JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 58, No. 6 AMERICAN WATER RESOURCES ASSOCIATION December 2022

Machine Learning Assisted Reservoir Operation Model for Long-Term Water Management Simulation

Sudershan Gangrade, Dan Lu, Shih-Chieh Kao, and Scott L. Painter

Research Impact Statement: Hybrid rule-based and data-driven models for reservoir operations are more accurate than either class of models alone, and can maintain physical consistency in long-term water management simulations.

ABSTRACT: This study explores strategies for long-term reservoir simulations by combining generic rule-based reservoir management model (RMM) and machine learning (ML) models for two major multipurpose reservoirs — Allatoona Lake and Lake Sidney Lanier in the southeastern United States. First, a standalone RMM is developed to simulate daily release and storage during Water Year 1981–2015. Next, using Long-Short Term Memory (LSTM) as the ML technique, a standalone LSTM model is trained based on reservoir inflow and meteorological observations to simulate reservoir release and estimate reservoir storage through water balance calculation. Three hybrid modeling strategies are developed, one using RMM output as an additional LSTM input (H1), another using LSTM as the initial release estimate in RMM (H2), and the third combining the first two strategies (H3). The Nash–Sutcliffe efficiency (NSE) for release (NSE-r), storage (NSE-s), and their mean (NSE-avg) are used for model evaluation. Overall, H1 improves NSE-r to 0.65 and 0.54 for Allatoona and Lanier, respectively, compared to standalone RMM (0.44 and 0.21); however, its storage trajectory did not produce a physically feasible solution, similar to LSTM. H2 and especially H3 show that they can retain the best features from RMM and LSTM, with H3 NSE-avg being 0.695 and 0.55 for Allatoona and Lanier outperforming RMM (0.615 and 0.29). The findings suggest a robust simulation capacity for large-scale water management in future studies.

(KEYWORDS: reservoir operation; machine learning; LSTM; hybrid modeling; long-term water management.)

INTRODUCTION

Reservoirs are an important surface water management infrastructure that regulate natural hydrologic variabilities to support a variety of human activities (Yang et al. 2016; Ehsani et al. 2017). Globally, there are around 60,000 large dams (i.e., >15 m height or >3 million m^3 storage) providing a combined storage that can hold around one-fifth of total annual natural runoff (Wisser et al. 2013; ICOLD 2020). Given their significant influence on streamflow variabilities, the accurate representation of reservoirs is vital to studies across multiple scales, such as reservoir-specific operational forecasting, basin-scale water resources planning, and national-scale hydroclimate projections. However, due to insufficient data and process representation, the characterization of human alterations to hydrologic systems through large-scale reservoir operation remains a

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy (DOE). The U.S. government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. 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 U.S. 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). Received January 30, 2022; accepted August 10, 2022. Paper No. JAWR-22-0009-P of the *Journal of the American Water Resources Association* (JAWR). **Discussions are open until six months from issue publication**.

Environmental Sciences Division (Gangrade, Kao, Painter), Oak Ridge National Laboratory Oak Ridge, Tennessee, USA; and Computational Sciences and Engineering Division (Lu), Oak Ridge National Laboratory Oak Ridge, Tennessee, USA (Correspondence to Gangrade: gangrades@ornl.gov).

Citation: Gangrade, S., D. Lu, S. Kao, and S.L. Painter. 2022. "Machine Learning Assisted Reservoir Operation Model for Long-Term Water Management Simulation." *JAWRA Journal of the American Water Resources Association* 58 (6): 1592–1603. https://doi.org/10.1111/1752-1688.13060.

major challenge (Yassin et al. 2019) in the Earth and water resources modeling communities.

Reservoir operations are conventionally represented by rule-based reservoir management models (RMMs) that mimic reservoir management decisions following a series of operational rules and water balance calculation. A decision to release or store the water is governed by pre-determined operational rule curves (e.g., targeted reservoir elevation at each month), or by other flood control, environmental, or legal compliance constraints. Detailed RMMs such as HEC-ResSim and RiverWare (Zagona et al. 2001; Klipsch and Hurst 2013) have been used to support operational decisions, but can only be set up when the "full" reservoir information (i.e., elevation-storage relationships, withdrawals, constraints, and operation rules) are available. Since many reservoir-related information is either safety- or business-sensitive, the full reservoir information is usually unavailable which makes the applications of detailed RMMs challenging. Simplified RMMs which rely on generic schemes can be utilized as an alternative in large-scale applications (Shin et al. 2019; Yassin et al. 2019; Dang et al. 2020), but may not reach the same accuracy required for long-term water resources planning. Additionally, the multipurpose reservoirs are often operated by experienced operators with their own professional judgments to account for some specific local constraints that cannot be precisely captured by the operation targets (Yang et al. 2019).

Alternatively, data-driven approaches can be used to represent reservoir operations. For instance, machine learning (ML) and deep learning (DL) techniques have gained attraction (Zhang et al. 2019; Yang et al. 2020) due to their ability to capture nonlinear relationships and detect patterns in the historic reservoir operation and release information. The ML-based techniques may also be used to derive operational rule curves which can then be used as inputs within RMMs (Coerver et al. 2018). ML techniques such as artificial neural network (ANN; Brêda et al. 2021; Zhang et al. 2018), recurrent neural network (RNN; Yang et al. 2019; Zhang et al. 2019), long short term memory (LSTM; Zhang et al. 2018), support vector machine, and classification and regression tree (Yang et al. 2016) have been utilized in several reservoir operation applications, mostly focusing on the prediction of reservoir release. However, while ML-based models may yield good performance in real-time or short-term release forecasting (Yang et al. 2019), given their purely data-driven nature, it remains unclear if they can also be used to simulate other reservoir functions (e.g., storage) to support long-term water resource planning. In particular, the lack of process representations may result in physically infeasible outputs, especially when inputs are outside the range of training data. These effects may even be more pronounced when ML-based models are used for multi-decadal projections, in which large changes in hydroclimate conditions can be expected.

At the intersection lies the hybrid modeling approach that can leverage the individual strengths of both process-based and ML-based models. The hybrid modeling approach has shown promises in various fields in hydrology such as streamflow simulation (Konapala et al. 2020; Lu et al. 2021) and lake water temperature modeling (Read et al. 2019; Jia et al. 2021). The hybrid models can be developed through different strategies, such as using ML to improve the raw predictions made by process-based models (Konapala et al. 2020), predict model residuals (Wan et al. 2018), or by incorporating a physics-based loss function during regularization (Khandelwal et al. 2020). In some instances, the ML-based models may be used to replace certain physical processes for a better and more efficient process representation. For instance, Ehsani et al. (2015) developed a general reservoir operation scheme using ANN coupled with a water balance model and demonstrated its applicability for climate change application in a follow-up study (Ehsani et al. 2017). However, even though the hybrid modeling approach has emerged in hydrologic studies, their potentials in simulating reservoir functions have not been fully explored. For long-term water resources planning, it is of interest to understand if the hybrid approach may provide a feasible and efficient solution at expanded temporal and spatial scales.

