



Long short-term memory models to quantify long-term evolution of streamflow discharge and groundwater depth in Alabama



Hossein Gholizadeh ^a, Yong Zhang ^{a,*}, Jonathan Frame ^b, Xiufen Gu ^c, Christopher T. Green ^d

^a Department of Geological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

^b Floodbase, New York, NY 10024, USA

^c School of Mathematics and Information Science, Yantai University, Yantai, Shandong 264005, China

^d U.S. Geological Survey, Water Resources Mission Area, Moffett Field, CA 94035, USA

ARTICLE INFO

Editor: Christian Herrera

Keywords:

Long short-term memory
Hydrographs
Groundwater depth
Long-term evolution

ABSTRACT

Long short-term memory (LSTM) models have been shown to be efficient for rainfall-runoff modeling, and to a lesser extent, for groundwater depth forecasting. In this study, LSTMs were applied to quantify the spatiotemporal evolution of surface and subsurface hydrographs in Alabama in the Southeastern United States, where water sustainability has not been fully quantified across spatiotemporal scales. First, the surface water LSTM model with extensive dynamic (precipitation and other weather variables) and static (basin characteristics) inputs predicted the main characteristics of streamflow for six years at 19 gauged basins in Alabama. The model tended to underestimate extremely high streamflow but adding drainage density as an input feature slightly improved the predictions of extreme events. Second, to predict the groundwater depth evolution, a groundwater LSTM (GW-LSTM) model was proposed and applied using static inputs capturing the aquifers' hydrogeological properties and dynamic inputs of meteorological information. Three precipitation scenarios were also explored to evaluate the groundwater hydrograph evolution in the next two decades. The GW-LSTM model predicted the general trend of daily groundwater depth fluctuations (at 21 wells distributed across Alabama from 1990 to 2021) including most extremely high groundwater levels, and recovered groundwater depth for locations withheld from model training and validation. This study, therefore, extended the application of LSTMs in quantifying the spatiotemporal evolution of surface water and groundwater, two manifestations of a single integrated resource.

1. Introduction

Long short-term memory (LSTM) is an artificial neural network (ANN) proposed by Hochreiter and Schmidhuber (1997) to improve the standard feedforward neural networks by capturing a short-term memory for a recurrent neural network that can last many timesteps. LSTM's application to hydrology has been expanding rapidly. LSTM has also been applied to obtain satisfactory time series predictions in many other scientific areas (Gers et al., 2002; Zaytar and El Amrani, 2016; Ma et al., 2015; Shi et al., 2015; Song et al., 2020; among many others), and is becoming a focus of deep learning, according to the extensive review by Yu et al. (2019). LSTM also exhibits remarkable success in predicting rainfall-runoff response (Kratzert et al., 2018a, 2018b, 2019, 2022; Gauch et al., 2021; Frame et al., 2021; Nearing et al., 2022; Gholizadeh et al., 2022), probably due to its promising capability in handling both

short-term impulses and long-term dependencies embedded in many hydrologic processes. For example, Nevo et al. (2022) used LSTM for Google's operational flood forecasting system, which is a reliable source for flood extent prediction (Hamidi et al., 2022). Boulmaiz et al. (2020) found that LSTM only required a relatively small amount of data to "learn" the behavior of rainfall–runoff relationship. Recently, Frame et al. (2022) showed that, although LSTM remains relatively accurate in predicting extreme events compared to both a conceptual model (i.e., the Sacramento Model developed by Burnash et al. (1973)) and a process-based model (i.e., United States (U.S.) National Water Model (<https://water.noaa.gov/about/nwm>) developed by National Oceanic and Atmospheric Administration (NOAA)), it still underpredicts extremely high events. Further efforts are still needed to explore LSTM's modeling capabilities and accuracy for surface water flows especially for extreme events, which motivated this study.

* Corresponding author.

E-mail address: yzhang264@ua.edu (Y. Zhang).

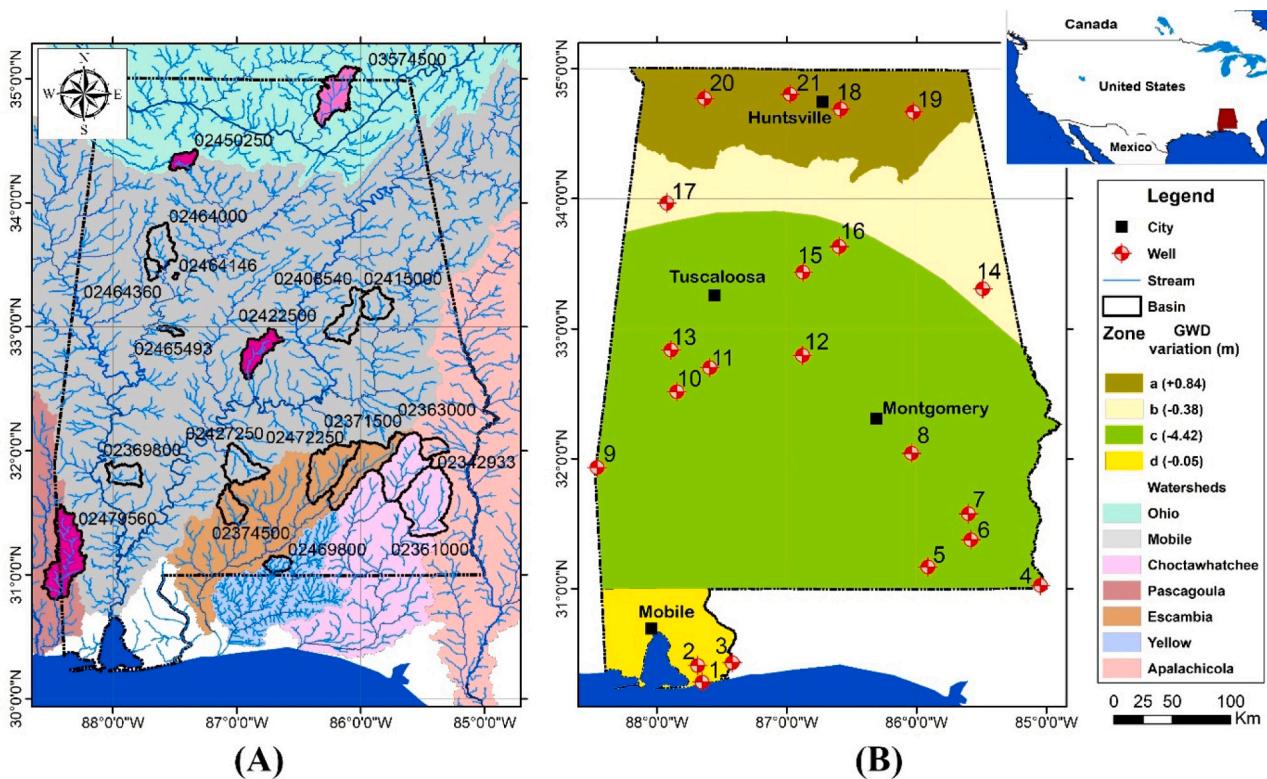


Fig. 1. Study area – Alabama state map with 19 gauged basins adapted from CAMELS dataset (the four pink filled polygons are four representative basins) (A) and 21 wells (B). Four areas with different colors (a, b, c, and d) shown in plot (b) are four different zones categorized in this study by the average groundwater depth (GWD) variations during the last 30 years. A positive GWD variation means rising of groundwater head, and a negative one represents falling groundwater head in the last 30 years. The number (1 to 21) in plot (B) near each well represents the well ID, and the number near each basin in plot (A) represents the basin CAMELS ID (the base map is adapted from USGS).

Table 1
List of LSTM model parameters.

No	LSTM parameters	Value
1	Number of LSTM layers	2
2	cell/hidden state length	20
3	Initial forget bias	3
4	Dropout rate	0.4
5	Learning rate	0.01
6	Batch size	256
7	Optimizer	Adam
8	Number of training epochs	50
9	Sequence length	365

LSTM has also been recently applied to quantify groundwater flow, although to a lesser extent. Groundwater sustainability has been seriously challenged worldwide by various factors including climate change, agricultural and industrial contamination, and over exploitation (Bouwer, 2000; Gleeson et al., 2012). Many studies have been focused on modeling regional/continental scale groundwater flow (Zhou and Li, 2011; Maxwell, 2013; Maxwell et al., 2015; Maxwell and Condon, 2016; De Graaf et al., 2020; Gholizadeh et al., 2020; Condon et al., 2021; Wu et al., 2021) and forecasting groundwater levels using ANNs (Nayak et al., 2006; Banerjee et al., 2009; Yoon et al., 2011; Adamowski and Chan, 2011; Sahoo et al., 2017; Bowes et al., 2019; Rahman et al., 2020; Ahmadi et al., 2022). Recently, Tao et al. (2022) applied LSTMs to predict groundwater level. Zhang et al. (2018) developed an LSTM based model for predicting water table depth in agricultural areas in northern China. Chen et al. (2020) simulated groundwater dynamics in north-western China using various machine learning algorithms. Most of these studies did not consider the impacts of aquifer physical properties, such

Table 2
Inputs of the surface water LSTM models.

Surface water LSTM inputs	Type	Source
Precipitation	Dynamic	CAMELS (https://gdex.ucar.edu/dataset/camels.html)
Solar radiation		
Maximum temperature		
Minimum temperature		
Vapor pressure		
Mean basin elevation	Static	
Mean basin slope		
Basin area		
Forest fraction		
Maximum leaf area index (LAI)		
Minimum LAI		
LAI difference		
Maximum green vegetation fraction (GVF)		
Minimum GVF		
Soil depth		
Soil porosity		
Soil conductivity		
Maximum water content		
Sand fraction		
Silt fraction		
Clay fraction		
Carbonate rocks fraction		
Permeability		
Mean potential evapotranspiration (PET)		
Aridity		
Snow fraction		

Table 3
Inputs of the GW-LSTM models.

