Journal of Hydrology 616 (2023) 128857

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Long-term river water temperature reconstruction and investigation: A case study of the Dongting Lake Basin, China

Feng Huang [a,*], Bao Qian [b], Carlos G. Ochoa [c]

[a] College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China
[b] Bureau of Hydrology, Changjiang Water Resources Commission of the Ministry of Water Resources, Wuhan 430010, China
[c] College of Agricultural Sciences - Ecohydrology Lab, Oregon State University, Corvallis, OR 97331, USA

ARTICLE INFO

Keywords:
Riverine thermal regime
Framework
Deep learning
Data reconstruction
Long-term dynamics
Yangtze River

ABSTRACT

Water temperature is a fundamental property of river habitats and aquatic ecosystems. In some cases, assessing water temperature changes is restrained by the availability of historical observation. This study proposed a framework to address the water temperature data scarcity in thermal regime assessment by incorporating the long short-term memory (LSTM) network and the time series technologies. The factors affecting water temperature included air temperature, evaporation, and streamflow. Driven by the historical observation of these factors, the LSTM network produced long-term water temperature series. The time series technologies unraveled the temporal dynamics of the reconstructed water temperature. The trend, periodicity, and complexity were investigated using linear regression, trend-free pre-whitening Mann-Kendall test, Sen's slope, wavelet transform, and moving sample entropy. The framework was applied in the Dongting Lake Basin, China, which drains four tributaries: Xiangjiang, Zishui, Yuanjiang, and Lishui Rivers. The LSTM network reconstructed the monthly water temperature series from 1960 to 2020, which was good surrogate data for thermal regime evaluation. The annual water temperature of the four tributaries increased significantly, with an average warming rate of 0.15 ℃ per decade. Heterogeneity of tendency and changing rates existed across sub-basins and months, and the warming water temperature dominated the variation. A significant oscillation of a one-year period across six decades was identified in the Dongting Lake Basin. The water temperature complexity decreased significantly in the Xiangjiang, Zishui, and Yuanjiang Rivers, while it increased significantly in the Lishui River. The case study indicated that the framework was practicable and flexible with explicit structure. It could be applied in other basins to address water temperature data scarcity and support the comprehensive riverine thermal regime assessment.

1. Introduction

Water temperature is one of the critical environmental factors affecting water quality, biogeochemistry, metabolism, and aquatic organisms and ecosystems (Bonacina et al., 2022). The riverine thermal regime alterations can have essential effects on biological communities and the overall health of ecosystems (Tao et al., 2021). Aquatic organisms usually have a specific range of temperatures to survive, and variations in the river water temperature directly affect their growth rates, behavior, and distributions (Caissie, 2006). In a changing climate, typically characterized by a warming trend in air temperature, river water temperature increases due to heat exchange with the atmosphere (Kedra and Wiejaczka, 2018). There is compelling evidence that river water temperature is increasing globally under climate change, which is exacerbated by increased abstractions to satisfy human water requirements (Ouellet et al., 2020).

Given that the anomalies in riverine thermal regimes may have significant consequences for the structure and functioning of aquatic ecosystems, the foremost step is to unravel the long-term temporal dynamics of river water temperature (Chen et al., 2016). Investigating the characteristics of water temperature changes is the premise of evaluating the environmental and ecological effects of thermal regime alterations. Understanding the long-term thermal dynamics in terms of trend, periodicity, and complexity is vital to developing new environmental management and ecological protection strategies with an adaptation to climate change in an anthropogenic-driven era (Bonacina et al., 2022).

Extensive efforts have investigated the river water temperature

* Corresponding author.
E-mail addresses: huangfeng1987@hhu.edu.cn (F. Huang), jacber@163.com (B. Qian), Carlos.Ochoa@oregonstate.edu (C.G. Ochoa).

https://doi.org/10.1016/j.jhydrol.2022.128857

Available online 28 November 2022

changes in major basins of many countries, e.g., the US, India, the UK, France, Switzerland, Canada, and China (Beaufort et al., 2020; Chen et al., 2016; Hannah and Garner, 2015; Hebert et al., 2011; Michel et al., 2020; Rehana and Mujumdar, 2011; Wagner et al., 2017). These studies employed long-term water temperature measurements, which provided an essential foundation for riverine thermal regime assessment. For example, little evidence of seasonal changes was suggested in the river water temperatures of the basins in the eastern US, while more evidence of warming for river sites in the western US, particularly during the fall and winter seasons, was found (Wagner et al., 2017). Historical evidence suggested that the UK river water temperature increased in the latter part of the 20th century (Hannah and Garner, 2015). In the Yongan watershed in Eastern China over the 1980–2012 period, river water temperature increased between 0.29 and 0.46 ℃/10a, partially due to the 0.50 ℃/10a increase in air temperature (Chen et al., 2016). Observed water temperature data have been used in previous studies to shed light on the variations in river water temperature. However, knowledge gaps still exist in understanding how river water temperature changes in those basins, where the water temperature measurement just started in recent years, and long-term historical observation was lacking.

Although continuous and regular water temperature monitoring is essential for assessing riverine thermal alterations, these data are rarely available in many basins (Pohle et al., 2019). Water temperature models prove a helpful tool for reconstructing river water temperature over a long period, thereby providing valuable information to assess thermal regime alterations (DeWeber and Wagner, 2014). Water temperature models are generally classified into physically-based deterministic and stochastically data-driven models. Deterministic models employ an energy budget approach to predict river water temperature, whereas stochastical models typically rely on air-to-water temperature relationships (Caissie, 2006). Deterministic models are generally based on the heat exchange dynamics between the water body and the surrounding environments. They require a myriad of inputs, including fluvial topography, a complete set of meteorological variables, and the hydraulic properties of the river (Piccolroaz et al., 2016). Due to the intrinsic complexity of heat equations that require a lot of data inputs, the deterministic models' applicability in the case of data scarcity is limited (Tao et al., 2021). Flexible statistical models, particularly artificial intelligence algorithms, are promising alternatives (Qiu et al., 2020; Zhu and Piotrowski, 2020). Deep learning, a new generation of artificial neural network research, is especially suited for information extraction from sequential data. It serves functionalities to build models with more accurate simulation, far greater processing capability, and reduced requirements for human involvement and expertise (Shen, 2018). Long short-term memory (LSTM) network is a type of deep machine-learning method. Previous studies have demonstrated that the LSTM network has a satisfying potential to simulate river water temperature and quantify temporal oscillations in thermal regimes induced by climate change and anthropogenic activities (Feigl et al., 2021; Qiu et al., 2021).

Considering the abundance of hydrometeorological data and recent developments in artificial intelligence deep learning, this study proposed a framework to help bridge the knowledge gap for assessing riverine thermal alterations in water temperature data-scarce basins. The proposed framework employs the LSTM network to reconstruct and extend the river water temperature series and examines the long-term dynamics in terms of trend, periodicity, and complexity. The framework was applied in the Dongting Lake Basin, a sub-basin of the Yangtze River Basin in China. The Dongting Lake Basin was a representative basin where water temperature monitoring was initiated recently and lacked historical observation data, resulting in knowledge gaps in the long-term dynamics of riverine thermal regimes. The water temperature was monitored from 2007 to 2020, lacking long-term records for evaluating thermal variations. The streamflow and meteorological observation continued from 1960 to 2020, providing the inputs for the LSTM network to extend the water temperature series. This case study illustrated the framework's execution processes and practicality and provided implications for other basins suffering from data scarcity. For local interests, this application's results provided vital information for managers and policymakers to systematically assess the riverine thermal regime alteration and develop strategies to deal with this change in the Dongting Lake Basin. For worldwide interests, the framework proposed by this study provided the basins lacking historical water temperature observations with an effective tool for water temperature reconstruction and assessment, which is a vital component of fishery management and watershed preservation policy initiatives.

2. Case study site and data

2.1. Study area

Yangtze River is the longest river in China and the third-longest river in the world. It originates in the Qinghai-Tibet Plateau and flows 6300 km into the East China Sea, draining an area of 1.8 million km^2 (Fig. 1a). The basin lies in the monsoon region of East Asia's subtropical zone and has mean annual precipitation of 1090 mm (Jiang et al., 2006). Generally, spring lasts from March to May, summer lasts from June to August, autumn lasts from September to November, and winter lasts from December to February. The middle and lower reaches of the Yangtze River have thousands of freshwater lakes; among them, Poyang Lake and Dongting Lake are the two largest lakes and have excellent value for socio-economic development and environmental conservation. Dongting Lake is the second-largest freshwater lake in China and is located on the south bank of the middle Yangtze River (Fig. 1b). Four large tributaries feed Dongting Lake, namely Xiangjiang River, Zishui River, Yuanjiang River, and Lishui River. The published literature has provided a detailed introduction to the geography, meteorology, hydrology, and ecology of the Yangtze River Basin and its Dongting Lake sub-basin (Liang et al., 2021; Yang et al., 2020; Yu et al., 2018; Zhao et al., 2017).

2.2. Hydrological and meteorological data

Monthly-averaged water temperature (℃) and streamflow (m^3/s) data were collected from four hydrometric stations, Xiangtan, Taojiang, Taoyuan, and Shimen. These hydrometric stations are located at the lower reaches of the tributaries discharging into Dongting Lake (Fig. 1b). Xiangtan, Taojiang, Taoyuan, and Shimen stations monitor the streamflow processes of Xiangjiang, Zishui, Yuanjiang, and Lishui Rivers, respectively. Monthly average water temperature data spanned from 2007 to 2020, and monthly average streamflow data spanned from 1960 to 2020. The hydrological data were provided by the Changjiang Water Resources Commission and the Hydrological Bureau of Hunan Province. They have checked and firmly controlled the homogeneity and reliability of the hydrological data.

