

RESEARCH ARTICLE | AUGUST 21 2024

A deep learning perspective on meteorological droughts prediction in the Mun River Basin, Thailand

Usa Wannasingha Humphries ; Muhammad Waqas ; Phyothandar Hliang; Porntip Dechpichai; Angkool Wangwongchai 



AIP Advances 14, 085026 (2024)

<https://doi.org/10.1063/5.0209709>



View
Online



Export
Citation

Articles You May Be Interested In

Efficient and consistent adaptive mesh generation for geophysical models: A case study over the Gulf of Thailand

AIP Advances (May 2024)

Variation and prediction of rainy season in Thailand using ensemble neural model

AIP Conf. Proc. (December 2019)

Droughts and climate change

AIP Conference Proceedings (June 1992)



A deep learning perspective on meteorological droughts prediction in the Mun River Basin, Thailand

Cite as: AIP Advances 14, 085026 (2024); doi: 10.1063/5.0209709

Submitted: 22 March 2024 • Accepted: 14 June 2024 •

Published Online: 21 August 2024



View Online



Export Citation



CrossMark

Usa Wannasingha Humphries,¹ Muhammad Waqas,^{2,3} Phyothandar Hliang,^{2,3} Porntip Dechpichai,¹ and Angkool Wangwongchai^{1,a)}

AFFILIATIONS

¹ Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10140, Thailand

²The Joint Graduate School of Energy and Environment (JGSEE), King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10140, Thailand

³Center of Excellence on Energy Technology and Environment (CEE), Ministry of Higher Education, Science, Research and Innovation, Bangkok, Thailand

^{a)}Author to whom correspondence should be addressed: angkool.wan@kmutt.ac.th

ABSTRACT

Accurate drought prediction is crucial for enhancing resilience and managing water resources. Developing robust forecasting models and understanding the variables influencing their outcomes are essential. This study developed models that integrate wavelet transformation (WT) with advanced artificial intelligence (AI) models, increasing prediction accuracy. This study investigates the prediction of meteorological droughts using standalone bootstrapped random forest (BRF) and bi-directional long short-term memory (Bi-LSTM) models, compared to wavelet-decomposed hybrid models (WBRF, WBi-LSTM). These models were evaluated in the Mun River Basin, Thailand, utilizing monthly meteorological data (1993–2022) from the Thai Meteorological Department. The predictions were assessed using statistical metrics (R^2 , MAE, RMSE, and MAPE). For the Standardized Precipitation Index (SPI), the hybrid WBRF model consistently outperformed the standalone BRF across various metrics and timescales, demonstrating higher R^2 (0.89–0.97 for SPI-3) and lower error metrics (MAE: 0.144–0.21 for SPI-6, RMSE: 0.2–0.3 for SPI-12). Similarly, the hybrid WBi-LSTM model outperformed the standalone Bi-LSTM in SPI predictions, exhibiting higher R^2 (0.87–0.91 for SPI-3) and lower error metrics (MAE: 0.19–0.23 for SPI-6, RMSE: 0.27–0.81 for SPI-12) across all timescales. This trend was also observed for the China Z-index, Modified China Z-index, Hutchinson Drought Severity Index, and Rainfall Anomaly Index, where hybrid models achieved superior performance compared to standalone models. The WBi-LSTM model emerged as the preferred choice across different timespans. The integration of WT enhanced the predictive accuracy of hybrid models, making them effective tools for drought prediction.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0209709>

20 October 2024 06:20:00

I. INTRODUCTION

Drought is a natural occurrence characterized by continued periods of insufficient precipitation and discernible patterns, leading to substantial economic consequences. It primarily contributes to global economic losses attributed to natural calamities.¹ Droughts are exhibited across varying temporal scales, from weeks to decades. Their impact exceeds localities, affecting regions

from local to national scales and significantly affecting agricultural systems, water resources, and socioeconomic domains.² The implications of drought extend across diverse sectors, making formulating a universally applicable drought definition practically challenging. The complex impact of droughts causes complexities that confront a peculiar, comprehensive characterization.³ So, defining individual drought types based on their sector-specific impacts is

imperative.⁴ Throughout history, efforts have been directed toward categorizing drought definitions into meteorological, agricultural, hydrological, and socioeconomic classifications.⁵ Drought impacts are typically observed in agriculture and progressively expand to other water-dependent industries. Generally, drought is divided into three stages: meteorological, agricultural, and hydrological. Below-average rainfall defines meteorological drought during each period.⁶ A meteorological drought exists when precipitation is below average.⁷ Agricultural drought occurs when soil moisture is insufficient for a crop at a given period. Agricultural drought follows meteorological drought but before the hydrological drought.⁸ Surface and subsurface water shortages cause hydrological drought.⁹ Socioeconomic drought occurs when decreasing precipitation reveals water availability. Meteorological, hydrological, and agricultural droughts influenced the socioeconomic drought.¹⁰ Nevertheless, the scholarly discourse has arisen regarding the expansion of drought definitions to comprehend additional pivotal domains, including but not limited to groundwater,¹¹ ecological, and environmental aspects.^{4,12–14} This argument holds merit as it promises a more nuanced differentiation, enhancing our comprehension of drought propagation. Nevertheless, the consensus within the drought research community on including these additional aspects remains elusive.⁴ To effectively manage droughts, a proactive approach requires a nuanced understanding of past droughts and diverse conceptualizations. It involves analyzing the intricate relationships among precipitation, water demand, streamflow, temperature, reservoir capacity, and infrastructure management for optimized water distribution.¹⁵ A drought index is essential for evaluating and monitoring drought circumstances.⁵ Today's key drought indicators encompass the RVI (rainfall variability index), PDSI (Palmer drought severity index), SWSI (surface water supply index), effective drought index (EDI), and SSMAI (standardized soil moisture anomaly index). Despite their utility, these indices may fall short of offering comprehensive drought information due to the inherent variability in climatic conditions.¹⁶ Among meteorological drought indices are the SPI (standardized precipitation index),¹⁷ MCZI (the modified China Z index),¹⁸ and RAI (rainfall anomaly index).¹⁹ These indices monitor and assess multiscale drought scenarios, particularly in weather warming, by comprehensively pondering the mutual impacts of precipitation, runoff, and evapotranspiration.

Thailand stands out as one of the Asia-Pacific region's most heavily affected countries by drought, grappling with the persistent challenges of frequent drought events.^{20–22} Recurrent droughts severely affect Northeast Thailand's ecosystem and agriculture due to its inland location, which hinders rainfall, relies on few tropical cyclones, and faces water retention challenges with its sandy terrain.²³ Rice farming, a prominent agricultural activity in the region, covered 59 005 Km² in 2018. In the rice cultivation season, they demand significant water resources, emphasizing the crucial need for water management.²⁴ The significant impacts emphasize the constraint of creating robust forecasting drought models, facilitating the rapid development of strategies to mitigate the risks associated with drought.²⁵ Advancements in computing and artificial intelligence (AI) have facilitated the rapid development of machine and deep learning (ML and DL) models, which have demonstrated significant success in addressing intricate multivariate nonlinear problems.²⁶ This success is attributed to their inherent black-box characteristics.²⁷ In recent years, globally, various AI techniques,

including RNN (recurrent neural network), LSTM (long short-term memory), SVM (support vector machine), RF (random forest), DNN (deep neural networks), ANN (artificial neural network), CNN (convolutional neural network), and DT (decision tree), have been extensively employed in modeling of droughts.^{16,25,28–30} Notably, hybrid ML models exhibit superior performance, integrating variables decomposed through signal decomposition techniques such as wavelet transformation (WT).^{1,31} Despite efforts to categorize and understand droughts through various indices and classifications, the complexity of their impacts remains a hurdle.³² Existing drought indices offer valuable insights but may lack the comprehensive information needed due to climatic variability. Moreover, the intricate nature of drought propagation demands advanced forecasting models capable of handling multivariate nonlinear relationships.³³ Northeast Thailand, especially the Mun River Basin, faces recurrent drought events, significantly impacting its ecosystem and agriculture. There is a critical need to systematically characterize these droughts to better understand their dynamics and facilitate effective mitigation strategies.³⁴

A comprehensive understanding of drought characteristics and accurate forecasting models are essential for informed decision-making, enabling practical drought risk assessment, adaptation measures, and water resources planning.³⁵ This study introduces a new approach to drought prediction by integrating standalone and hybrid AI models, specifically bi-directional LSTM, bootstrapped RF, and hybrid models with wavelet decomposition, to forecast key drought indices (SPI, RAI, HDSI, CZI, and MZCI) over various timescales in the Mun River Basin, Northeast Thailand. This advanced methodology addresses the limitations of traditional drought indices by effectively capturing complex multivariate nonlinear relationships within the data. By incorporating advanced AI techniques and wavelet decomposition, our research significantly enhances prediction accuracy, offering a more nuanced understanding of drought dynamics. This comprehensive analysis bridges a critical knowledge gap, providing valuable insights for developing effective drought risk assessment and adaptation strategies. Finally, findings enable more informed decision-making for water resource management and agricultural planning in drought-prone regions, benefiting not only the Mun River Basin but also other areas facing similar environmental challenges. The study focuses on forecasting (3, 6, 9, and 12-month timescales), representing medium and long-term drought conditions. This will be helpful for a complete understanding of drought characteristics is crucial for informed decision-making, enabling practical drought risk assessment strategies for adaptation and planning of water resources in the Mun River Basin and areas worldwide sharing similar environmental conditions. The subsequent sections delve into the study area, data, and indices estimation method that provides a concise overview of the ML models combined with wavelet decomposition. Results are presented in Sec. III, while a detailed discussion and conclusion are presented in Sec. IV.