GW-LSTM inputs	Type	Source
Precipitation	Dynamic	NOAA (https://www.ncei.noaa.gov/cdo-web/datasets)
Temperature	Static	CAMELS (https://gdex.ucar.edu/dataset/camels.html)
Soil depth		
Porosity		
Maximum water content		
Leaf area index (LAI)		
Mean potential evapotranspiration (PET)		
Sand fraction		
Silt fraction		
Clay fraction		
Carbonate rocks fraction		
Hydraulic conductivity	Static	GSA (https://www.gsa.state.al.us/gsa/groundwater/wellrecords)

as conductivity, porosity, and soil properties, on groundwater depth change. Wunsch et al. (2021) compared LSTM, convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input for groundwater level forecasting, and they found that LSTMs can perform substantially better with longer datasets, for example four years or more. Afzaal et al. (2019) employed LSTM for groundwater estimation using dynamic physical hydrology inputs such as precipitation, temperature, etc. Vu et al. (2021) used LSTM for interpolating groundwater levels, but their only input was groundwater level data. Many of the previous studies used the previous time step of groundwater level data, which made the LSTMs autoregressive models due to the lack of inputs such as hydraulic conductivity and soil properties. Therefore, it

is necessary to evaluate the potential impact of typical aquifer properties (which may dominate groundwater dynamics) on LSTM modeling of groundwater level time series, so that one can better and reliably expand LSTM to compensate the traditional process-based, computational-demanding numerical models for simulating and predicting subsurface flow. In addition, prediction of groundwater depth evolution at locations without recorded data (i.e., ungauged setting), which is practically important, has not been addressed in the previous studies.

This work aims to fill the above-mentioned knowledge gaps and continue to expand LSTM applications for quantifying surface and subsurface hydrographs in three ways. First, we propose to apply catchment drainage density (defined as the total catchment stream length divided by the catchment area) data to improve runoff discharge in extreme events (higher than a certain streamflow discharge based on each catchment recorded data) calculated by LSTM for catchments across the state of Alabama. Second, we expand the application of LSTM by adding fundamental static/dynamic inputs (including precipitation, temperature, hydraulic conductivity, maximum water content, soil porosity, and soil depth data) to predict groundwater depth, and this expanded LSTM is called “GW-LSTM” (where “GW” stands for groundwater) in this study. This GW-LSTM model follows the ParFlow-ML model proposed by Tran et al. (2021), who added hydraulic conductivity and topography in LSTM to predict ParFlow solutions of pressure heads (notably, Tran et al. (2021) considered only ParFlow model results in a short period (~5 days) which makes it simply a surrogate model for ParFlow. Surrogate models are often developed to reduce computational cost of complex physical models, but they are then subject to the same predictive limitations of those physical models. Our model is trained to predict real data, and we consider more general conditions with real-world data over

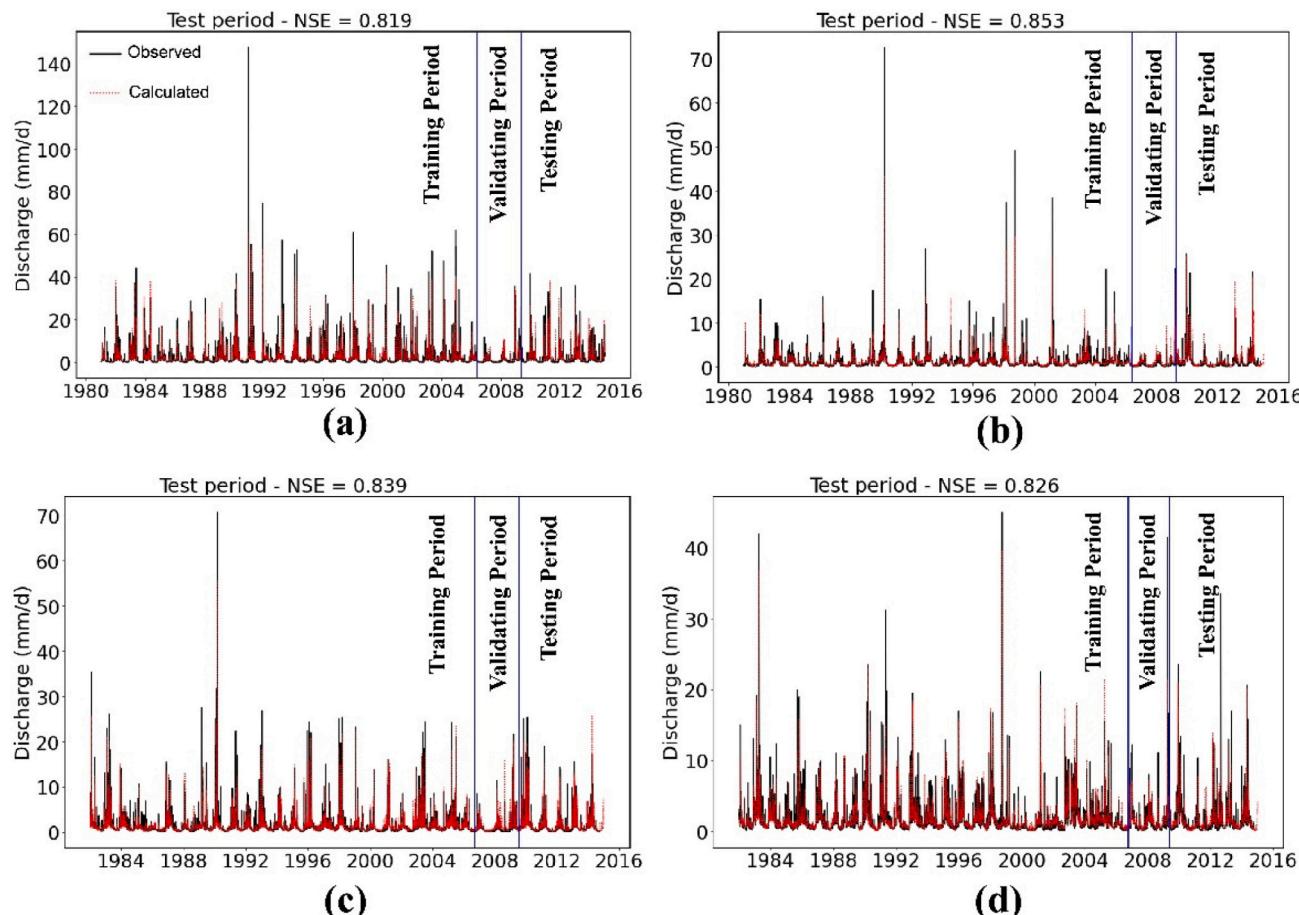


Fig. 2. Streamflow discharge calculated by the LSTM model (i.e., the Neural Hydrology network) (red lines) versus the observed hydrograph (black lines) at four representative basins a, b, c and d (basin ID: 03574500, 02450250, 02422500 and 02479560) shown by the four pink filled polygons in Fig. 1A.

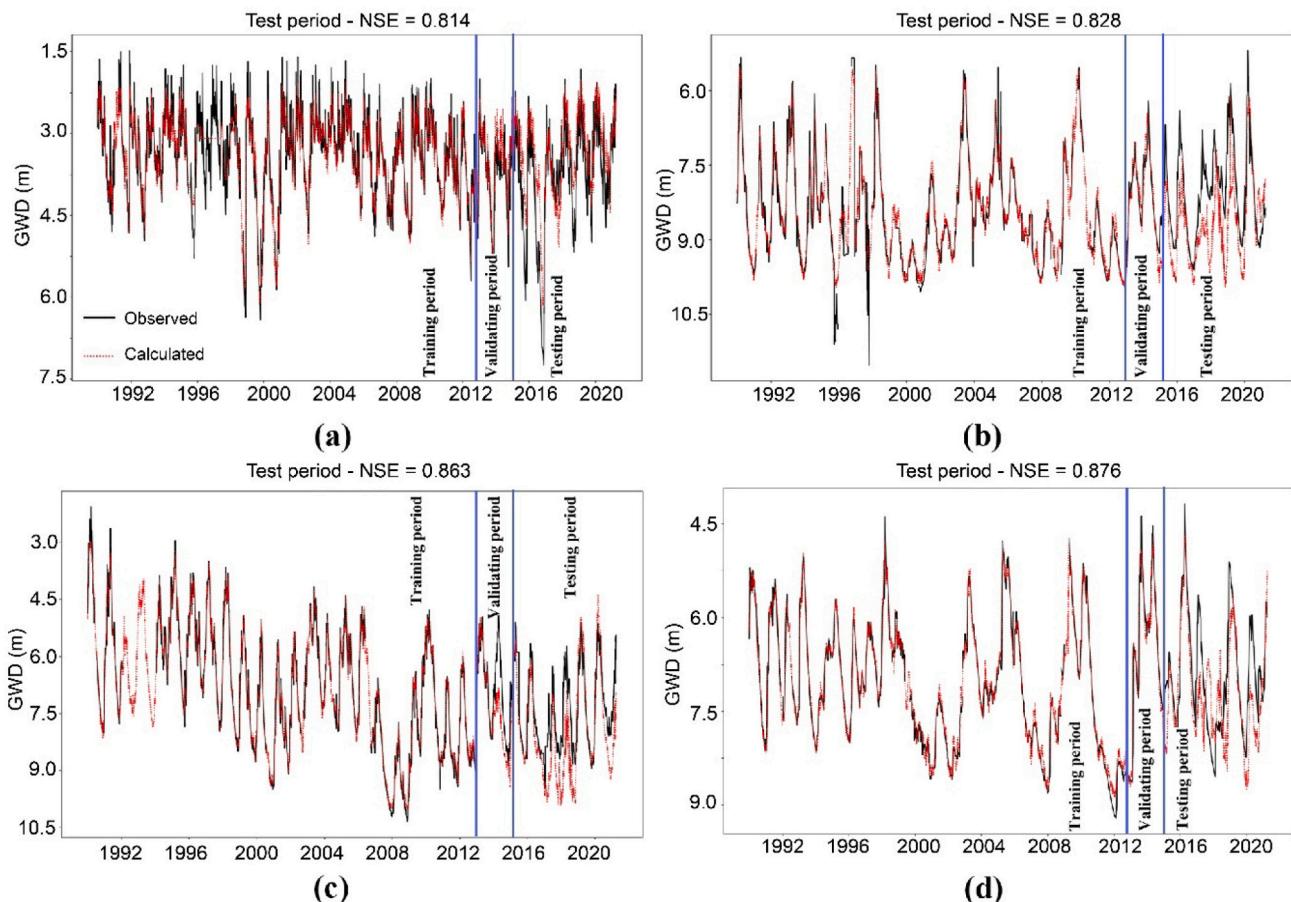


Fig. 3. Time series data of the groundwater depth (GWD) below the ground surface: the GW-LSTM model results (red lines) versus the observed time series data (black lines) for 4 representative wells in Alabama. Wells a, b, c, and d are wells No. 18, 17, 12, and 2 shown in Fig. 1B.

