Identifying Hydrometeorological Factors Influencing Reservoir Releases Using Machine Learning Methods

Ming Fan
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak ridge, USA
fanm@ornl.gov

Lujun Zhang
The School of Civil Engineering and Environmental Science
The University of Oklahoma
Norman, USA
lujun.zhang@ou.edu

Siyan Liu
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak ridge, USA
lius1@ornl.gov

Tiantian Yang
The School of Civil Engineering and Environmental Science
The University of Oklahoma
Norman, USA
tiantian.yang@ou.edu

Dan Lu
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak ridge, USA
lud1@ornl.gov

***Abstract*—Simulation of reservoir releases plays a critical role in social-economic functioning and our nation's security. However, it is challenging to predict the reservoir release accurately because of many influential factors from natural environments and engineering controls such as the reservoir inflow and storage. Moreover, climate change and hydrological intensification causing the extreme precipitation and temperature make the accurate prediction of reservoir releases even more challenging. Machine learning (ML) methods have shown some successful applications in simulating reservoir releases. However, previous studies mainly used inflow and storage data as inputs and only considered their short-term influences (e.g, previous one or two days). In this work, we use long short-term memory (LSTM) networks for reservoir release prediction based on four input variables including inflow, storage, precipitation, and temperature and consider their long-term influences. We apply the LSTM model to 30 reservoirs in Upper Colorado River Basin, United States. We analyze the prediction performance using six statistical metrics. More importantly, we investigate the influence of the input hydrometeorological factors, as well as their temporal effects on reservoir release decisions. Results indicate that inflow and storage are the most influential factors but the inclusion of precipitation and temperature can further improve the prediction of release especially in low flows. Additionally, the inflow and storage have a relatively long-term effect on the release. These findings can help optimize the water resources management in the reservoirs.**

***Index Terms*—Machine Learning, Long Short-Term Memory Network, Hydrometeorological Factor, Temporal Importance, Reservoir Release**

I. INTRODUCTION

A reservoir is an open storage area, which is usually formed by constructing dams to collect and release water from natural or artificial lakes. Most reservoirs in United States (US) are multi-functional including hydropower generation, agricultural irrigation, industrial and municipal water supply, ecosystem productivity, drought and flood protection, and recreation [1], [2]. Thus, accurate simulation of reservoir releases plays an important role in economic development and environment protection, and strongly affects water resources management [3]–[5].

Many factors impact reservoir release decisions, including reservoir inflow and storage, reservoir functionalities, upstream operations, as well as meteorological and climate forcing [6], [7]. Some factors are static such as reservoir functionalities; some factors are more decision related such as upstream operations; and some others are more dynamic with time series observations such as inflow, storage and meteorological forcing. To simulate the influence of these factors on reservoir releases, reservoir operators use empirical models to estimate reservoir yields and quantify system behaviors based on pre-defined operating rules [1]. The release rules are usually formulated based on historical data or design scenarios, and are defined in a general way meeting many situations and operational goals. These rule-based simulation models are transparent and thus understandable to the operators. On the other hand, as these pre-defined rules are hard model

This manuscript has been authored by UT-Battelle LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

2375-9259/22/$31.00 ©2022 IEEE
DOI 10.1109/ICDMW58026.2022.00143

constraints, the models are not necessarily suitable to the current situation with changing inflow and storage, and may not cover various conditions from changing climates.

Recently, machine learning (ML) models have been increasingly being applied for the reservoir release simulation. The ML models learn storage-release relationship based on historical data and estimate the current release from the inputs of previous days and current day. Several ML models have been applied, including linear regression models (e.g. ridge and lasso regressions), support vector regression, k-nearest neighbors (KNN) regression, decision trees, random forest, and multi-layer perceptron (MLP) neural networks [1], [4]. These ML models directly simulate input-output relationship, but not necessarily consider temporal dependence if the inputs and outputs are time-series observations. To consider influence of lag observations of inputs on the current outputs, the lag observations must be presented as additional input features [8].

Long short-term memory (LSTM) networks (a variant of recurrent neural networks (RNNs)), which is specially designed for time-series prediction, have also been applied in the reservoir simulation. The LSTM model enables learning system patterns associated with the observed dynamical system behaviors from input and output sequences. LSTM takes the sequence data as the model input, unlike aforementioned ML models where lag observations must be presented as input features. Additionally, LSTM directly support multiple parallel input sequences for multivariate inputs, unlike aforementioned ML models where multivariate inputs are presented in a flat structure. These two advantages of LSTM enable it to effectively extract the input time-series information and efficiently learn the relationship between the multi-inputs and reservoir release. Furthermore, LSTM can learn long-term dependence in time-series observations, making it particularly suitable for daily reservoir release simulation where lag times between inputs (including inflow and precipitation) and release can be up to several days.

Several studies have demonstrated successful applications of LSTM in reservoir modeling. For example, Zhang et al. (2018) [9] used a range of ML methods to simulate reservoir operations using hourly, daily, and monthly observations, and showed the superior performance of LSTM in predicting the reservoir release at different time scales. Zhang et al. (2019) [10] tested the applicability of RNNs in optimizing reservoir operations for different flow regimes and demonstrated that RNN can help generate the operation plan and achieve the goal of flood control and power generation. Yang et al. (2019) [11] combined RNNs with a distributed hydrological model to develop a real-time operation system for water resources management. More recently, Yang et al. (2021) [1] applied the LSTM model to predict the controlled reservoir release from the historical data, learn the reservoir operation rules, and assist reservoir release decision making. In these previous studies, they mainly considered the reservoir inflow and storage as the inputs and used one- or two-time step lag observations of inputs to predict reservoir release on current time. Given the increase of extreme weather events in intensity and frequency, the precipitation and temperature play more and more important roles in affecting the reservoir release [12], [13]. And due to the complexity of surrounding watershed and reservoir operation rules, lag information of the hydrometeorological inputs can have a relatively long-term impact on the release prediction.