II. METHODS AND MATERIALS

A. Study area and data preprocessing

The Mun River Basin, situated in northeastern Thailand, spans latitudes 14° N to 16° N and longitudes 101°30' E to 105°30' E, covering an estimated drainage area of about 71 000 km² (see Fig. 1).³⁶

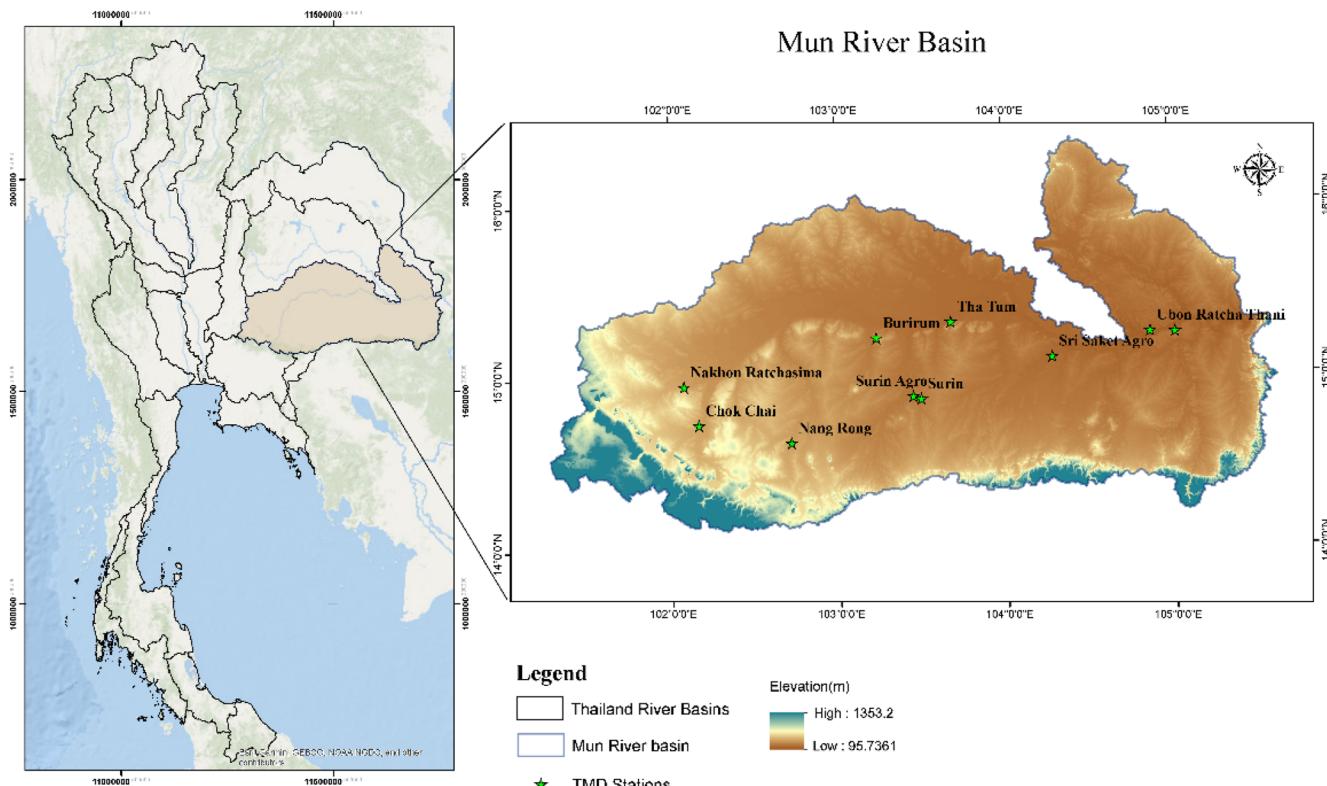


FIG. 1. Observation station in Mun River Basin territory in northeastern Thailand.

This basin features a diverse topography, with steep terrains and a plateau range located in the southwest, transitioning to expansive plains in its central and eastern regions.³⁷ The River Basin experiences a subtropical humid monsoon climate, characterized by distinct wet and dry seasons. Annual precipitation ranges between 1300 and 1500 mm, with ~90% of the total rainfall occurring during the wet season, primarily from May to October. The dry season, from November to April, accounts for the remaining 10% of the rainfall. Rainfall distribution within the basin increases progressively from west to east, with peak flows typically observed between September and October.³⁸ Temperature patterns in the basin are relatively stable, with the average yearly temperature consistently above 18 °C. The hottest month is April, while the coldest temperatures are usually recorded in January. The region's climate significantly influences agricultural practices, with rice, cassava, and sweet potato being the dominant crops. However, crop yields are highly susceptible to variations in monsoon patterns.³⁹ Data from ten meteorological stations within and surrounding the Mun River Basin, selected by the Thai Meteorology Department (TMD), were utilized in this study. The dataset includes daily average temperature (Tave), precipitation (PPT), and relative humidity (RH) from 1993 to 2022.

The overall data quality checking and preprocessing procedure is presented in Fig. 2. This study utilized the Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) method, as proposed by Wangwongchai (2023), to address any missing data.⁴⁰

Precipitation data can be compromised due to repeated observational or processing errors within a hydrological year, necessitating stringent checks for data consistency. To maintain data integrity, data from stations with more than ten consecutive days of missing values or exhibiting outliers were excluded from the analysis. When examining historical data, it is essential to consider the potential impacts of climate change, which can introduce irregular precipitation patterns that traditional statistical methods might classify as anomalies.⁴¹ To identify significant deviations from the typical

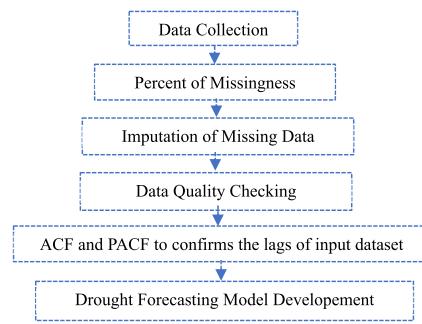


FIG. 2. Quality checking and data preprocessing procedure.

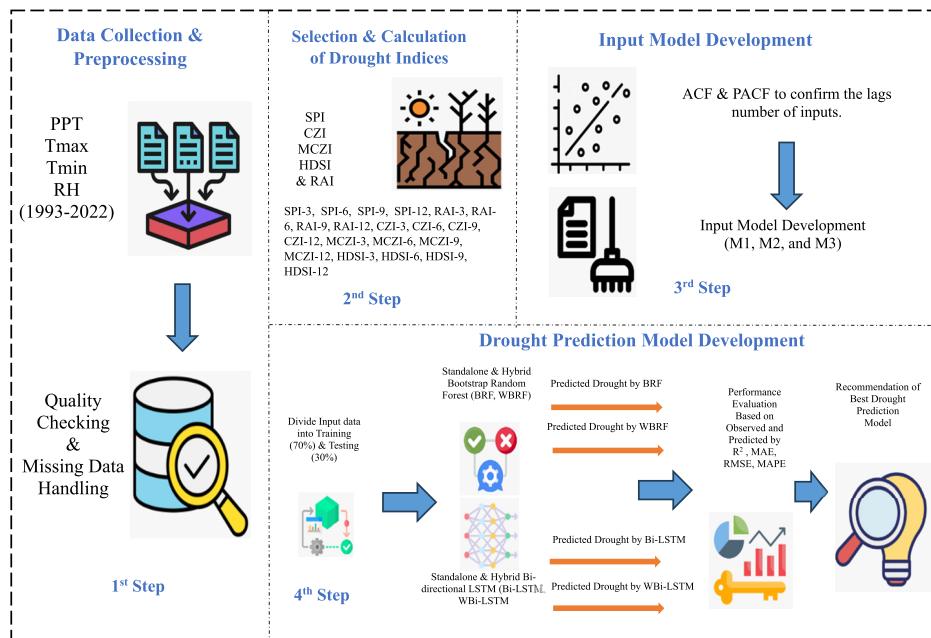


FIG. 3. The overall methodology employed in this study.

dataset range, this study employed the Grubbs and Beck technique.⁴² Thorough quality assessment and control procedures were applied to ensure data integrity, which involved removing extreme values, addressing data continuity issues, and interpolating missing data using methods proposed by relevant studies.⁴⁰ These procedures help maintain the accuracy and reliability of the meteorological indices used in the analysis. The overall methodology employed for predicting meteorological-based droughts is illustrated in Fig. 3.

B. Precipitation-based meteorological droughts determination

1. Standardized precipitation index (SPI)

SPI was introduced by McKee *et al.*; it is an instrumental mechanism for characterizing and monitoring drought phenomena.¹⁷ Notably, the SPI is the most commonly employed index among various drought assessment tools worldwide due to its reliance solely on precipitation data, computational simplicity, and high level of reliability.⁴³ SPI for different timesteps is calculated by employing cumulative monthly precipitation aggregates across various time intervals (like 3, 6, 9, and 12 months). For example, to compute the SPI for a three-month interval, the precipitation sums from the month (j to $j-2$) are attributed and aggregated of month j . Notably, the initial two months of the data time series are absent for this time scale. Subsequently, a normalization process ensues, wherein an appropriate probability density function is fitted to the long-term time series of accumulated precipitation. Ultimately, the cumulative distribution of data points, transformed into standardized normal variates, is determined using the fitted function. This sequential method is iteratively applied across all required time scales. The

calculated SPI for the Mun River basin for the years 1994–2022 is presented in Fig. 4.

2. Rainfall anomaly index (RAI)

The RAI, initially formulated¹⁹ and adopted by Ref. 44, incorporates a ranking process to give positive and negative anomalous magnitudes. It is divided into two phases: positive precipitation anomalies and negative precipitation anomalies. The application of the RAI proves valuable in mitigating drought effects on water resources, agriculture, and other sectors. RAI's flexibility is evident as it can be analyzed across diverse timescales, enhancing its adaptability to different contexts and facilitating comprehensive assessments of precipitation anomalies.⁴⁵ The RAI computed and assessed the occurrence or intensity of wet and dry years within the study regions. It can be expressed through the following equation:

$$RAI = \mp 3 \left[\frac{P_i - \bar{P}}{\bar{E} - \bar{P}} \right], \quad (1)$$

where P_i represents the sequence of measured precipitation at the time, \bar{P} signifies the precipitation average, and \bar{E} denotes the average of the ten extremes. The prefix ± 3 establishes the lower and upper bounds, delineating the range of anomalies. Specifically, this prefix imposes constraints to ensure that anomalies are confined within three standard deviations from the average, providing a standardized measure for assessing the deviation of precipitation values from the historical norm.⁴⁶ Figure 5 illustrates the estimated RAI indices for RAI-3, RAI-6, RAI-9, and RAI-12 over the past three decades within the basin.