decades in this study). Third, we evaluate the GW-LSTM for predicting groundwater depth in wells with historical data, and at locations where there are no groundwater level measurements to train the model. Both surface water and groundwater (located in the same catchment or the most adjacent catchment under the similar meteorological condition) are considered in this study, since surface water and groundwater form two manifestations of a single resource (Winter, 1999). We evaluate the potential discrepancy in the LSTM modeling applicability and extreme event responses in surface and subsurface dynamics by comparing the LSTM predictions of both surface and subsurface hydrographs for the same basin and precipitation inputs.

The rest of this article is organized as follows. Section 2 introduces the methodology and study site. Surface and subsurface water resources are evaluated by surface water LSTM model and GW-LSTM model using the Neural Hydrology (NH) Python library (Kratzert et al., 2018a). Three scenarios of future precipitation time series are then considered to assess the spatiotemporal evolution of groundwater depth variation. Section 3 shows the application results of the two LSTM models. Section 4 discusses the ability of surface water LSTM and GW-LSTM to predict the time series of streamflow discharge and groundwater depth, respectively, in various basins in Alabama. The importance of each input factor in the GW-LSTM model is also ranked, and both the improved surface water model (with drainage density as an additional input LSTM) and the GS-LSTM model behaviors are assessed in quantifying extreme events. Section 5 presents the main conclusions. Section A in Supplementary Material (SM) briefly reviews LSTM. Section B in SM shows the additional information for wells and basins in Alabama used in the main text, as well as the LSTM modeling results for the basins not listed in the

main text. Section C in SM applies wavelet analysis to assess the response of groundwater level fluctuations to precipitation and sea level changes, to support the LSTM results shown in the main text.

We emphasize again that LSTM has been widely applied by various researchers in hydrology and other research areas. This study aims to add three more components to the LSTM models and improve LSTM applications in hydrology. First, drainage density is added to the surface water LSTM model as a physiographic characteristic factor, resulting in a LSTM model called SW-DD-LSTM. As shown below, this new input slightly improves the model performance and can guide future efforts to predict extreme events of streamflow. Second, a groundwater LSTM model, GW-LSTM, is proposed using core hydrogeologic inputs (missed by many previous applications) including hydraulic conductivity and core soil properties that dominate subsurface response to surface recharge. Third, our LSTM models improve previous autoregression LSTM models in predicting groundwater depth at locations without recorded data.

2. Study site and LSTM models

2.1. Surface water analysis using the Neural Hydrology (NH)

NH is a Python library developed by Kratzert et al. (2018a) and is used in this study to calculate runoff for 19 gauged basins throughout Alabama with sufficient (30 years period of data) recorded streamflow data (Fig. 1A). The catchment information was obtained from the Catchment Attributes and Meteorological Large Sample (CAMELS) dataset presented by the U.S. National Center for Atmospheric Research

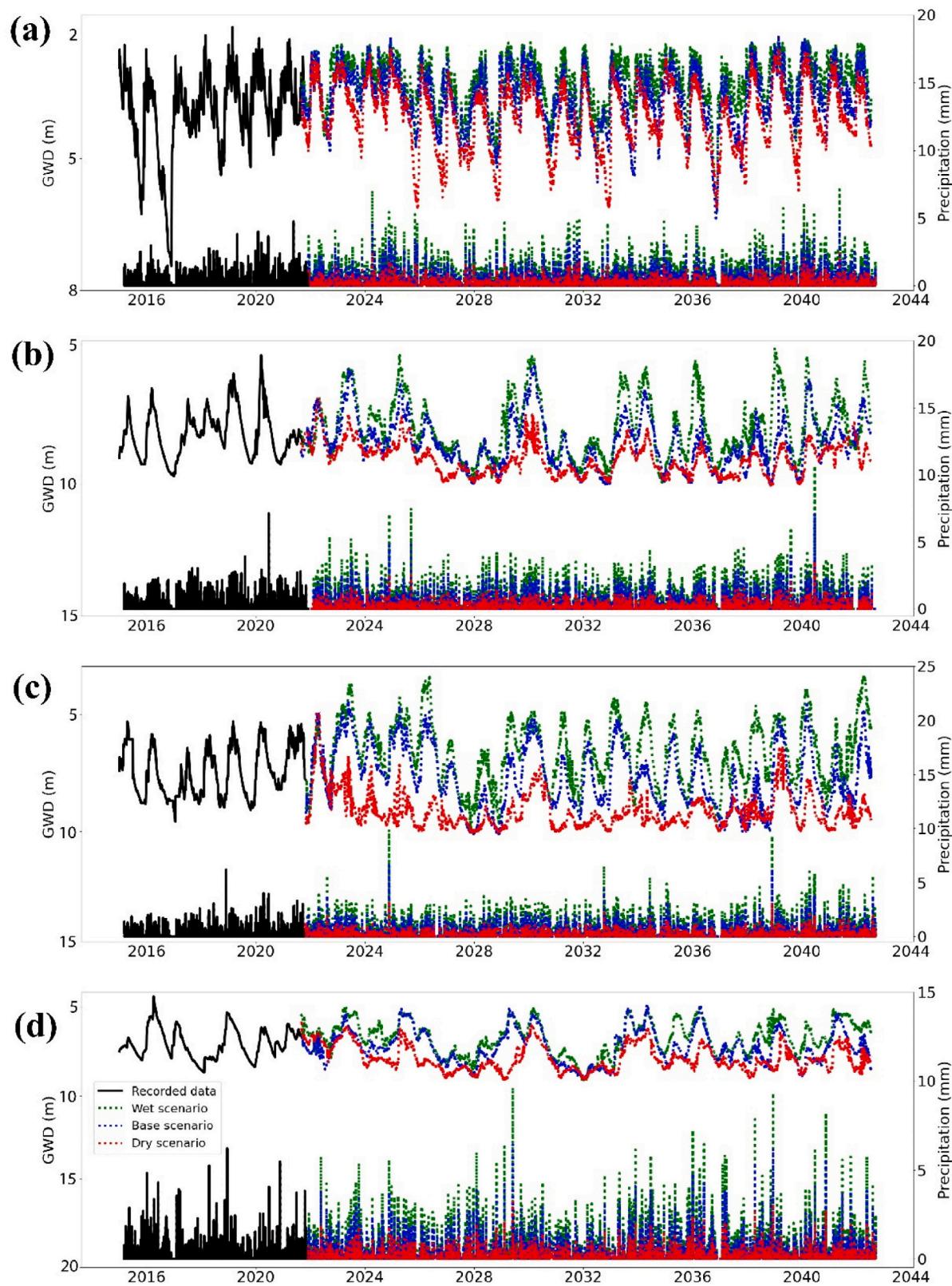


Fig. 4. Response of groundwater depth (from the land surface) to future climate change: evolution of GWD predicted by the GW-LSTM model for the next twenty years under three climate scenarios for zone (a), (b), (c), and (d) (wells No. 18, 17, 12, and 2). In each plot, the bottom curves represent precipitation for the three climate scenarios, and the top curves show the predicted GWD.

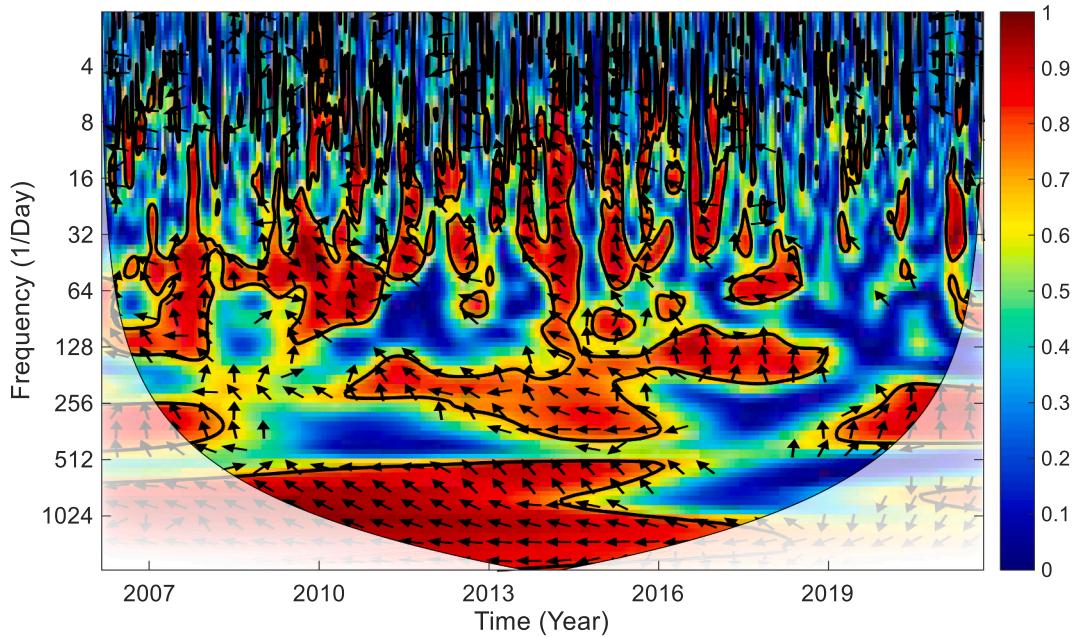


Fig. 5. Wavelet coherence spectrum between precipitation and groundwater depth for well #1 shown in Fig. 1B. The thick black contour designates the 95% confidence level. Bivariate wavelet coherence (BWC) outside the solid circular line cannot be used in the analysis due to the boundary effect. Arrows indicate the relative phase relationship: arrows pointing to the right exhibit the same phase, and arrows pointing to the left show antiphase. Shading indicates the strength of coherence. The color bar (with the legend) denotes the correlation coefficient of the two time series data.