In this effort, we use LSTM networks for reservoir release prediction based on four input variables including inflow, storage, precipitation, and temperature and consider their long-term influences. We apply the LSTM model to 30 reservoirs in Upper Colorado River Basin, US. We analyze the prediction performance using six statistical metrics. Specifically, we investigate the following three questions: (1) accuracy of LSTM in simulating the human-controlled reservoir release based on the historical reservoir storage and inflow, and the meteorological forcing of temperature and precipitation; (2) effectiveness of incorporating the meteorological data in assisting the reservoir release decision-making; and (3) influence of the lag length of previous input observations on the improvement of reservoir release prediction accuracy.

The main contributions of this paper are:

- We apply LSTM networks for predicting reservoir releases based on multiple hydrometeorological inputs and by considering their long-term influence on the release prediction.
- We evaluate the prediction performance using multiple statistics and identify the influential hydrometeorological factors and their temporal influence on the release prediction.
- We perform the analysis on 30 diverse reservoirs and discuss the impact of reservoir features on the analysis.

II. METHODOLOGY

A. Study area and data

The Upper Colorado River Basin is comprised of four states, including Colorado, New Mexico, Utah, and Wyoming. The vast majority of fresh water contributed to the Colorado River Basin is collected from the Upper Basin States, primarily through winter snowpack and spring runoff. Because of the impacts caused by the climate change, the water availability of Upper Colorado River Basin faces many challenges from the amount of snowpack in winter, the timing of spring runoff, and the persistent drought. The reservoir systems in the Upper Colorado River Basin provide water supply to nearly 40 million people and serves multiple roles in managing water resources, such as hydropower generation, flood control, agricultural irrigation, as well as recreational opportunities from skiing to fishing and boating. The reservoir inflow, storage, and release records in the basin can be downloaded from the U.S. Bureau of Reclamation water operation archive (https://www.usbr.gov/rsvrWater/HistoricalApp.html). In this study, we consider 30 reservoirs in the Upper Colorado River Basin based on the criterion that there are no more than ten days missing data in the record, which may significantly reduce the prediction accuracy within this time period. Fig.

1 shows locations of these 30 reservoirs. They are non-uniformly distributed in the basin with a quite large difference in elevations. Table I summarizes the basic information of these 30 reservoirs including their names, data record length, elevation and storage. As we can see, the reservoirs have a varying elevation and storage. Also, some reservoirs have a long record of data up to 30 years and some have a relatively short record with only 13 years of data. The data size have a significant impact on prediction performance when the data-driven ML models are used. Moreover, the reservoir features such as its functionality and storage affect the release patterns, which could also affect the prediction performance. We discuss these factors' influence in detail in Section III.

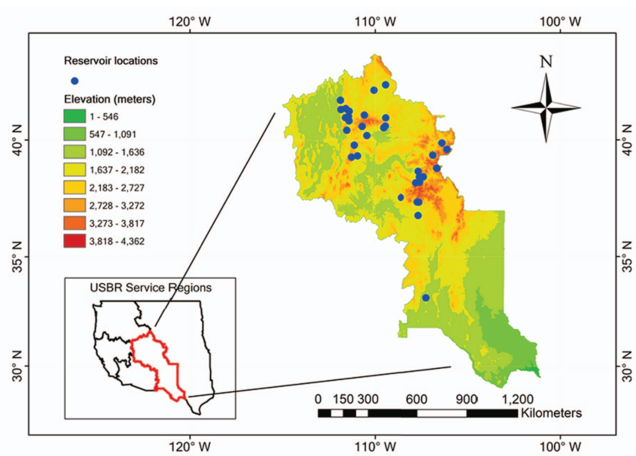

Fig. 1. The 30 reservoirs (blue dots) in Upper Colorado River Basin, US are considered in this study. The figure is modified from [1].

Besides the hydrological data of inflow and storage, we additionally consider the influence of meteorological data on reservoir releases. The daily average precipitation and temperature data are retrieved from the AN81d dataset generated from Parameter-elevation Regressions on Independent Slopes Model (PRISM). The gridded PRISM has a spatial resolution of 4 km (about 0.04 degree) and merges surface measurement with an elevation model [14]. Both PRISM precipitation and temperature are retrieved from the overlapping pixels of the corresponding study reservoirs in the period between 01/01/1982 and 12/31/2011. We then trim the length of precipitation and temperature data in each reservoir to match its inflow and storage record shown in Table I.

B. Long short-term memory network

We use LSTM networks to simulate reservoir releases from hydrometeorological observations and investigate their influences. LSTM, a special type of recurrent neural networks, is structured to learn long-term dependence in time series prediction. In daily reservoir release modeling, we use previous t days of hydrometeorological observations as inputs to predict release on the current day. LSTM learns a mapping for the inputs over time to an output. Thus, it knows what observations it has seen previously are relevant and how they are relevant to the prediction, which enables a dynamical learning of temporal dependence.

The LSTM cell uses four functions and three of them are acting as regulatory gates to control the information flow extracted from the inputs. Furthermore, LSTM introduces an additional cell state $\mathbf{c}_t$ to add and store information and then transfer the stored information to the hidden state $\mathbf{h}_t$. Specifically, LSTM first uses a sigmoid (σ) function $\mathbf{f}_t$ that acts as a forget gate to decide what information should be thrown away from the old cell state. Then, it uses another two functions $\mathbf{g}_t$ and $\mathbf{i}_t$ to decide what new information should be stored in the cell state, where the hyperbolic tangent (tanh) function $\mathbf{g}_t$ first creates a vector of new candidate values that could be added to the cell state, and a sigmoid function $\mathbf{i}_t$ that performs as an input gate then decides which candidate values need to be updated. Next, the two input functions are combined with the forget gate to update the old cell state $\mathbf{c}_{t-1}$ to a new cell state $\mathbf{c}_t$. Finally, LSTM uses the fourth function $\mathbf{o}_t$ that acts as an output gate to decide what parts of the cell state should be exported to update the hidden state. And lastly we use the input information saved in the hidden state $\mathbf{h}_t$ to predict reservoir release y.