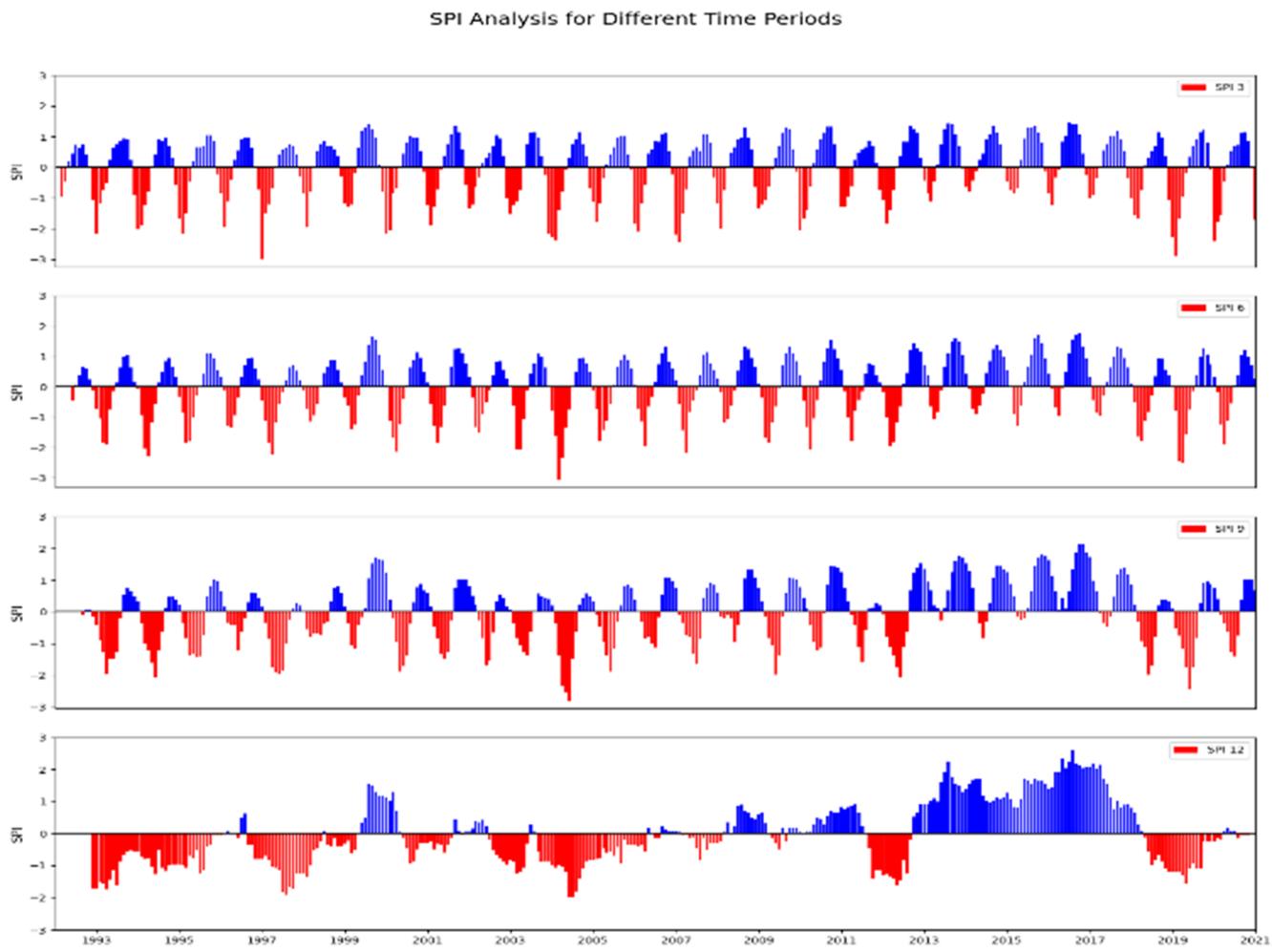


FIG. 4. Standardized Precipitation Index (SPI) analysis for 3, 6, 9, and 12-month time scales for the Mun River Basin.

C. China Z index (CZI) and modified China Z index (MCZI)

The CZI was introduced to monitor and assess drought severity in China. Initially designed for a one-month time scale,¹⁸ The CZI serves as a metric to gauge drought conditions in China, providing valuable information on the extent and intensity of drought events over a specific timeframe. CZI operates under the assumption that precipitation data adhere to the Pearson Type III distribution. This distribution is linked to the Wilson–Hilferty cube-root transformation,⁴⁷ which facilitates the conversion from a chi-square variable to the Z-scale, as elucidated by Ref. 48. It can be calculated by following formulas:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2)$$

$$C_s = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n * \sigma^3}, \quad (3)$$

$$\text{China Z Index} = \frac{6}{C_s} \left(\frac{C_s}{2} (Z - Score) - 1 \right)^{\frac{1}{3}} - \frac{6}{C_s} + \frac{C_s}{6}, \quad (4)$$

where “ C_s ” represents the coefficient of skewness, “ σ ” is the deviation standard of precipitation data, and “ n ” is the observation numbers. The calculation of the MCZI closely resembles that of the CZI, but instead of using the mean for the precipitation data, the median is utilized in the calculation. This modification reflects a shift in the central tendency measure, providing an alternative perspective on drought assessment by incorporating the median value. Figures 6 and 7 present the CZI indices and MCZI indices for the basin.

D. Hutchinson drought severity index (HDSI)

The HDSI stands out due to its specific calibration for agricultural drought.⁴⁹ This agricultural and meteorological index relies exclusively on data precipitation and computes drought conditions’

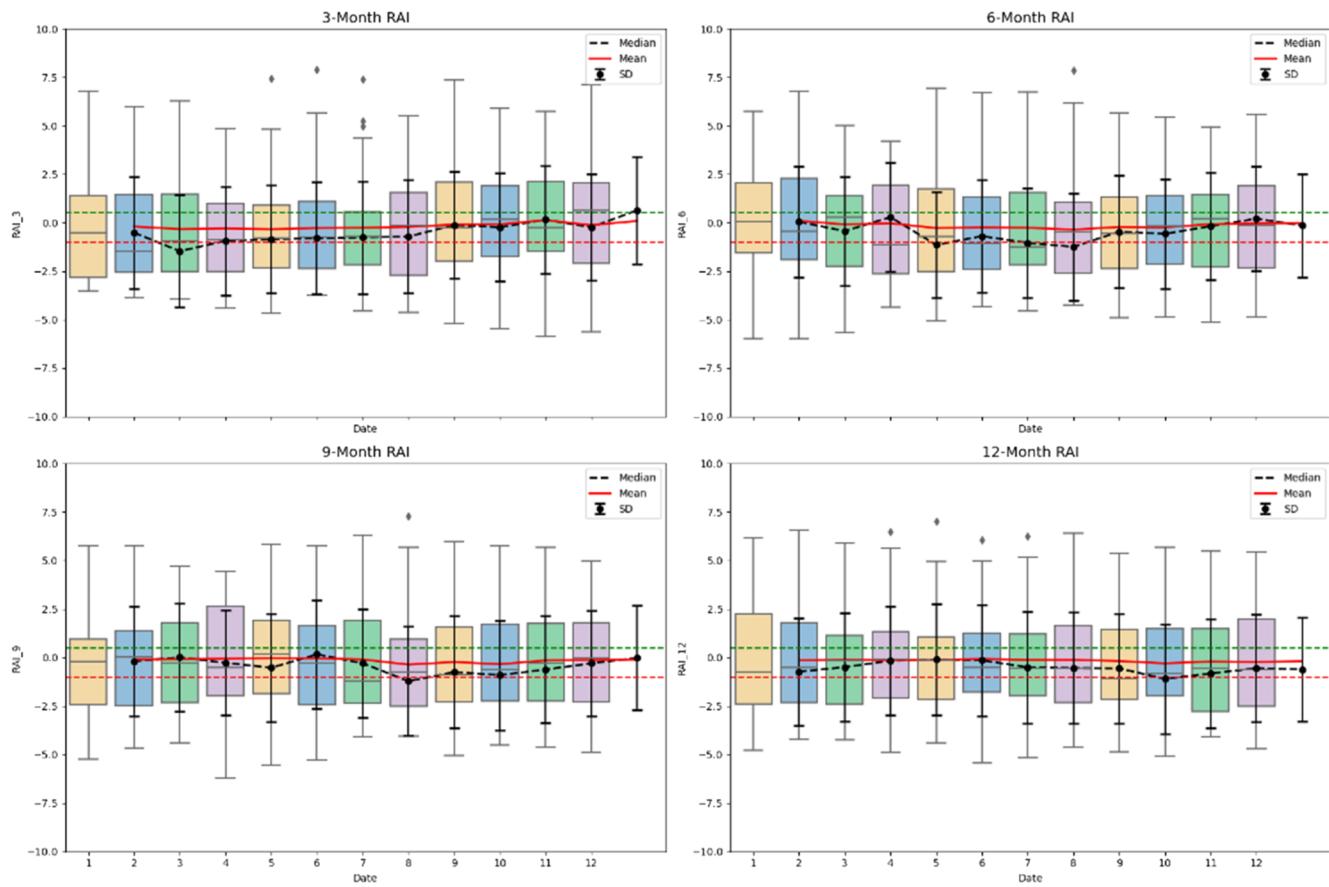


FIG. 5. Rainfall Anomaly Index (RAI) analysis for 3, 6, 9, and 12-month time scales for the Mun River Basin.

intensity and duration. As outlined by Smith *et al.* in 1992, the HDSI produces a pair of indicators, namely the Hutchinson “sum” and “count,” both derived from the calculation of the Hutchinson-score. It involves computing rainfall sums within specific windows at monthly intervals and expressing them in percentiles when associated with the rainfall totals for the corresponding windows sequence across the long-term record.⁴⁹ The calculated HDSI for different timespans (3, 6, 9, and 12 months) are illustrated in Fig. 8.

In this study, the partial autocorrelation function (PACF) and autocorrelation function (ACF) are employed to ascertain the number of inputs correlating with different antecedent values.⁵⁰ The ACF is a statistical metric for assessing the correlation between a time series and its delayed iterations. A statistically significant relationship is detected when the ACF graph surpasses the 95% strength limits. Figure 9 presents ACF and PACF graphs for SPI indices across 3-, 6-, 9-, and 12-time scales. As depicted in Fig. 9, the maximum autocorrelation coefficient is consistently observed at one lag time for all examined indicators and time scales. Mainly, there is an observable increase in the autocorrelation coefficient as the time scale extends from 3 to 12 months, rising from 0.1 to 1.0.

Similarly, ACF and PACF analyses were conducted for other hydrological indices, yielding insights into the temporal dependencies and relationships within the datasets.

Table I provides a comprehensive overview of the finalized optimal lagged input combinations. The first model, M1, entails the incorporation of lagged datasets representing various hydrological indices, including the SPI, RAI, CZI, MCZI, and HDSI. These datasets are integrated at different time lags, ranging from (t-1) to (t-10), to explore their influence on target indices. Model M2 extends the analysis by integrating the current month precipitation (PPT) alongside lagged precipitation datasets [P(t-1), P(t-2), P(t-3)], enabling a more nuanced understanding of precipitation dynamics and their impact on hydrological variables. Finally, Model M3 encompasses a combination of the (t-1) lagged dataset for the target index, along with current month precipitation (PPT), average temperature (T_{ave}), and relative humidity (RH). This multifaceted approach allows for a comprehensive examination of the interplay between various meteorological and hydrological variables, facilitating more accurate drought predictions.