In this study, we explore the potential of hybrid RMM-LSTM in representing reservoir dynamics (i.e., both reservoir storage and release) for long-term water management simulation and its usefulness for multi-decadal projections. The main objectives are to (1) explore the suitability and challenges of RMM- and LSTM-only models in representing dynamics of complex multipurpose reservoirs, and (2) explore if the hybrid modeling approach can improve the reservoir representation for applications in integrated modeling systems over multi-decadal timescales. We utilized an existing RMM from Gangrade (2019) adopted to two major multipurpose reservoirs in the southeastern United States (U.S.), Allatoona Lake, and Lake Sidney Lanier serving the greater Atlanta metropolitan area. The standalone RMM was developed to simulate daily release and storage during Water Year (WY) 1980–2015 using long-term reservoir inflow observations and reservoir pertinent information (physical attributes of dam). We developed a standalone LSTM based on reservoir inflow and meteorological observations to simulate reservoir release, and then used water balance calculation to estimate the corresponding reservoir storage. We then considered three hybrid models by combining RMM and LSTM, one using RMM output as an additional LSTM input, another using LSTM as the initial release estimate in RMM, and the third combining the two strategies. We evaluate and compare the performance of these different modeling approaches to understand their strengths and limitations for further expanded applications.

STUDY AREA AND DATA

We focus on two multipurpose reservoirs in the southeast U.S., Allatoona Lake and Dam (Allatoona) and Lake Lanier and Buford Dam (Lanier) within Alabama Coosa Tallapoosa (ACT) and Apalachicola-Chattahoochee-Flint (ACF) River Basin, respectively (Figure 1). The reservoirs are owned and operated by U.S. Army Corps of Engineers (USACE). The upstream drainage area of both reservoirs is similar (i.e., ~2,850 km^2 for Allatoona and ~2,700 km^2 for Lanier). The conservation pool capacity for Lanier is ~1,283 million m^3 (1,040,400 acre-ft), which is roughly three times larger than Allatoona. Jointly, Allatoona, and Lanier provide multiple benefits such as flood risk management, hydropower generation, navigation, water supply, water quality, fish and wildlife conservation, and recreation to the community. These reservoirs also provide municipal and industrial water supply to the Atlanta metropolitan and surrounding urban areas (USACE 2015, 2017). Given their importance in the region and the ongoing water allocation conflict between the states of Alabama, Georgia, and Florida, they make a suitable and an interesting test case for this study.

The reservoir pertinent information is available through the master water control manuals for both ACT and ACF river basins (USACE 2015, 2017). The long-term inflow, release, and storage observations are obtained from Duke University's USACE Database (https://nicholasinstitute.duke.edu/reservoir-data/) where peer-reviewed daily observation is available until September 2015. For further information, readers are referred to (Patterson and Doyle 2018).

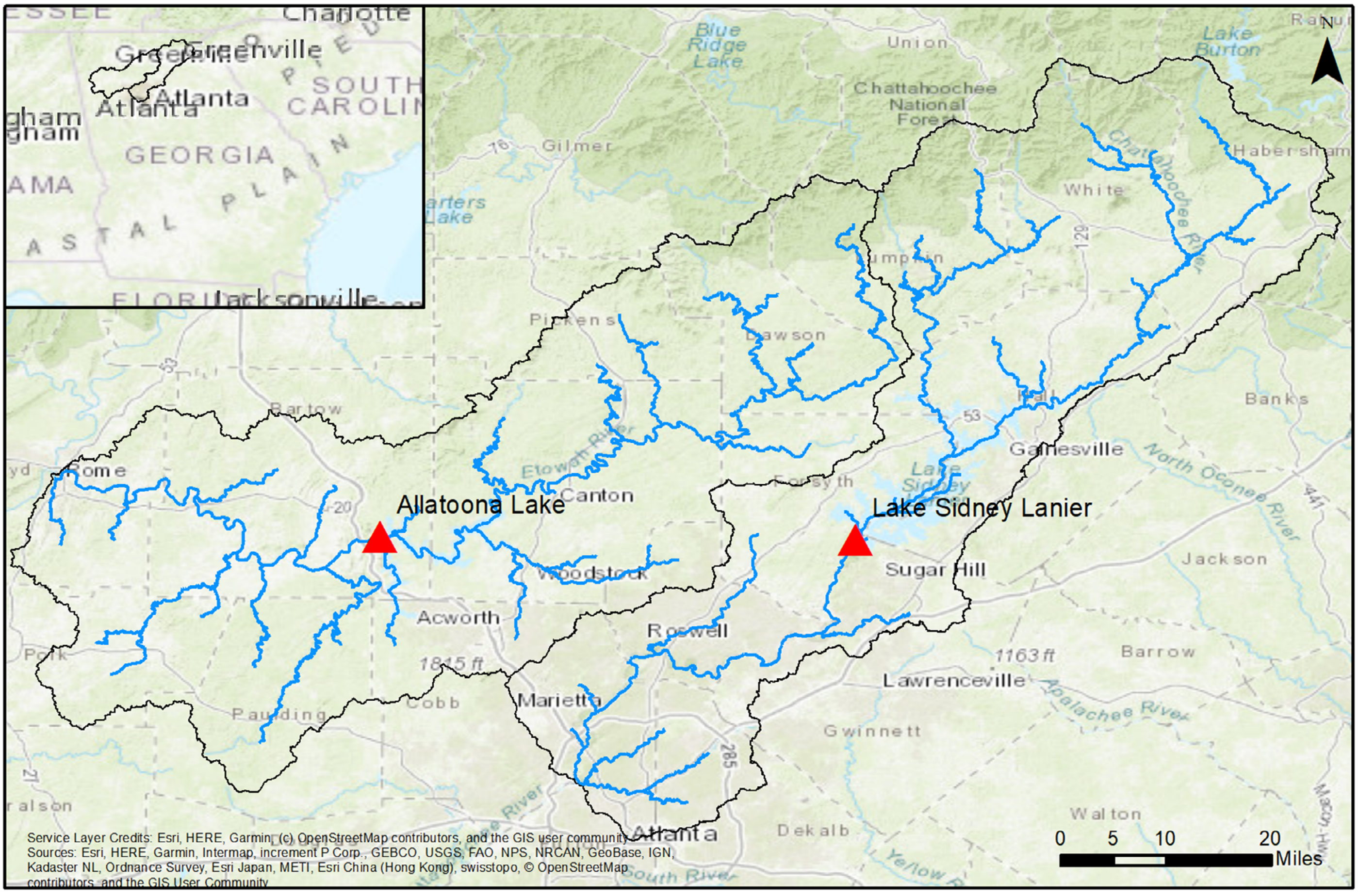

FIGURE 1. Lakes Allatoona and Lanier near Atlanta metropolitan area in the southeast United States.

The meteorological observations including daily precipitation and temperature are obtained from Daymet (https://daymet.ornl.gov; Thornton et al. 2021), which is a widely used gridded meteorological data product from 1980 to present for the entire North America at 1 km horizontal resolution. An area-weighted average timeseries for the upstream catchments of both reservoirs was derived for WY 1981–2015 using the latest Daymet V4 data which provides several enhancements such as reduced bias in timing among others.

