.98	6.34	3.66
0% (Direct transfer)	−0.69	−0.76	74.06	71.71
2%	0.64	0.28	34.31	30.96
4%	0.92	0.83	16.26	10.41
6%	0.94	0.90	13.91	8.06
8%	0.96	0.89	11.60	5.90
10%	0.97	0.91	10.56	5.12
20%	0.99	0.95	6.65	2.73



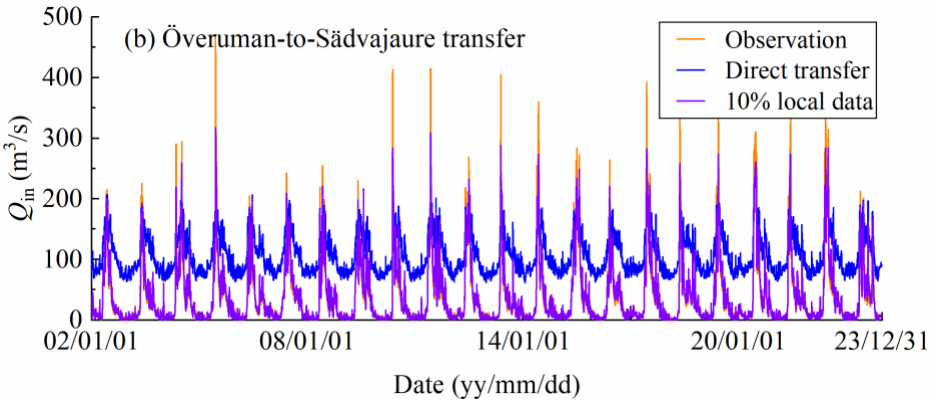


Figure 8. Cross-reservoir generalization using transfer learning: (a) Sädvajaure-to-Överuman and (b) Överuman-to-Sädvajaure transfer.

4.4 EFFECT OF LEAD TIMES

The predictive performance of the LSTM model at different lead times (LTs) is summarized in Table 5. As expected, model accuracy gradually decreases with increasing lead time for both reservoirs, reflecting the cumulative uncertainty inherent in longer-term inflow forecasting. For Överuman, the model maintains high skill for short-term prediction, but its performance declines notably beyond 3 days, with the NSE dropping below 0.7 and the RMSE rising above 20 m³/s. Beyond approximately 10 days, the forecasts lose substantial accuracy (NSE < 0.5), indicating the limit of reliable lead-time prediction for this reservoir.

For Sädvajaure, the model sustains strong predictive capability for a longer horizon, with NSE > 0.8 and KGE > 0.9 up to 4 days ahead. Even at a 7-day lead time, NSE remains above 0.75, highlighting the model's superior temporal memory in this reservoir. The more stable hydrological regime and larger inflow magnitude at Sädvajaure likely enhance the persistence of temporal patterns. This allows the model to capture delayed system responses more effectively.

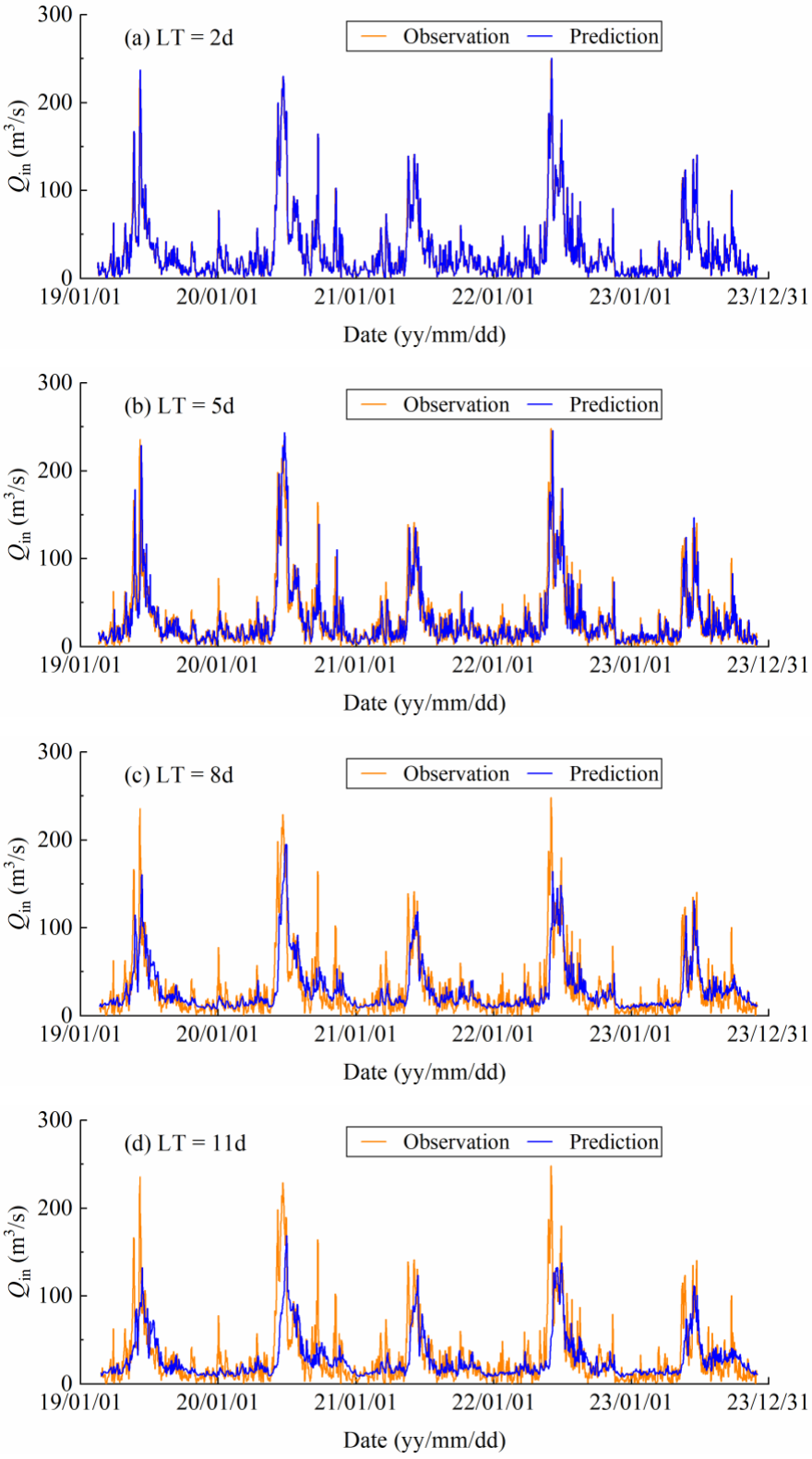
The temporal evolution of model performance under different lead times is illustrated in Figure 9 and Figure 10. The model performs well at short lead times for both reservoirs, accurately capturing the magnitude and temporal variability of inflows. For Överuman, the predicted inflows closely follow the observed hydrograph during shorter horizons. For Sädvajaure, the model also reproduces peak timing and amplitude with high fidelity. However, as the forecast horizon extends, the model's ability to represent detailed fluctuations diminishes, and the predictions increasingly reflect only the overall trend of inflow variation. This behavior stems from the progressive loss of temporal dependency between current inputs and future inflows, as well as the accumulation of uncertainty from unmodeled hydro-meteorological processes.