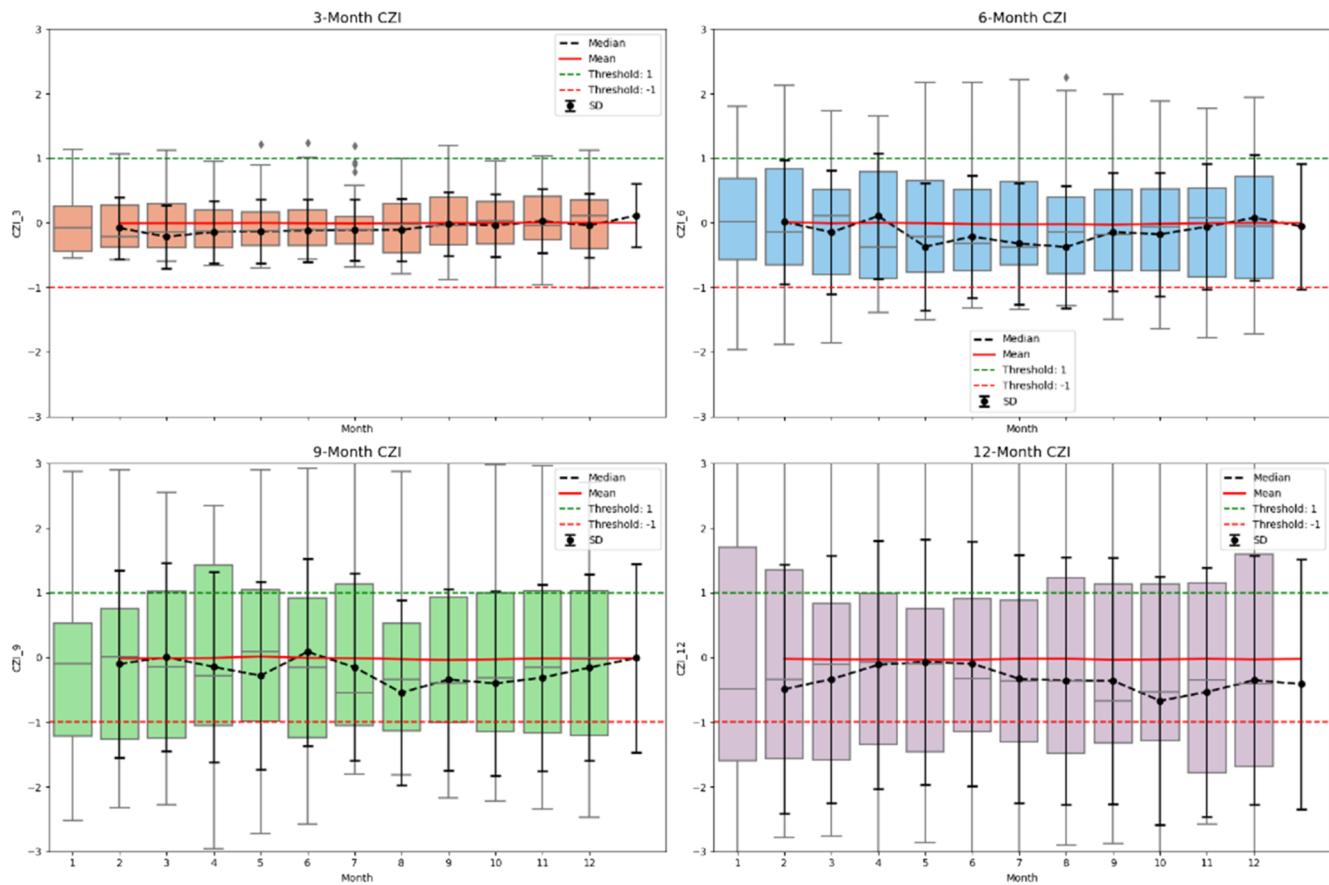


FIG. 6. China Z index (CZI) analysis for 3, 6, 9, and 12-month time scales for the Mun River Basin.

20 October 2024 06:20:00

E. Wavelet transformation

Many researchers have utilized Wavelet Transformations (WTs) to extract additional insights from time series datasets, introduced by Grossmann and Morlet in 1984.⁵¹⁻⁵⁹ WT enables analysis by providing both time and frequency information, offering a time-frequency representation.⁵¹ Unlike the Fourier transform, which utilizes infinite duration sinusoids, WT employs window functions with finite durations, making it suitable for non-stationary data analysis due to its varying scale. This variability allows for accurate analysis in both time and frequency domains, facilitating the decomposition of time series data into components with different resolutions. The mother wavelet, represented as $\psi(t)$, possesses finite energy and is pivotal in the mathematical representation of WT,

$$\int_{-\infty}^{\infty} \psi(t) dt = 0, \quad (5)$$

$$\text{Wavelet function} = \psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right). \quad (6)$$

The WT relies on two real number variables, “a” and “b,” representing scale or frequency and translation parameters, respectively.

The wavelet function $\psi(t)$ is modulated by these variables, with “a” controlling dilation or contraction across scales and “b” governing temporal displacement. Temporal data analysis involves translation, repositioning the wavelet along the temporal axis, and transformation, adjusting its scale to process data. Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) are employed for signal analysis, with CWT accommodating variable window sizes and DWT computing coefficients for dyadic scales and translations, ensuring efficient data analysis. The CWT overcomes the window size limitation of Short-Term Fourier Transformation (STFT), utilizing dilation and translation parameters “a” and “b” to analyze the signal frequency and temporal data.⁶⁰

$$\text{CWT}_{a,b}(t) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{a}} \psi^*\left(\frac{t-b}{a}\right) dt. \quad (7)$$

Its mathematical expression involves integrating the product of the wavelet function and the original data across the temporal range. Scale parameter “a” determines the wavelet scale, with smaller scales capturing high-frequency components and larger scales representing low-frequency components. CWT coefficients are visualized as contour maps, depicting similarity values from 0 to 1.

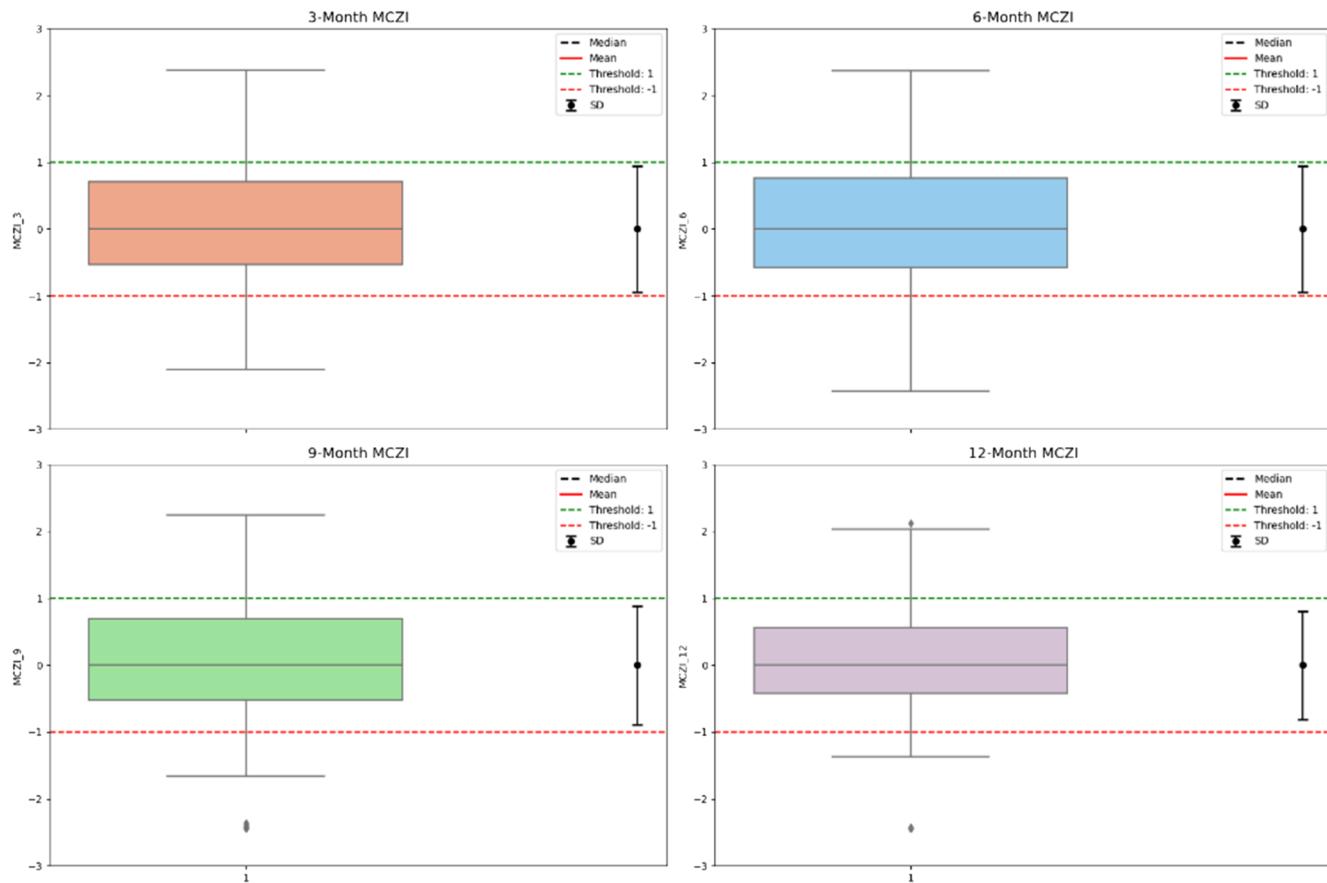


FIG. 7. Modified China Z index (CZI) analysis for 3, 6, 9, and 12-month time scales for the Mun River Basin.

20 October 2024 06:20:00

Conversely, the DWT decomposes the signal into approximation and detail sub-signals at different resolution levels,⁵⁶

$$\text{DWT}(m, n) = 2^{-\frac{m}{2}} \sum_{t=0}^{N-1} \psi * (2^{-m} - 2) xf(t). \quad (8)$$

Its mathematical expression involves convolving the wavelet function with the original data. The DWT utilizes integers "m" and "n" to represent scale and translation and decomposes the signal into large-scale, low-frequency components and low-scale, high-frequency components. Various wavelet families, including Haar, Daubechies, Coiflets, Symlets, and Meyer, offer different characteristics for wavelet analysis.⁵⁶ The Haar wavelet, known for its step-like behavior, is suitable for sudden changes in time series data but is non-differentiable due to its discontinuous nature.⁶¹ The Daubechies family comprises compact, orthonormal wavelets with different vanishing moments, while Coiflets are characterized by computational efficiency in transformation.^{53,54} Symlets offer minimal phase asymmetry, and Meyer wavelets are defined in the frequency domain, providing additional options for signal processing applications.^{62,63}

F. Standalone and hybrid Bi-directional LSTM

LSTM, proposed by Hochreiter and Schmidhuber in 1997,⁶⁴ has gained significant popularity, especially in addressing time series prediction tasks. As a variant of recurrent neural networks (RNNs), LSTM offers versatility across a wide range of problems and has seen extensive adoption.⁶⁵ Its effectiveness stems from its capability to handle the challenge of retaining information over long periods by integrating gate units and memory cells into the neural network architecture. These memory cells maintain cell states that store recently encountered data.²⁷ The LSTM model manipulates and updates the cell state upon input, retaining information for subsequent predictions. Its design enables the effective handling of tasks with long-term dependencies, distinguishing it from other recurrent neural network architectures. Within LSTM, gates control data flow into the cell state, combining sigmoid functions with multiplication processes to regulate data passage. Various gates, including input, forget, and output gates, facilitate data transfer between cells, ensuring effective information processing and retention.^{27,40}

Whereas, Bi-LSTMs represent an advancement over conventional LSTMs, incorporating both past and future states to enhance prediction accuracy.^{66,67} While ordinary LSTMs solely consider past observations, bidirectional LSTMs integrate information from both