Mathematically, the learning process of LSTM networks can be summarized below.

$$\mathbf{f}_t = \sigma\big(\mathbf{W}_f\mathbf{x}_t + \mathbf{U}_f\mathbf{h}_{t-1} + \mathbf{b}_f\big) \tag{1}$$

$$\mathbf{g}_t = tanh\big(\mathbf{W}_g\mathbf{x}_t + \mathbf{U}_g\mathbf{h}_{t-1} + \mathbf{b}_g\big) \tag{2}$$

$$\mathbf{i}_t = \sigma\big(\mathbf{W}_i\mathbf{x}_t + \mathbf{U}_i\mathbf{h}_{t-1} + \mathbf{b}_i\big) \tag{3}$$

$$\mathbf{c}_t = \mathbf{f}_t \times \mathbf{c}_{t-1} + \mathbf{i}_t \times \mathbf{g}_t \tag{4}$$

$$\mathbf{o}_t = \sigma\big(\mathbf{W}_o\mathbf{x}_t + \mathbf{U}_o\mathbf{h}_{t-1} + \mathbf{b}_o\big) \tag{5}$$

$$\mathbf{h}_t = \mathbf{o}_t \times tanh\big(\mathbf{c}_t\big) \tag{6}$$

$$y = \mathbf{W}_d\mathbf{h}_t + \mathbf{b}_d \tag{7}$$

where $\mathbf{x}_t$ presents previous t days of hydrometeorological observations, and y presents the reservoir release on the current day. $\mathbf{W}$ and $\mathbf{U}$ are weight matrices and $\mathbf{b}$ is the bias vector that need to be calculated during LSTM training.

C. Numerical experiments

For each reservoir, we use 80% of its data record as a training set to construct and calibrate the LSTM model, and then use the rest of 20% data as a test set to evaluate the model prediction performance. The data record is split sequentially and use a unit of years. For example, reservoir CRY has 30 years of data from 1982, then the first 24 years of data (i.e., 01/01/1982–12/31/2005) are in the training period and the subsequent 6 years of data (i.e., 01/01/2006–12/31/2011) are in the test period.

LSTM models have several hyperparameters that need to be determined before they can be used for prediction, such as those parameters related to network architectures, learning rates, and the look-back window size t of the input sequence.

TABLE I
INFORMATION OF THE RESERVOIRS CONSIDERED IN THIS STUDY AND THE LSTM MODEL PREDICTION PERFORMANCE ON THE RESERVOIR RELEASE. THE HIGHLIGHTED RESERVOIRS ARE ANALYZED IN DETAIL IN FIG. 3, 4, AND 5.

Initials	Names	Data Start Date	Data Length (years)	Elevation (m)	Storage (acre feet)	Prediction performance					
						CORR	NSE	KGE	RMSE	RSR	PBIAS
BSR	Big Sandy Reservoir	1990	22	2060	38300	0.98	0.94	0.93	0.85	0.24	4.83
CAU	Causey Reservoir	1999	13	1745	8970	0.98	0.96	0.92	1.04	0.20	2.22
CRY	**Crystal Reservoir**	**1982**	**30**	**2251**	**26000**	**1.00**	**0.99**	**0.98**	**2.19**	**0.10**	**0.25**
DCR	Deer Creek Reservoir	1987	25	1653	152000	0.92	0.82	0.78	3.32	0.42	9.87
DIL	Dillon Reservoir	1985	27	2751	257304	0.93	0.86	0.86	3.09	0.37	-0.61
ECH	Echo Reservoir	1982	30	1691	73900	0.97	0.94	0.91	1.97	0.24	-0.31
ECR	East Canyon Reservoir	1992	20	1749	49500	0.95	0.90	0.89	0.68	0.32	0.05
FGR	Flaming Gorge Reservoir	1982	30	1828	3788900	0.98	0.96	0.93	5.40	0.20	5.43
FON	Fontenelle Reservoir	1990	22	1976	345360	1.00	0.99	0.97	4.17	0.10	2.47
GMR	Green Mountain Reservoir	1982	30	2406	153000	0.94	0.88	0.88	3.44	0.35	0.38
HNR	Huntington North Reservoir	1999	13	1774	5420	0.89	0.76	0.66	0.21	0.49	22.98
HYR	Hyrum Reservoir	1999	13	1427	18685	0.97	0.93	0.89	1.82	0.26	3.23
JOR	Jordanelle Reservoir	1997	15	1636	320000	0.91	0.81	0.79	3.53	0.44	6.62
JVR	Joes Valley Reservoir	1996	16	2129	62460	0.95	0.89	0.86	1.30	0.33	4.03
LEM	Lemon Reservoir	1982	30	2478	40146	0.98	0.95	0.95	0.48	0.22	2.22
MCP	**Mcphee Reservoir**	**1991**	**21**	**2073**	**381000**	**0.90**	**0.71**	**0.52**	**4.62**	**0.54**	**-29.48**
MCR	Meeks Cabin Reservoir	1998	14	2647	32470	0.96	0.92	0.85	2.66	0.28	7.45
NAV	Navajo Reservoir	1986	26	1801	1397495	0.91	0.83	0.81	9.35	0.41	3.05
PIN	Pineview Reservoir	1990	22	1495	110000	0.98	0.95	0.94	1.85	0.22	-3.84
RFR	Red Fleet Reservoir	1989	23	1721	26000	0.97	0.92	0.85	0.30	0.28	-10.78
RID	Ridgway Reservoir	1990	22	2101	85000	0.99	0.98	0.94	0.80	0.14	-1.15
ROC	Rockport Reservoir	1982	30	1807	60900	0.98	0.97	0.97	0.93	0.18	-0.79
RUE	Ruedi Reservoir	1982	30	2349	102000	0.96	0.91	0.89	0.83	0.30	1.54
SCO	**Scofield Reservoir**	**1996**	**16**	**2338**	**73600**	**0.83**	**0.69**	**0.65**	**1.40**	**0.56**	**-0.49**
SJR	Silver Jack Reservoir	1992	20	2725	13520	0.99	0.97	0.92	0.94	0.17	3.17
STA	Starvation Reservoir	1982	30	1700	167310	0.96	0.92	0.91	1.12	0.28	-0.80
STE	Steinaker Reservoir	1982	30	1655	33400	0.97	0.94	0.90	0.28	0.25	5.44
TPR	Taylor Park Reservoir	1982	30	2847	1115000	0.99	0.99	0.98	0.40	0.11	-1.69
USR	Upper Stillwater Reservoir	1991	21	2445	32500	0.89	0.77	0.61	2.66	0.48	19.83
VAL	Vallecito Reservoir	1986	26	2318	129700	0.99	0.97	0.96	1.26	0.16	0.15