From an operational perspective, the results indicate that the model provides highly reliable short-term forecasts and reasonably accurate medium-range predictions, with forecast quality progressively decreasing as the lead time increases. Incorporating the model within a rolling-update framework, in which

predictions are continually refined using new observations, would further enhance its reliability for real-time operational decision-making.

Table 5. Predictive performance at different lead times

Reservoir	LT (day)	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
Överuman	1	0.94	0.94	8.97	5.96
	2	0.88	0.94	12.67	8.15
	3	0.69	0.84	20.17	13.02
	4	0.55	0.78	24.28	15.43
	5	0.57	0.78	23.92	14.61
	6	0.61	0.75	22.65	13.75
	7	0.60	0.70	23.15	14.08
	8	0.56	0.67	24.07	14.41
	9	0.50	0.60	25.86	15.30
	10	0.47	0.58	26.48	15.59
	11	0.45	0.54	26.97	15.50
	12	0.39	0.49	28.38	16.28
	13	0.37	0.49	28.87	16.87
	14	0.32	0.44	30.10	17.53
Sädvajaure	1	0.99	0.98	6.34	3.66
	2	0.97	0.98	11.64	5.83
	3	0.90	0.95	20.09	9.61
	4	0.81	0.91	27.02	12.61
	5	0.76	0.88	30.58	14.02
	6	0.75	0.85	31.52	14.91
	7	0.75	0.83	32.03	15.01
	8	0.71	0.77	33.99	16.31
	9	0.66	0.73	36.76	17.56
	10	0.64	0.72	37.75	17.82
	11	0.59	0.69	40.23	19.54
	12	0.56	0.65	41.79	19.14
	13	0.50	0.62	44.58	21.17
	14	0.53	0.62	43.07	20.34


(a) LT = 2d
(b) LT = 5d
(c) LT = 8d
(d) LT = 11d
Observation
Prediction
Q_{in} (m^3/s)
0
100
200
300
19/01/01
20/01/01
21/01/01
22/01/01
23/01/01
23/12/31
Date (yy/mm/dd)

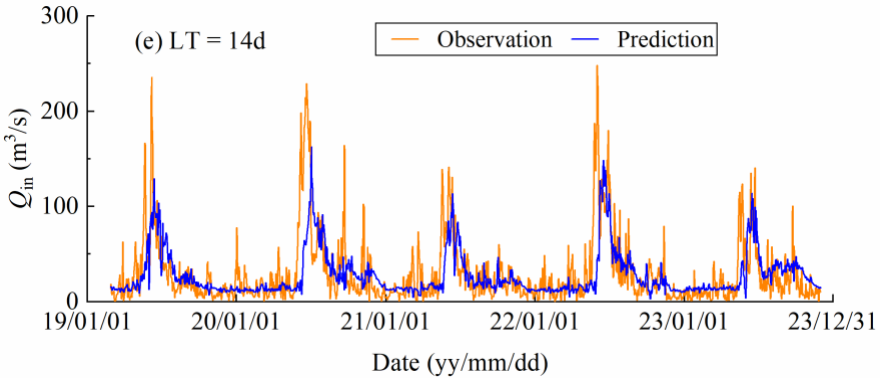
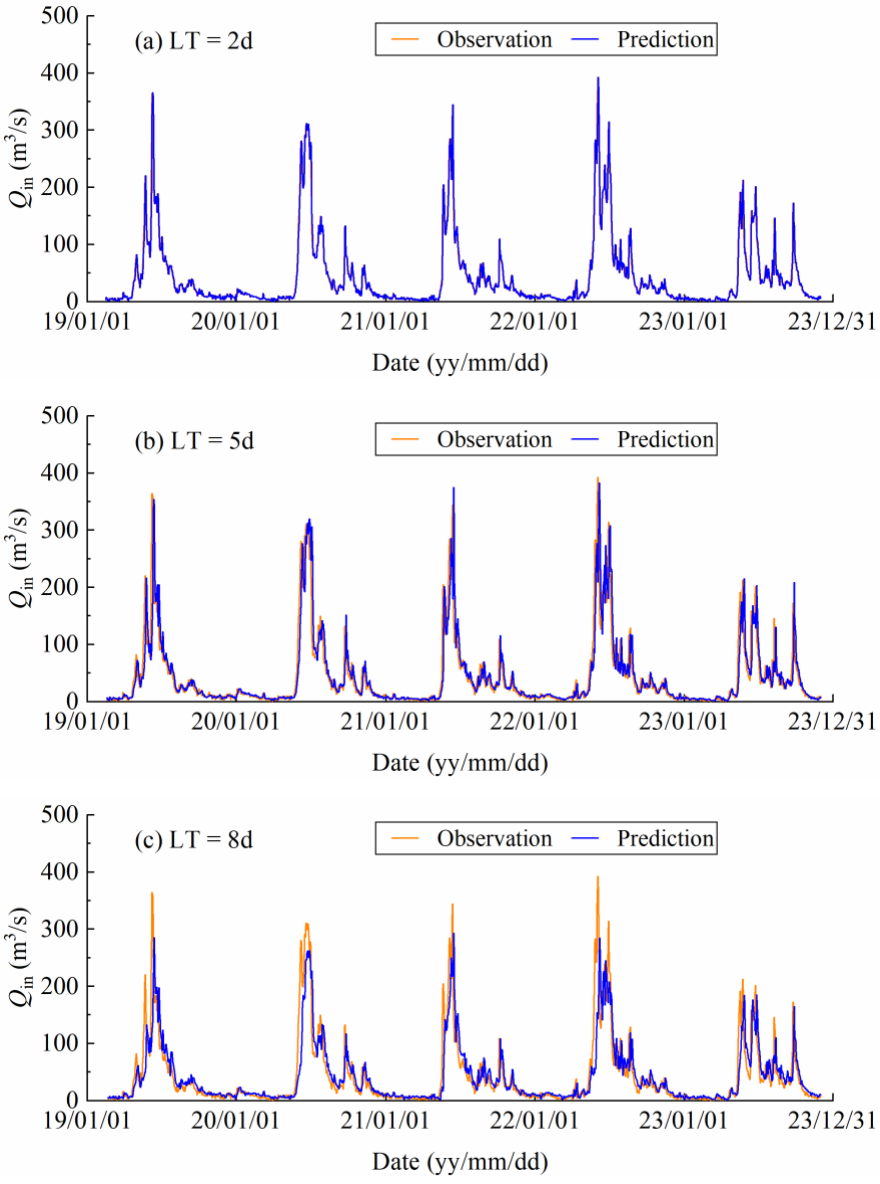