(NCAR) (Newman et al., 2014; Addor et al., 2017), which contains the hydrometeorological time series and attributes for 671 basins in the U.S. The surface water LSTM model uses four types of data, which are the daily precipitation, solar radiation, the maximum and minimum temperatures, and vapor pressures, as dynamic inputs in LSTM modeling. Static inputs include (long-term) catchment attributes and climate data (21 total input variables), including the mean basin elevation, mean basin slope, basin area, forest fraction, maximum leaf area index (LAI), minimum LAI, LAI difference, maximum green vegetation fraction (GVF), minimum GVF, soil depth, soil porosity, soil conductivity, maximum water content, sand fraction, silt fraction, clay fraction, carbonate rocks fraction, permeability, mean potential evapotranspiration (PET), aridity, and snow fraction. Hence, the static inputs in LSTM built in CAMELS consider a wide range of attributes that may affect catchment behavior and the corresponding hydrological processes (i.e., runoff response to rainfall), providing a relatively complete dataset (using only the Alabama CAMELS sites) for this study.

To evaluate the modeling results, Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) is used (as well as root-mean-squared-error (RMSE)). NSE is the ratio of the error variance of the modeled time-series and the variance of the observed time-series subtracted from one:

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_0^t - Q_m^t)^2}{\sum_{t=1}^T (Q_0^t - \bar{Q}_0)^2},$$

where \bar{Q}_0 denotes the mean observed streamflow, Q_m^t and Q_0^t denote the modeled and observed streamflow at time t , respectively, and T represents the total time. For a perfect model with a zero-estimation error variance, the resulting NSE equals 1. Therefore, NSE values closer to one suggest that the model fitting/prediction matches better the real data. Often, a $NSE > 0.8$ is used as a criterion for acceptable model predictions (Moriasi et al., 2015).

In this study, we used a LSTM architecture consisting of three LSTM layers. Each of the first two layers has 20 hidden state lengths, and the third layer is a dense one connecting LSTM output at the last time step to

a single output neuron with linear activation. Another hyperparameter is the length of the input sequence, which defines the number of days of meteorological input data for LSTM to predict the next streamflow value. We kept the sequence length constant at 365 (days) and assumed a full year's worth of seasonality is the longest memory that the LSTM needs, so that we can capture a full annual cycle. To predict the daily streamflow and groundwater depth, we provide the last 365 time steps of meteorological observations as LSTM inputs. The initial learning rate is 1×10^{-2} , and the neural network is trained for 50 epochs to minimize the RMSE. The hyperparameters adopted in this study are similar to those tuned by Kratzert et al. (2019). The LSTM model parameters are reported in Table 1.

In our LSTM model, we added dropout between layers. This technique can prevent the model from overfitting (Srivastava et al., 2014). Particularly, dropout sets a certain percentage of random neurons to zero during training to force the network into a more robust feature learning. To avoid the effect of overfitting of the network on the training data, we identified the number of epochs for having the highest NSE for the validation period. Our preliminary experiments showed that the highest mean NSE for training period was achieved after 50 epochs.

The main source of uncertainty in the LSTM modeling is the optimizer type. We used the Adam optimizer (Kingma and Ba, 2014), one of the strong optimization algorithms, for training time series data. In addition, like the other data-based models, the data uncertainty also relates to the reliability of the quality and quantity of the data used to train the LSTM model. We selected reliable databases, such as CAMELS which contains basin scale hydrometeorological forcing data for 671 basins in the United States Geological Survey's Hydro-Climatic Data Network 2009 (HCDN-2009) (Lins, 2012) conterminous U.S. basin subset to achieve the best performance. We also normalized the streamflow discharge data by dividing the streamflow discharge of each basin by its area to avoid possibly incorrect pattern learning of the model. In addition, model evaluations performed for ungauged settings (as one of the model application extensions) could obviate the data uncertainty.

Because we obtained streamflow data from CAMELS and groundwater level data from the Geological Survey of Alabama (GSA) well

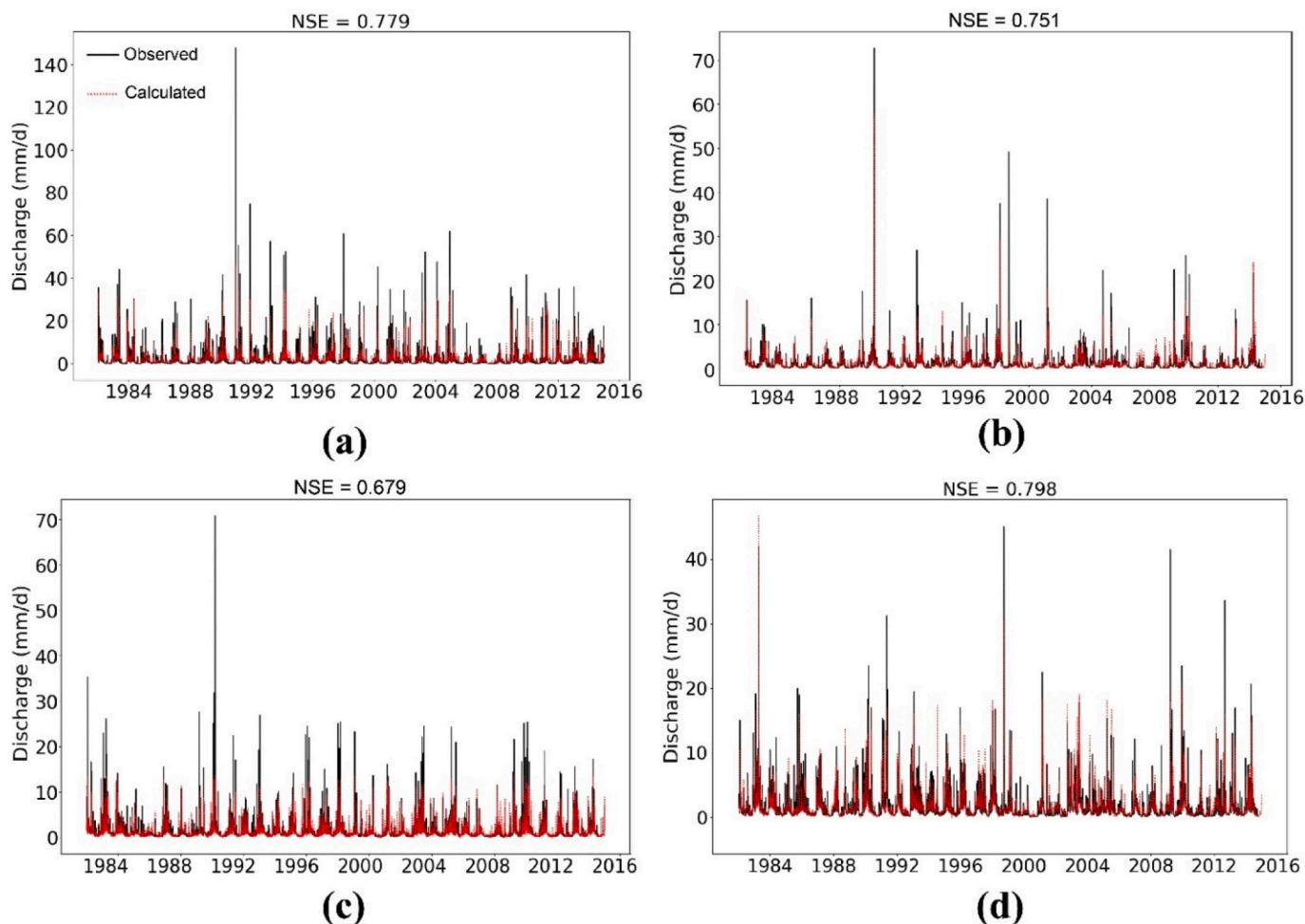


Fig. 6. Streamflow discharge prediction by the surface water LSTM model for basins held out from training in four different zones in Alabama (marked as pink polygons in Fig. 1A).

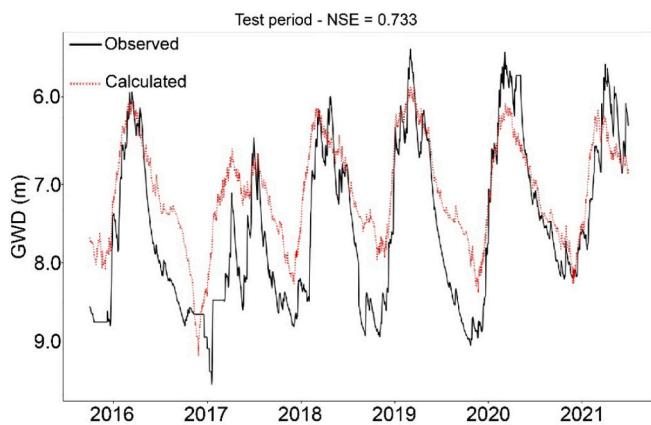


Fig. 7. Application of GW-LSTM for predicting GWD for well held out from training: GWD predictions using the GW-LSTM model for well #15 (marked in Fig. 1B) without observed data for training and validating.

records database (<https://www.gsa.state.al.us/gsa/groundwater/well-records>), the main limitation of the LSTM model is the period of the time series data. The hydraulic properties of aquifers are another limitation due to their complexity, although we employed local borehole data as the updated hydrogeological data for the selected sites. As shown below, the groundwater models developed in this study can lead to reliable

Table 4

Importance ranking (from high to low) of factors affecting the surface water LSTM model and the GW-LSTM model outputs.