METHODOLOGY

Rule-Based Management Model

The RMM used in the study is a simplified and decoupled version of an intermediate complexity multipurpose reservoir module from DHSVM-Res (Zhao et al. 2016). It employs a storage-release scheme by dividing the total reservoir storage into several management pools. The reservoir release volume (Q_t) at any given time (t) is determined based on the current reservoir storage (S_t) with respect to different management pool volumes. Similar practice is also adopted by USACE for Allatoona and Lanier, making this a suitable approach. During the setup of RMM, we only use top of the conservation pool as the operational target, while in reality Allatoona and Lanier are operated on more complex set of rules. Although this model is relatively simple compared to USACE's operational model, it is set up with publicly available information and has showed reasonable performance reported by Zhao et al. (2016) and Gangrade (2019). Further details and equation are provided in Supporting Information.

Long Short-Term Memory Network

Long-Short Term Memory network is a subset of RNN and has shown promise in hydrological/reservoir operation forecasting due to its ability to learn long-term nonlinear relationship in the temporal data. The LSTM architecture includes an additional cell state compared to RNN, and three gates including forget, input, and output gates to control the flow of information between cell states and hidden states (Kratzert et al. 2018). Further details and equation are provided in Supporting Information. In this study we utilize a single-layer LSTM model to learn the N-to-1 relationship, that is, using inputs from previous N timesteps to predict output at the next time step. The details about the input variables are described later. We conducted a grid-based search to identify the following hyperparameters: (1) length of previous N time steps (i.e., lookback window size) from the input timeseries (values used: 90-, 180-, and 365-day), (2) hidden size: number of hidden units of the recurrent layer (values used: 5, 10, 20, and 40), and (3) learning rate: (values used: 0.01, 0.001, 0.005, and 0.0005). A constant dropout rate of 0.1 was utilized to avoid overfitting in the network training. The calibration period used by RMM (WY 1981–1998) is split into 90/10 ratio as training and validation data for LSTM network training and hyperparameter tuning. The following network architecture and learning rate perform best for the validation data: 365-day lookback window, 40 hidden states, and a learning rate of 0.005 (for Allatoona) and 0.01 (for Lanier). These hyperparameter values were then kept unchanged to provide a fair evaluation among different hybrid experiments as described in later sections. During the experiments, the LSTM networks are trained on full calibration period (WY 1981–1998) to minimize the loss value (mean squared error) between the observed and predicted release until no further improvement is seen for loss function in validation period (WY 1999–2015) for consistency with RMM experiments. All LSTM experiments are conducted using PyTorch library. The LSTM code was adapted from publicly available LSTM implementation for rainfall-runoff modeling from Kratzert et al. (2018).

Experimental Setup

We present an overview of different approaches to simulate reservoir dynamics (i.e., both release and storage) through the following simulations. The experimental setup is also presented as a schematic in Figure 2.

Simulation 1 — Standalone RMM. We setup a standalone RMM model using historical timeseries of inflow, outflow, storage, operational target for both Allatoona and Lanier at daily time step. We simulate timeseries of both reservoir release and storage which are compared against the observations. Using the available input data (such as inflow, release, storage, operational target, and other physical attributes of reservoir.) from WY 1981–2015, the model is setup to simulate release and storage. As the records of water-demand are not available, an average net flux volume (F_t) is determined based on the difference between historical inflow and release for both reservoirs. We also specify the initial reservoir storage looked up from historical observations. The first half of data (WY 1981–1998) is used for model calibration, while the second half (WY 1999–2015) is used for model validation. During calibration, the daily Nash–

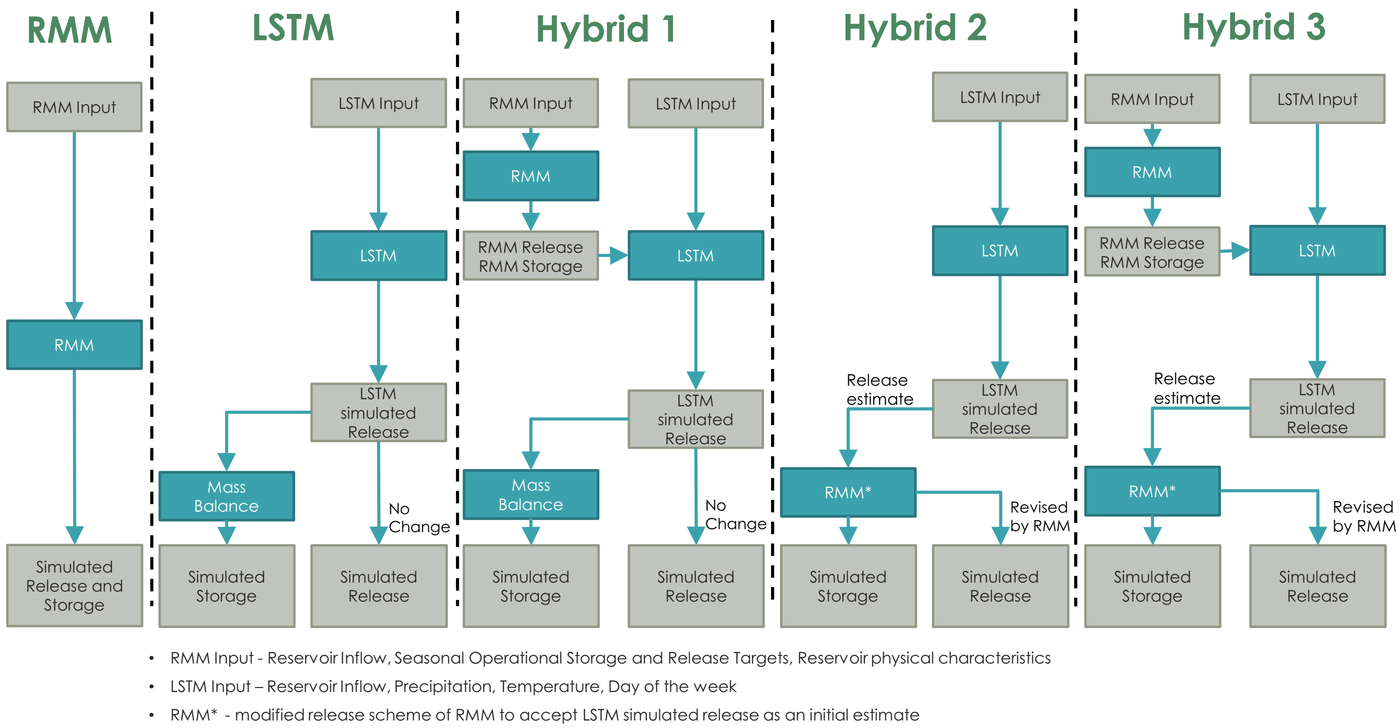

FIGURE 2. Schematic of five experimental setup. RMM, reservoir management model; LSTM, Long-Short Term Memory.

Sutcliffe efficiency (NSE) is maximized for release and storage simultaneously using shuffled complex evolution algorithm (SCE; Duan et al. 1994), by adjusting selected parameters as suggested by Zhao et al. (2016).