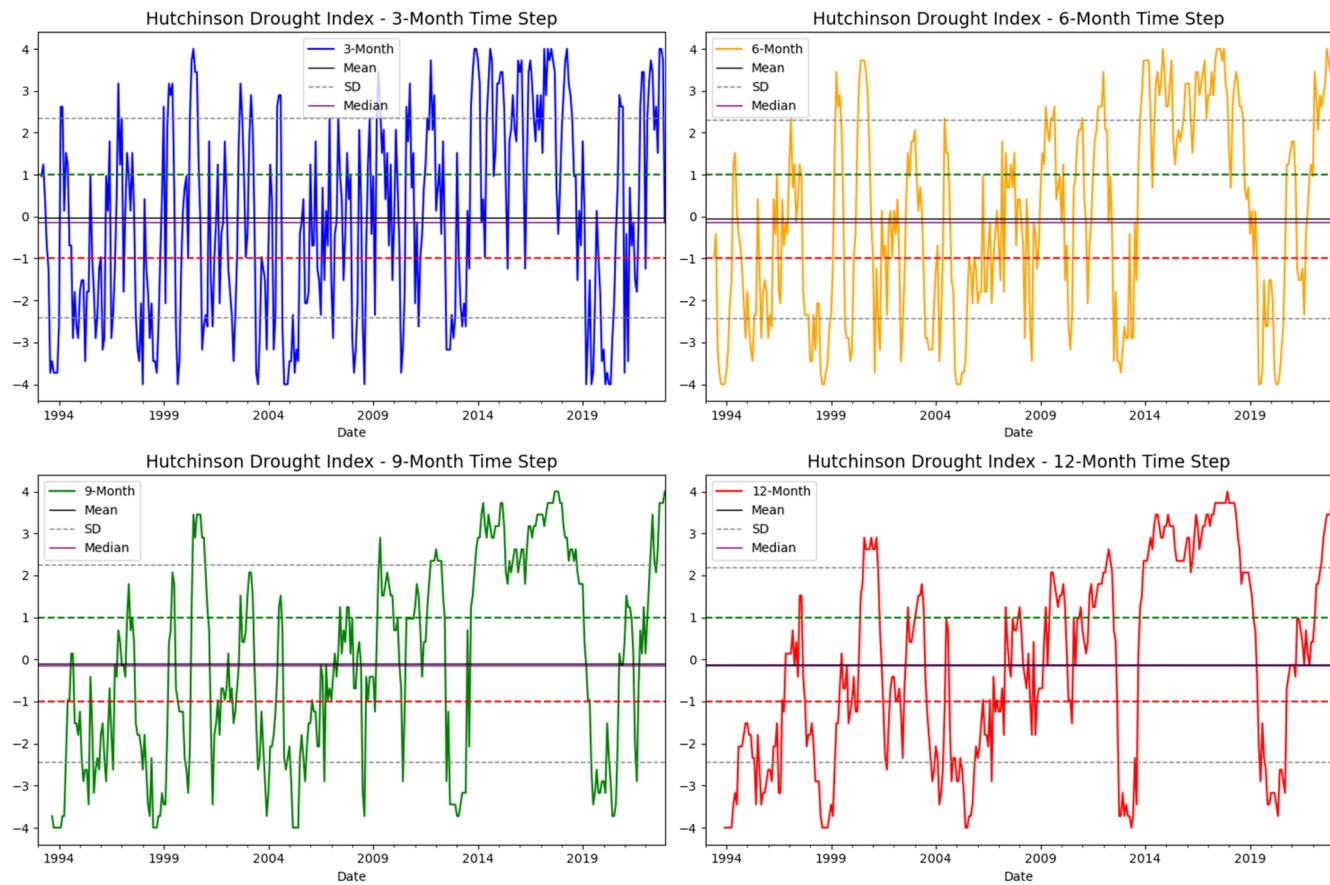


FIG. 8. Hutchinson drought severity index (HDSI) analysis for 3, 6, 9, and 12-month time scales for the Mun River Basin.

20 October 2024 06:20:00

past and future observations. Consequently, Bi-LSTMs analyze data in both forward and backward directions, unlike their ordinary LSTM counterparts, which process information solely in one direction.⁶⁸ The architecture of a bidirectional LSTM network featuring two hidden layers is illustrated in Fig. 10. In a Bi-LSTM setup, two LSTM networks are trained: the first network processes the input sequence in the forward direction, while the second network processes it in reverse with a mirrored copy of the input sequence. Consequently, Bi-LSTM networks offer superior learning capabilities. At each time step, the outputs of the forward and backward layers are combined. Bi-LSTMs are typically more efficient than ordinary LSTMs due to their ability to extract additional features during the backward pass, thereby enhancing the overall learning process.

In this study, the WT signal processing algorithm is leveraged for data decomposition to enhance model prediction accuracy through data denoising. Specifically, the DWT with Haar wavelet is employed after a complete evaluation of all wavelet families to decompose the input data, yielding one approximation component and multiple detail components based on specified levels. Bi-LSTM

networks, renowned for their ability to forecast time series, are integrated with a deep LSTM architecture featuring 100 neurons per hidden layer, and the “tanh” activation function was used. Employing Bi-LSTM, both low and high-frequency decomposition components are predicted. The resulting forecasting outcomes from each sub-series are aggregated to provide final projections, thereby enhancing reliability for practical applications. The systematic strategy for optimizing the decomposition-based model involves data splitting, preprocessing with decomposition techniques, model selection, and evaluation using statistical error indicators such as MAE, RMSE, MAPE, and R².

TensorFlow serves as the primary deep-learning framework in this study. Initially, the data undergoes preprocessing using min-max feature scaling. Subsequently, the processed dataset is partitioned into training and testing sets, with ~70% allocated for training and 30% for testing. Both WBi-LSTM and Bi-LSTM models are trained using various tuning parameters to predict drought. The predicted dataset is then compared with the actual dataset to evaluate prediction accuracy, as shown in Fig. 11.

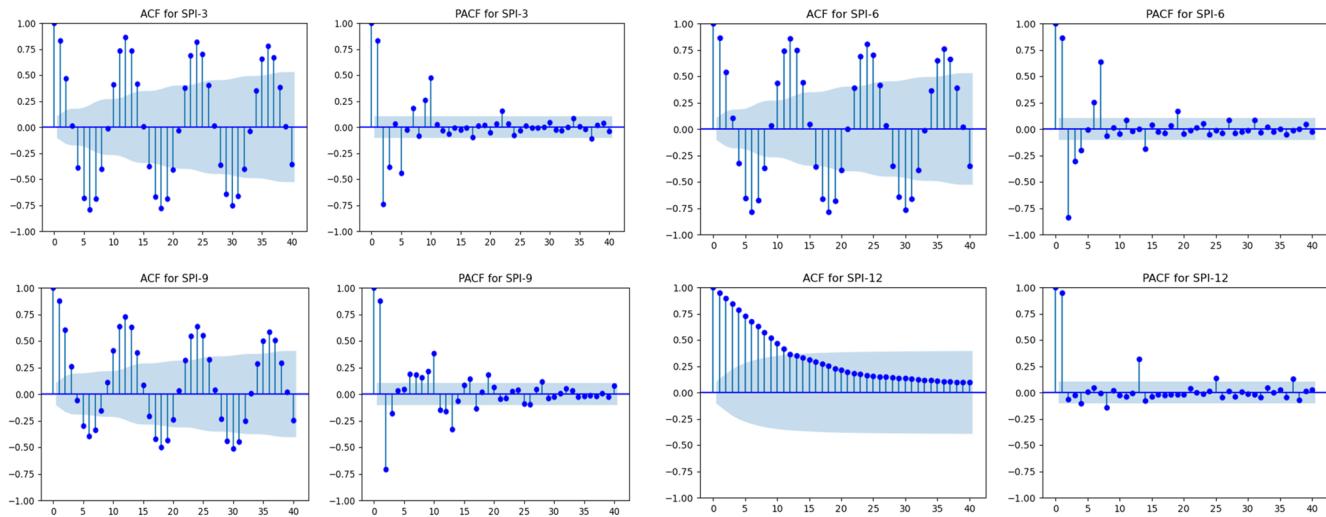


FIG. 9. Autocorrelation function (ACF) and the partial autocorrelation function (PACF) analysis for SPI in the selection of optimal lagged input data for the prediction models.

TABLE I. Final selected combinations for the predictions of SPI, RAI, CZI, MCZI, HDSI, and DI for the prediction models (standalone and hybrid AIs).

Combinations	Input variables	Targets for each index
M1	(t-1), (t-2), (t-3), (t-4), (t-5), (t-6), (t-7), (t-8), (t-9), (t-10)	SPI-3, SPI-6, SPI-9, SPI-12, RAI-3, RAI-6, RAI-9, RAI-12, CZI-3, CZI-6, CZI-9, CZI-12, MCZI-3, MCZI-6, MCZI-9, MCZI-12, HDSI-3, HDSI-6, HDSI-9, HDSI-12
M2	(t-1), (t-2), (t-3), (t-4), (t-5), (t-6), (t-7), (t-8), (t-9), (t-10), PPT, P(t-1), P(t-2), P(t-3)	
M3	(t-1), PPT (t), Tave (t), RH (t)	

G. Standalone and hybrid bootstrapped random forest

Random Forests (RFs) stand as a widely embraced machine learning paradigm, initially formulated by Breiman.⁶⁹ RF serves as a robust ensemble learning technique, notably enhancing the efficacy of classification and regression trees across diverse domains, including weather forecasting and hydrological analysis.^{26,27,70} Its capacity to adeptly manage high-dimensional data and intricate nonlinear relationships is coupled with its capability to furnish importance scores for input features.²⁷ In regression scenarios, RF constructs a training set from time series data by employing bootstrap sampling to generate multiple regression trees. The evaluation of tree performance is facilitated through out-of-bag data, with final predictions aggregated from all trees and accuracy assessed using mean squared error.⁷¹ The bootstrapped RF (BRF) integrates uncertainty into critical aspects such as data sampling and feature selection during the construction of individual decision trees within its ensemble. This process initiates with the random sampling of subsets from the dataset, followed by the construction of decision trees employing the Classification and Regression Tree (CART) algorithm. In the context of this study, meteorological data with two distinct input variable combinations were employed. Additionally, after applying

WT to the input data and DWT using the Haar wavelet family with three-level decomposition, each decision tree underwent meticulous hyperparameter tuning via grid search to ascertain optimal parameters. The operational mechanism of WBRF is delineated in Fig. 12. The comparison between the actual and predicted SPI-3 by the developed hybrid WBRF is presented in Fig. 13, highlighting the model's capability to capture well the relationship between the observed and the predicted drought; similar patterns were generated in other cases.

H. Evaluation metrics

This study utilized the R^2 , RMSE, MAE, and MAPE evaluation metrics to assess the capability of the prediction models by following formulas:

$$R^2 = 1 - \sum \frac{(I_{obs} - I_{pre})^2}{(I_{obs} - I_{avg})^2}, \quad (9)$$

$$MAE = \sum \frac{|I_{pre} - I_{obs}|}{N}, \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n |I_{obs} - I_{pre}| * 100, \quad (11)$$