Figure 9. Comparison of observed and predicted inflow at the Överuman Reservoir under different lead times.



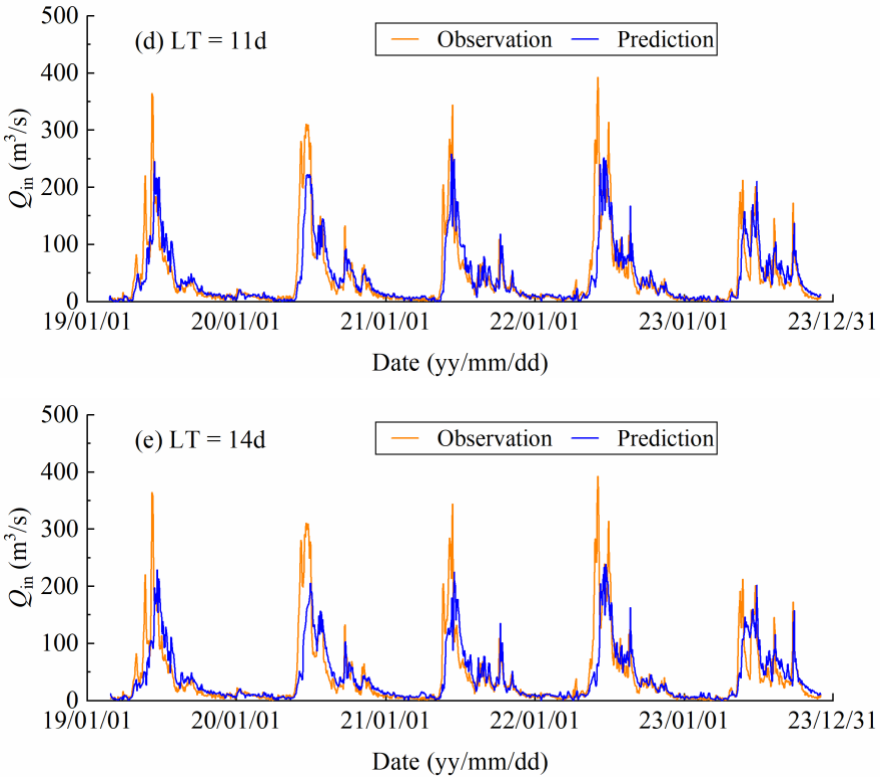


Figure 10. Comparison of observed and predicted inflow at the Sädvajaure Reservoir under different lead times.

4.5 COMPARISON WITH HBV

The HBV model serves as the operational hydrological model in Sweden and is widely used by hydropower companies. To contextualize the DL model's performance, its predictive skill is compared with that of the HBV model. The HBV model is calibrated by the hydropower companies using a longer observational record than the DL model. Table 6 summarizes the performance metrics for both LSTM and HBV at the two reservoirs. At Överuman, the LSTM achieves an NSE of 0.94 and an RMSE of 8.97 m³/s, while HBV attains an NSE of 0.80 with an RMSE of 16.71 m³/s. The DL model yields lower prediction errors and greater overall agreement with observations than the HBV model. At Sädvajaure, the LSTM again shows higher predictive accuracy, with an NSE of 0.99 and an RMSE of 6.34 m³/s. The HBV model reaches an NSE of 0.93 and an RMSE of 16.43 m³/s. Across both reservoirs, the results demonstrate that the LSTM consistently outperforms the HBV model. This highlights the strong predictive capability of data-driven approaches for short-term inflow forecasting.

Figure 11 presents a direct comparison between the LSTM and HBV models for both reservoirs. At Överuman, both models reproduce the overall temporal pattern of the observed inflow, including the seasonal peaks and low-flow periods. The LSTM simulation (Figure 11a) closely follows the observed curve throughout the hydrograph, with smaller deviations during both the rising and falling limbs. The HBV simulation (Figure 11b) accurately captures the timing of the peaks but

exhibits larger discrepancies from the observations, particularly in the magnitude of several high-flow events.