Simulation 2 — Standalone LSTM. We develop a standalone LSTM model to predict reservoir release using daily timeseries of reservoir inflow, daily precipitation, temperature, and weekday vs. weekend as inputs. The same calibration and validation periods (with standalone RMM) are used to train and test the LSTM network. Since the LSTM network here does not calculate reservoir storage explicitly, a storage trajectory is calculated by using the LSTM simulated reservoir release as Q_t in the water balance equation (Equation S1).

Simulation 3 — Hybrid 1 (H1). Here, we test first hybrid approach, which uses process-based RMM model outputs to train a LSTM network. In particular, we train LSTM network to predict daily release using daily timeseries of reservoir inflow, daily precipitation, temperature, weekday vs. weekend, and RMM simulated release and storage as inputs. Similar to Simulation 2, the final reservoir storage is calculated using the water balance equation (Equation S1).

Simulation 4 — Hybrid 2 (H2). Next, we test another way to leverage the benefits of both LSTM and RMM. Here, the RMM release scheme is modified to accept the predicted release ($Q_{\mathrm{pred},(t)}$) from the standalone LSTM from Simulation 2 as an initial estimate of release. A water balance check is then performed to ensure that the reservoir storage can remain physically feasible with the predicted release. If the storage will be largely below the conversion pool target ($S_t < \mathrm{OT}_t$), the final release is reduced by $Q_t = r2 \times Q_{\mathrm{pred},t}$. Otherwise, if the storage will be largely above the flood control pool target, the final release is enlarged by $Q_t = r3 \times Q_{\mathrm{pred},t}$. Here, $r2$ and $r3$ are empirical factors to impose water balance constrain specific to a reservoir and can be further calibrated. Such correction is helpful to maintain reasonable water balance over time. The RMM with modified release scheme is referred as RMM*.

Simulation 5 — Hybrid 3 (H3). Lastly, we test a slight variation that combine both H1 and H2. Starting from H1, we first use the process-based RMM model outputs as an additional input in the LSTM training to generate the initial estimate of release ($Q_{\mathrm{pred},t}$). The water balance check in H2 is then performed to revise Q_t based the projected reservoir storage.

Model Evaluation Statistics

To evaluate the performance of these five different approaches, we use the following measures: relative root mean squared error (RRMSE), RMSE-observation standard deviation ratio (RSR), percent bias (PBIAS), and NSE. Further, we also use percent bias for reservoir release during high flow volume (FHV defined as top 2% of release) and low flow volume (HLV defined as bottom 30% of release) (Ouyang et al. 2021) to evaluate model performance during peak and low release regimes. Further details and equation are provided in Supporting Information.

For all simulations, after calibration/validation of RMM or training/testing of LSTM, the simulations are conducted for the entire time period (WY 1981–2015). The model evaluation statistics are then reported for WY 1982–2015, by dropping the first year for model spin up.

RESULTS AND DISCUSSION

Performance of RMM and LSTM

RMM Performance. The RMM simulated reservoir release and storage for both reservoirs are illustrated in Figures 3, 4a, and 4b. While the NSE of daily release (NSE-r) is lower as 0.44 and 0.21, the NSE of daily storage (NSE-s) is higher as 0.79 and 0.37 (Figures 3b and 4b; Table 1). Both reservoirs show an underestimation in release under high flow conditions, and an overestimation in release under low flow conditions, as evident by PBIAS for FHV and FLV (Table 2). Overall, RMM performs better for Allatoona than Lanier, which may be explained by their different historic operations (Figures 3b and 4b). While the historic reservoir storage of Allatoona was kept close to the top of conservation pool, the storage of Lanier involved larger variability. In terms of reservoir storage, it can be attributed to the fact that the conservation pool of Lanier is roughly three times of Allatoona, and therefore Lanier has greater operational flexibility than Allatoona, making it more difficult to be represented through RMM.

In terms of release, while the selected RMM can provide good performance at weekly and monthly scales (not showed), the performance at daily scale is much weaker. Upon further evaluation, it was found that the release can be quite different during weekday vs. weekend, possibly due to the generation of hydropower. The current release scheme employed in RMM also tends to preserve water by allowing a release at a minimum rate and therefore contribute to the lower performance of Lanier release. Nevertheless, the current NSE-r and NSE-s are acceptable and comparable to other generic rule-based RMMs as reported in the literature (Voisin et al. 2013; Shin et al. 2019; Yassin et al. 2019; Ouyang et al. 2021). It demonstrates while a low-to-medium complexity release scheme (as employed here by DHSVM-Res) can capture the reservoir regulation effects on release and storage sufficiently, they may be inadequate to represent day-to-day operation. We note that Lake Allatoona and Lanier are operated on a much more complex set of rules. Capturing the full reservoir operation requires highly customizable software and detailed operational information that are not fully available for open scientific research.

LSTM Performance. The LSTM simulated release (Figures 3c and 4c) and associated performance statistics indicate a significant improvement in modeling daily release. The NSE-r values are raised to 0.62 and 0.58 for Allatoona and Lanier, respectively. The results also suggest roughly 19%–27% improvement in RMSE for both reservoirs. Further, the percent bias in FHV and FLV are significantly improved as reported in Table 2. This shows that LSTM is able to learn and mimic the daily release patterns based on the inputs of reservoir inflow, precipitation, temperature, and weekday vs. weekend. A sensitivity test was conducted to understand the importance of using weekday vs. weekend as an input for LSTM training (Table S1). The results indicated that such a differentiation may boost daily NSE-r values from 0.49 to 0.62 for Allatoona and 0.46 to 0.58 for Lanier, highlighting its importance in predicting reservoir release potentially due to hydropower generation patterns. On a closer look at the release patterns, we also identify certain instances of negative release values, which can be considered physically inconsistent in the current reference.

We intentionally did not train on reservoir storage because that information is often not available and one of our goals was to explore methods that are able to learn reservoir release patterns based on widely available information. Moreover, we did not include historical observations of outflow as input variables in our LSTM-predicted releases. Although inclusion of observed outflow and storage may significantly improve the accuracy of predicted release in short-term forecasts (Yang et al. 2019; Zhang et al. 2019), our aim is to eventually use this model for multidecadal hydroclimate studies where that information is not available. For that reason, we limited our input variables to quantities that are provided by hydroclimatic projections (i.e., atmospheric forcing and reservoir inflows). Further, a comparison of storage trajectory indicates that LSTM predicted storage did not provide physically feasible solutions in the long term, as the errors continued to accumulate (Figures 3d and 4d). This is a common limitation of ML-based techniques also reported by other studies (Yang et al. 2016, 2019).