Monthly mean air temperature (℃) and monthly accumulated evaporation (mm) data were obtained from twenty meteorological stations scattered over the Dongting Lake Basin (Fig. 1b). These data were provided by the China Meteorological Data Service Centre (https://data.cma.cn/en). The meteorological records spanned from 1960 to 2020, covering sixty-one years. Each sub-basin's spatial mean air temperature and evaporation values were calculated by averaging the available data from meteorological stations in Xiangjiang, Zishui, Yuanjiang, and Lishui River Basins.

3. Methods

3.1. Methodological framework

Fig. 2 illustrates the methodological framework used to simulate the long-term water temperature and investigate its variations. Data collection and pre-processing, e.g., air temperature, evaporation,

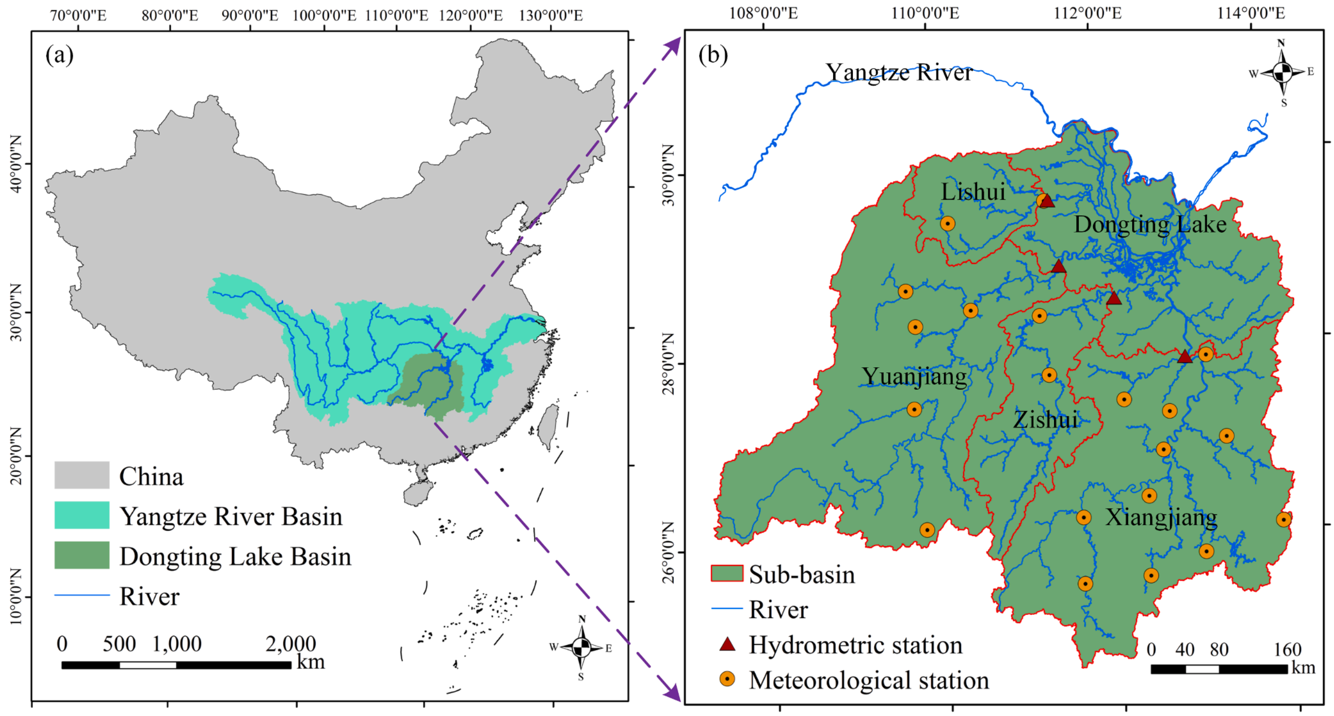

Fig. 1. Sketch of the Yangtze River Basin and Dongting Lake: (a) geographic location of the Yangtze River and Dongting Lake Basin; and (b) the sub-basins of the Dongting Lake Basin and the distribution of hydrometeorological stations.

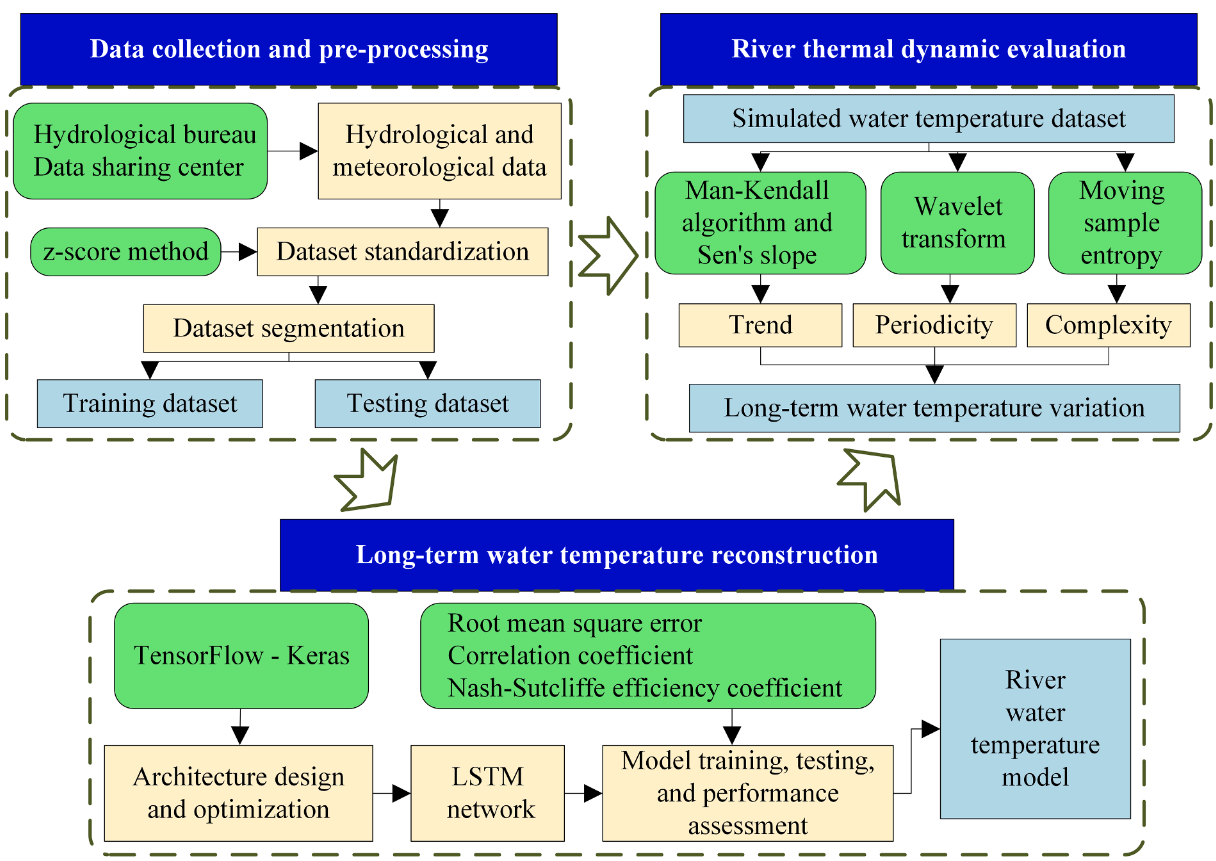

Fig. 2. A framework for modeling and investigating long-term river water temperature.

streamflow, and water temperature, provide the information basis for constructing the water temperature model, extending the water temperature series, and evaluating the thermal alterations. The original datasets are standardized and segmented for model construction and water temperature simulation. LSTM network is suggested to reconstruct long-term water temperature from its affecting factors, e.g., air temperature, evaporation, and streamflow.

Over the same period as river water temperature data are available, the linkage between water temperature and its affecting factors is depicted using the LSTM network. The LSTM network can be constructed, optimized, trained, and tested using Keras, a deep learning application programming interface written in Python, running on the machine-learning platform TensorFlow (https://keras.io/). The validated water temperature model, driven by the long-term observed hydrometeorological data, i.e., air temperature, evaporation, and streamflow, produces long-term water temperature. Investigating the trend, periodicity, and complexity of the reconstructed water temperature series helps in understanding the temporal dynamics of the riverine thermal regimes.

3.2. LSTM model for modeling river water temperature

LSTM network is a special kind of recurrent neural network, the unit of which is composed of a cell, an input gate, an output gate, and a forget gate (Hochreiter and Schmidhuber, 1997). The LSTM architecture consists of a set of recurrently connected sub-networks known as memory blocks. These blocks maintain their states over time and regulate the information flow. The output of the block is recurrently connected back to the block input and all of the gates (Van Houdt et al., 2020). One reason for the LSTM network's success lies in its ability to handle the exploding and vanishing gradient problem, which stands as a difficult issue to be circumvented when training recurrent or very deep neural networks (Van Houdt et al., 2020). Previous studies have proved the practicability and effectiveness of the LSTM network to simulate sea, river, and lake surface water temperatures (Kim et al., 2020; Qiu et al., 2021; Rahmani et al., 2021; Read et al., 2019). Given the satisfying performance of the LSTM network with a simple structure, an LSTM network encompassing an input layer, two LSTM layers, one fully connected layer, and an output layer was used to model the river water temperature in the Dongting Lake Basin. The LSTM formulation for modeling the river water temperature with a monthly scale is expressed by:

$$T_{W,t} = f(T_{A,t},\ E_t,\ Q_t) \tag{1}$$

where $T_{W,t}$ is the river water temperature at the current month t, unit ℃; $T_{A,t}$ is the spatially averaged air temperature of a sub-basin at the current month t, unit ℃; E_t is the spatially averaged evaporation of a sub-basin at the current month t, unit mm; and Q_t is the tributary streamflow at the current month t, unit m^3/s.