At Sädvajaure, both models align well with the observations and capture the recurrent seasonal inflow peaks. The LSTM results (Figure 11c) closely track the observed inflow across the full range of conditions, including peak flows. The HBV results (Figure 11d) reproduce the general pattern but show larger differences in peak magnitude and exhibit a wider spread around the observed curve.

Figure 12 provides a point-by-point comparison of observed inflow with predictions from both the LSTM and HBV models for the two reservoirs. For Överuman (Figure 12a), the LSTM predictions are more tightly clustered around the 1:1 line, showing smaller dispersion across the full inflow range. The HBV predictions also follow the general trend but display a wider spread, particularly at mid-range flows between approximately 50 and 150 m^3/s. Deviations from the 1:1 line are visibly larger for HBV, indicating less consistent matching of observed magnitudes. For Sädvajaure (Figure 12b), both models align reasonably well with the observations. The LSTM again exhibits a narrower distribution around the 1:1 line. The HBV results show greater scatter, with more points falling below or above the reference line. The difference between the two models is similar to that observed at Överuman: both follow the linear trend, but the LSTM maintains closer agreement with the observations across the entire inflow range.

In summary, these comparisons show that while HBV reproduces the overall inflow patterns at both reservoirs, the LSTM model provides consistently closer agreement with observations. The LSTM delivers higher predictive fidelity than the HBV model for the studied datasets.

Although both the HBV and LSTM models are useful, they differ in the way inflow forecasts are generated in the configurations evaluated in this study. HBV is a conceptual, process-based hydrological model that uses externally supplied meteorological forecasts, typically precipitation and temperature, to drive its internal routines and produce inflow estimates over the forecast horizon. In contrast, the LSTM model implemented here is a data-driven approach trained to predict future inflow from historical observations of hydrological and meteorological variables. Using a sliding window of past conditions, the model learns temporal patterns and produces forecasts for one or more days ahead.

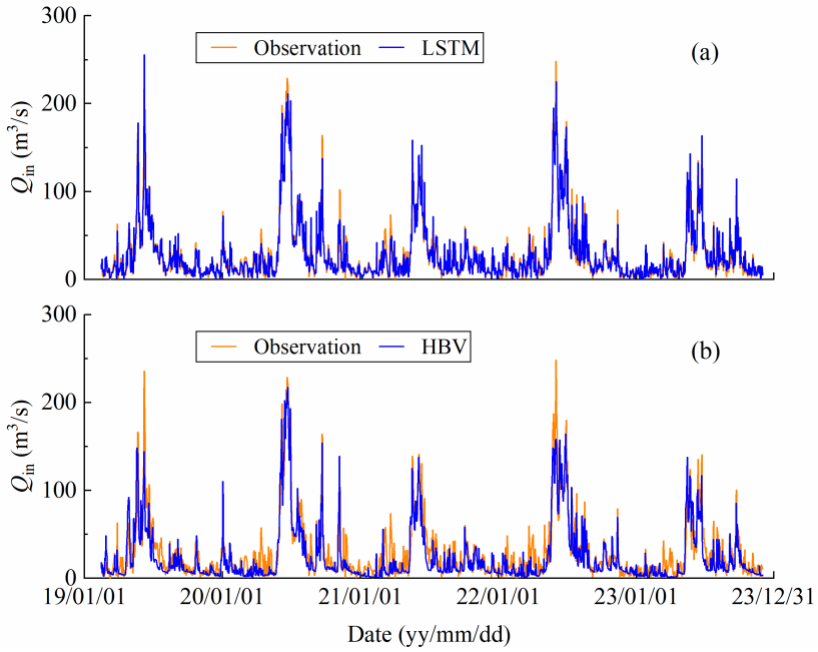
It should be emphasized that the present LSTM setup does not rely on future meteorological forecasts, but this is a modeling choice rather than a fundamental limitation of the method. In an operational setting, the LSTM could be implemented in a rolling daily forecasting framework in which the most recent observations are ingested as they become available, and updated inflow forecasts are generated for the upcoming days. Future developments could also extend this framework by incorporating meteorological forecasts, such as precipitation and temperature forecasts, as additional input variables. In such a hybrid data-driven forecasting setup, predicted inflows may be recursively used as part of the input sequence while forecasted meteorological variables provide external information about future forcing conditions. This could help extend the practical forecast

horizon, although additional validation would be required to assess uncertainty propagation and operational robustness.

These differences reflect two complementary forecasting philosophies: HBV explicitly represents hydrological processes driven by meteorological forcing, whereas LSTM-based approaches infer inflow dynamics from observed data and can, if designed accordingly, also be coupled with meteorological forecasts in hybrid operational frameworks.

Table 6. Performance metrics of LSTM and HBV models

Reservoir	Model	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
Överuman	LSTM	0.94	0.94	8.97	5.96
	HBV	0.80	0.77	16.71	11.25
Sädvajaure	LSTM	0.99	0.98	6.34	3.66
	HBV	0.93	0.93	16.43	9.50