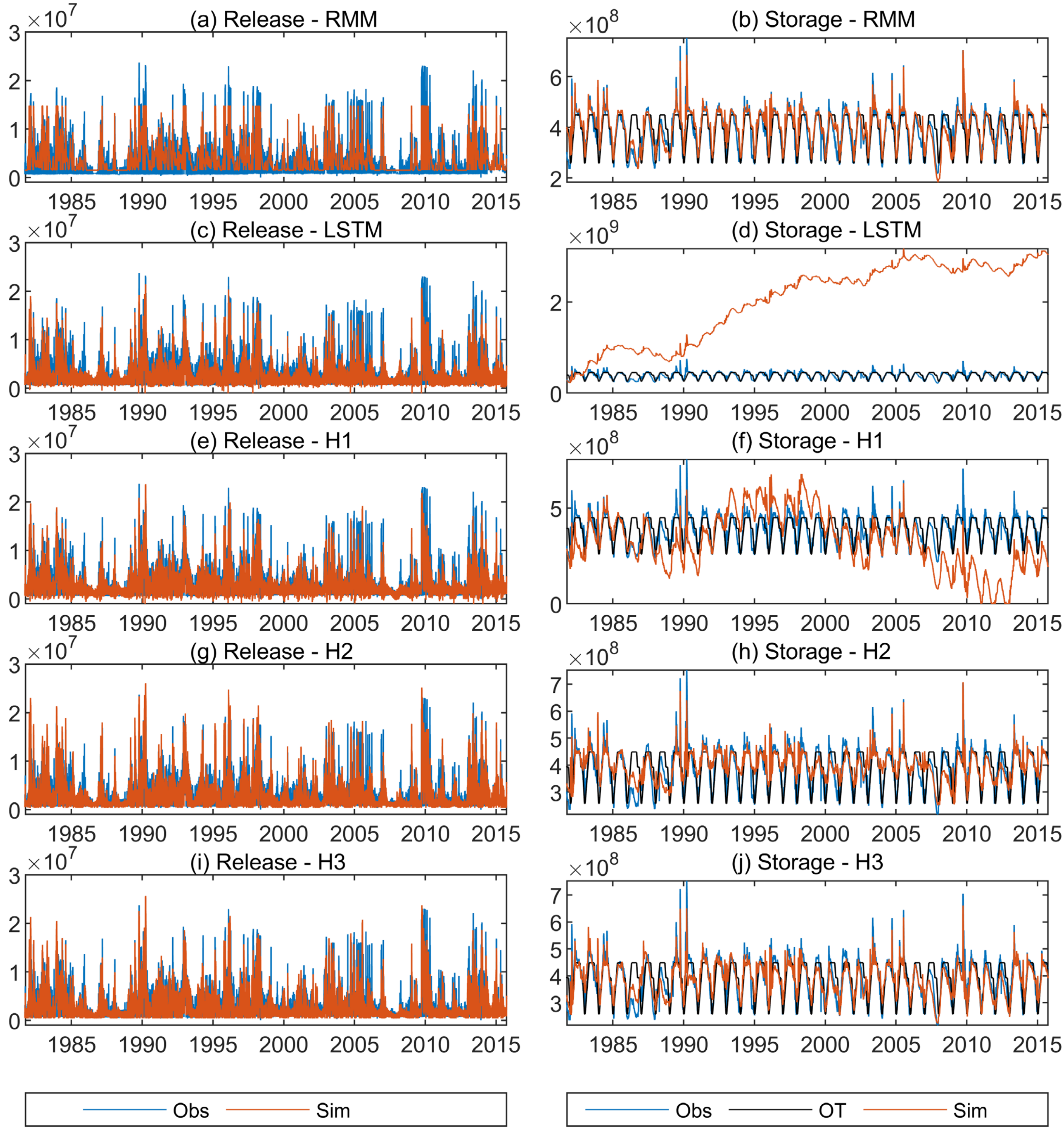

FIGURE 3. Observed and simulated daily reservoir release volume (m^3, panels a, c, e, g and i) and storage (m^3, panels b, d, f, h and j) for Water Year (WY) 1982–2015 for Lake Allatoona.

Performance of Hybrid Models

H1 Performance. The first hybrid strategy is motivated by the approach where ML models are trained using additional outputs from a process-based model (Konapala et al. 2020). We test this approach through the H1 experiment. In terms of release, H1 simulated release (Figures 3e and 4e) and statistics

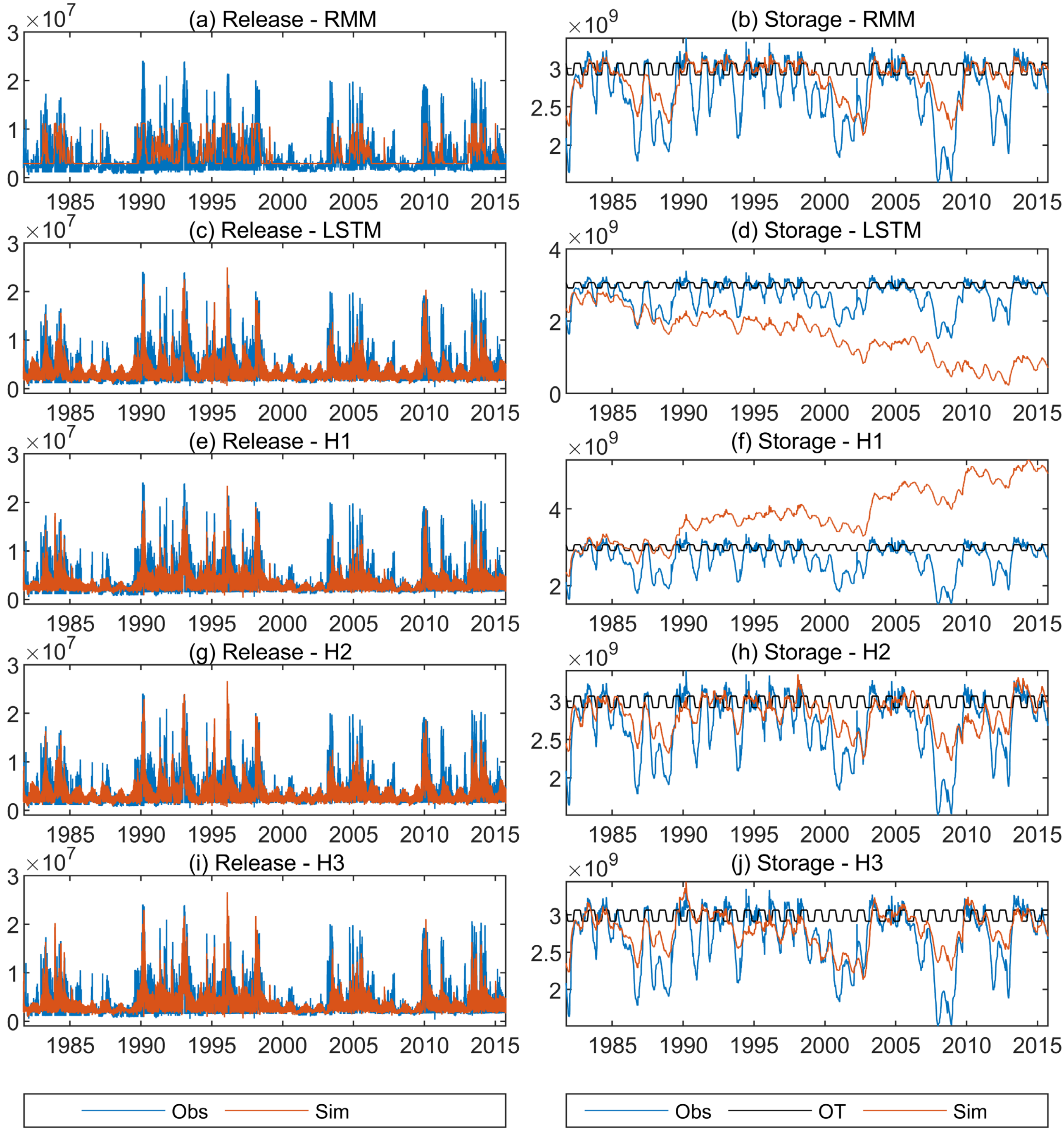

FIGURE 4. Observed and simulated daily reservoir release volume (m^3, panels a, c, e, g and i) and storage (m^3, panels b, d, f, h and j) for Water Year (WY) 1982–2015 for Lake Lanier.

show a good performance with NSE-r of 0.65 and 0.54 for Allatoona and Lanier. While it showed a large improvement when comparing to the standalone RMM, there was no clear improvement when comparing to the standalone LSTM (or even lower performance in Lanier). This indicates that RMM outputs for Lanier are not providing meaningful information for LSTM. While it is considered that

TABLE 1. Summary statistics for NSE, PBIAS, RRMSE, and RSR.