All computations were conducted using Python 3.8.10 software. The Keras running on TensorFlow was used to program the LSTM network. The input layer received air temperature, evaporation, and streamflow. It transferred the inputting information to the LSTM and fully connected layers for processing, and the output layer produced the water temperature simulation results. The LSTM layers used the hyperbolic tangent function as the activation function, and the fully connected layer used the linear function as the activation function. The key to the LSTM network is the hidden neurons in the LSTM layer called memory cells, which receive information from the input layer and perceive the information at the previous moment. The number of hidden neurons was set using the "keras_tuner" function, suggesting 30 memory cells for each LSTM layer to model water temperature. The adaptive moment estimation was used to optimize the network during the learning process, and the early stopping method was applied to prevent over-fitting.

When constructing the water temperature models for the Xiangjiang, Zishui, Yuanjiang, and Lishui Rivers, the observed data from 2011 to 2020 were used to train the LSTM networks. The observed data from 2007 to 2010 were used to test the models. Because the differences in absolute values of air temperature, evaporation, streamflow, and water temperature negatively affect the model's learning ability, the hydrological and meteorological data were standardized using the z-score method. The model performance was evaluated using root mean square error (*RMSE*), correlation coefficient (*R*), and Nash-Sutcliffe efficiency coefficient (*NSE*), which are commonly used performance evaluators (Huang et al., 2021a; Huang et al., 2021b; Qiu et al., 2021). Given air temperature, evaporation, and streamflow data spanned from 1960 to 2020, water temperature data were reconstructed from 1960 to 2020 using the validated LSTM network.

3.3. Trend, periodicity, and complexity analysis

3.3.1. Trend detection

The trend embedded in the water temperature series was identified by the linear regression analysis combined with the trend-free pre-whitening Mann-Kendall (TFPW-MK) and Sen's slope algorithms. Linear regression analysis illustrated the trend, and the TFPW-MK test ascertained whether the variation was significant. Sen's slope (s) quantified the magnitude of increasing or decreasing tendency in the water temperature series. The TFPW-MK test addresses serial autocorrelation issues by detrending the series before pre-whitening and provides a more accurate assessment of the trend's significance. The diagram of the TFPW-MK test is illustrated in Fig. 3 (Yue et al., 2002b). The published literature provided detailed information on the algorithms of the TFPW-MK test and Sen's slope (Sen, 1968; Yue et al., 2002a; Yue et al., 2002b). The Python package "pymannkendall" (https://pypi.org/project/pymannkendall/) supported trend detection, and the significance level (p) was set as 0.05 (Hussain and Mahmud, 2019).

3.3.2. Periodicity identification

The periodicity embedded in the water temperature series was analyzed using wavelet transform, a standard tool for detecting localized power variations within a time series (Torrence and Compo, 1998). Wavelet transform uses generalized local base functions named wavelets that can be stretched and translated with a flexible resolution in frequency and time. The flexible windows are adaptive to the entire time–frequency domain, known as the wavelet domain, which narrows while focusing on high-frequency signals and widens while searching the low-frequency background (Lau and Weng, 1995). The wavelet analysis results interpret the dominant modes of variability and how those modes vary in time. The diagram of the wavelet transform is illustrated in Fig. 4, and a detailed introduction to the algorithms can be found in the literature (Lau and Weng, 1995; Torrence and Compo, 1998). The wavelet analysis was conducted using the Python package "PyCWT" (https://pypi.org/project/pycwt/). The complex Morlet wavelet method was used because of its reliability in detecting signal periodicity (Liang et al., 2013). Detrending is necessary to remove the impact of tendency on periodicity detection. The detrending was performed by fitting a one-degree polynomial function and subtracting it from the original series. Before the wavelet analysis, the detrended water temperature series was standardized using the z-score method.

3.3.3. Complexity investigation

Water temperature complexity analyzed in this study could be defined as the degree of uncertainty or the rate of information production of water temperature series (Sivakumar and Singh, 2012). In light of the sample entropy's effectiveness and reliability, it was used to quantify the water temperature complexity (Huang et al., 2021b). The sample entropy was an unbiased estimator of the conditional probability that two similar sequences of m consecutive data points (m was the template length) would remain similar when one more consecutive point was included. A detailed description of the sample entropy algorithms can be

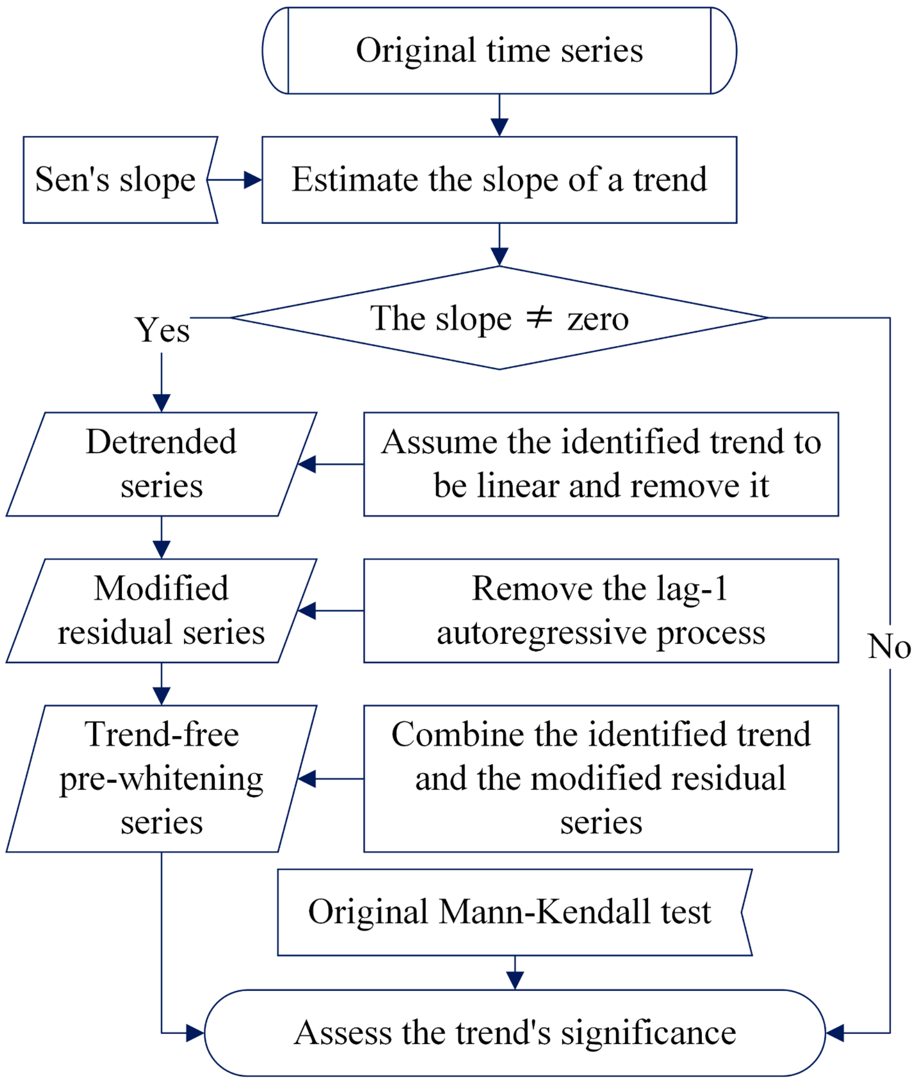

Fig. 3. Diagram of the trend-free pre-whitening Mann-Kendall test.

found in the literature (Huang et al., 2011; Wu et al., 2013).

The moving sample entropy was calculated to unravel how the water temperature complexity changed over six decades. A ten-year-long window moved with a one-year step divided the 1960–2020 series into fifty-two subseries: 1960–1969, 1961–1970, 1962–1971, …, 2009–2018, 2010–2019, and 2011–2020. Calculating sample entropy for each subseries produced a sample entropy series covering 1960 to 2011. The linear regression and TFPW-MK test were employed to identify the trend in the sample entropy series. The computation of water temperature complexity used Python programming. When calculating sample entropy, previous studies have suggested setting the template length at $m = 2$ and the tolerance criterion at $r = 0.15\delta$; where δ denotes the standard deviation of the original water temperature series (Li and Zhang, 2008; Wu et al., 2013).

4. Results and discussion

4.1. Long-term river water temperature simulation

Fig. 5 illustrates the performance of the LSTM network to simulate the river water temperature in the Dongting Lake Basin. During the model training and testing, the scatters of water temperature were distributed alongside the 1:1 line. The simulated water temperature matched well with the observed one, with satisfactory goodness-of-fit statistics of the LSTM model performance (Table 1). For training the model, the evaluator *RMSE* ranged from 1.56 to 1.81 ℃, with a mean value of 1.64 ℃, the evaluator *R* ranged from 0.96 to 0.98, with a mean value of 0.97, and the evaluator *NSE* ranged from 0.92 to 0.95, with a mean value of 0.94. For testing the model, the evaluator *RMSE* ranged from 1.21 to 2.35 ℃, with a mean value of 1.51 ℃, the evaluator *R* ranged from 0.95 to 0.99, with a mean value of 0.97, and the evaluator *NSE* ranged from 0.90 to 0.98, with a mean value of 0.95. The goodness-of-fit statistics suggested that the LSTM network adequately simulated the river water temperature and could be applied to reconstruct long-term water temperature series.