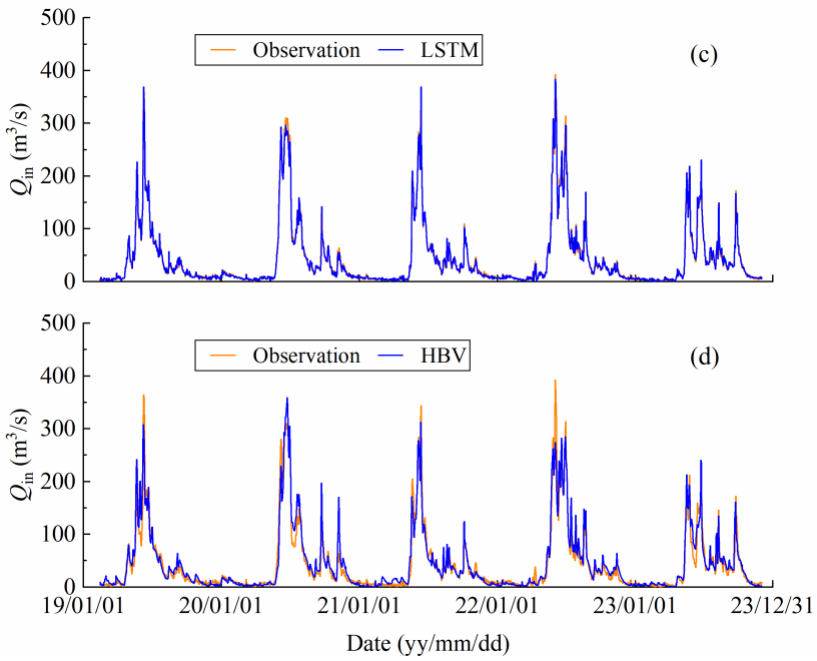
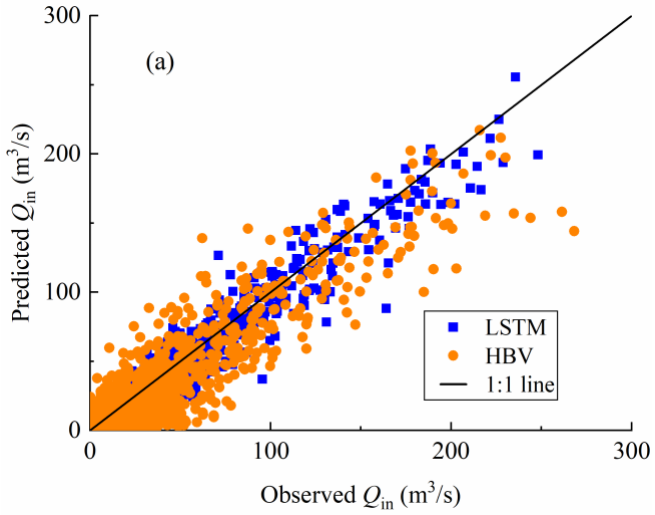


Figure 11. Time series comparison of LSTM and HBV results for (a, b) the Överuman and (c, d) Sädvajaure Reservoirs.



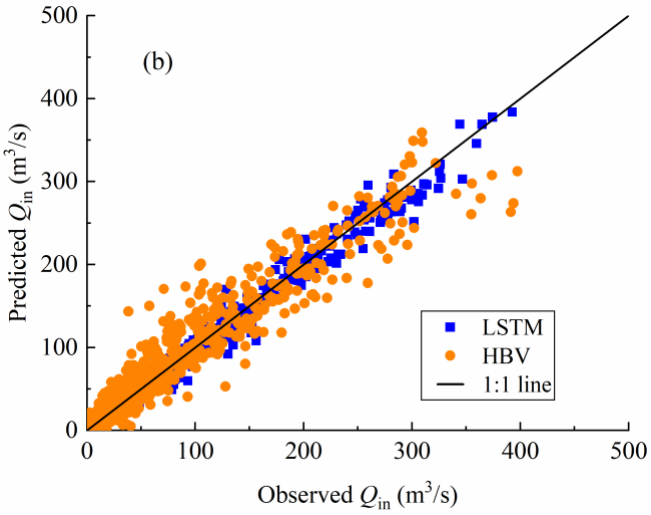


Figure 12. Scatter plot comparison of LSTM and HBV results for (a) the Överuman and (b) Sädvajaure Reservoirs: observation vs. prediction.

In operational hydropower forecasting, long-term inflows are often produced using the Ensemble Streamflow Prediction (ESP) method within the HBV model. ESP uses the current hydrological state of the catchment to simulate future inflows by forcing the model with historical meteorological sequences, thereby generating an ensemble of possible future outcomes. This approach is beneficial for seasonal planning, such as assessing the potential for spring flooding. To enable comparable long-term scenario generation with a data-driven approach, we extend the LSTM model to an ESP-style forecasting framework. In this setup, the LSTM is initialized using the most recent observed conditions and is then driven forward using historical meteorological sequences from past years. At the same time, the inflow component is updated daily with observed values to maintain consistency with the current hydrological state. By iterating this process for each historical year, the LSTM produces an ensemble of long-term inflow trajectories that are analogous to those of HBV-ESP. In the LSTM model, the ensemble consists of meteorological forcing sequences from 2000 to 2022, with each year forming a separate ensemble member. The January–July (210 days) period is selected to capture the spring flood season.

Figure 13 presents the ensemble forecasts of the spring flood for the Överuman and Sädvajaure Reservoirs. The ensemble mean captures the seasonal rise and recession of inflow at both reservoirs, while the 10–90% range reflects the spread induced by interannual meteorological variability during the snowmelt period. For Överuman, the observed 2023 inflow generally follows the ensemble mean but exhibits sharper short-term fluctuations that are partially smoothed in the ensemble prediction. In contrast, Sädvajaure shows a more pronounced and concentrated spring flood, with a wider ensemble spread during peak flow, indicating stronger sensitivity to climatic forcing.