Reservoir	Simulation	NSE	PBIAS	RRMSE	RSR
Allatoona — release	RMM	0.44	0.17	0.75	0.75
	LSTM	0.62	−5.48	0.61	0.61
	H1	0.65	0.40	0.59	0.59
	H2	0.59	0.06	0.64	0.64
	H3	0.63	0.07	0.60	0.60
Allatoona — storage	RMM	0.79	2.40	0.09	0.46
	LSTM	−619.10	419.55	4.73	24.90
	H1	−3.03	−16.12	0.38	2.01
	H2	0.56	2.12	0.13	0.66
	H3	0.76	−0.74	0.09	0.49
Lanier — release	RMM	0.21	−0.67	0.75	0.89
	LSTM	0.58	3.50	0.55	0.65
	H1	0.54	−4.34	0.57	0.68
	H2	0.57	−0.43	0.55	0.65
	H3	0.55	−0.23	0.57	0.67
Lanier — storage	RMM	0.37	6.57	0.12	0.79
	LSTM	−8.87	−39.76	0.47	3.14
	H1	−11.01	44.79	0.52	3.47
	H2	0.47	5.29	0.11	0.73
	H3	0.56	4.03	0.10	0.66

Notes: NSE, Nash–Sutcliffe efficiency; PBIAS, percent bias; RRMSE, relative root mean squared error; RSR, RMSE-observation standard deviation ratio.

TABLE 2. Summary Statistics for FHV and FLV.

Reservoir	Simulation	PBIAS-FHV	PBIAS-FLV
Allatoona — release	RMM	−29.0	279.9
	LSTM	−25.9	114.3
	H1	−23.7	115.5
	H2	−11.4	109.2
	H3	−17.9	108.7
Lanier — release	RMM	−45.1	141.0
	LSTM	−24.1	69.3
	H1	−32.2	64.2
	H2	−22.4	60.5
	H3	−24.4	68.2

Notes: FHV, high flow volume; FLV, low flow volume.

"more data are usually better" for ML models, including additional input vectors may hinder the efficient training of LSTM. The role and importance of input sequences require for time-series prediction application and especially for long input sequences requires further exploration, which is also highlighted by Qin et al. (2017).

In terms of storage, while H1 showed a strong improvement than the standalone LSTM (Figures 3f and 4f), the same issue remains, that is, the errors in predicated release still accumulate through time and yield unreasonable long-term storage prediction. This physically inconsistent storage values can be due to the lack of mechanisms to regulate the flow conditions and/or enforce physical constraints in the overall reservoir storage.

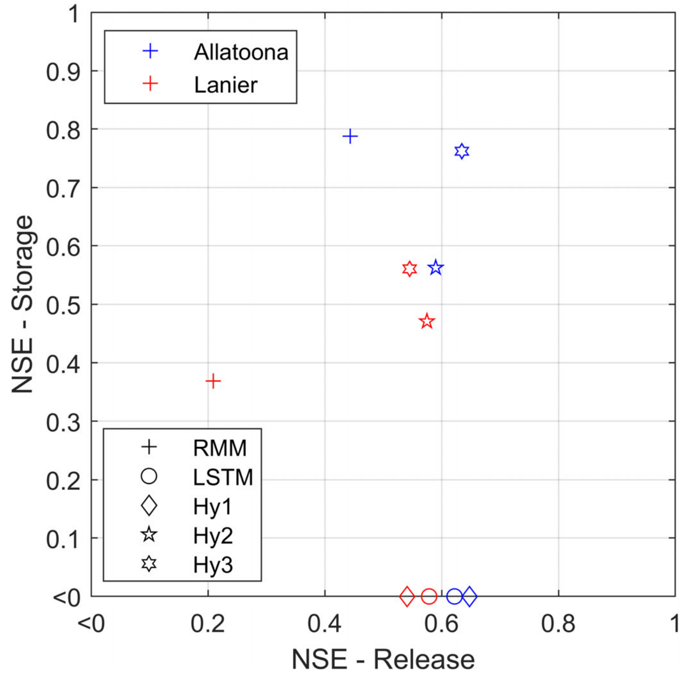

FIGURE 5. Summary of daily NSE for release and storage for Allatoona (blue) and Lanier (red) from all simulations during WY 1982–2015.

H2 and H3 Performance. The second hybrid strategy (which entails both H2 and H3 experiments) aims at using ML to provide an initial estimate of release. Such an estimate is then revised considering the overall storage, similar to how ML may be used by an operator to support decision making. In other words, both H2 and H3 releases were empirically adjusted to ensure physical consistency. As a result, the performance of storage prediction has been largely improved (Figures 3 and 4; Table 1). Based on the performance of release (NSE-r), we also observed that H2 and H3 can retain the improvements made by the standalone LSTM (Tables 1 and 2; Figure 5). Evaluation of Allatoona for both release and storage (i.e., daily mean NSE (NSE-avg = (NSE-r + NSE-s)/2) suggest that compared to the standalone RMM, H3 (NSE-avg = 0.695) outperform RMM (NSE-avg = 0.615) while H2 (NSE-avg = 0.575) may show some deterioration mainly due to storage trajectory). In case of Lanier, H2 (NSE-avg = 0.52) and H3 (NSE-avg = 0.55) both show an improvement over RMM only simulation (NSE-avg = 0.29).

Implications

We test different hybrid models (H1, H2, and H3) to capture reservoir dynamics with a focus on long-term water management applications. Our results show encouraging performance (H2 and H3) for two selected major multipurpose reservoirs under different historic operations. Although the findings are based a one selected RMM (DHSVM-Res), the approach is general and can also be applied to other alternative models (e.g., Gutenson et al. 2020; Wu et al. 2020). The extension of these techniques to other multipurpose reservoirs with diverse priorities (e.g., irrigation, hydropower, flood control, water supply, etc.) in various geographical regions will be required in the future studies. Also, we tested these approaches for the two selected reservoirs where historical observations of release and storage were available. Extension to reservoirs with even less data will remain a major challenge. Applications for cascading reservoir system or coordinated reservoirs will also require further consideration and evaluation. Although we show that hybrid strategy may perform satisfactorily for long-term water management simulations, the suitability of these approaches in nonstationary climate conditions or in reservoirs with changing operation rules, is yet to be explored.