Ranking	Surface water LSTM	GW-LSTM
1	Precipitation	Precipitation
2	Permeability	Hydraulic conductivity
3	Maximum temperature	Soil depth
4	Solar radiation	Maximum water content
5	Vapor pressure	Leaf area index (LAI)
6	Maximum water content	
7	Minimum temperature	
8	Soil depth	
9	Carbonate rocks fraction	
10	Forest fraction	

predictions with a NSE value typically above 0.8.

We trained a single LSTM model using all sites' data for both surface water and GW models because ungauged predictions can only be made when using a single model for all sites, rather than individual LSTMs. This gives us an LSTM model that has learned a general process representation that can be applied in locations without data. We trained the LSTM model using two types of inputs: dynamic and static inputs for surface water and groundwater LSTM models (Table 2 and Table 3). In the LSTM model, each iteration step in training works with a batch of available training data. The number of samples per batch was defined as 256, a common size used by other researchers (e.g., Srivastava et al., 2020; Yin et al., 2021). Each of these samples consists of one target value

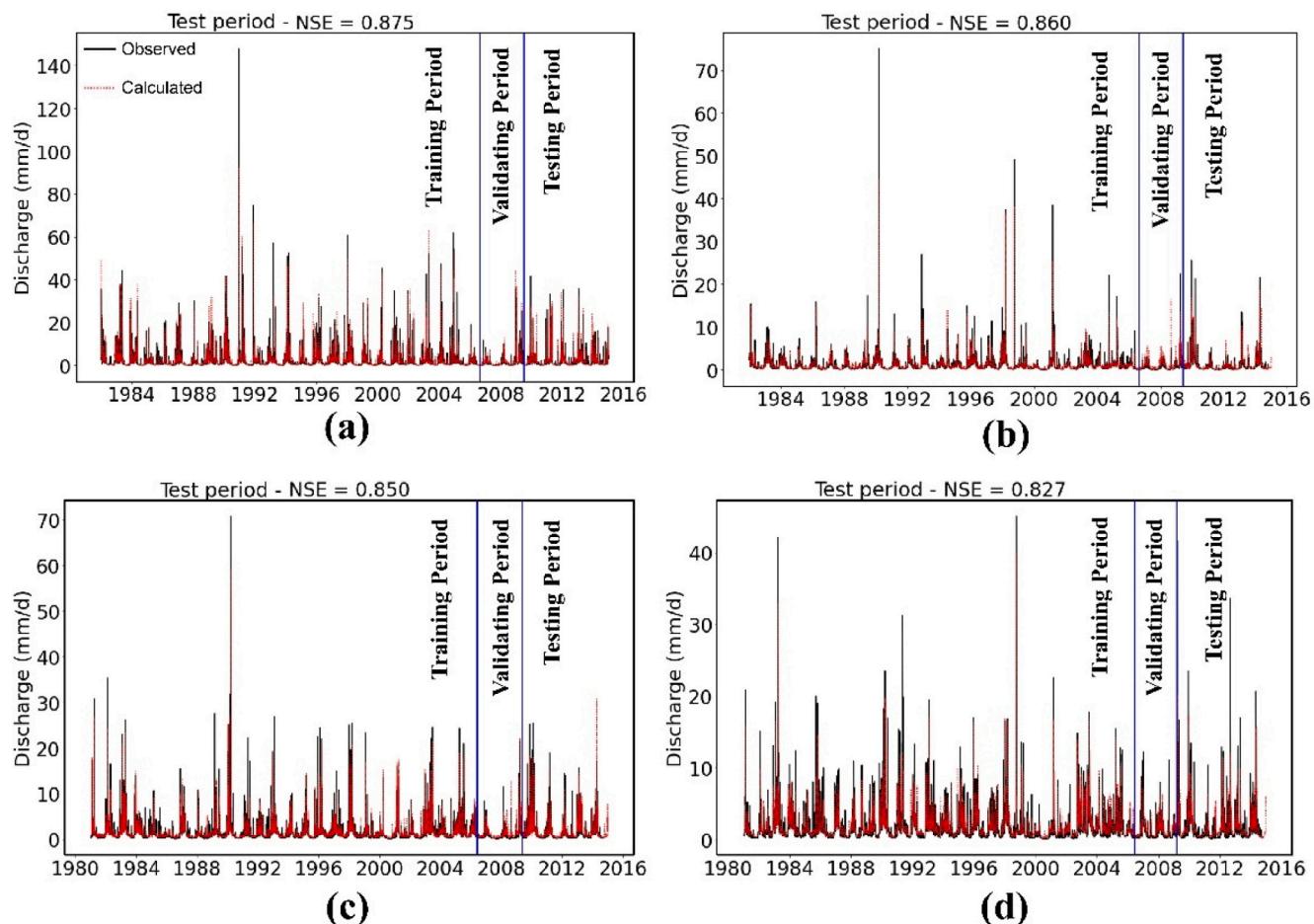


Fig. 8. Application of SW-DD-LSTM at four representative basins (marked in Fig. 1B): Streamflow discharge calculated by the SW-DD-LSTM model compared to the observed data.

of a given day and the dynamic input of the n preceding days. The model calculates the average of the RMSE between the simulated and observed values as the loss function in every iteration step for these 256 samples. The best performance for the training set was achieved after 50 epochs, where epoch is the period in which each training sample is used once for updating the model parameters.

Alabama has a humid subtropical climate with an average annual temperature of 18 °C. The southern part of the state is warmer than the northern part. The average annual precipitation of the state is 1400 mm (adopted from NOAA), and ~10% of the precipitation enters the subsurface and recharges aquifers (adopted from Geological Survey of Alabama, Alabama's Water, Educational Series 11, <https://www.gsa.state.al.us/gsa/groundwater/waterinfo>). There are four major physiographic provinces in Alabama, including 1) Appalachian and interior low plateaus in the north that consist of Ordovician limestone, 2) Valley and Ridge in central Alabama that are formed by sandstone, chert beds, shale, and carbonated rocks, 3) Piedmont in the eastern part that consists of metamorphic rocks including low-grade greenschist and migmatite facies, and 4) the coastal plain of Alabama in the southern and western parts of the state formed by Mesozoic and Cenozoic unconsolidated sediments (Ebersole et al., 2019). Based on the report of Alabama Department of Economic and Community Affairs (<https://adeca.alabama.gov/water-management/publications-reports-and-information/>), water use in Alabama was about 8239 million gallons per day (Mgal/day) in 2015. Total surface-water withdrawals were about 7743 Mgal/day (94% of the total withdrawn) and the remaining 496 Mgal/day were from groundwater. The daily meteorological and streamflow time series data for the last 30 years and catchment attributes are taken from NCAR.

The daily meteorological and streamflow time series data for insert data range (30-year time period) and catchment attributes are from NCAR (<https://rda.ucar.edu/>). The daily groundwater level data of insert date range (30-year time period) and aquifer properties are taken from GSA, which assesses groundwater in Alabama with real-time and periodic monitoring programs (<https://www.gsa.state.al.us/gsa/groundwater>).

2.2. Groundwater resources and GW-LSTM model

We extended the classical LSTM applications by developing the GW-LSTM model for exploring the spatiotemporal evolution of groundwater depth in Alabama. The dynamics data of groundwater depth at 21 wells with complete data for 30 years distributed across Alabama (marked in Fig. 1B) were acquired from GSA (see above). Dynamic inputs in GW-LSTM including daily precipitation and temperature data were obtained from NOAA. In addition, static inputs in GW-LSTM include the long-term average of PET, LAI, and the major aquifer and soil hydrological properties including hydraulic conductivity, porosity, maximum water content, and soil depth which connect precipitation and groundwater recharge. These inputs were obtained from NCAR (CAMELS) and GSA's groundwater assessment program (described also above).

3. Model results

3.1. Surface water hydrograph evaluation

We used 80% of the hydrograph data for training (75%) and validation (5%) and kept the remaining 20% for testing, a common split