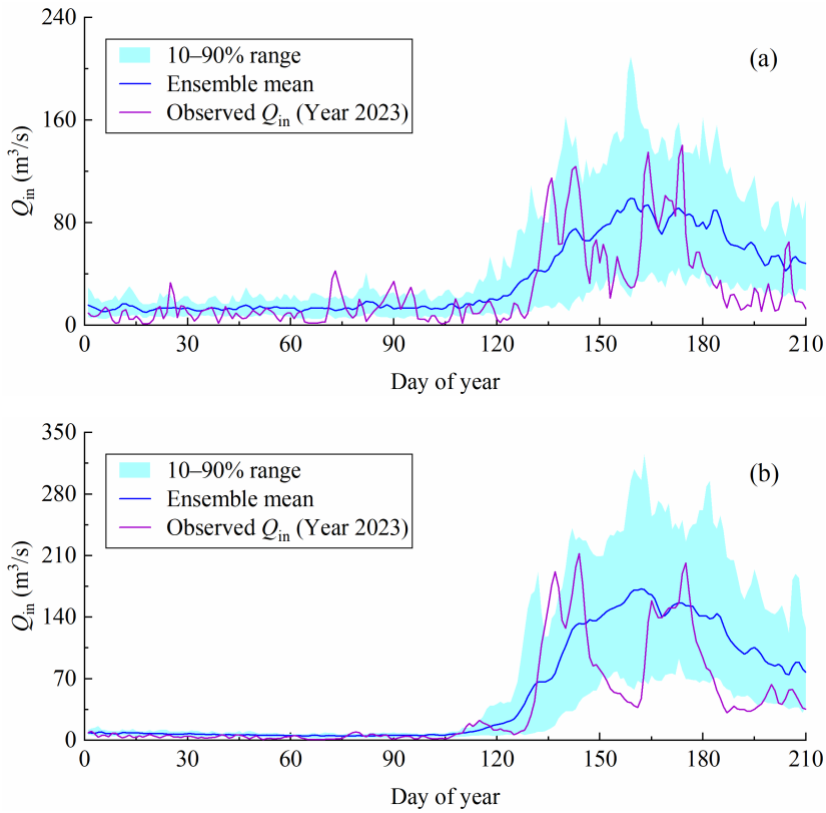


Figure 13. Ensemble forecasting of spring flood for (a) the Överuman and (b) Sädvajaure Reservoirs.

4.6 EFFECT OF INPUT FEATURES

To examine the influence of input features on model performance, a series of LSTM models are trained and tested. The comparative analysis of input features (Table 7 and Table 8) reveals that model performance is strongly dependent on the selected predictors for both reservoirs. For Överuman, the historical inflow is the most influential input, and its combination with meteorological variables (*P* and *T*) further enhances model performance, achieving NSE ≈ 0.94 and RMSE < 10 m³ /s. The addition of water level provides limited benefit once historical inflow is included, reflecting the high correlation between these two hydraulic variables. For Sädvajaure, a similar but slightly stronger dependence on meteorological predictors is observed. Models incorporating historical inflow with precipitation and/or temperature (e.g., M9–M15) yield excellent accuracy (NSE ≈ 0.99, RMSE ≈ 6–8 m³ /s), with the best performance obtained when all four variables (WL, P, T, Q) are combined. This suggests that inflow memory alone is insufficient to fully describe the system, and that precipitation and temperature signals are crucial for capturing the reservoir's climatic and seasonal variability.

The dominance of historical inflow arises from its ability to encode the catchment's hydrological memory. Physically, inflow reflects the integrated response of precipitation, snowmelt, soil moisture, and groundwater storage, making it a direct indicator of the system's current and antecedent hydrological state. Consequently,

past inflow inherently carries information about basin wetness, runoff potential, and delayed subsurface contributions that persist over time. From a data-driven perspective, the LSTM network is specifically designed to learn temporal dependencies within sequential data. Past inflow provides a strong, temporally correlated signal that the model can exploit to predict future inflows. The water level is also hydraulically linked to inflow through reservoir storage and discharge relationships. However, once inflow is included, water level contributes limited additional predictive power, but the former responds more directly to variations in inflow. Precipitation and temperature serve as external driving forces, providing crucial information on the generation of surface runoff and snowmelt. Although each parameter individually exhibits high short-term variability, their inclusion alongside historical inflow improves the model's ability to anticipate changes caused by climatic and seasonal dynamics.

The inflow distributions corresponding to different input-feature combinations are shown in Figure 14. The results are consistent with the quantitative measures presented in Table 7 and Table 8. When past inflow is incorporated, either alone or in combination with other variables, the interquartile range narrows markedly, and the median aligns closely with the observed inflow distribution.

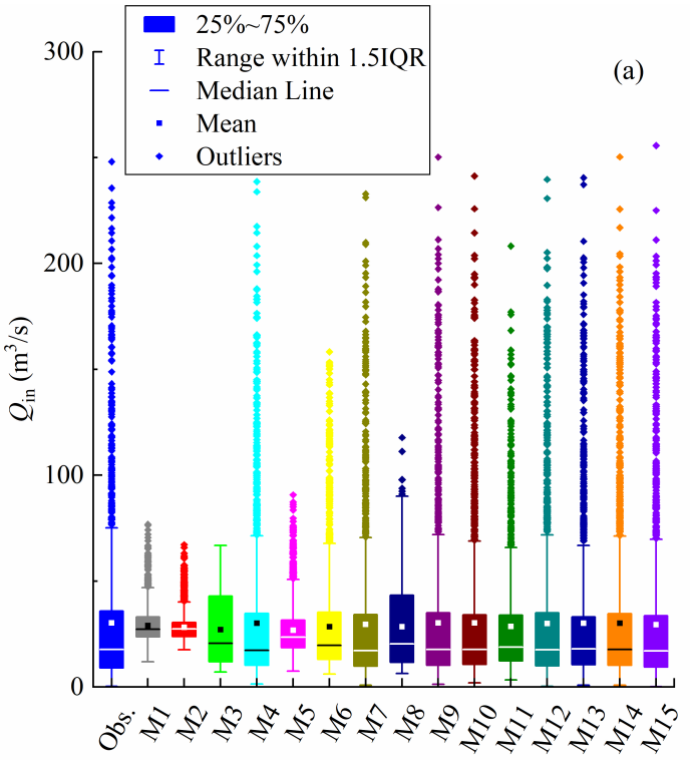
Table 7. Model performance with different inputs for the Överuman Reservoir