The data-driven LSTM model is powerful in feature extraction and input–output relationship learning. Additionally, it considers the influence of temporal dependence of input sequence on the output prediction. As LSTM model is data hungry; additional, relevant, good-quality data usually can help in the learning process and allow more accurate and meaningful predictions. For instance, in this work, we can see that considering RMM simulation data as another set of LSTM inputs, improves the prediction performance. However, for reservoir operations, even with the most detailed operation rules, the operators can still exercise expert judgment and make adjustments depending on local conditions and prior experience that cannot be fully captured by rule-based models. Therefore, LSTM has the advantage to identify and mimic those site- and personal-specific features to improve the accuracy of the predictions based on the data.

Regarding the LSTM implementation, while we utilized a commonly used, off-the-shelf LSTM implementation in this study, other more sophisticated LSTM models may also be designed for the purpose of reservoir operation simulation. For instance, novel architectures such as Mass-Conserving LSTM (MC-LSTM; Hoedt et al. 2021) which allow the incorporation of physical constraints within the LSTM architecture may be suitable to simulate reservoir operation. Alternatively, other LSTM architectures utilizing robust physics guided loss functions and other physical mechanism to autoregulate the solution may also be leveraged to further constrain the solution. Lastly, we explored the application of LSTM as our ML technique in this study, however, applicability of other ML techniques such as reinforcement learning in lieu of LSTM may also be evaluated.

All-in-all, given the rapid development of ML methods in the recent decade, it is possible that a new ML model tailored for long-term water management simulation can be designed. Nevertheless, given the need of process-based understandings (especially for hydroclimate studies), our future ML models should also be explainable, which has been one biggest limitation of data-driven ML approaches. We believe that the hybrid modeling approach, effectively integrate physical modeling and ML, entails the strength of both standalone models remain one most efficient and credible modeling choice.

SUMMARY AND CONCLUSIONS

Reservoir management models can reasonably represent long-term reservoir dynamics (release and storage) in a hydrologic system, but their application is limited due to challenges associated with limited reservoir pertinent data availability and expertise necessary to accurately incorporate human decisions. Alternatively, ML techniques can efficiently learn reservoir operation patterns from historical data and perform reasonably well in the short-term forecasting. However, the pure ML-based technique may result in physically inconsistent outputs especially in the long-term hydroclimate simulations where errors may continue to accumulate due to lack of physical mechanisms. To overcome the individual challenges of both techniques, hybrid strategies may be suitable to leverage the best of both. Using both RMM and LSTM, this study evaluates three RMM-LSTM hybrid models against their standalone counterparts to simulate reservoir storage and release for the purpose of long-term water management simulations. For the two selected multipurpose reservoirs with different operations, we found that the hybrid approach can efficiently leverage the benefits of both RMM and LSTM and can robustly reproduce both release and storage at the multidecadal timescales.

While the applicability to other types of reservoirs in different geographical regions is to be tested, given ML's impressive capabilities in identifying hidden patterns from available historic observations, the broader applications can be expected. Since many reservoir-related information may never be openly released to the public, the ML-enabled modeling approach may be our best choice to simulate large-scale, regulated streamflow under the influence of multiple reservoir systems for the purpose of continental-scale streamflow forecasting, water resource planning, and hydroclimate projections.

DATA AVAILABILITY STATEMENT

The meteorological data used in the study are open access and can be obtained from Daymet website (https://daymet.ornl.gov/). The reservoir-related information is also in public domain gathered from USACE website and Duke University's Database (https://nicholasinstitute.duke.edu/reservoir-data/). The LSTM code is adapted from Kratzert et al. (2018). The RMM is adapted from DHSVM-RES reservoir module which can be obtained from contacting the authors Zhao et al. (2016). The SCE optimization algorithm was used from public MATLAB repository (https://www.mathworks.com/matlabcentral/fileexchange/7671-shuffled-complex-evolution-sce-ua-method) from Duan (2021).

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article: The Supporting Information includes details about RMM, LSTM, and experiments.

ACKNOWLEDGMENT

This research was supported by the AI initiative project through Oak Ridge National Laboratory (ORNL) and the ExaSheds project through the Office of Biological and Environmental Research in the Department of Energy (DOE) Office of Science. The research used resources of the Oak Ridge Leadership Computing Facility at ORNL, which is a DOE Office of Science User Facility. ORNL is managed by UT-BATTELLE for DOE under Contract DE-AC05-00OR22725. Accordingly, the U.S. government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. 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 U.S. Government purposes.

AUTHOR CONTRIBUTIONS

Sudershan Gangrade: Conceptualization; investigation; software; visualization; writing – original draft; writing – review and editing. Dan Lu: Conceptualization; investigation; software; writing – original draft; writing – review and editing. Shih-Chieh Kao: Conceptualization; funding acquisition; investigation; writing – original draft; writing – review and editing. Scott L. Painter: Conceptualization; funding acquisition; investigation; writing – original draft; writing – review and editing.

LITERATURE CITED

Brêda, J.P.L.F., R.C.D.D. Paiva, O.C. Pedrollo, O.A. Passaia, and W. Collischonn. 2021. "Modeling Coordinated Operation of Multiple Hydropower Reservoirs at a Continental Scale Using Artificial Neural Network: The Case of Brazilian Hydropower System." *Revista Brasileira de Recursos Hídricos* 26: 105–17.

Coerver, H.M., M.M. Rutten, and N.C. Giesen. 2018. "Deduction of Reservoir Operating Rules for Application in Global Hydrological Models." *Hydrology and Earth System Sciences* 22 (1): 831–51.

Dang, T.D., A. Chowdhury, and S. Galelli. 2020. "On the Representation of Water Reservoir Storage and Operations in Large-Scale Hydrological Models: Implications on Model Parameterization and Climate Change Impact Assessments." *Hydrology and Earth System Sciences* 24 (1): 397–416.

Duan, Q. 2021. "Shuffled Complex Evolution (SCE-UA) Method." https://www.mathworks.com/matlabcentral/fileexchange/7671-shuffled-complex-evolution-sce-ua-method.

Duan, Q., S. Sorooshian, and V.K. Gupta. 1994. "Optimal Use of the SCE-UA Global Optimization Method for Calibrating Watershed Models." *Journal of Hydrology* 158 (3–4): 265–84.

Ehsani, N., B.M. Fekete, C.J. Vörösmarty, and Z.D. Tessler. 2015. "A Neural Network Based General Reservoir Operation Scheme." *Stochastic Environmental Research and Risk Assessment* 30 (4): 1151–66. https://doi.org/10.1007/s00477-015-1147-9.

Ehsani, N., C.J. Vorosmarty, B.M. Fekete, and E.Z. Stakhiv. 2017. "Reservoir Operations under Climate Change: Storage Capacity Options to Mitigate Risk." *Journal of Hydrology* 555: 435–46. https://doi.org/10.1016/j.jhydrol.2017.09.008.