The validated LSTM network driven by the long-term observed air temperature, evaporation, and streamflow resulted in long-term monthly water temperature covering January-1960 to December-2020 (Fig. 6). The monthly mean variation in water temperature was dominated by the sinusoidal seasonal pattern in which maximum temperatures occur over the summer months and minimum over the winter months. Annual mean water temperature series were calculated, and each month's water temperature series were extracted for the following trend analysis.

4.2. Temporal dynamics of annual and monthly river water temperature

Linear regression illustrated prominent increasing trends in the

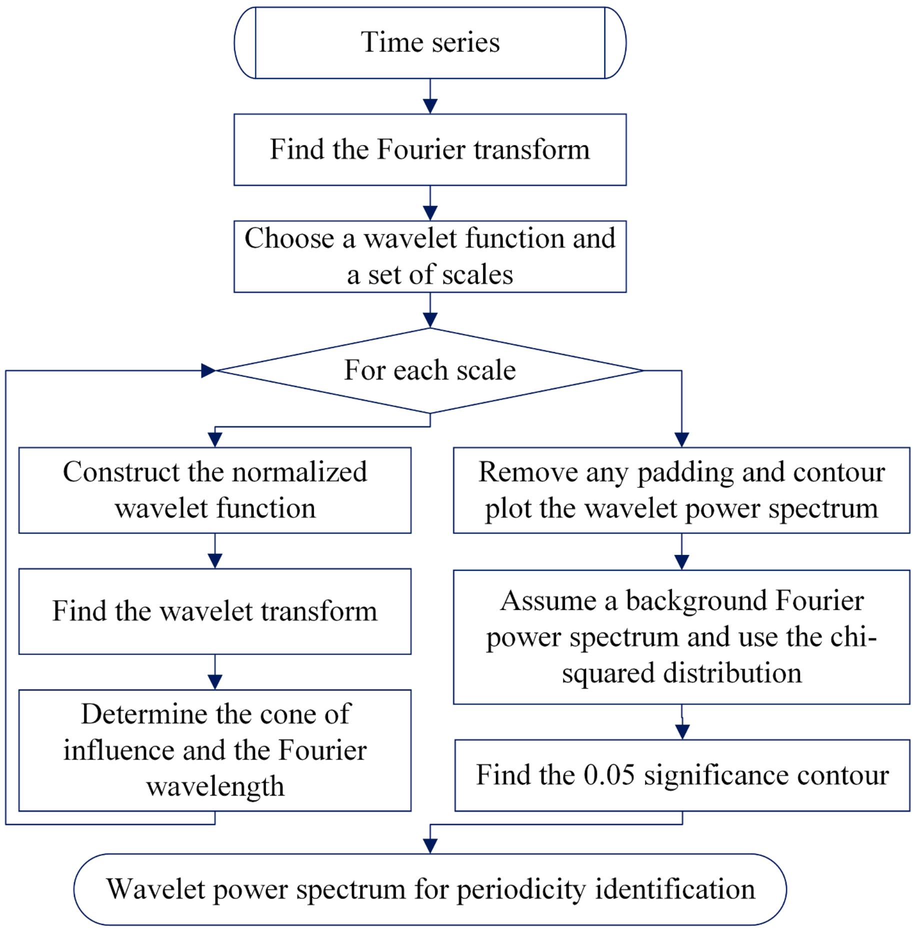

Fig. 4. Diagram of the wavelet transform.

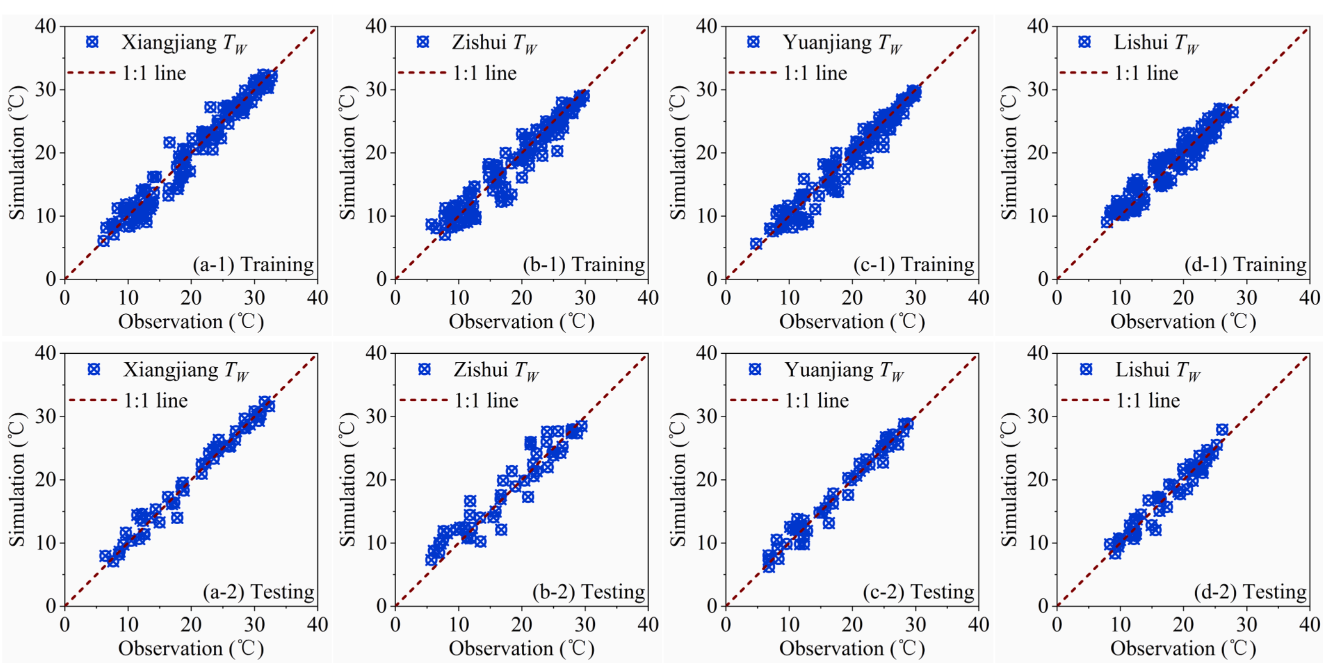

Fig. 5. Evaluating the LSTM network's performance in modeling river water temperature.

annual averaging water temperatures in the Dongting Lake Basin (Fig. 7). The river water temperature fluctuated and increased gradually over six decades. The TFPW-MK test confirmed that these warming trends of river water temperature were statistically significant. The warming rate of annual mean water temperature in the Xiangjiang River, Zishui River, Yuanjiang River, and Lishui River was 0.16, 0.10,

Table 1
Statistical evaluators of the LSTM network's performance in modeling river water temperature.

Metric	Model training				Model testing			
	Xiangjiang	Zishui	Yuanjiang	Lishui	Xiangjiang	Zishui	Yuanjiang	Lishui
RMSE(℃)	1.62	1.81	1.58	1.56	1.21	2.35	1.22	1.27
R	0.98	0.97	0.97	0.96	0.99	0.95	0.98	0.97
NSE	0.95	0.93	0.95	0.92	0.98	0.90	0.97	0.94

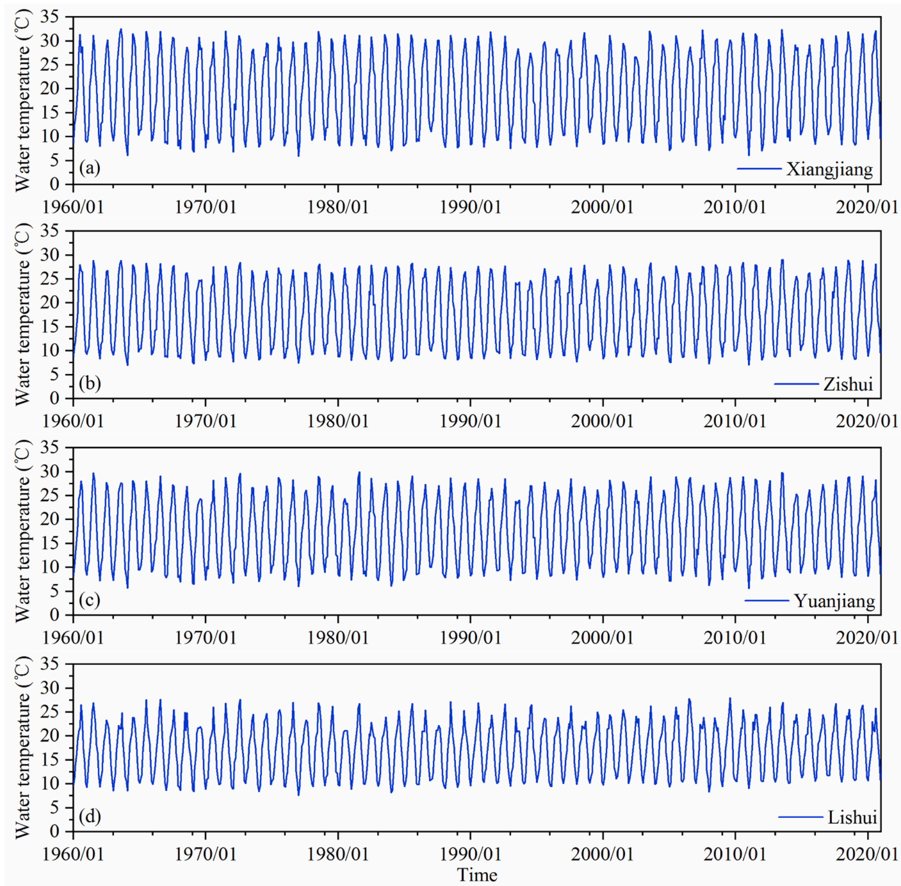

Fig. 6. River water temperature simulation using the validated LSTM model.