Model no.	Input	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
M1	*WL*	0.24	0.15	31.60	20.44
M2	*P*	0.01	−0.20	36.24	22.65
M3	*T*	0.24	0.24	31.76	18.64
M4	*Q*	0.91	0.93	10.62	6.93
M5	*WL*, *P*	0.29	0.22	30.65	18.31
M6	*WL*, *T*	0.61	0.60	22.72	14.11
M7	*WL*, *Q*	0.92	0.91	10.36	6.80
M8	*P*, *T*	0.29	0.34	30.67	17.98
M9	*P*, *Q*	0.92	0.94	10.00	6.63
M10	*T*, *Q*	0.92	0.92	10.09	6.71
M11	*WL*, *P*, *T*	0.68	0.67	20.60	12.40
M12	*WL*, *P*, *Q*	0.92	0.92	9.95	6.50
M13	*WL*, *T*, *Q*	0.93	0.93	9.65	6.51
M14	*P*, *T*, *Q*	0.94	0.94	9.19	6.16
M15	*WL*, *P*, *T*, *Q*	0.94	0.94	8.97	5.96

Table 8. Model performance with different inputs for the Sädvajaure Reservoir

Model no.	Input	NSE	KGE	RMSE (m^3/s)	MAE (m^3/s)
M1	*WL*	0.44	0.33	47.04	31.02
M2	*P*	0.04	-0.01	61.23	38.34
M3	*T*	0.38	0.42	49.45	27.24
M4	*Q*	0.98	0.98	8.01	4.35
M5	*WL*, *P*	0.35	0.30	50.62	33.13
M6	*WL*, *T*	0.74	0.74	31.65	16.90

M7	*WL*, *Q*	0.98	0.97	7.74	4.26
M8	*P*, *T*	0.38	0.42	49.48	26.57
M9	*P*, *Q*	0.99	0.98	7.33	4.14
M10	*T*, *Q*	0.99	0.98	7.42	4.08
M11	*WL*, *P*, *T*	0.76	0.75	30.50	16.16
M12	*WL*, *P*, *Q*	0.99	0.97	7.22	4.06
M13	*WL*, *T*, *Q*	0.99	0.98	6.84	3.85
M14	*P*, *T*, *Q*	0.99	0.98	6.59	3.80
M15	*WL*, *P*, *T*, *Q*	0.99	0.98	6.34	3.66



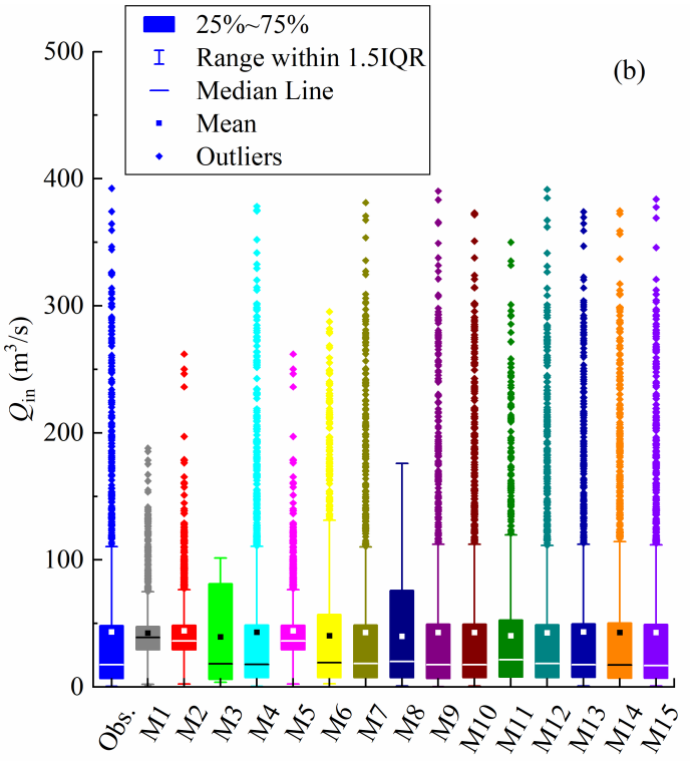


Figure 14. Inflow forecasting with different input features for the (a) Överuman and (b) Sädvajaure Reservoirs.

5 Concluding remarks

This study develops and evaluates DL-based models for reservoir inflow forecasting using long-term hydro-meteorological data from two cold-climate Swedish reservoirs, Överuman and Sädvajaure. Two DL architectures, LSTM and CNN-LSTM, are implemented to capture the nonlinear and temporal dependencies of inflow processes. The results demonstrate that DL-based models can achieve high predictive skill (LSTM: NSE = 0.94–0.99, RMSE = 6.34–8.97 m^3/s) and effectively reproduce both the magnitude and timing of inflow variations, outperforming conventional physics-based approaches.

Between the two models, the LSTM exhibits superior accuracy and stability, particularly for Överuman. The model also demonstrates high robustness under Monte Carlo-based uncertainty analysis, with narrow confidence intervals even at a 10% input perturbation (median uncertainty ≈ 1 m^3/s). This confirms the LSTM network's reliability for operational inflow prediction under noisy conditions.

The study also explores transfer learning, where a model trained on one reservoir is adapted to another using limited local data. Direct model transfer between Överuman and Sädvajaure yields poor performance (NSE < 0), indicating that hydrological and climatic differences between basins hinder the direct generalization of learned patterns. When the pre-trained LSTM is fine-tuned using a limited amount of local data, model accuracy improves markedly. For example, with only 2–4% of local data, NSE rises above 0.6, and with 10–20%, the performance matches or even surpasses that of locally trained models (NSE ≈ 0.95–0.99). However, the model still tends to underestimate peak flows and overestimate low flows, especially at Överuman. Therefore, although transfer learning provides a practical means to extend DL-based inflow forecasting to data-scarce reservoirs, the use of locally observed data, whenever available, is recommended to ensure site-specific reliability.