The t value is particularly important [10] which determines the lag length of input observations used to generate the model output, i.e., the t value represents the period over which the influence of hydrometeorological inputs is taken into account to calculate the reservoir release. However, as a data-driven model, the t value in LSTM cannot be determined a priori with physical basis. Sometimes t is determined heuristically or empirically. For example, some studies thought past one and two days of hydrometeorological information have strong correlations to the current reservoir release and used a small t (i.e., t=1 or 2) in their ML models [1], [15].

In this study, we investigate the influence of t on LSTM model prediction. We first perform a hyperparameter tuning to select the value of t that gives good prediction performance on the validation data and then in numerical experiments we examine the reasonableness and the impact of the selected t on reservoir release prediction. We use 20% of the training set as the validation data to tune the hyperparameters of the LSTM model. We consider a range of t from 1 to 15 and evaluate the prediction performance on the validation data for each t. We define six metrics for performance evaluation which are shown in Section II-D. Results indicate that t=7 days give the best prediction of the reservoir release. Take the reservoir CAU as an example, Fig. 2 illustrates that the NSE value (defined in Equation (9)) gradually increases as the increase of t at the beginning and achieves the highest value when the lag length t is 7 and then the performance drops again. We then use t=7 for all the 30 reservoirs and find that it gives the best prediction performance on the validation data for most of the reservoirs. Therefore, in the numerical experiments of this study, we use previous 7 days of hydrometeorological observations for reservoir release prediction. Other hyperparameters are also tuned as summarized below. For the LSTM model structure, we use one hidden layer with different number of nodes for different reservoirs; in general more nodes are used for the reservoirs with a longer data record. We use the Adam optimizer with the learning rate of 0.001 and use the batch size of 20. The training ends at the epoch of 200. Additionally, all the four input variables are scaled into the range of [0,1] to ensure a stable training.

D. Model evaluation metrics

We employ six statistical metrics widely used in hydrological community [1], [16] to evaluate the LSTM model prediction performance, which include the Correlation Coefficient (CORR), the Root Mean Squared Error (RMSE), the Nash-Sutcliffe Model Efficiency Coefficient (NSE), Kling-Gupta Efficiency (KGE), RMSE-observation standard deviation ratio (RSR), and Percent bias (PBIAS). They are defined as follows.

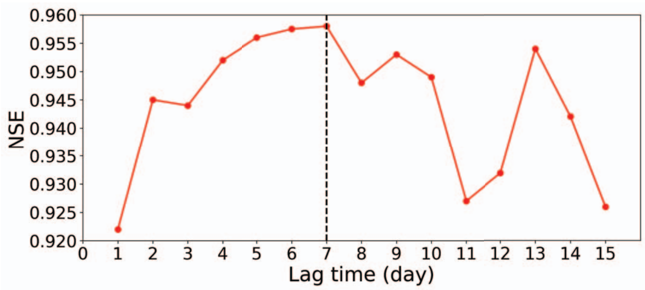

Fig. 2. The NSE values in predicting reservoir release for the different lag length of hydrometeorological observations at reservoir CAU. The LSTM model yields the highest prediction accuracy when the lag length is 7.

$$CORR = \frac{\sum_{i=1}^{n}\left((S_i - \overline{S_i})(O_i - \overline{O_i})\right)}{\sqrt{\sum_{i=1}^{n}\left(S_i - \overline{S_i}\right)^2 \sum_{i=1}^{n}\left(O_i - \overline{O_i}\right)^2}} \quad (8)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(S_i - O_i\right)^2}{\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right)^2} \quad (9)$$

$$KGE = 1 - \sqrt{(1-CORR)^2 + (1-\sigma_S/\sigma_O)^2 + (1-\mu_S/\mu_O)^2} \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{n}} \quad (11)$$

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}\left(O_i - S_i\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right)^2}} \quad (12)$$

$$PBIAS = \frac{\sum_{i=1}^{n}\left(O_i - S_i\right) \times 100}{\sum_{i=1}^{n}\left(O_i\right)} \quad (13)$$

where S_j and O_j are the predicted and observed values, respectively, and $\overline{S_j}$ and $\overline{O_j}$ are their corresponding averages; σ_S and μ_S are the mean and standard deviation of the predicted values; σ_O and μ_O are the mean and standard deviation of the observed values; and n is the total number of observations.