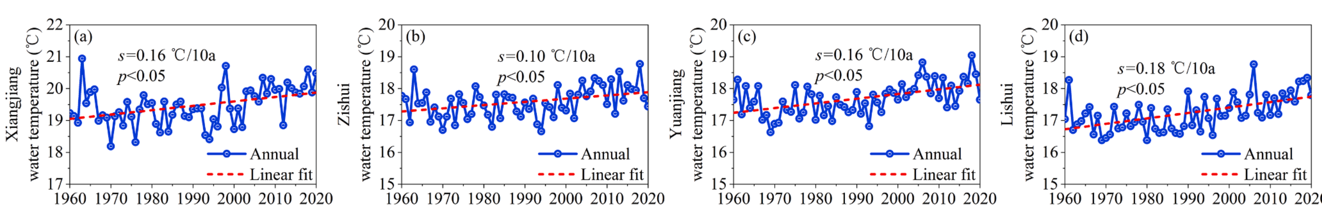

Fig. 7. Temporal dynamics of annual mean river water temperature in the Dongting Lake Basin.

0.16, and 0.18 ℃ per decade, respectively. They resulted in an average warming rate of 0.15 ℃ per decade. The results were in line with the previous findings and added a piece of case evidence supporting the idea that river water temperature was increasing globally under climate change (Ouellet et al., 2020).

The monthly mean water temperature in the Xiangjiang River experienced different variations from 1960 to 2020 (Fig. 8). Linear regression and Sen's slope analysis indicated a decreasing water temperature trend in July; however, the TFPW-MK test rejected the significance of this decreasing trend. In other months except for July, the water temperatures increased with various Sen's slops, ranging from 0.01 to 0.37 ℃ per decade. The TFPW-MK test identified significant warming water temperatures in the winter month of February and the spring month of April. In April, the water temperature's warming rate reached 0.37 ℃ per decade, the highest value compared with the warming rates in other months. In February, the water temperature's warming rate was 0.30 ℃ per decade. Therefore, the rising annual mean water temperature in the Xiangjiang River might be partially attributed to February and April's significantly increased water temperature.

The long-term reconstructed data suggested different temporal

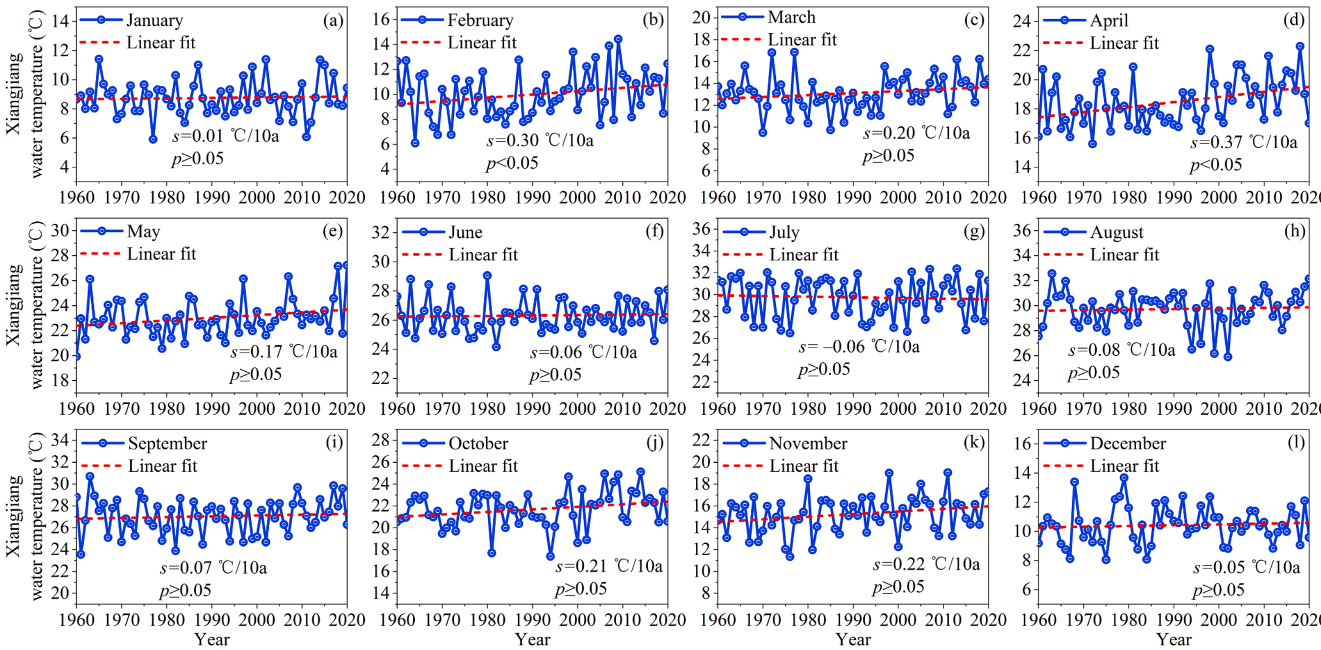

Fig. 8. Temporal dynamics of monthly mean river water temperature in the Xiangjiang River.

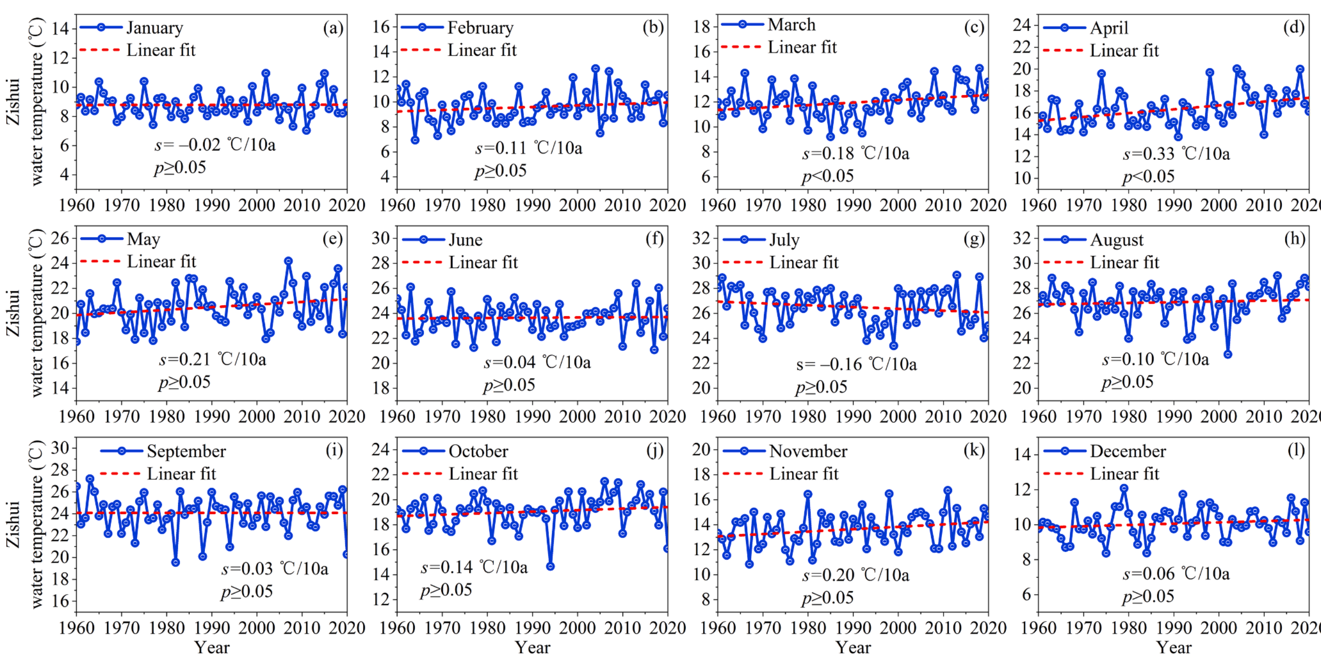

Fig. 9. Temporal dynamics of monthly mean water temperature in the Zishui River.

dynamics of monthly mean water temperature in the Zishui River (Fig. 9). In January and July, the water temperature dropped slightly with a decreasing rate of 0.02 and 0.16 ℃ per decade, respectively. However, these decreasing trends had not passed the TFPW-MK test yet. In other months except for January and July, the water temperatures increased. The water temperature warming rates ranged from 0.03 to 0.33 ℃ per decade, with a mean value of 0.14 ℃ per decade. The increasing water temperature trends in the two spring months of March and April were significant, as confirmed by the TFPW-MK test. The substantial increase in water temperature in these two months contributed to the substantial increase in the annual mean water temperature of the Zishui River.

The monthly water temperature variations in the Yuanjiang River were dominated by the increasing tendency in the past six decades, indicated by the linear regression and Sen's slope analysis (Fig. 10). A slight and insignificant decreasing trend was detected in only one month, July. The water temperatures in other months increased, and the warming rates ranged from 0.07 to 0.36 ℃ per decade, with a mean value of 0.18 ℃ per decade. As suggested by the TFPW-MK test, the increasing water temperature trends in the two spring months of March and April and the two autumn months of October and November were significant, accounting for the substantial increase in the annual mean water temperature of the Yuanjiang River.