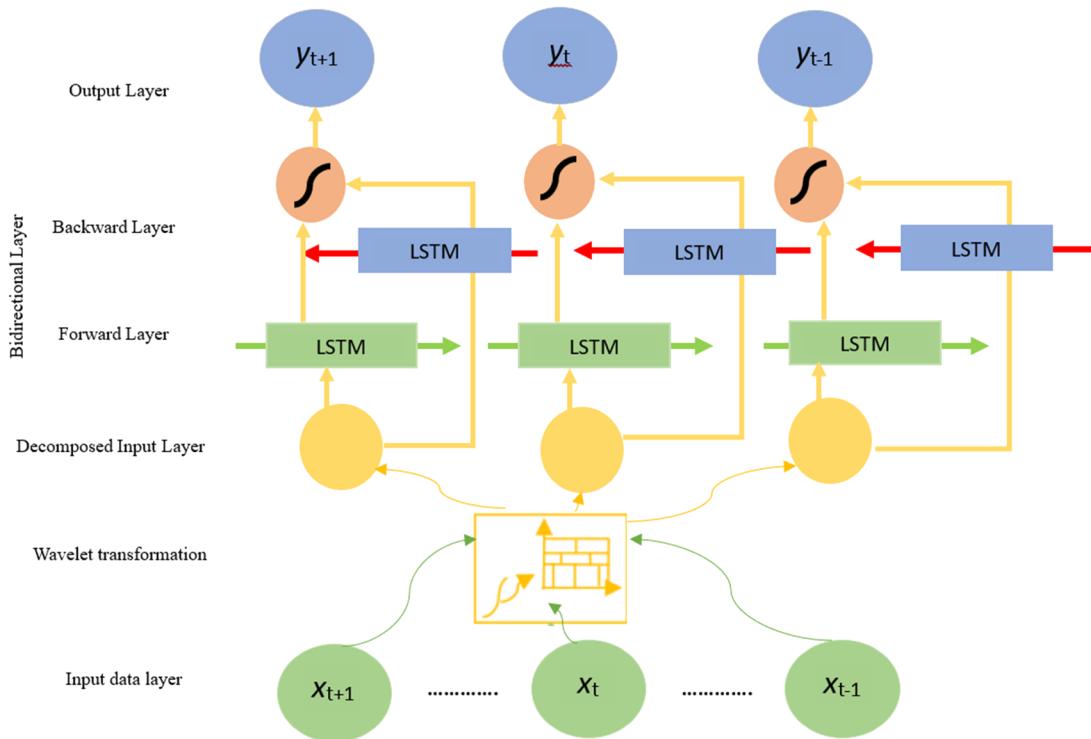


FIG. 10. Working mechanism of bidirectional LSTM.

$$\text{RMSE} = \frac{\sqrt{\sum_{i=1}^N (I_{\text{obs}} - I_{\text{pre}})^2}}{N}, \quad (12)$$

where I_{obs} indicates the observed, I_{pre} denotes the predicted index, and I_{ave} describes the observed index's average. The R^2 lies between 0 and 1, representing the strength of the link between predicted and observed values, with a value closer to 1 suggesting a more effective model. MAE indicates the proximity of calculated values to measured values, whereas MAPE measures predictive model accuracy, which typically ranges from 0% for a perfect forecast to more than 100% for significant forecast errors, with lower values indicating superior forecast accuracy.⁴⁰ The RMSE represents the average prediction errors in units of the dependent variable, ranging from 0 to ∞ , regardless of error direction.^{52,72}

III. RESULTS AND DISCUSSION

This study employed standalone bootstrapped random forest (BRF) and bi-directional long short-term memory (Bi-LSTM) for the prediction of different meteorological-based droughts (SPI, CZI, MCZI, HDSI, and RAI) and then compared them to wavelet decomposed hybrid models (WBRF, WB-LSTM). Their predictions were assessed by statistical metrics (R^2 , MAE, RMSE, and MAPE).

A. Standardized precipitation index (SPI)

First, the results from the analysis indicate that the hybrid model (WBRF) consistently outperforms the standalone BRF model

across various metrics and timespans for predicting the SPI, as shown in Fig. 14. For instance, for SPI-3 predictions, the R^2 for the hybrid model ranges from 0.89 to 0.97 across different input combinations, whereas the standalone BRF model achieves a lower R^2 ranging from 0.71 to 0.73. This trend persists across other timespans, indicating the superior ability of the hybrid model to capture the underlying patterns in precipitation data. Moving to MAE, which quantifies the average magnitude of errors, the hybrid WBRF consistently exhibits lower MAE values than the standalone BRF. For example, in SPI-6 predictions, the MAE values for the hybrid model range from 0.144 to 0.21, while the standalone BRF model yields MAE values ranging from 0.17 to 0.18. These results suggest that the hybrid WBRF model tends to make more accurate predictions, with smaller discrepancies between predicted and observed SPI values. Similarly, RMSE, which penalizes larger errors more heavily, reflects the superior predictive accuracy of the hybrid WBRF model. Across different timespans, the hybrid model consistently yields lower RMSE values compared to the standalone BRF model. For instance, in SPI-12 predictions the RMSE values for the hybrid model range from 0.2 to 0.3, while the standalone BRF model produces RMSE values ranging from 0.23 to 0.24. Finally, MAPE, which assesses prediction accuracy relative to observed values, shows that the hybrid WBRF model generally performs well, although there are instances where it exhibits higher MAPE values compared to the standalone BRF. For example, in SPI-9 predictions with input combination M-3, the hybrid model yields a MAPE value of 305.07, while the standalone BRF model achieves a lower MAPE value of

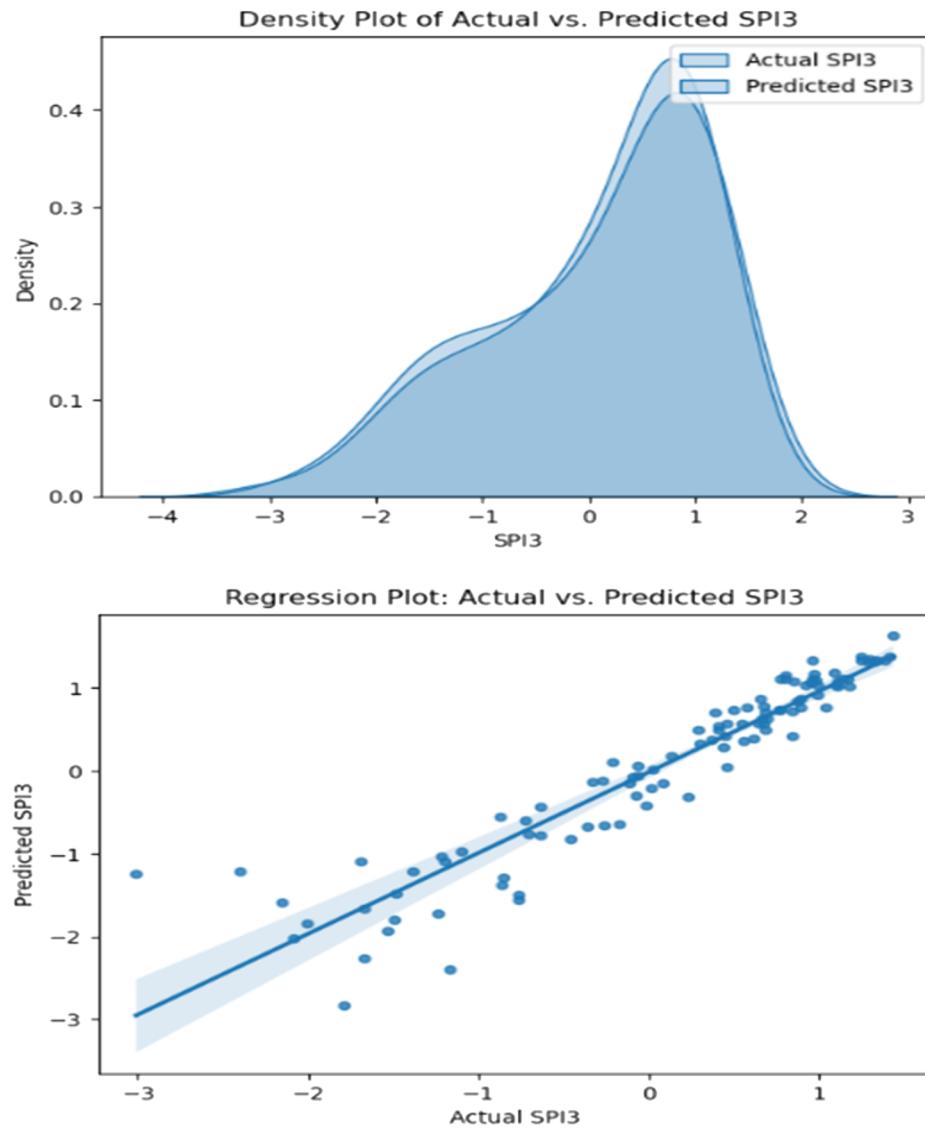


FIG. 11. Example between actual and predicted SPI-3 by hybrid WBi-LSTM.

222.14. Meanwhile, the hybrid WBRF model demonstrates superior performance in terms of R^2 , MAE, and RMSE across most scenarios.

Figure 15 presents the results for Bi-LSTM and WBi-LSTM. The performance metrics provide a quantitative assessment of the effectiveness of the standalone Bi-LSTM and hybrid WBi-LSTM models in predicting the SPI at different timescales. Across all timescales, the hybrid model consistently outperforms the standalone Bi-LSTM, as evidenced by higher R^2 and lower values of MAE, RMSE, and MAPE. For instance, at SPI-3, the R^2 for the hybrid model ranges from 0.87 to 0.91, while for the standalone Bi-LSTM, they range from 0.7 to 0.72. Similarly, the MAE values for the hybrid model are notably lower, ranging from 0.19 to 0.23, compared to the standalone Bi-LSTM, which ranges from 0.2 to 0.32.

This trend persists across all timescales, indicating the superior predictive performance of the hybrid model. The higher R^2 observed for the hybrid model suggests that it explains a greater proportion of the variance in SPI compared to the standalone Bi-LSTM. Moreover, the lower MAE, RMSE, and MAPE values for the hybrid model indicate reduced errors in SPI predictions compared to the standalone Bi-LSTM. For example, at SPI-6, the MAPE for the hybrid model ranges from 506.64 to 807.94, whereas for the standalone Bi-LSTM, it ranges from 158.11 to 469.97. Conversely, the standalone Bi-LSTM exhibits higher error metrics, particularly at longer timescales such as SPI-9 and SPI-12. For instance, at SPI-12, the R^2 for the hybrid model ranges from 0.35 to 0.89, while for the standalone Bi-LSTM, they range from 0.12 to 0.71.

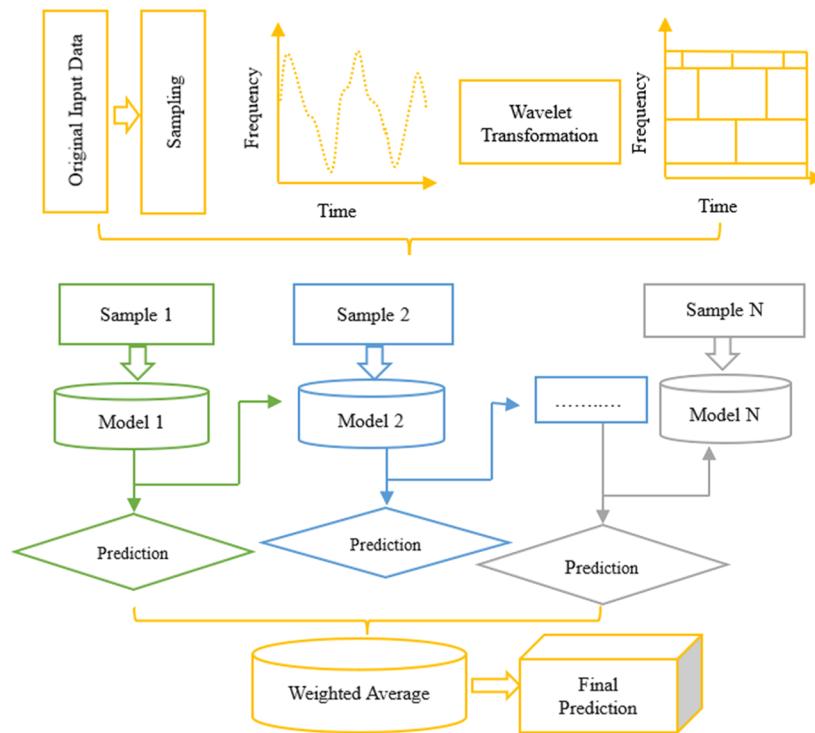


FIG. 12. Working mechanism of hybrid bootstrapped random forest (WBRF).