The measurements of CORR, RMSE, and NSE are widely used statistical measures to quantify how the simulated reservoir release matches the observed release. CORR measures how the simulated time series data vary with the corresponding observations. RMSE quantifies the accumulated biases between the prediction and observation. As a combined statistic of both CORR and RMSE, NSE reflects both the temporal variation and bias between the prediction and observation. KGE can amend some shortcomings of the NSE measurments by decomposing NSE values into linear correlation, bias, and variability components. PBIAS quantifies the percentage of biases between the prediction and observation. Based on above definitions, the criteria of RMSE, RSR, and PBIAS with a value of zero, and the metric of CORR, NSE, and KGE having a value of 1 indicate the best prediction accuracy. According to previous studies [1], the model performance and the NSE/KGE scores have the following correspondence: $NSE/KGE > 0.75$: Very good prediction; $0.65 < NSE/KGE \leq 0.75$: Good prediction; $0.5 < NSE/KGE \leq 0.65$: Satisfactory prediction; $0.4 < NSE/KGE \leq 0.5$: Acceptable prediction; and $NSE/KGE \leq 0.4$: Unsatisfactory prediction.

III. Results and Discussions

In this section, we first analyze the LSTM model prediction performance in the 30 reservoirs based on four hydrometeorological inputs (i.e. inflow, storage, precipitation, and temperature) using their 7-day lag observations. Next, we discuss the influence of the four factors and their temporal importance on the reservoir release simulation.

A. LSTM model prediction performance

We summarize the prediction performance in the test period of the 30 reservoirs in Table I using the six evaluation metrics. Overall, the reservoirs have the metric values in the following ranges: $0.83 \leq CORR \leq 1.00$, $0.69 \leq NSE \leq 0.99$, $0.52 \leq KGE \leq 0.98$, $0.21 \leq RMSE \leq 9.35$, $0.10 \leq RSR \leq 0.56$, and $-34.15\% \leq PBIAS \leq 22.98\%$. More than 93% of the reservoirs have a NSE above 0.75, which suggests a very good prediction of LSTM according to the NSE categorization defined in Section II-D. Other five evaluation metrics indicate a consistently good performance as the NSE. This remarkably good prediction performance on these diversely-featured reservoirs demonstrates the accuracy and reliability of the LSTM model in simulating the reservoir release. Next, we pick three reservoirs with different features (e.g., functionality, elevation, and storage) for detailed analysis. We will analyze the model prediction performance from both perspectives of the data size and reservoir feature.

Fig. 3 shows the observed and LSTM-predicted daily release at the reservoir CRY in both training and test periods. According to Table I, this reservoir has the best prediction performance with the NSE of 0.99 in both training and test data. And we can see that the predicted release closely matches the observations despite the irregular release patterns. Reservoir CRY is multi-functional and provides hydropower generation. When providing hydroelectric supplies to meet the energy demands during peak hours, the water is diverted to the powerhouse and this water is accounted into the total release observations. Thus, we see high peak flows and non-smooth variations at low flow regimes in Fig. 3 caused by the reservoir's hydropower generation. In spite of the variations, LSTM can accurately simulate the release. One reason could come from the long record of observation data. Reservoir CRY has 30 years of data where 24 years of data are used for training the LSTM model. This long record of data enables LSTM to capture the underlying dynamics and make the accurate release decision.

On the other hand, reservoirs SCO and MCP show the worst prediction performance, but still having the NSE of 0.69 and 0.71, respectively, according to Table I. Fig. 4 depicts

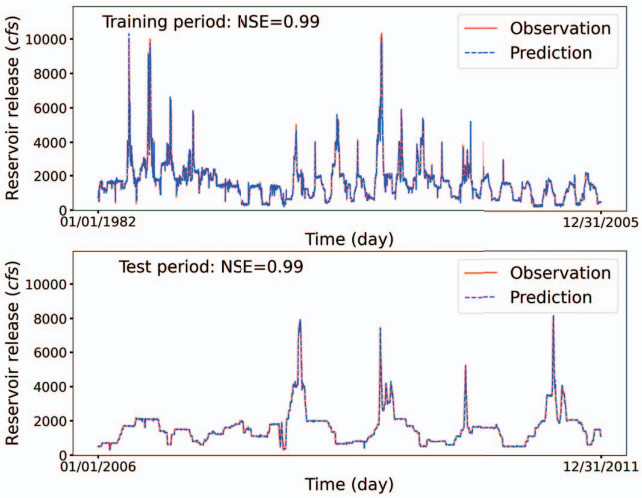

Fig. 3. Observed and LSTM-predicted daily release in reservoir CRY. The LSTM model provides an accurate prediction with high NSEs, although the reservoir CRY has an irregular release pattern resulted from the hydropower generation.

the observations and predictions of release in reservoir SCO. Although the LSTM model misses some peak flows, it is able to accurately capture the general patterns of the release and the timing of the peak flows. Reservoir SCO has a relatively small number of observations, where we use 13 years of its total 16 years of data for training. We expect that with additional observations, the LSTM model can make a better prediction, since data quantity is a key factor for data-driven ML models.

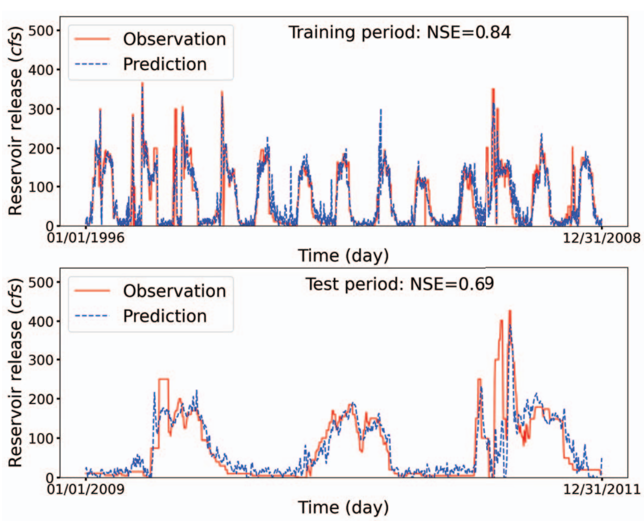

Fig. 4. Observed and LSTM-predicted daily release at reservoir SCO. The LSTM model accurately captures the patterns of the released flow and the timing of peak flows despite the relatively short record of training data.