In the Lishui River, the reconstructed monthly data suggested a dominantly rising water temperature from 1960 to 2020, with a minimum warming rate of 0.11 ℃ per decade and a maximum warming rate

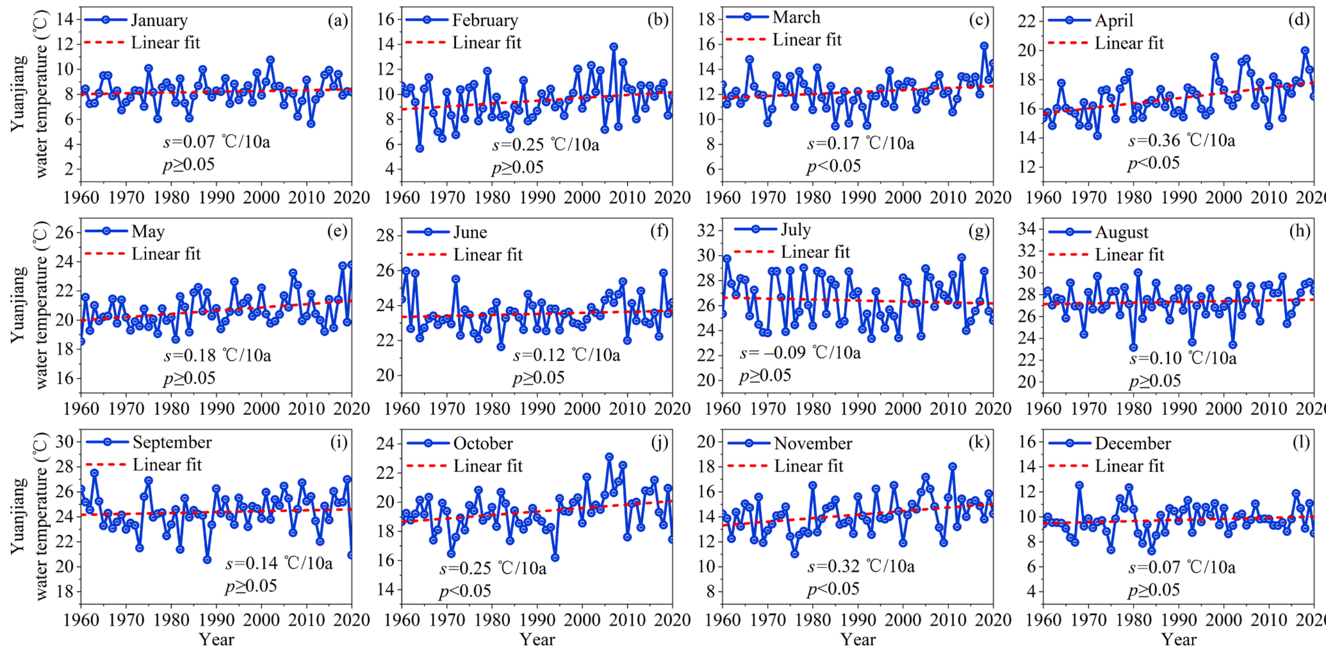

Fig. 10. Temporal dynamics of monthly mean water temperature in the Yuanjiang River.

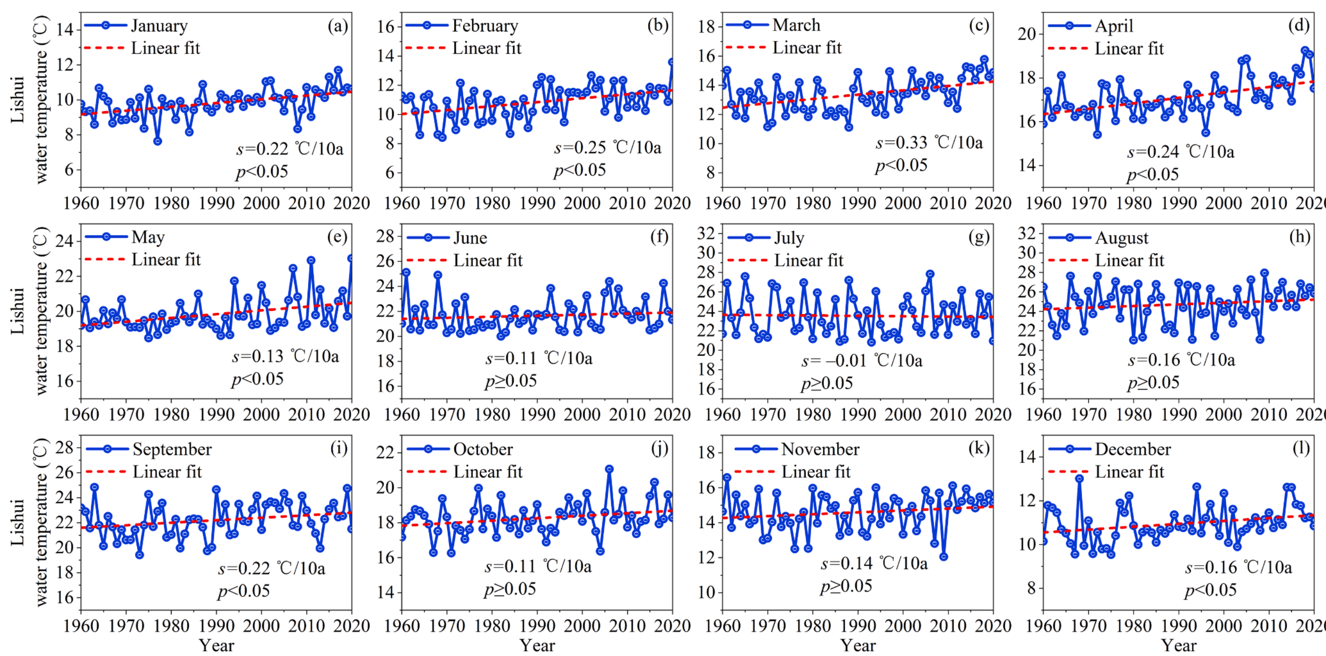

Fig. 11. Temporal dynamics of monthly mean water temperature in the Lishui River.

of 0.33 ℃ per decade (Fig. 11). The TFPW-MK test detected significant increasing trends in the water temperatures for January, February, March, April, May, September, and December. In these seven months, the average warming rate of water temperature reached 0.22 ℃ per decade, contributing to the noticeable increase in the annual mean water temperature of the Lishui River. The water temperature variations in the remaining months, i.e., June, July, August, October, and November, were not significant, confirmed by the TFPW-MK test.

These results suggested spatial and temporal heterogeneity in river water temperature variations, consistent with the findings from other regions (Wagner et al., 2017). Although a common characteristic was the warming water temperature as a more pronounced variation, the changing rates of water temperature varied across the four tributaries in the Dongting Lake Basin. Except for the heterogeneity across sub-basins, heterogeneity of tendency and changing rates existed across months. Some negative Sen's slope values indicated the cooling water temperature, but the TFPW-MK test rejected their significance. A homogeneity in water temperature variations was the significantly warming water temperature in the spring month of April in the four tributaries, although the warming rates differed. The results were inconsistent with the seasonal water temperature changes for river sites in the western US, where the water temperature notably increased during the fall and winter seasons (Wagner et al., 2017). The seasonal variations also differed from those in fifty-two catchments in Switzerland, where water temperature has significantly increased over the past five decades, with positive trends for all four seasons (Michel et al., 2020).

4.3. Periods embedded in monthly river water temperature series

Given the well-defined trends in the river water temperature series of the Dongting Lake Basin, the long-term reconstructed data were detrended and then analyzed using the wavelet transform method, resulting in the wavelet power spectrums (Fig. 12). In Fig. 12, cross-hatched regions on either end indicated the cone of influence, where edge effects became important. Higher wavelet power at a time scale implied more robust oscillations at the corresponding time scale and vice versa. A high wavelet power was identified at the one-year time scale, and this high power spectrum continued through the six decades. The thick dashed red contour enclosed less than 0.05 significance level regions, indicating the significant one-year period embedded in the monthly mean water temperature series. The river water temperature had pronounced seasonal components, with low water temperature occurring in winter when the air temperature reached a minimum and high water temperature occurring in summer (Fig. 6). This intrinsic intra-annual fluctuation induced a significant one-year period as identified by the wavelet transform. Except for the prominent one-year period, the wavelet transform did not suggest strong power at greater time scales. There was no indication of a significant period other than one year despite the considerable inter-annual variation of the monthly water temperature series.

4.4. Complexity changes of monthly river water temperature series

Fig. 13 displays the temporal evolution of water temperature complexity quantified using sample entropy. A higher sample entropy value indicates more complexity in water temperature series characterized by disorderliness, randomness, and irregularity. In contrast, a lower sample entropy value indicates more orderliness, regularity, and self-similarity. Linear regression and Sen's slopes suggested declined sample entropies of water temperatures in the Xiangjiang River, Zishui River, and Yuanjiang River. The TFPW-MK test identified the significance of these decreasing trends. There was a significant increasing trend in the sample entropy of water temperature in the Lishui River, as confirmed by the TFPW-MK test. The river water temperature complexity in the Dongting Lake Basin experienced inter-annual variations from 1960 to 2020. Significantly lowered water temperature complexity was detected in the Xiangjiang, Zishui, and Yuanjiang Rivers, while significantly enhanced water temperature complexity was suggested in the Lishui River.

4.5. Advantages, limitations, and future research

Water temperature is a vital attribute of riverine habitat, and assessing its alterations is one of the focal objectives of river engineering and management. This study proposed a framework to address the limitation of data scarcity in comprehensively evaluating riverine thermal regime alterations. The framework incorporated the LSTM network to simulate and extend the water temperature series and the time series analysis technologies to reveal the temporal dynamics of the thermal regime. The framework's application in the Dongting Lake Basin verified its practicality and suggested its potential in other basins. Given the development and prosperity of artificial intelligence, other deep learning technologies, except for the LSTM network, can also be employed to construct the water temperature models when applying the proposed framework in other basins. These machine-learning algorithms include but are not limited to a random forest, extreme gradient boosting, feed-forward neural networks, and wavelet neural networks (Abdi et al., 2021; Feigl et al., 2021; Piotrowski et al., 2015).