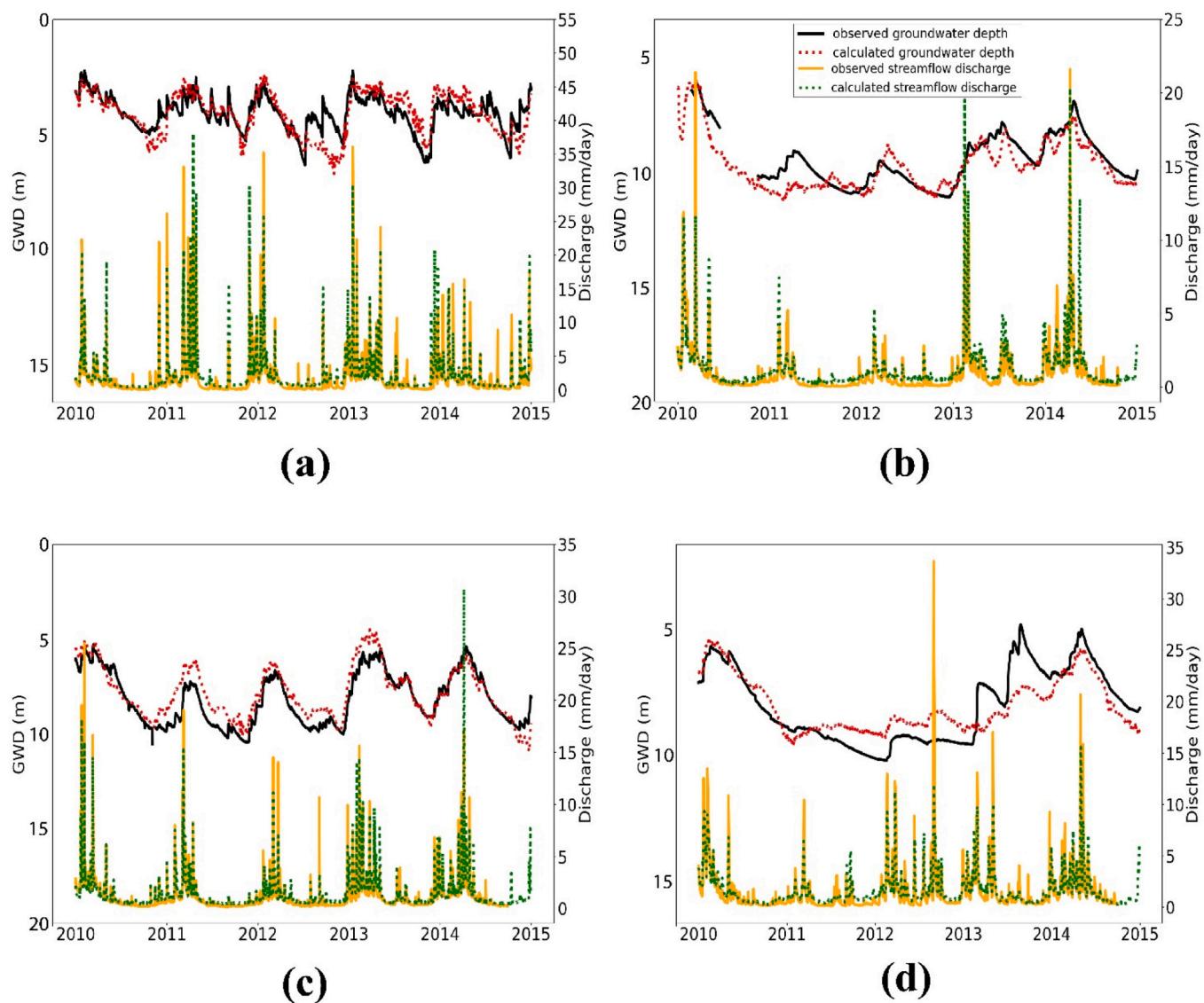


Fig. 9. Comparison between surface and subsurface outputs in extreme events for four basins and wells marked in [Fig. 1](#).

used for LSTM (Géron, 2022). In the training stage, the weights and biases in LSTM are updated based on the given loss function of each iteration step, where the mean-squared error (MSE) is used as an objective criterion. The loss (between the observed and model predicted discharges) is then calculated and used to update the model parameters. The surface water LSTM model uses the Adam optimization algorithm (Kingma and Ba, 2014) to tune the weights and biases to improve the model performance. For illustration purposes, four basins were selected (marked in [Fig. 1A](#) by pink-filled polygons), one from each of the geographic zones, distinguished by varying trends of groundwater depth (GWD) over the last 30 years (discussed in [Section 3.3](#)). The observed and LSTM modeled hydrographs for these four basins are plotted in [Fig. 2](#). The LSTM solutions for the other 15 basins are shown in Appendix B. These results show that the LSTM model captures the overall pattern of all the observed hydrographs. Most of the calculated NSEs for the testing period only are >0.8 ([Fig. 2](#)), and hence the LSTM model can reasonably predict hydrographs in Alabama's gauged basins after training and validation.

3.2. Groundwater depth prediction by the GW-LSTM model

To predict the temporal evolution of groundwater depth, the GW-

LSTM model first uses groundwater depth time series data from 1990 to 2013 for training, and from 2013 to 2015 for validation. The testing period is selected as the subsequent 5 years (2015–2021) to incorporate possible climate change effects in recent years. The sequence length is 365, and we set 50 training epochs in the GW-LSTM model. Twenty-one wells (introduced in [Section 2.2](#) and marked in [Fig. 1B](#)) with long period data and distributed across Alabama are selected for checking the applicability of the GW-LSTM model. Results for four representative wells (corresponding to each of the four basins plotted in [Fig. 1](#)) are shown in [Fig. 3](#). The GW-LSTM model results show an overall close match with the observed GWD ($\text{NSE} > 0.8$), indicating that this model can likely fill in missing observations over short periods.

3.3. Spatiotemporal analysis of groundwater depth response to climate change

As mentioned above, based on the historical groundwater depth data and their trends, we delineated 4 zones (denoted as Zone A, B, C, and D) for the 19 gauged basins in Alabama (shown in [Fig. 1](#)). Zone A is in the Ohio watershed, and it is characterized by a 0.84 m increase in the groundwater level in the last 30 years (1992–2022) (the only region with a rising groundwater table in Alabama). Zone B is located in the

northern part of the Mobile watershed and has a higher elevation than zones C and D. Zone C covers the broad area of central Alabama, has an average decline of 4.42 m (the greatest decline of all four zones) in the groundwater level during the most recent 30 years, and is therefore more vulnerable than the other three zones to potential future droughts. Zone D covers the southern coastal area with a substantial amount of annual rainfall and has the lowest GWD variation, probably because it receives more precipitation (1724 mm annually based on CAMELS precipitation data) than the northern regions.

To further evaluate the impact of future climate change on groundwater sustainability in Alabama, we apply GW-LSTM to predict the groundwater depth change under three climate scenarios in the next 20 years: base scenario, wet scenario, and dry scenario (Fig. 4).

For the base scenario, artificial precipitation data in the next 20 years (2023–2043) are generated by repeating the same precipitation patterns observed in the last twenty years (2002–2022) for all four zones (Fig. 4). The GW-LSTM model results follow a similar trend as the last 20 years obtained in Section 3.2.

For the wet scenario, the precipitation rate increases by 50% for all future years (Fig. 4). Results of this scenario show that most wells exhibit an elevated water table relative to the base scenario (due to the increasing precipitation), while the groundwater depth in zone A remains relatively close to the base scenario, probably because the average groundwater depth is very close to the land surface in this zone. Evaporation and drainage from shallow groundwater may constrain the groundwater depth fluctuation in this zone.

The dry scenario (where precipitation is decreased by 50% for all future years) simulates an overall lower precipitation rate in the future (somehow related, for example, to climate change in southeastern U.S.) (Kunkel et al., 2013). Forecasted results using GW-LSTM are plotted in Fig. 4, showing that zone C is the most vulnerable region in Alabama to the potential drought in the next two decades. This might be due to the strong dependence on precipitation for GWD at zone C wells. As shown in Fig. 1B, zone C had a significant response to the recent droughts by having the greatest groundwater depth increase in all four zones. Fig. B1 in Appendix B shows the groundwater hydrographs for 4 wells screened at different depths, showing that the zone C groundwater table has dropped regardless of the well depth.

3.4. Wavelet analysis of surface/subsurface time series data

We also apply wavelet analysis (introduced in Appendix C) to assess the response of groundwater depth to precipitation and sea level change. Two wells (#1 and #2 marked in Fig. 1B) in Zone D are considered, since they are located near the Gulf of Mexico in Baldwin County, Alabama. Although the wavelet analysis results do not show a meaningful correlation between the sea level and groundwater depth (Appendix C), the wavelet coherence power spectrum between precipitation and groundwater depth reveals a large area above the significant test level, especially for the temporal scale around 64 days (Fig. 5). The patterns of the power spectrum show negative correlations (red coloring and left-pointing arrows) across many frequencies and dates (Fig. 5), also indicating damped fluctuation and delayed rise of the groundwater table in response to precipitation. This finding identifies the different responses of extreme events in the groundwater level hydrographs to precipitation, which will then be used in Section 4.4.

4. Discussion

4.1. Surface water with out-of-sample predictions

A key potential application of the LSTM models is to predict streamflow for basins without streamflow data (Worland et al., 2019a, 2019b). Dynamic inputs used for the LSTM to predict runoff discharge for ungauged basins include 4 types of data listed in Section 2.1, and static inputs include the 21 catchment attributes/climate data listed also

in Section 2.1. By using cross validation techniques, the four selected basins with runoff data are now removed from the training and validating processes, and represent the ungauged basins and model trained on the remaining basins. The resultant surface water LSTM model predictions (for the whole period) are depicted in Fig. 6, showing that the model has a good prediction of runoff discharge for ungauged basins in Alabama.

4.2. Groundwater with out-of-sample predictions

We use GW-LSTM to predict groundwater depth (the depth to the water table below the land surface) variations at locations without wells. To achieve this goal, we perform a cross validation: train the model on some batch of wells, and then test the model on a well that is not included in the training set. Model inputs needed for predicting groundwater depth variations without recorded data include daily precipitation and temperature data as dynamic inputs, and the hydraulic conductivity, porosity, maximum water content, and soil depth as static inputs, the same as those listed in Section 2.2. Fig. 7 shows the result of out-of-sample in space predictions for the one of the four wells discussed above. The GW-LSTM model can reasonably predict groundwater depth variations ($NSE = 0.733$, which means good prediction) at locations where there are no wells.

4.3. Ranking factors affecting hydrographs

We apply feature importance analysis to identify the effect of input factors on the results of the surface water LSTM and GW-LSTM models (Table 4). We used the integrated gradient method, which is a technique for attributing the importance of each input feature to the output of an LSTM model. The ordered magnitudes of the integrated gradients are equal to the order of importance of the inputs. The first ten important factors identified for surface water LSTM and the first five important factors affecting the GW-LSTM modeling results (considering that GW-LSTM inputs are less than those of the surface LSTM model) are ranked (from high to low) in this table. These rankings are consistent with known effects of hydrological processes. For example, among these factors, precipitation as a dynamic input exhibits the most important impact on the LSTM results for both models because it is the greatest source for producing streamflow and groundwater recharge. In the GW-LSTM mode, the soil water content exhibits a higher impact than the LAI in affecting the groundwater table.