Likewise, reservoir MCP also has a relatively short record of data, i.e., 17 years of data, for training. Additionally, this reservoir has a large storage and a larger reservoir is usually associated with more complex constraints, natural environmental variabilities, and operating criteria, which jointly challenge the predictions of releases. As we can see in Fig. 5, at most of the time, the release at reservoir MCP is close to zero, and when it releases water, the flow achieves the peak in a short period of time and then quickly drops. This erratic release pattern is difficult to predict. However, the LSTM model predictions can match the observations pretty well giving a good performance with the NSE of 0.71 in test data and 0.96 in the training data. Furthermore, it can reasonably simulate the patterns of release and capture the peak flow timing. Accurate prediction of the peak flow timing of the reservoir release is crucial for flood control and water resources management.

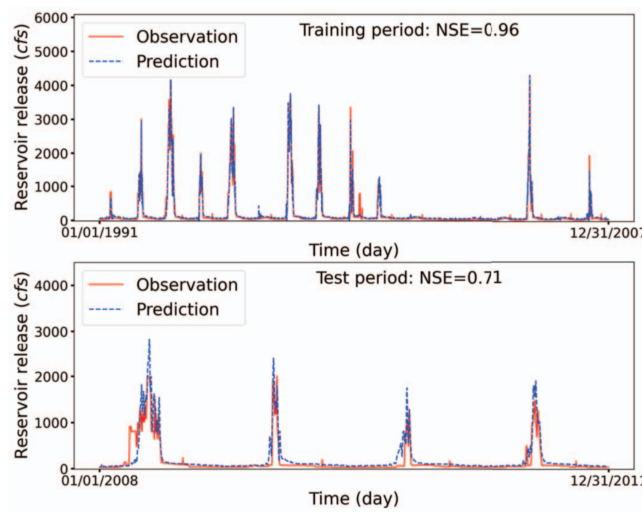

Fig. 5. Observed and LSTM-predicted daily release at reservoir MCP. The LSTM model can make an accurate prediction with a high NSE, despite of its erratic release patterns caused by the large storage and large reservoir size.

Yang et al. (2021) [1] also applied LSTM to predict the release on these 30 reservoirs along with other 11 ML models (e.g., linear regression models, support vector regression models, KNN regression model, decision tree-based models, and MLP models). Different from our work, they considered storage and inflow as inputs only and used their 1-day and 2-day lag observations for release prediction. In their study, the averaged evaluation metrics across the ML models are: $0.48 \le CORR \le 0.99$, $0 \le NSE \le 0.97$, $0.29 \le KGE \le 0.94$, $0.35 \le RMSE \le 39.19$, $0.15 \le RSR \le 1$, and $-32.36\% \le PBIAS \le 28.89\%$. According to the NSE categorization in Section II-D, overall, about 20% of reservoirs obtain a Very Good performance with the NSE exceeding 0.75, which shows an inferior performance to our study where about 93% reservoirs having the NSE over 0.75 (Table I).

We use multi-step lag observations of inflow, storage, precipitation and temperature to predict releases in the 30 reservoirs of the Upper Colorado River Basin, and demonstrate a high prediction accuracy of our LSTM model which also presents a superior performance to the previous study in [1]. We are interested to know, in comparison to the study in

[1], whether the incorporation of additional precipitation and temperature data improves the prediction of reservoir releases and whether their relatively long-term lag observations bring new information to enhance the LSTM learning. Particularly, we want to investigate, in the following section, which hydrometeorological input variables are most influential to the release prediction and whether this influence is from their short-term or long-term contributions.

B. Discussion on the influential factors

We first investigate the influence from the meteorological forcing of temperature and precipitation. We take the experiment with all four input variables as the baseline, and additionally perform another three experiments, where experiment I considers storage and inflow only, experiment II considers storage, inflow, and temperature, and experiment III considers storage, inflow, and precipitation. All four experiments use the 7-day lag observations as inputs. Fig. 6 summarizes the NSE values of these four experiments in predicting the release of the 30 reservoirs in their test period. Overall, we do not observe remarkable difference between the four experiments, which suggests that the temperature and precipitation may not bring significantly influential information for the release prediction.

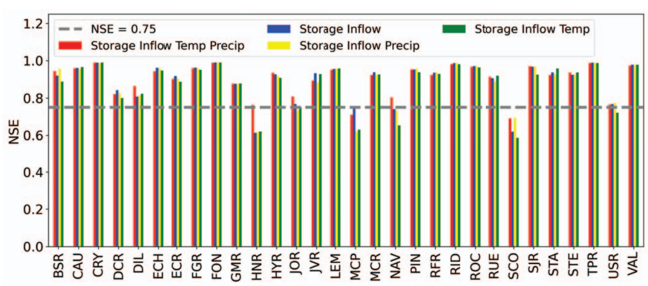

Fig. 6. NSE scores of predicting reservoir release based on different hydrometeorological factors, where the legend (Storage Inflow Temp Precip) represents the four factors, i.e., storage, inflow, temperature, and precipitation, are considered. The NSEs greater than 0.75 suggest a very good prediction.