Gutenson, J.L., A.A. Tavakoly, M.D. Wahl, and M.L. Follum. 2020. "Comparison of Generalized Non-Data-Driven Lake and Reservoir Routing Models for Global-Scale Hydrologic Forecasting of Reservoir Outflow at Diurnal Time Steps." *Hydrology and Earth System Sciences* 24: 2711–29. https://doi.org/10.5194/hess-24-2711-2020.

Gangrade, S. 2019. Assessing Hydrologic Vulnerability and Resilience of Critical Energy-Water Infrastructures in a Changing Environment. PhD diss., University of Tennessee. https://trace.tennessee.edu/utk_graddiss/5779

Hoedt, P.-J., F. Kratzert, D. Klotz, C. Halmich, M. Holzleitner, G. Nearing, S. Hochreiter, and G. Klambauer. 2021. "MC-LSTM: Mass-Conserving LSTM." arXiv Preprint arXiv:210105186.

ICOLD. 2020. "World Register of Dams." https://www.icold-cigb.org/GB/world_register/general_synthesis.asp.

Jia, X., J. Willard, A. Karpatne, J.S. Read, J.A. Zwart, M. Steinbach, and V. Kumar. 2021. "Physics-Guided Machine Learning for Scientific Discovery: An Application in Simulating Lake Temperature Profiles." *ACM/IMS Transactions on Data Science* 2 (3): 1–26.

Khandelwal, A., S. Xu, X. Li, X. Jia, M. Stienbach, C. Duffy, J. Nieber, and V. Kumar. 2020. "Physics Guided Machine Learning Methods for Hydrology." arXiv Preprint arXiv:201202854.

Klipsch, J.D., and M.B. Hurst. 2013. *HEC-ResSim, Reservoir System Simulation User's Manual, Version 3.1*. US Army Corps of Engineers, Institute for Water Resources & Hydrologic Engineering Center (HEC).

Konapala, G., S.-C. Kao, S.L. Painter, and D. Lu. 2020. "Machine Learning Assisted Hybrid Models Can Improve Streamflow

Simulation in Diverse Catchments across the Conterminous US." *Environmental Research Letters* 15 (10): 104022.

Kratzert, F., D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger. 2018. "Rainfall-Runoff Modelling Using Long Short-Term Memory (LSTM) Networks." *Hydrology and Earth System Sciences* 22 (11): 6005–22. https://doi.org/10.5194/hess-22-6005-2018.

Lu, D., 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): 1421–38. https://doi.org/10.1175/Jhm-D-20-0082.1.

Ouyang, W., K. Lawson, D. Feng, L. Ye, C. Zhang, and C. Shen. 2021. "Continental-Scale Streamflow Modeling of Basins with Reservoirs: Towards a Coherent Deep-Learning-Based Strategy." *Journal of Hydrology* 599: 126455.

Patterson, L.A., and M.W. Doyle. 2018. "A Nationwide Analysis of U.S. Army Corps of Engineers Reservoir Performance in Meeting Operational Targets." *Journal of the American Water Resources Association* 54 (2): 543–64. https://doi.org/10.1111/1752-1688.12622.

Qin, Y., D. Song, H. Chen, W. Cheng, G. Jiang, and G. Cottrell. 2017. "A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction." arXiv Preprint arXiv:170402971.

Read, J.S., X.W. Jia, J. Willard, A.P. Appling, J.A. Zwart, S.K. Oliver, A. Karpatne, et al. 2019. "Process-Guided Deep Learning Predictions of Lake Water Temperature." *Water Resources Research* 55 (11): 9173–90. https://doi.org/10.1029/2019wr024922.

Shin, S., Y. Pokhrel, and G. Miguez-Macho. 2019. "High-Resolution Modeling of Reservoir Release and Storage Dynamics at the Continental Scale." *Water Resources Research* 55 (1): 787–810. https://doi.org/10.1029/2018wr023025.

Thornton, P.E., R. Shrestha, M. Thornton, S.-C. Kao, Y. Wei, and B.E. Wilson. 2021. "Gridded Daily Weather Data for North America with Comprehensive Uncertainty Quantification." *Scientific Data* 8 (1): 1–17.

USACE. 2015. "Master Water Control Manual, Alabama-Coosa-Tallapoosa (ACT) River Basin." https://www.sam.usace.army.mil/Missions/Planning-Environmental/ACT-Master-Water-Control-Manual-Update/ACT-Document-Library/.

USACE. 2017. "Master Water Control Manual, Apalachicola-Chattahoochee-Flint (ACF) River Basin." https://www.sam.usace.army.mil/Missions/Planning-Environmental/ACF-Master-Water-Control-Manual-Update/ACF-Document-Library/.

Voisin, N., H. Li, D. Ward, M. Huang, M. Wigmosta, and L. Leung. 2013. "On an Improved Sub-Regional Water Resources Management Representation for Integration into Earth System Models." *Hydrology and Earth System Sciences* 17 (9): 3605–22.

Wan, Z.Y., P. Vlachas, P. Koumoutsakos, and T. Sapsis. 2018. "Data-Assisted Reduced-Order Modeling of Extreme Events in Complex Dynamical Systems." *PLoS One* 13 (5): e0197704. https://doi.org/10.1371/journal.pone.0197704.

Wisser, D., S. Frolking, S. Hagen, and M.F. Bierkens. 2013. "Beyond Peak Reservoir Storage? A Global Estimate of Declining Water Storage Capacity in Large Reservoirs." *Water Resources Research* 49 (9): 5732–39.

Wu, J.W., H. Yen, J.G. Arnold, Y.C.E. Yang, X.M. Cai, M.J. White, C. Santhi, C.Y. Miao, and R. Srinivasan. 2020. "Development of Reservoir Operation Functions in SWAT plus for National Environmental Assessments." *Journal of Hydrology* 583: 124556. https://doi.org/10.1016/j.jhydrol.2020.124556.

Yang, S., 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: 124229. https://doi.org/10.1016/j.jhydrol.2019.124229.

Yang, T., 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): 1626–51. https://doi.org/10.1002/2015wr017394.

Yang, T., 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): 04019072.

Yassin, F., S. Razavi, M. Elshamy, B. Davison, G. Sapriza-Azuri, and H. Wheater. 2019. "Representation and Improved Parameterization of Reservoir Operation in Hydrological and Land-Surface Models." *Hydrology and Earth System Sciences* 23 (9): 3735–64.

Zagona, E.A., T.J. Fulp, R. Shane, T. Magee, and H.M. Goranflo. 2001. "Riverware: A Generalized Tool for Complex Reservoir System Modeling 1." *Journal of the American Water Resources Association* 37 (4): 913–29.

Zhang, D., 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: 720–36.

Zhang, D., Q.D. Peng, J.Q. Lin, D.S. Wang, X.F. Liu, and J.B. Zhuang. 2019. "Simulating Reservoir Operation Using a Recurrent Neural Network Algorithm." *Water* 11 (4): 865. https://doi.org/10.3390/w11040865.

Zhao, G., H.L. Gao, B.S. Naz, S.C. Kao, and N. Voisin. 2016. "Integrating a Reservoir Regulation Scheme into a Spatially Distributed Hydrological Model." *Advances in Water Resources* 98: 16–31. https://doi.org/10.1016/j.advwatres.2016.10.014.