Based on a basin's physiographic characteristics and data availability, model inputs to be considered can include air temperature, precipitation, evaporation, global radiation, streamflow, landform attributes, and land cover attributes, to mention a few (DeWeber and Wagner, 2014; Feigl et al., 2021). The research objective and the time scale of available data determine the time scale of model output, e.g., hourly, daily, monthly, and annual simulations. It is worth noting that more complex and advanced models do not necessarily outperform simple and popular neural networks. The performance significance and choice of models should be verified (Piotrowski et al., 2015). The parameterized and validated water temperature model produces long-term water temperature data, an essential supplement to the missing historical data. The reconstructed data provide a critical basis for assessing thermal regime alterations regarding trend, periodicity, and complexity and exploring the driving mechanisms behind the thermal regime alterations.

River water temperature is a product of multiple environmental factors and is also impacted by anthropogenic perturbations, such as thermal pollution and deforestation (Caissie, 2006; Tao et al., 2020). The present study was limited in not incorporating anthropogenic factors into water temperature simulation and failed to distinguish the impact of climate change and human activities on water temperature variations. The LSTM network was trained and tested using the parallel water temperature data and its driving factors. The LSTM network's data requirement restrained the framework's applicability in the basins where the water temperature data are unavailable. As essentially a "black box" model, the LSTM network could not explain the physical

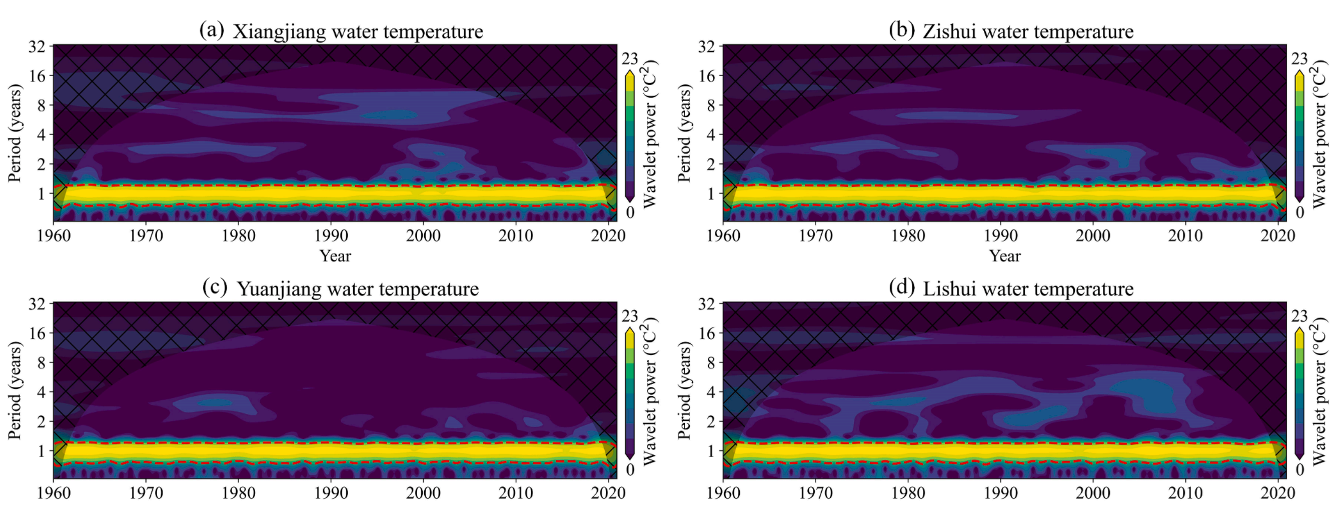

Fig. 12. Wavelet power spectrum of monthly mean river water temperature in the Dongting Lake Basin.

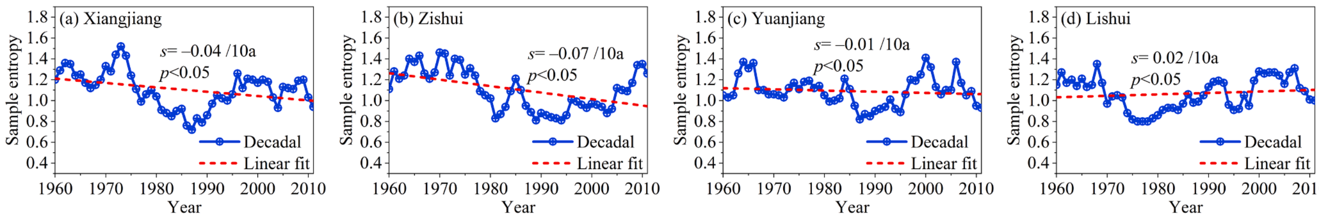

Fig. 13. Temporal variations in the complexity of monthly mean river water temperature in the Dongting Lake Basin.

mechanisms behind the water temperature evolution. Future efforts will improve water temperature simulation, identify riverine thermal regime alterations, and analyze causative mechanisms. Changes in river water temperature are anticipated to affect thermal habitat and fish population directly. For example, in the Fraser River Basin of British Columbia, the increase in average summer water temperature has doubled the number of days exceeding 20 ℃, the water temperature that, if exceeded, potentially increases the physiological stress of salmon during migration (Islam et al., 2019).

5. Conclusions

The proposed framework for reconstructing and investigating long-term water temperature data was practical in addressing the issue of observed data scarcity for riverine thermal assessment. The incorporation of the LSTM network and the trend, periodicity, and complexity algorithms resulted in an explicit framework structure whose components could be flexibly substituted by other algorithms of machine learning and time series analysis. The framework structure was clear, and its performance was satisfactory, as suggested by the application in the Dongting Lake Basin, China.

In the Dongting Lake Basin, the LSTM network parameterized using the 2007 to 2020 data produced the river water temperature series from 1960 to 2020. The produced water temperature series can be a good surrogate for investigating long-term riverine thermal dynamics. The combination of linear regression, the TFPW-MK test, and Sen's slope revealed the trend in water temperature over six decades. In the four tributaries, namely Xiangjiang River, Zishui River, Yuanjiang River, and Lishui River, the annual water temperature increased significantly, with an average warming rate of 0.15 ℃ per decade. Heterogeneity of tendency and changing rates existed across sub-basins and months, with the warming water temperature demonstrating a more pronounced variation. In the spring month of April, the river water temperature commonly increased significantly, while there was no indication of a significant declining trend in other months. For the four tributaries, wavelet transform unraveled a significant oscillation over one year from 1960 to 2020 and did not suggest prominent inter-annual periodic fluctuations. Moving sample entropy combined with the linear regression and TFPW-MK test indicated lowered water temperature complexity in the Xiangjiang, Zishui, and Yuanjiang Rivers. It suggested enhanced water temperature complexity in the Lishui River. The practicality and flexibility of the proposed framework proved its potential to be applied in other basins for long-term river water temperature simulation and analysis.

CRediT authorship contribution statement

Feng Huang: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Bao Qian:** Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization. **Carlos G. Ochoa:** Methodology, Resources, Software, Supervision, Validation, Writing – review & 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

The authors do not have permission to share data.

Acknowledgments

This research was funded by the Water Conservancy Science and Technology Project of Hunan Province, grant number XSKJ2021000-03.

Data availability statement

The water temperature and streamflow data that support the findings of this study are available from the Changjiang Water Resources Commission and the Hydrological Bureau of Hunan Province. The air temperature and evaporation data that support the findings of this study are available from China Meteorological Data Service Centre (http://data.cma.cn/en). Restrictions apply to the availability of these data, which were used under license for this study.

References

Abdi, R., Rust, A., Hogue, T.S., 2021. Development of a multilayer deep neural network model for predicting hourly river water temperature from meteorological data. Front. Environ. Sci. 9, 738322 https://doi.org/10.3389/fenvs.2021.738322.

Beaufort, A., Moatar, F., Sauquet, E., Loicq, P., Hannah, D.M., 2020. Influence of landscape and hydrological factors on stream-air temperature relationships at regional scale. Hydrol. Process. 34 (3), 583–597. https://doi.org/10.1002/hyp.13608.

Bonacina, L., Fasano, F., Mezzanotte, V., Fornaroli, R., 2022. Effects of water temperature on freshwater macroinvertebrates: a systematic review. Biol. Rev. https://doi.org/10.1111/brv.12903.

Caissie, D., 2006. The thermal regime of rivers: a review. Freshw. Biol. 51 (8), 1389–1406. https://doi.org/10.1111/j.1365-2427.2006.01597.x.

Chen, D.J., Hu, M.P., Guo, Y., Dahlgren, R.A., 2016. Changes in river water temperature between 1980 and 2012 in Yongan watershed, eastern China: magnitude, drivers and models. J. Hydrol. 533, 191–199. https://doi.org/10.1016/j.jhydrol.2015.12.005.

DeWeber, J.T., Wagner, T., 2014. A regional neural network ensemble for predicting mean daily river water temperature. J. Hydrol. 517, 187–200. https://doi.org/10.1016/j.jhydrol.2014.05.035.

Feigl, M., Lebiedzinski, K., Herrnegger, M., Schulz, K., 2021. Machine-learning methods for stream water temperature prediction. Hydrol. Earth Syst. Sci. 25 (5), 2951–2977. https://doi.org/10.5194/hess-25-2951-2021.

Hannah, D.M., Garner, G., 2015. River water temperature in the United Kingdom: changes over the 20th century and possible changes over the 21st century. Progr. Phys. Geogr.-Earth Environ. 39 (1), 68–92. https://doi.org/10.1177/0309133314550669.