For a deeper investigation, we pick one reservoir (STE) and analyze the predicted release from the four experiments in detail. Although Fig. 6 shows that the NSEs of the four experiments have similar values in predicting the release at reservoir STE, Fig. 7(a) illustrates that additionally incorporating precipitation and temperature data can slightly improve the prediction of release especially addressing the overestimation of the low flows. Fig. 7(b) shows the temperature and precipitation data used to predict the release in the test period of Fig. 7(a). Focusing on the regions bounded by the two dashed black lines, we can see that without considering temperature and precipitation, the reservoir release suffers from an overprediction. And in this region, the temperature is below zero, the precipitation is solid snow, and there is no reservoir release. Incorporating this meteorological information assists LSTM to capture the reservoir inflow-release dynamics and improve the prediction accuracy.

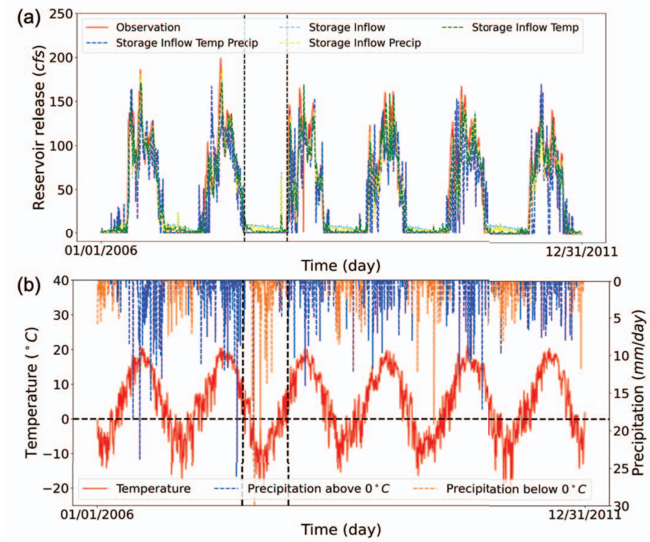

Fig. 7. (a) The predicted daily release of reservoir STE for the four experiments with considering different hydrometeorological inputs. The legend (Storage Inflow Temp Precip) represents the experiment considering all four inputs of storage, inflow, temperature, and precipitation. (b) The temperature and precipitation data used to predict the release at reservoir STE in the test period. Incorporating temperature and prediction improves prediction of release especially in low flow regimes, e.g., the area between the two dashed black lines.

From above discussions, we find that although precipitation and temperature are not the main factors influencing reservoir releases (which is consist with findings from previous studies), they can bring useful information especially in accurately simulating the low flows of release.

In the following, we investigate the influence of lag length of input observations on the LSTM's predictions. We compare the prediction performance from two experiments, one considering 7-day lag input observations and the other using 2-day lag observations. In both experiments, we consider inflow and storage as inputs only as these two factors are influential. More importantly, previous work mainly used short-term lag observations of these two inputs for release prediction, thus we can make a fair comparison and specifically investigate the temporal influence. Fig. 8 presents the NSE scores of these two experiments for the 30 reservoirs. The figure demonstrates that overall the use of 7-day lag observations improves the reservoir release prediction. The average prediction accuracy improvement is 7.98% for the 30 reservoirs. For some reservoirs such as ECR, HNR, JOR, and NAV, the improvement is outstanding, which boosts the prediction performance from acceptable/satisfactory, even unsatisfactory, to very good according to the NSE score categorization in Section II-D.

Take reservoir ECR as an example, we analyze its prediction performance between these two experiments in detail. Fig. 9(a) illustrates the predicted reservoir release using 7-day and 2-day lag observations along with the observed release. We can see that incorporating 7-day lag observations greatly improves the prediction performance where the predicted releases closely

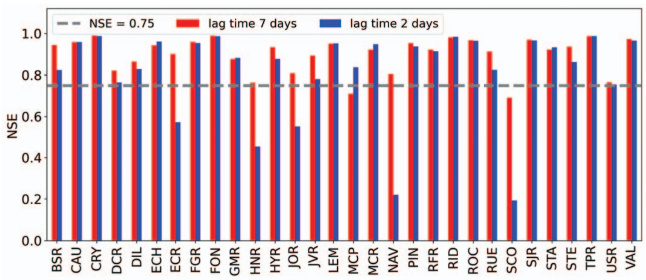

Fig. 8. NSE scores of predicting reservoir release based on different lag times, where the legend (lag time 7 days) represents using previous 7 days of storage and inflow data to predict the reservoir release on the current day. The NSEs greater than 0.75 suggest a very good prediction.

match their observed values. In contrast, the 2-day lag observations do not seem provide sufficient information for a good prediction where the predicted releases underestimate the peak flow, overestimate the low flow, as well as have a large fluctuation. A correlation analysis between the lag observations of inputs and the reservoir release indicates the inflow and storage have a relatively long-term correlation to the release. As shown in Fig. 9(b), the correlation between the 7th day of inflow in the past and the release on the current day can be as high as 0.67 although the correlation has a slight decrease with an increase of the lag length. This confirms that storage and inflow have a long-term contribution to the release and using a longer lag of their observations can improve the LSTM model prediction.

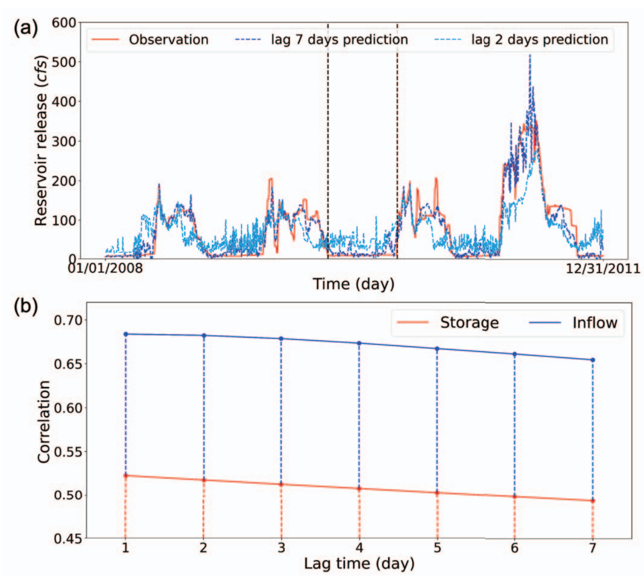

Fig. 9. (a) Observed and LSTM-predicted daily release in reservoir ECR using a different length of lag observations of storage and inflow as inputs (i.e., previous 7 days or 2 days). (b) Correlations between the previous 7 days' storage and inflow and the reservoir release on the current day.