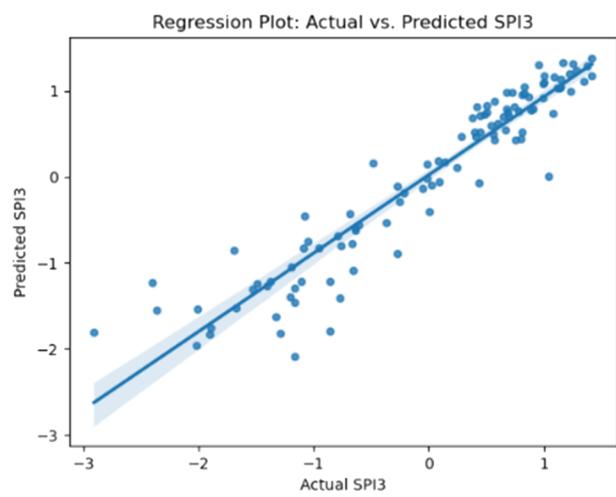
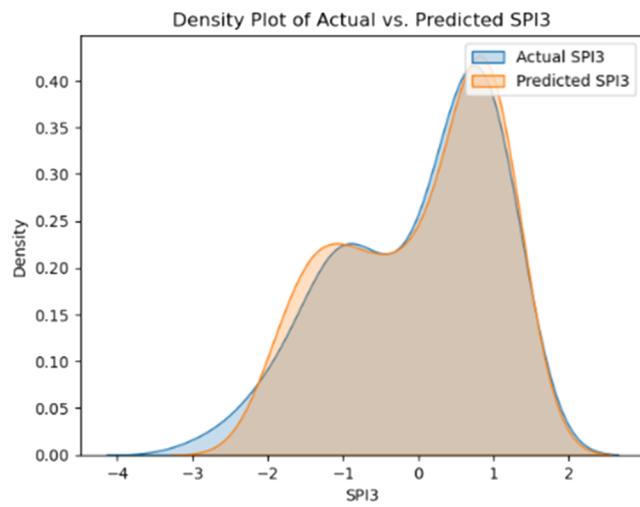


FIG. 13. Example between actual and predicted SPI-3 by hybrid WBRF.

Overall, from Figs. 14 and 15, it can be observed that the hybrid WBRF and WBi-LSTM models outperform the standalone BRF and Bi-LSTM models due to their ability to capture multi-scale temporal patterns in precipitation data. Wavelet decomposition enables the hybrid model to adaptively analyze high-frequency fluctuations and low-frequency trends, enhancing its

predictive accuracy across various timescales. Conversely, the standalone BRF model lacks this adaptive capability, resulting in inferior performance, particularly at longer timescales. The integration of wavelet decomposition with Bi-LSTM facilitates the modeling of intricate temporal dependencies, leading to more robust SPI predictions.

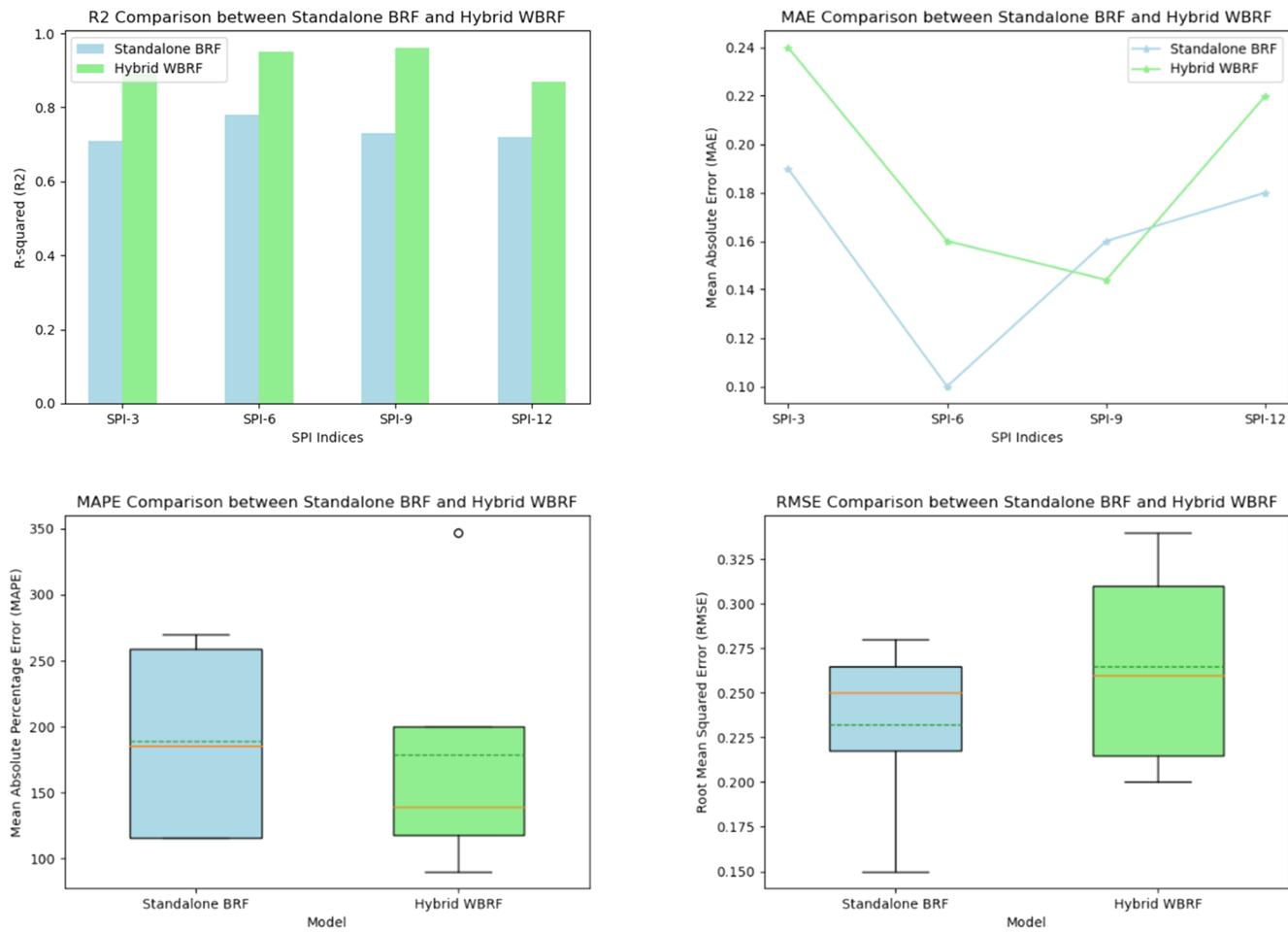


FIG. 14. Comparison of SPI prediction results for monthly time scales (SPI-3, SPI-6, SPI-9, and SPI-12) by wavelet decomposition (WBRF) and the standalone BRF models.

20 October 2024 06:20:00

B. China Z-index (CZI)

Like SPI, hybrid WBRF and standalone BRF models were used to predict the CZI for similar timespans, as shown in Fig. 16. For the CZI-3 timespan, the hybrid WBRF model achieves R^2 (0.41–0.84), indicating a moderate to strong fit of the model to the data. In contrast, the standalone BRF model exhibits lower R^2 (0.35–0.71), suggesting a weaker fit. Furthermore, the hybrid WBRF model consistently yields lower MAE (0.55–0.78), compared to the standalone BRF model with MAE (0.69–0.98). Similarly, the hybrid WBRF model produces a lower RMSE (0.67–0.88), compared to the standalone BRF model with RMSE (0.84–1.06). For the CZI-6 timespan, the hybrid WBRF model continues to outperform the standalone BRF model across all metrics. Specifically, the hybrid WBRF model achieves R^2 from 0.71 to 0.81, indicating a strong fit of the model to the data. In contrast, the standalone BRF model exhibits slightly lower R^2 (0.60–0.69). Additionally, the hybrid WBRF model demonstrates lower MAE values, ranging from 0.34 to 0.41, compared to the standalone BRF model with MAE (0.43–0.51). Similarly, the hybrid WBRF model produces lower

RMSE (0.43–0.50), compared to the standalone BRF model with RMSE values ranging from 0.54 to 0.63. In the prediction of CZI-9 and CZI-12 timespans, the hybrid WBRF model maintains its superiority over the standalone BRF model. Specifically, the hybrid WBRF model achieves higher R^2 values, lower MAE values, and lower RMSE values across both timespans compared to the standalone BRF model. For instance, in the CZI-9 timespan, the hybrid WBRF model achieves R^2 (0.80–0.82), while the standalone BRF model achieves R^2 (0.72–0.74). Similarly, in the CZI-12 timespan, the hybrid WBRF model achieves R^2 (0.75–0.78), while the standalone BRF model achieves R^2 (0.67–0.69). These consistent trends across different timespans affirm the robustness and reliability of the hybrid WBRF model in accurately predicting CZI for different timespans.

On the other hand, the hybrid WBi-LSTM model achieves R^2 (0.29–0.73), indicating a moderate to strong fit of the model to the data (Fig. 17). Comparatively, the standalone Bi-LSTM model exhibits slightly lower R^2 (0.26–0.66). Moreover, the hybrid WBi-LSTM model consistently yields lower MAE and RMSE ranges from