4.4. Extreme events

For the surface water model, the results shown in Section 2 revealed that, although the LSTM model has satisfactory outputs for non-extreme events of streamflow, the model results are less reliable for predicting extreme events (see Fig. 2). As a preliminary test to address this issue, here we add the drainage density as a new physiographic characteristic (which is not a variable included in the CAMELS dataset, and has not yet been studied thoroughly as a model input) and input to the surface water LSTM model. The drainage density, which is a measurement of the channel lengths per unit area of catchment, is indicative of infiltration and permeability of the drainage basin because it is controlled mostly by geological properties and therefore is related to both climate and physical characteristics of the basin. We assume that the drainage density may have a significant impact on stream discharge during extreme precipitation events, such as acting as an amplifier of the peak discharge. Our test results show that the surface water model with drainage density as an additional input (called “SW-DD-LSTM”) slightly promotes the streamflow discharge predictions (notably, the NSE of the SW-DD-LSTM results (listed in Fig. 8) is slightly larger than that for the standard LSTM (listed in Fig. 2)). The root mean square error for the predictions of extreme events also decreased by 0.6, 0.55, 0.2, and 0.04 mm/day for basin a, b, c, and d, respectively. The model, however, is still unable to

predict the exact extremely high values (Fig. 8). Further effort is needed to improve the extreme event prediction by for example identifying the other physiographic characteristics such as Horton-Strahler ratio (Bamufleth et al., 2020), using the Synthetic Minority Over-sampling Technique (SMOTE) presented by Chawla et al. (2002) and including some dynamic inputs such as soil moisture.

To test predictability of extreme events with GW-LSTM and for visual clarity, we changed the time scale for the LSTM results of the groundwater depth to explore the model outputs of extreme events, e.g., the streamflow discharge higher than 20 mm/day in zone A (Fig. 9). Results indicate that our proposed GW-LSTM model can predict GWD response to extreme rainfall events, probably due to (i) the lag time between precipitation and groundwater table change, and (ii) the damping of surface signals when they propagate into the subsurface.

5. Conclusion

This study applied LSTM models to investigate the spatiotemporal evolution of surface water and groundwater resources in Alabama. Model applications and analysis of results led to the following four main conclusions.

First, the LSTM model (without drainage density) made a satisfactory prediction of streamflow discharge at 19 gauged basins in Alabama, but it cannot fully capture extreme events of stream discharge. The inefficiency in modeling extreme discharges might be because the physiographic data that have vital role in peak streamflow during extreme events were not sufficiently captured by the CAMELS static attributes. Additional inputs, such as the drainage density (resulting in the SW-DD-LSTM model), may be needed to improve the LSTM model's predictability for extreme events it's one of many options for improving high flows. Our preliminary tests showed that SW-DD-LSTM slightly improved the surface water LSTM model predictions, and further work is still needed to predict extreme events by identifying additional, relevant physiographic characteristics of hydrological basins.

Second, the groundwater LSTM model (GW-LSTM) proposed by this study predicted the groundwater depth evolution for 21 wells across Alabama. Unlike previous studies which used groundwater data at a previous time step (which is unsuitable for any location without a monitoring well), GW-LSTM accepted input features including hydraulic conductivity, soil depth, soil porosity, and maximum water content, which can capture the main hydrogeological properties of aquifers and therefore explain why GW-LSTM can predict the general trend of the daily groundwater level fluctuations.

Third, the LSTM model can predict subsurface hydrographs well for capturing extreme events, probably due to the damped fluctuation and delayed response of groundwater table (to surface inputs), as shown by the wavelet analysis.

Fourth, the GW-LSTM model adequately predicts groundwater depth for aquifers where no wells or samples are available. The same conclusion was found for the surface water LSTM model, which can reasonably predict runoff discharge for ungauged basins in Alabama. This predictive skill can be useful for assessing the groundwater table or stream discharge in locations without measurements or filling the sampling gaps.

CRediT authorship contribution statement

Hossein Gholizadeh: Investigation, Data processing, Software, Writing Original draft preparation.

Yong Zhang: Conceptualization, Supervision, Writing-Reviewing and editing.

Jonathan Frame: Software, Writing-Reviewing.

Xiufen Gu: Software, Data processing.

Christopher T. Green: Writing-Reviewing and editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This project was paid for in part with federal funding from the Department of the Treasury under the Resources and Ecosystems Sustainability, Tourist Opportunities, and Revived Economies of the Gulf Coast States Act of 2012 (RESTORE Act). The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the Department of the Treasury or ADCNR. H.G. was also partially supported by Geological Sciences Advisory Board (GSAB) from University of Alabama. H.G. also thanks Bahareh Karimidermani for her help in literature review. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitenv.2023.165884>.

References

- Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. *J. Hydrol.* 407 (1–4), 28–40. <https://doi.org/10.1016/j.jhydrol.2011.06.013>.
- Addor, N., Newman, A.J., Mizukami, N., Clark, M.P., 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrol. Earth Syst. Sci.* 21 (10), 5293–5313. <https://doi.org/10.5194/hess-21-5293-2017>.
- Afzaal, H., Farooque, A.A., Abbas, F., Acharya, B., Esau, T., 2019. Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning. *Water* 12 (1), 5. <https://doi.org/10.3390/w12010005>.
- Ahmadi, A., Olyaei, M., Heydari, Z., Emami, M., Zeynolabedin, A., Ghomlaghi, A., Daccache, A., Fogg, G.E., Sadegh, M., 2022. Groundwater level modeling with machine learning: a systematic review and meta-analysis. *Water* 14 (6), 949. <https://doi.org/10.3390/w14060949>.
- Bamufleth, S., Al-Wagdany, A., Elfeki, A., Chaabani, A., 2020. Developing a geomorphological instantaneous unit hydrograph (GIUH) using equivalent Horton-Strahler ratios for flash flood predictions in arid regions. *Geomat. Nat. Hazards Risk* 11 (1), 1697–1723. <https://doi.org/10.1080/19475705.2020.1811404>.
- Banerjee, P., Prasad, R.K., Singh, V.S., 2009. Forecasting of groundwater level in hard rock region using artificial neural network. *Environ. Geol.* 58 (6), 1239–1246. <https://doi.org/10.1007/s00254-008-1619-z>.
- Boulmaiz, T., Guermoui, M., Boutaghane, H., 2020. Impact of training data size on the LSTM performances for rainfall-runoff modeling. *Model. Earth Syst. Environ.* 6 (4), 2153–2164.
- Bouwer, H., 2000. Integrated water management: emerging issues and challenges. *Agric. Water Manag.* 45 (3), 217–228. <https://doi.org/10.1007/s40808-020-00830-w>.
- Bowes, B.D., Sadler, J.M., Morsy, M.M., Behl, M., Goodall, J.L., 2019. Forecasting groundwater table in a flood prone coastal city with long short-term memory and recurrent neural networks. *Water* 11 (5), 1098. <https://doi.org/10.3390/w11051098>.
- Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A Generalize Streamflow Simulation System: Conceptual Modeling for Digital Computers. Joint Federal-State River Forecast Center, Sacramento, CA.
- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357. <https://doi.org/10.1613/jair.953>.
- Chen, C., He, W., Zhou, H., Xue, Y., Zhu, M., 2020. A comparative study among B and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Sci. Rep.* 10 (1), 1–13. <https://doi.org/10.1038/s41598-020-60698-9>.
- Condon, L.E., Kollet, S., Bierkens, M.F., Fogg, G.E., Maxwell, R.M., Hill, M.C., Fransen, H.J.H., Verhoef, A., Van Loon, A.F., Sulis, M., Abesser, C., 2021. Global groundwater modeling and monitoring: opportunities and challenges. *Water Resour. Res.* 57 (12), e2020WR029500 <https://doi.org/10.1029/2020WR029500>.
- De Graaf, I., Condon, L., Maxwell, R., 2020. Hyper-resolution continental-scale 3-D aquifer parameterization for groundwater modeling. *Water Resour. Res.* 56 (5), e2019WR026004 <https://doi.org/10.1029/2019WR026004>.