IV. Conclusions

In this work, we use the LSTM models to predict the reservoir release at 30 diverse reservoirs in the Upper Colorado River Basin, US from four hydrometeorological inputs in consideration of their long-term lag observations. We evaluate the prediction performance, as well as investigate the input variable importance and their temporal influence on the release prediction. The main conclusions are summarized as follows:

1) The LSTM model can accurately simulate the reservoir release from hydrometeorological observations with a high prediction performance. For the 30 reservoirs, we obtain the following statistics, $0.83 \leq CORR \leq 1.00$, $0.69 \leq NSE \leq 0.99$, $0.52 \leq KGE \leq 0.98$, $0.21 \leq RMSE \leq 9.35$, $0.10 \leq RSR \leq 0.56$, and $-34.15\% \leq PBIAS \leq 22.98\%$. Even the worst NSE value is above 0.69 demonstrating a good prediction performance.
2) Inflow and storage are the most influential factors for reservoir release. Although temperature and precipitation are less important, incorporation of them can improve the prediction of release especially for the low flow estimation.
3) The storage and inflow have a relatively long-term effect on the reservoir release. Incorporating of their 7-day lag observations can greatly improve the prediction accuracy in comparison to the case with only 2-day lag observations considered.

These findings can provide guidance for reservoir operators to make reasonable decisions and also optimize the water resources management.

Acknowledgment

This research was supported by the Artificial Intelligence Initiative as part of the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US DOE under contract DE-AC05-00OR22725. It is also sponsored by the Data-Driven Decision Control for Complex Systems (DnC2S) project funded by the US DOE, Office of Advanced Scientific Computing Research. The work is also partially supported by the National Science Foundation under Grant No. OIA-1946093 and its subaward No. EPSCoR-2020-3, and the National Science Foundation under Grant No. NSF1802872.

References

[1] T. Yang, L. Zhang, T. Kim, Y. Hong, D. Zhang and Q. Peng, 2021. A large-scale comparison of Artificial Intelligence and Data Mining (AI&DM) techniques in simulating reservoir releases over the Upper Colorado Region. Journal of Hydrology, 602, p.126723.
[2] D. Zhang, D. Wang, Q. Peng, J. Lin, T. Jin, T. Yang, S. Sorooshian and Y. Liu, 2022. Prediction of the outflow temperature of large-scale hydropower using theory-guided machine learning surrogate models of a high-fidelity hydrodynamics model. Journal of Hydrology, 606, p.127427.
[3] T. Yang, X. Liu, L. Wang, P. Bai and J. Li, 2020. Simulating hydropower discharge using multiple decision tree methods and a dynamical model merging technique. Journal of Water Resources Planning and Management, 146(2), p.04019072.

[4] T. Yang, X. Gao, S. Sorooshian and X. Li, 2016. Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme. Water Resources Research, 52(3), pp.1626-1651.
[5] M. Zarei, O. Bozorg-Haddad, S. Baghban, M. Delpasand,E. Goharian and H.A. Loáiciga, 2021. Machine-learning algorithms for forecast-informed reservoir operation (FIRO) to reduce flood damages. Scientific reports, 11(1), pp.1-21.
[6] T. Yang, A.A. Asanjan, E. Welles, X. Gao, S. Sorooshian and X. Liu, 2017. Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. Water Resources Research, 53(4), pp.2786-2812.
[7] N. Ehsani, C.J. Vörösmarty, B.M. Fekete and E.Z. Stakhiv, 2017. Reservoir operations under climate change: Storage capacity options to mitigate risk. Journal of Hydrology, 555, pp.435-446.
[8] D. Lu, G. Konapala, S.L. Painter, S.C. Kao and S. Gangrade, 2021. Streamflow simulation in data-scarce basins using bayesian and physics-informed machine learning models. Journal of Hydrometeorology, 22(6), pp.1421-1438.
[9] D. Zhang, J.Lin, Q.Peng, D.Wang, T.Yang, S. Sorooshian, X. Liu and J. Zhuang, 2018. Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. Journal of Hydrology, 565, pp.720-736.
[10] D. Zhang, Q.Peng, J.Lin, D.Wang, X. Liu and J. Zhuang, 2019. Simulating reservoir operation using a recurrent neural network algorithm. Water, 11(4), p.865.
[11] S. Yang, D. Yang, J. Chen and B. Zhao, 2019. Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. Journal of Hydrology, 579, p.124229.
[12] P. Fan, D. Lu and D. Rastogi, 2022. Multimodel Ensemble Predictions of Precipitation using Bayesian Neural Networks. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States).
[13] M. Fan, D. Lu, D. Rastogi and E.M. Pierce, 2022. A Spatiotemporal-Aware Climate Model Ensembling Method for Improving Precipitation Predictability. https://arxiv.org/abs/2208.04390.
[14] C. Daly and K.Bryant, 2013. The PRISM climate and weather system—an introduction. Corvallis, OR: PRISM climate group, 2.
[15] T. Zhao, D. Yang, X. Cai, J. Zhao and H. Wang, 2012. Identifying effective forecast horizon for real-time reservoir operation under a limited inflow forecast. Water Resources Research, 48(1).
[16] H.V. Gupta, H. Kling, K.K. Yilmaz and G.F. Martinez, 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of hydrology, 377(1-2), pp.80-91.