Hebert, C., Caissie, D., Satish, M.G., El-Jabi, N., 2011. Study of stream temperature dynamics and corresponding heat fluxes within Miramichi River catchments (New Brunswick, Canada). Hydrol. Process. 25 (15), 2439–2455. https://doi.org/10.1002/hyp.8021.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Huang, F., Xia, Z.Q., Zhang, N., Zhang, Y.D., Li, J., 2011. Flow-complexity analysis of the upper reaches of the Yangtze River, China. J. Hydrol. Eng. 16 (11), 914–919. https://doi.org/10.1061/(asce)he.1943-5584.0000392.

Huang, F., Ochoa, C.G., Chen, X., Zhang, D.R., 2021a. Modeling oasis dynamics driven by ecological water diversion and implications for oasis restoration in arid endorheic basins. J. Hydrol. 593, 125774 https://doi.org/10.1016/j.jhydrol.2020.125774.

Huang, F., Ochoa, C.G., Guo, L.D., Wu, Y., Qian, B., 2021b. Investigating variation characteristics and driving forces of lake water level complexity in a complex river-lake system. Stoch. Env. Res. Risk A. 35 (5), 1003–1017. https://doi.org/10.1007/s00477-020-01907-9.

Hussain, M.M., Mahmud, I., 2019. pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. J. Open Source Software 4 (39), 1556. https://doi.org/10.21105/joss.01556.

Islam, S.U., Hay, R.W., Dery, S.J., Booth, B.P., 2019. Modelling the impacts of climate change on riverine thermal regimes in western Canada's largest Pacific watershed. Sci. Rep. 9, 11398. https://doi.org/10.1038/s41598-019-47804-2.

Jiang, T., Zhang, Q., Zhu, D.M., Wu, Y.J., 2006. Yangtze floods and droughts (China) and teleconnections with ENSO activities (1470–2003). Quat. Int. 144, 29–37. https://doi.org/10.1016/j.quaint.2005.05.010.

Kedra, M., Wiejaczka, L., 2018. Climatic and dam-induced impacts on river water temperature: assessment and management implications. Sci. Total Environ. 626, 1474–1483. https://doi.org/10.1016/j.scitotenv.2017.10.044.

Kim, M., Yang, Y., Kim, J., 2020. Sea surface temperature and high water temperature occurrence prediction using a long short-term memory model. Remote Sens. (Basel) 12 (21), 3654. https://doi.org/10.3390/rs12213654.

Lau, K.-M., Weng, H.Y., 1995. Climate signal detection using wavelet transform: how to make a time series sing. Bull. Am. Meteorol. Soc. 76 (12), 2391–2402. https://doi.org/10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.

Li, Z.W., Zhang, Y.K., 2008. Multi-scale entropy analysis of mississippi river flow. Stoch. Env. Res. Risk A. 22 (4), 507–512. https://doi.org/10.1007/s00477-007-0161-y.

Liang, L.Q., Li, L.J., Liu, C.M., Cuo, L., 2013. Climate change in the Tibetan Plateau Three Rivers Source Region: 1960–2009. Int. J. Climatol. 33 (13), 2900–2916. https://doi.org/10.1002/joc.3642.

Liang, J., Yi, Y.R., Li, X.D., Yuan, Y.J., Yang, S.H., Li, X., Zhu, Z.Q., Lei, M.Q., Meng, Q.F., Zhai, Y.Q., 2021. Detecting changes in water level caused by climate, land cover and dam construction in interconnected river-lake systems. Sci. Total Environ. 788, 147692 https://doi.org/10.1016/j.scitotenv.2021.147692.

Michel, A., Brauchli, T., Lehning, M., Schaefli, B., Huwald, H., 2020. Stream temperature and discharge evolution in Switzerland over the last 50 years: annual and seasonal behaviour. Hydrol. Earth Syst. Sci. 24 (1), 115–142. https://doi.org/10.5194/hess-24-115-2020.

Ouellet, V., St-Hilaire, A., Dugdale, S.J., Hannah, D.M., Krause, S., Proulx-Ouellet, S., 2020. River temperature research and practice: recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. Sci. Total Environ. 736, 139679 https://doi.org/10.1016/j.scitotenv.2020.139679.

Piccolroaz, S., Calamita, E., Majone, B., Gallice, A., Siviglia, A., Toffolon, M., 2016. Prediction of river water temperature: a comparison between a new family of hybrid models and statistical approaches. Hydrol. Process. 30 (21), 3901–3917. https://doi.org/10.1002/hyp.10913.

Piotrowski, A.P., Napiorkowski, M.J., Napiorkowski, J.J., Osuch, M., 2015. Comparing various artificial neural network types for water temperature prediction in rivers. J. Hydrol. 529, 302–315. https://doi.org/10.1016/j.jhydrol.2015.07.044.

Pohle, I., Helliwell, R., Aube, C., Gibbs, S., Spencer, M., Spezia, L., 2019. Citizen science evidence from the past century shows that Scottish rivers are warming. Sci. Total Environ. 659, 53–65. https://doi.org/10.1016/j.scitotenv.2018.12.325.

Qiu, R.J., Wang, Y.K., Wang, D., Qiu, W.J., Wu, J.C., Tao, Y.W., 2020. Water temperature forecasting based on modified artificial neural network methods: two cases of the Yangtze River. Sci. Total Environ. 737, 139729 https://doi.org/10.1016/j.scitotenv.2020.139729.

Qiu, R.J., Wang, Y.K., Rhoads, B., Wang, D., Qiu, W.J., Tao, Y.W., Wu, J.C., 2021. River water temperature forecasting using a deep learning method. J. Hydrol. 595, 126016 https://doi.org/10.1016/j.jhydrol.2021.126016.

Rahmani, F., Shen, C.P., Oliver, S., Lawson, K., Appling, A., 2021. Deep learning approaches for improving prediction of daily stream temperature in data-scarce, unmonitored, and dammed basins. Hydrol. Process. 35 (11), e14400.

Read, J.S., Jia, X.W., Willard, J., Appling, A.P., Zwart, J.A., Oliver, S.K., Karpatne, A., Hansen, G.J.A., Hanson, P.C., Watkins, W., Steinbach, M., Kumar, V., 2019. Process-guided deep learning predictions of lake water temperature. Water Resour. Res. 55 (11), 9173–9190. https://doi.org/10.1029/2019wr024922.

Rehana, S., Mujumdar, P.P., 2011. River water quality response under hypothetical climate change scenarios in Tunga-Bhadra river, India. Hydrol. Process. 25 (22), 3373–3386. https://doi.org/10.1002/hyp.8057.

Sen, P.K., 1968. Estimates of regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63 (324), 1379–1389.

Shen, C.P., 2018. A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resour. Res. 54 (11), 8558–8593. https://doi.org/10.1029/2018wr022643.

Sivakumar, B., Singh, V.P., 2012. Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. Hydrol. Earth Syst. Sci. 16 (11), 4119–4131. https://doi.org/10.5194/hess-16-4119-2012.

Tao, Y.W., Wang, Y.K., Wang, D., Ni, L.L., Wu, J.C., 2020. A probabilistic modeling framework for assessing the impacts of large reservoirs on river thermal regimes - a case of the Yangtze River. Environ. Res. 183, 109221 https://doi.org/10.1016/j.envres.2020.109221.

Tao, Y.W., Wang, Y.K., Wang, D., Ni, L.L., Wu, J.C., 2021. A C-vine copula framework to predict daily water temperature in the Yangtze River. J. Hydrol. 598, 126430 https://doi.org/10.1016/j.jhydrol.2021.126430.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2.

Van Houdt, G., Mosquera, C., Napoles, G., 2020. A review on the long short-term memory model. Artif. Intell. Rev. 53 (8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1.

Wagner, T., Midway, S.R., Whittier, J.B., DeWeber, J.T., Paukert, C.P., 2017. Annual changes in seasonal river water temperatures in the Eastern and Western United States. Water 9 (2), 90. https://doi.org/10.3390/w9020090.

Wu, S.D., Wu, C.W., Lin, S.G., Wang, C.C., Lee, K.Y., 2013. Time series analysis using composite multiscale entropy. Entropy 15 (3), 1069–1084. https://doi.org/10.3390/e15031069.

Yang, L., Wang, L.C., Yu, D.Q., Yao, R., Li, C.A., He, Q.H., Wang, S.Q., Wang, L.Z., 2020. Four decades of wetland changes in Dongting Lake using Landsat observations during 1978–2018. J. Hydrol. 587, 124954 https://doi.org/10.1016/j.jhydrol.2020.124954.

Yu, Y.W., Mei, X.F., Dai, Z.J., Gao, J.J., Li, J.B., Wang, J., Lou, Y.Y., 2018. Hydromorphological processes of Dongting Lake in China between 1951 and 2014. J. Hydrol. 562, 254–266. https://doi.org/10.1016/j.jhydrol.2018.05.015.

Yue, S., Pilon, P., Cavadias, G., 2002a. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol. 259 (1–4), 254–271. https://doi.org/10.1016/s0022-1694(01)00594-7.

Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002b. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol. Process. 16 (9), 1807–1829. https://doi.org/10.1002/hyp.1095.

Zhao, Y.F., Zou, X.Q., Liu, Q., Yao, Y.L., Li, Y.L., Wu, X.W., Wang, C.L., Yu, W.W., Wang, T., 2017. Assessing natural and anthropogenic influences on water discharge and sediment load in the Yangtze River, China. Sci. Total Environ. 607, 920–932. https://doi.org/10.1016/j.scitotenv.2017.07.002.

Zhu, S.L., Piotrowski, A.P., 2020. River/stream water temperature forecasting using artificial intelligence models: a systematic review. Acta Geophys. 68 (5), 1433–1442. https://doi.org/10.1007/s11600-020-00480-7.