- Ebersole, S.E., Guthrie, G.M., VanDervoort, D.S., 2019. An update to the physiographic districts of Alabama. In: Alabama Geological Survey Open-file Report 1901 (20p).
- Frame, J.M., Kratzert, F., Raney, A., Rahman, M., Salas, F.R., Nearing, G.S., 2021. Post-processing the national water model with long short-term memory networks for streamflow predictions and model diagnostics. *J. Am. Water Resour. Assoc.* 57 (6), 885–905. <https://doi.org/10.1111/1752-1688.12964>.
- Frame, J.M., Kratzert, F., Klotz, D., Gauch, M., Shelev, G., Gilon, O., Qualls, L.M., Gupta, H.V., Nearing, G.S., 2022. Deep learning rainfall-runoff predictions of extreme events. *Hydrol. Earth Sys. Sci.* 26 (13), 3377–3392. <https://doi.org/10.5194/hess-26-3377-2022>.
- Gauch, M., Mai, J., Lin, J., 2021. The proper care and feeding of CAMELS: how limited training data affects streamflow prediction. *Environ. Model. Softw.* 135, 104926. <https://doi.org/10.1016/j.envsoft.2020.104926>.
- Geological Survey of Alabama, Alabama's waters, Educational Series 11, <https://www.gsa.state.al.us/gsa/groundwater/waterinfo>.
- Géron, A., 2022. Hands-on Machine Learning With Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.
- Gers, F.A., Eck, D., Schmidhuber, J., 2002. Applying LSTM to time series predictable through time-window approaches. In: N.A. Nets WIRN Vietri-01. Springer, London, pp. 193–200.
- Gholizadeh, H., Behrouj Peely, A., Karney, B.W., et al., 2020. Assessment of groundwater ingress to a partially pressurized water-conveyance tunnel using a conduit-flow process model: a case study in Iran. *Hydrogeol. J.* 28, 2573–2585. <https://doi.org/10.1007/s10040-020-02213-y>.
- Gholizadeh, H., Zhang, Y., Frame, J.M., 2022. Long short-term memory networks to explore the spatiotemporal evolution of surface and subsurface water sources in Alabama State. In: AGU Fall Meeting Abstr., vol. 2022, pp. H33B-06.
- Gleeson, T., Wada, Y., Bierkens, M.F., Van Beek, L.P., 2012. Water balance of global aquifers revealed by groundwater footprint. *Nature* 488 (7410), 197–200. <https://doi.org/10.1038/nature11295>.
- Hamidi, S.M.E., Peter, B.G., Muñoz, D.F., Moftakhar, H., Moradkhani, H., 2022. Fast flood mapping with synthetic aperture radar data using Google Earth engine. In: AGU Fall Meeting Abstr., vol. 2022, pp. H55M-0739.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *N.A. Comput.* 9 (8), 1735–1780.
- Alabama Department of Economic and Community Affairs report. <https://adeca.alabama.gov/water-management/publications-reports-and-information/>, 2015.
- Geological Survey of Alabama (GSA) well records database. <https://www.gsa.state.al.us/gsa/groundwater/wellrecords>.
- Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization, arXiv preprint. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). 10.48550/arXiv.1412.6980.
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018a. Rainfall-runoff modelling using long short-term memory (LSTM) networks. *Hydrol. Earth Sys. Sci.* 22 (11), 6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>.
- Kratzert, F., Klotz, D., Herrnegger, M., Hochreiter, S., 2018b. A glimpse into the unobserved: runoff simulation for ungauged catchments with LSTMs. In: Workshop on Modeling and Decision-Making in the Spatiotemporal Domain, 32nd Conference on Neural Information Processing Systems (NeuRIPS 2018), Montréal, Canada.
- Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A.K., Hochreiter, S., Nearing, G.S., 2019. Toward improved predictions in ungauged basins: exploiting the power of machine learning. *Water Resour. Res.* 55 (12), 11344–11354. <https://doi.org/10.1029/2019WR026065>.
- Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., 2022. Caravan—a global community dataset for large-sample hydrology. *Sci. Data* 10 (1), 61. <https://doi.org/10.1038/s41597-023-01975-w>.
- Kunkel, K.E., Stevens, L.E., Stevens, S.E., Sun, L., Janssen, E., Wuebbles, D., Hilberg, S.D., Timlin, M.S., Stoecker, L., Westcott, N., Dobson, J.G., 2013. Regional climate trends and scenarios for the US National Climate Assessment part 3: Climate of the Midwest US. NOAA Technical Report NESDIS, pp. 142–143.
- Lins, H.F., 2012. USGS hydro-climatic data network 2009 (HCDN-2009). In: U.S. Geological Survey Fact Sheet, vol. 3047(4).
- Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y., 2015. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. In: *Trans. Res. C – Emer. Tech.* 54, pp. 187–197. <https://doi.org/10.1016/j.trc.2015.03.014>.
- Maxwell, R.M., 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Adv. Water Resour.* 53, 109–117. <https://doi.org/10.1016/j.advwatres.2012.10.001>.
- Maxwell, R.M., Condon, L.E., 2016. Connections between groundwater flow and transpiration partitioning. *Science* 353 (6297), 377–380. <https://doi.org/10.1126/science.aaf7891>.
- Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geosci. Model Dev.* 8 (3), 923–937. <https://doi.org/10.5194/gmd-8-923-2015>.
- Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: performance measures and evaluation criteria. *Trans. ASABE* 58 (6), 1763–1785. <https://doi.org/10.13031/trans.58.10715>.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol.* 10, 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Nayak, P.C., Rao, Y.R., Sudheer, K.P., 2006. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resour. Manag.* 20 (1), 77–90. <https://doi.org/10.1007/s11269-006-4007-z>.
- Nearing, G.S., Klotz, D., Sampson, A.K., Kratzert, F., Gauch, M., Frame, J.M., Shalev, G., Nevo, S., 2022. Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks. *Hydrol. Earth Syst. Sci.* 26 (21), 5493–5513. <https://doi.org/10.5194/hess-26-5493-2022>.
- Nevo, S., Morin, E., Gerzi Rosenthal, A., Metzger, A., Barshai, C., Weitzner, D., Voloshin, D., Kratzert, F., Elidan, G., Drot, G., Begelman, G., 2022. Flood forecasting with machine learning models in an operational framework. *Hydrol. Earth Syst. Sci.* 26 (15), 4013–4032. <https://doi.org/10.48550/arXiv.2111.02780>.
- Newman, A.J., Sampson, K., Clark, M.P., Bock, A., Viger, R.J., Blodgett, D., 2014. A Large-sample Watershed-scale Hydrometeorological Dataset for the Contiguous USA. UCAR/NCAR, Boulder, CO. <https://doi.org/10.5065/D6MW2F4D>.
- Rahman, A.S., Hosono, T., Quilty, J.M., Das, J., Basak, A., 2020. Multiscale groundwater level forecasting: coupling new machine learning approaches with wavelet transforms. *Adv. Water Resour.* 141, 103595. <https://doi.org/10.1016/j.advwatres.2020.103595>.
- Sahoo, S., Russo, T.A., Elliott, J., Foster, I., 2017. Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. *Water Resour. Res.* 53 (5), 3878–3895. <https://doi.org/10.1002/2016WR019933>.
- Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., Woo, W.C., 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *Adv. Neur. Infor. Process. Sys.* 28.
- Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L., Cheng, Z., 2020. Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *J. Pet. Sci. Eng.* 186, 106682. <https://doi.org/10.1016/j.petrol.2019.106682>.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.
- Srivastava, A., Wang, T.Y., Zhang, P., De Rose, C.A.F., Kannan, R., Prasanna, V.K., 2020. MemMAP: compact and generalizable meta-LSTM models for memory access prediction. In: Lauw, H., Wong, R.W., Ntoulas, A., Lim, E.P., Ng, S.K., Pan, S. (Eds.), Advances in Knowledge Discovery and Data Mining, PAKDD 2020, Lecture Notes in Computer Science, vol. 12085. Springer, Cham. https://doi.org/10.1007/978-3-03-047436-2_5.
- Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Salim, H., Sungwon, K., Sulaiman, S.O., Tan, M.L., Sa'adi, Z., Mehr, A.D., Allawi, M.F., 2022. Groundwater level prediction using machine learning models: a comprehensive review. *Neurocomputing* 489 (7), 271–308. <https://doi.org/10.1016/j.neucom.2022.03.014>.
- Tran, H., Leonarduzzi, E., De la Fuente, L., Hull, R.B., Bansal, V., Chennault, C., Gentile, P., Melchior, P., Condon, L.E., Maxwell, R.M., 2021. Development of a deep learning emulator for a distributed groundwater-surface water model: Parflow-ML. *Water* 13 (23), 3393. <https://doi.org/10.3390/w13233393>.
- Vu, M.T., Jardani, A., Massei, N., Fournier, M., 2021. Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network. *J. Hydrol.* 597, 125776. <https://doi.org/10.1016/j.jhydrol.2020.125776>.
- Winter, T.C., 1999. Ground Water and Surface Water: A Single Resource, vol. 1139. Diane Publishing. <https://doi.org/10.3133/cir1139>.
- Worland, S.C., Knight, R.R., Asquith, W.H., 2019a. Observed and modeled daily streamflow values for 74 U.S. Geological Survey streamgage locations in the Trinity and Mobile-Tombigbee River basins in the southeast United States: 2000–2009. In: U.S. Geological Survey Data Release. <https://doi.org/10.5066/P92F1ROU>.
- Worland, S.C., Steinschneider, S., Farmer, W., Asquith, W., Knight, R., 2019b. Copula theory as a generalized framework for flow-duration curve based streamflow estimates in ungauged and partially gaged catchments. *Water Resour. Res.* 55 (11), 9378–9397. <https://doi.org/10.1029/2019WR025138>.
- Wu, C., Zhang, X., Wang, W., Lu, C., Zhang, Y., Qin, W., Tick, G.R., Liu, B., Shu, L., 2021. Groundwater level modeling framework by combining the wavelet transform with a long short-term memory data-driven model. *Sci. Total Environ.* 783, 146948. <https://doi.org/10.1016/j.scitotenv.2021.146948>.
- Wunsch, A., Liesch, T., Broda, S., 2021. Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrol. Earth Syst. Sci.* 25 (3), 1671–1687. <https://doi.org/10.5194/hess-25-1671-2021>.
- Yin, J., Han, J., Xie, R., Wang, C., Duan, X., Rong, Y., Zeng, X., Tao, J., 2021. Mc-lstm: real-time 3d human action detection system for intelligent healthcare applications. *IEEE T. Biomed. Cir.* 15 (2), 259–269. <https://doi.org/10.1109/TBCAS.2021.3064841>.
- Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., Lee, K.K., 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* 396 (1–2), 128–138. <https://doi.org/10.1016/j.jhydrol.2010.11.002>.
- Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. *N.A. Comput.* 31 (7), 1235–1270. https://doi.org/10.1162/neco_a_01199.
- Zaytar, M.A., El Amrani, C., 2016. Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *Int. J. Comput. Appl.* 143 (11), 7–11. <https://doi.org/10.5120/ijca2016910497>.
- Zhang, J., Zhu, Y., Zhang, X., Ye, M., Yang, J., 2018. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* 561, 918–929. <https://doi.org/10.1016/j.jhydrol.2018.04.065>.
- Zhou, Y., Li, W., 2011. A review of regional groundwater flow modeling. *Geosci. Front.* 2 (2), 205–214. <https://doi.org/10.1016/j.gsf.2011.03.003>.