The analysis of forecast horizons reveals that model accuracy decreases gradually as the forecast horizon increases. Reliable short-term forecasts (1–3 days) and acceptable medium-term predictions (~7 days) are achieved. For longer forecast horizons, the model captures only the overall trend in inflow, and short-term peaks and variations become less distinct. Feature-sensitivity analysis confirms that historical inflow is the most informative predictor, and precipitation and temperature enhance the model's responsiveness to climatic and seasonal variability.

Overall, this work demonstrates that DL provides a robust and transferable framework for inflow forecasting in regulated reservoirs. The combination of data-driven modeling and transfer learning offers an efficient pathway for scaling inflow prediction across multiple hydropower systems with minimal local calibration. Future work should focus on integrating physical hydrological knowledge into DL architectures and on incorporating meteorological forecast products into hybrid ML-based forecasting frameworks for real-time applications in adaptive, climate-resilient hydropower management.

6 Acknowledgement

This project is carried out within the HUVA (Hydrologiskt utvecklingsarbete) research program and is financially supported by Energiforsk. The authors gratefully acknowledge this support, which made it possible to explore the potential of deep learning techniques for hydrological forecasting within the Swedish hydropower sector.

This project comprises a reference group, including (in alphabetical order) Beatriz Quesada Montano (Uniper), Björn Norell (Vattenregleringsföretagen), Imenne Ahlen (Fortum), Magnus Rylander (Holmen Energi), Mikael Sundby (Vattenfall), and Susanne Nyström (Vattenfall). At a later stage, Alexandra Hamm (Vattenfall) joined as a replacement. The authors express their sincere appreciation to all reference group members for their insightful feedback, constructive comments, and valuable discussions throughout this work.

Thanks are extended to Björn Norell (Vattenregleringsföretagen), Lars Pettersson (Skellefteälvens Vattenregleringsföretag), and Lisa Forsgren (Skellefteälvens Vattenregleringsföretag) for kindly sharing hydrological and meteorological datasets for the Överuman and Sädvajaure Reservoirs. The authors acknowledge Qianqian Li (SMHI) for insightful discussions on HBV modeling, and Kevin Xie (AFRY) for contributions to technical discussions and model development.

During the preparation of this report, the authors utilized ChatGPT to assist with language editing and proofreading. After using this tool, the authors reviewed and edited the content as needed, taking full responsibility for the final version.

7 References

Alhussein, M., Aurangzeb, K. & Haider, S.I., 2020. Hybrid CNN-LSTM model for short-term individual household load forecasting. IEEE Access, 8, 180544-180557.

Betts, M.G., Yang, Z., Hadley, A.S., Smith, A.C., Rousseau, J.S., Northrup, J.M., Nocera, J.J., Gorelick, N. & Gerber, B.D., 2022. Forest degradation drives widespread avian habitat and population declines. Nature Ecology & Evolution, 6(6), 709-719.

DiPietro, R. & Hager, G.D., 2020. Handbook of Medical Image Computing and Computer Assisted Intervention, pp. 503-519, Elsevier.

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z. & Azim, M.A., 2022. Transfer learning: a friendly introduction. Journal of Big Data, 9(1), 102.

Katipoğlu, O.M., 2023. Combining discrete wavelet decomposition with soft computing techniques to predict monthly evapotranspiration in semi-arid Hakkâri province, Türkiye. Environmental Science and Pollution Research, 30(15), 44043-44066.

LeCun, Y., Kavukcuoglu, K. & Farabet, C., 2010. Convolutional networks and applications in vision, pp. 253-256, IEEE.

Li, S. & Yang, J., 2022. Modelling of suspended sediment load by Bayesian optimized machine learning methods with seasonal adjustment. Engineering Applications of Computational Fluid Mechanics, 16(1), 1883-1901.

Li, S., Yang, J. & Teng, P., 2024. Data-driven prediction of cylinder-induced unsteady wake flow. Applied Ocean Research, 150, 104114.

Norrman, J., 1989. Power plant in the Skellefteälven. Silver from Nasafjäll: archaeology at Silbojokk, 13-18.

SMHI, 2026. Damm- och sjöregister. https://vattenwebb.smhi.se/svarwebb/.

Yao, Y., Zhao, Y., Li, X., Feng, D., Shen, C., Liu, C., Kuang, X. & Zheng, C., 2023. Can transfer learning improve hydrological predictions in the alpine regions? Journal of Hydrology, 625, 130038.

Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S. & Xu, Y., 2022. Forecasting monthly gas field production based on the CNN-LSTM model. Energy, 260, 124889.

DEEP LEARNING-BASED RESERVOIR INFLOW FORECASTING

Accurate inflow forecasting is essential for efficient and climate-resilient hydropower management. This study presents a deep learning-based framework for reservoir inflow prediction using long-term data from two Swedish hydropower reservoirs: Överuman and Sädvajaure.

The developed deep learning models achieve high predictive accuracy (NSE = 0.94~~0~~.99, RMSE = 6.34~~8~~.97 m3/s) and robust performance across different hydrological conditions. The results demonstrate reliable short- to medium-term forecasts that effectively capture both seasonal variability and extreme inflow events.

By applying transfer learning, models trained on one reservoir can be efficiently adapted to another using only limited local data. This enables scalable forecasting solutions for data-scarce sites, reducing calibration time and costs.

The findings highlight the potential of deep learning as a promising complement to traditional hydrological models such as HBV, supporting HUVA's goal of improving forecast precision and operational flexibility in the Swedish hydropower system.

A new step in energy research

The research company Energiforsk initiates, coordinates, and conducts energy research and analyses, as well as communicates knowledge in favor of a robust and sustainable energy system. We are a politically neutral limited company that reinvests our profit in more research. Our owners are industry organisations Swedenergy and the Swedish Gas Association, the Swedish TSO Svenska kraftnät, and the gas and energy company Nordion Energi.

Energiforsk