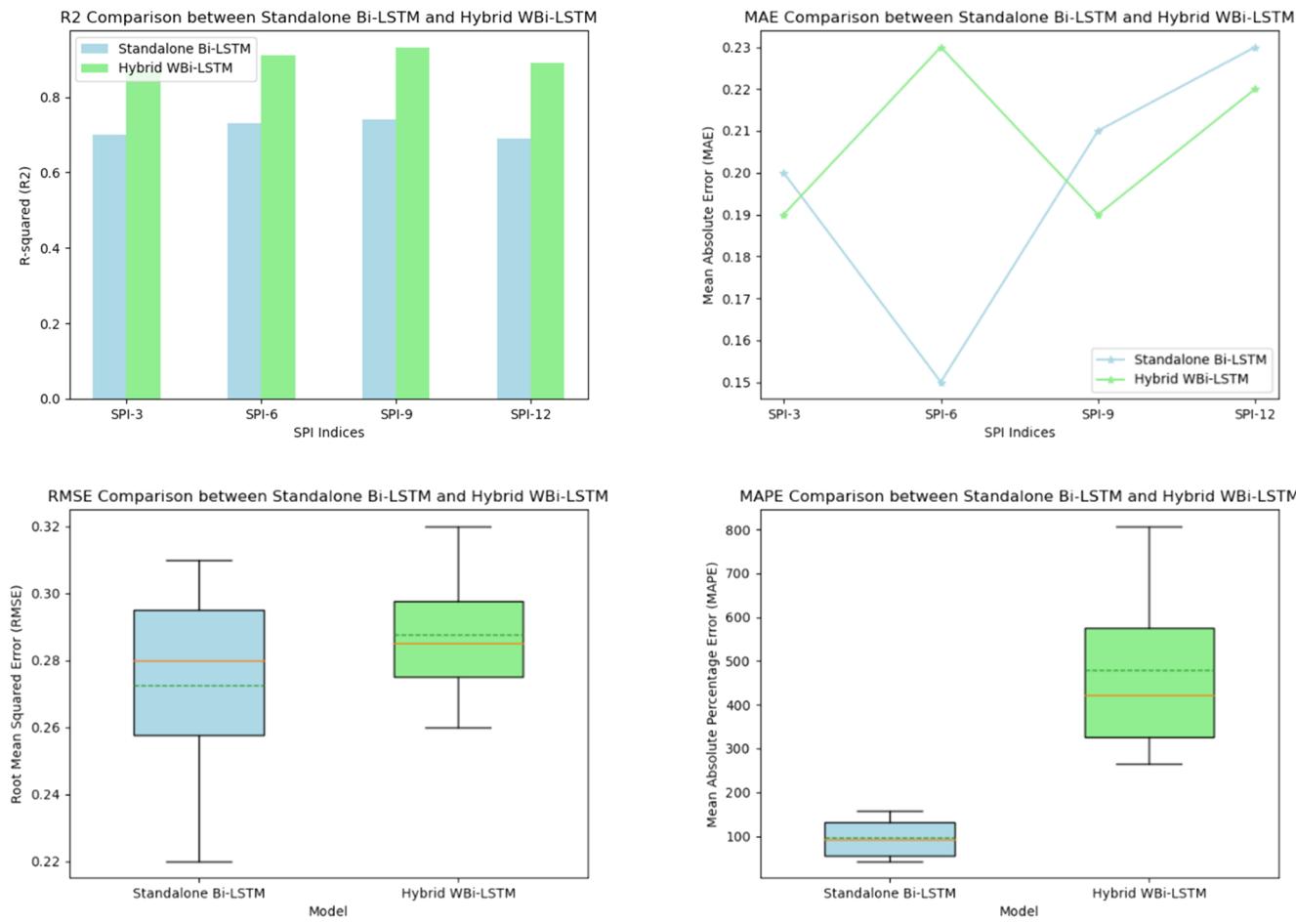


FIG. 15. Comparison of SPI prediction results for monthly time scales (SPI-3, SPI-6, SPI-9, and SPI-12) by wavelet decomposition (WBi-LSTM) and the standalone Bi-LSTM models.

0.21 to 0.34 and 0.26–0.42, respectively, compared to the standalone Bi-LSTM model with MAE and RMSE ranging from 0.26 to 0.43 and 0.33–0.53, respectively. Moving to the CZI-6 timespan, the hybrid WBi-LSTM model maintains its superiority over the standalone Bi-LSTM model across various performance metrics. Specifically, the hybrid WBi-LSTM model achieves higher R^2 , lower MAE, and lower RMSE compared to the standalone Bi-LSTM model. For instance, the hybrid WBi-LSTM model achieves R^2 (0.24–0.84), whereas the standalone Bi-LSTM model achieves R^2 (0.22–0.76). Additionally, the hybrid WBi-LSTM model produces lower MAE (0.27–0.36) and lower RMSE (0.27–0.81) compared to the standalone Bi-LSTM model with MAE 0.34–0.83 and RMSE (0.34–1.01). For the CZI-9 and CZI-12 timespans, the hybrid WBi-LSTM model continues to exhibit superior performance over the standalone Bi-LSTM model. Notably, the hybrid WBi-LSTM model achieves higher R^2 and lower MAE and RMSE across both timespans compared to the standalone Bi-LSTM model. For instance, in the CZI-9 timespan, the hybrid WBi-LSTM model achieves R^2 (0.12–0.91), while the standalone Bi-LSTM model achieves R^2 (0.11–0.82). Similarly, in the CZI-12

timespan, the hybrid WBi-LSTM model achieves R^2 (0.29–0.93), while the standalone Bi-LSTM model achieves R^2 (0.25–0.79). These consistent trends across different timespans highlight the robustness and reliability of the hybrid WBi-LSTM model in accurately predicting CZI, demonstrating its potential as an effective tool.

C. Modified China Z-index (MCZI)

The results for the MCZI indicate that the hybrid WBRF model consistently outperforms the standalone BRF model across all timespans (Fig. 18). For example, in the case of MCZI-3, the hybrid WBRF model achieves R^2 of 0.42, 0.68, and 0.55 for M-1, M-2, and M-3, respectively, compared to R^2 of 0.336, 0.544, and 0.44 for the standalone BRF model. This trend persists across other timespans as well. In terms of MAE, the hybrid WBRF model consistently produces lower errors compared to the standalone BRF model. For instance, for MCZI-6, the hybrid model yields MAE values of 0.38, 0.38, and 0.37 for M-1, M-2, and M-3, respectively, while the standalone model yields MAE values of 0.304, 0.304, and 0.296 for

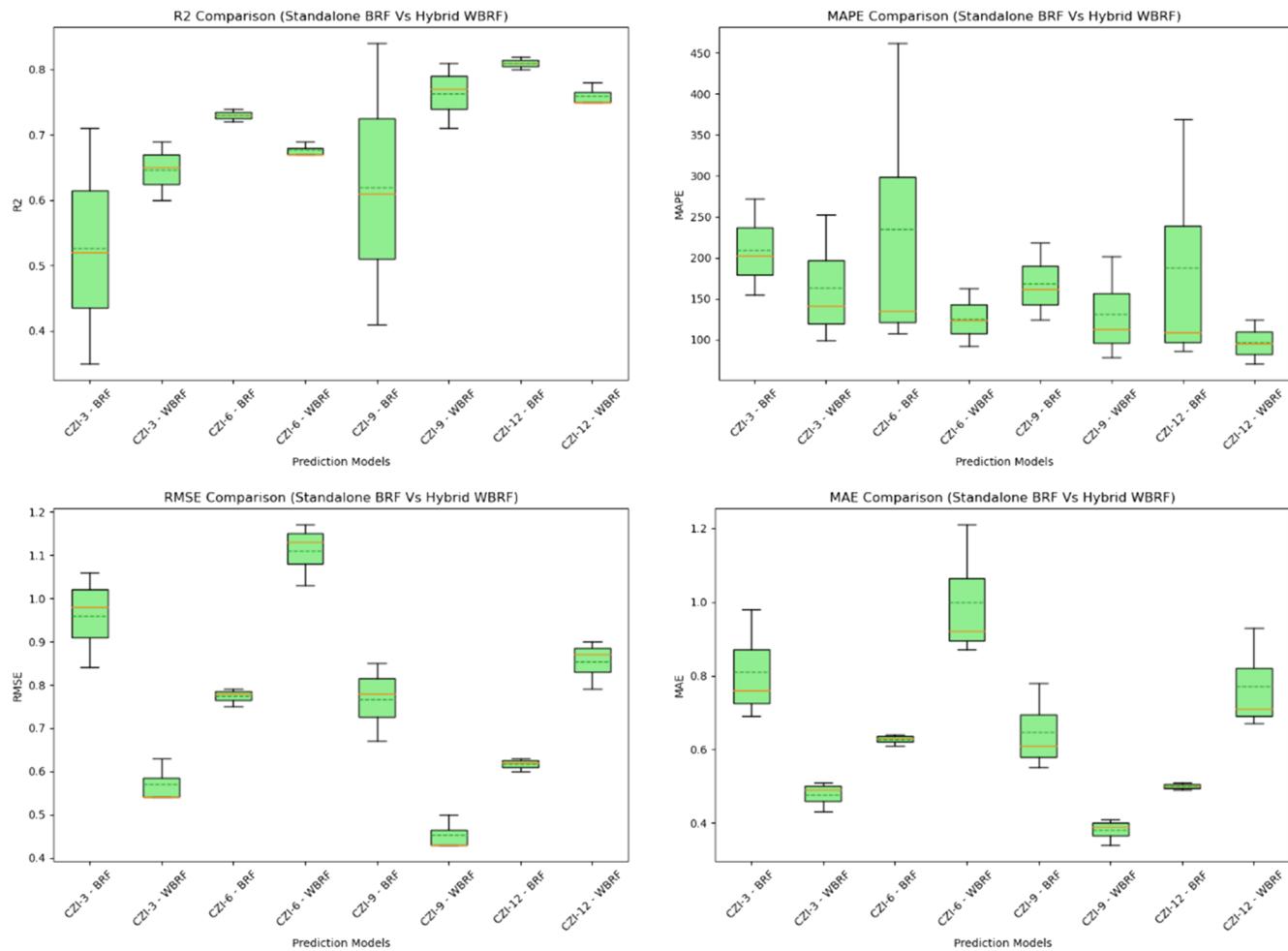


FIG. 16. Comparison of CZI prediction results for monthly time scales (CZI-3, CZI-6, CZI-9, and CZI-12) by wavelet decomposition (WBRF) and the standalone BRF models.

the same timespans. Similarly, the RMSE values demonstrate the superior performance of the hybrid WBRF model. In MCZI-9, for example, the hybrid model achieves RMSE values of 0.43, 0.37, and 0.39 for M-1, M-2, and M-3, respectively, compared to RMSE values of 0.344, 0.296, and 0.312 for the standalone model. However, it is worth noting that the MAPE values show some variability. While the hybrid WBRF model generally performs better in terms of MAE and RMSE, its performance in terms of MAPE varies. In certain scenarios, such as MCZI-3 for M-2, the standalone BRF model exhibits a lower MAPE, suggesting it may provide more accurate relative predictions in specific cases. Overall, the results suggest that incorporating wavelet decomposition enhances the predictive power of the model, leading to improved accuracy in predicting MCZI across various timespans compared to the standalone BRF model.

For the MCZI-3 timespan (Fig. 19), the hybrid WBi-LSTM model achieves R^2 from 0.27 to 0.75 across different input combinations (M-1, M-2, M-3), with MAE values ranging from 0.52 to 0.72, RMSE values ranging from 0.64 to 0.88, and MAPE values

ranging from 346.89 to 609.15. In comparison, the standalone Bi-LSTM model achieves R^2 ranging from 0.24 to 0.68, MAE from 0.65 to 0.90, RMSE (0.80–1.10), and MAPE from 277.51 to 487.32. Overall, the hybrid WBi-LSTM model demonstrates higher R^2 values and lower error metrics compared to the standalone Bi-LSTM model. For MCZI-6 and MCZI-9 timespans, the hybrid WBi-LSTM model consistently outperforms the standalone Bi-LSTM model in terms of R^2 , MAE, and RMSE. For example, in the MCZI-6 timespan, the hybrid WBi-LSTM model achieves R^2 from 0.13 to 0.82, MAE values ranging from 0.32 to 0.68, RMSE values ranging from 0.41 to 0.83, and MAPE values ranging from 286.91 to 440.18, while the standalone Bi-LSTM model achieves R^2 from 0.12 to 0.74, MAE from 0.40 to 0.85, RMSE values ranging from 0.51 to 1.04, and MAPE values ranging from 229.53 to 352.14. In the MCZI-12 timespan, both models face challenges in capturing the variability in the data. The hybrid WBi-LSTM model achieves R^2 from 0.06 to 0.95, MAE values ranging from 0.15 to 0.54, RMSE values ranging from 0.19 to 0.72, and MAPE values ranging from 246.45 to 756.69, while the

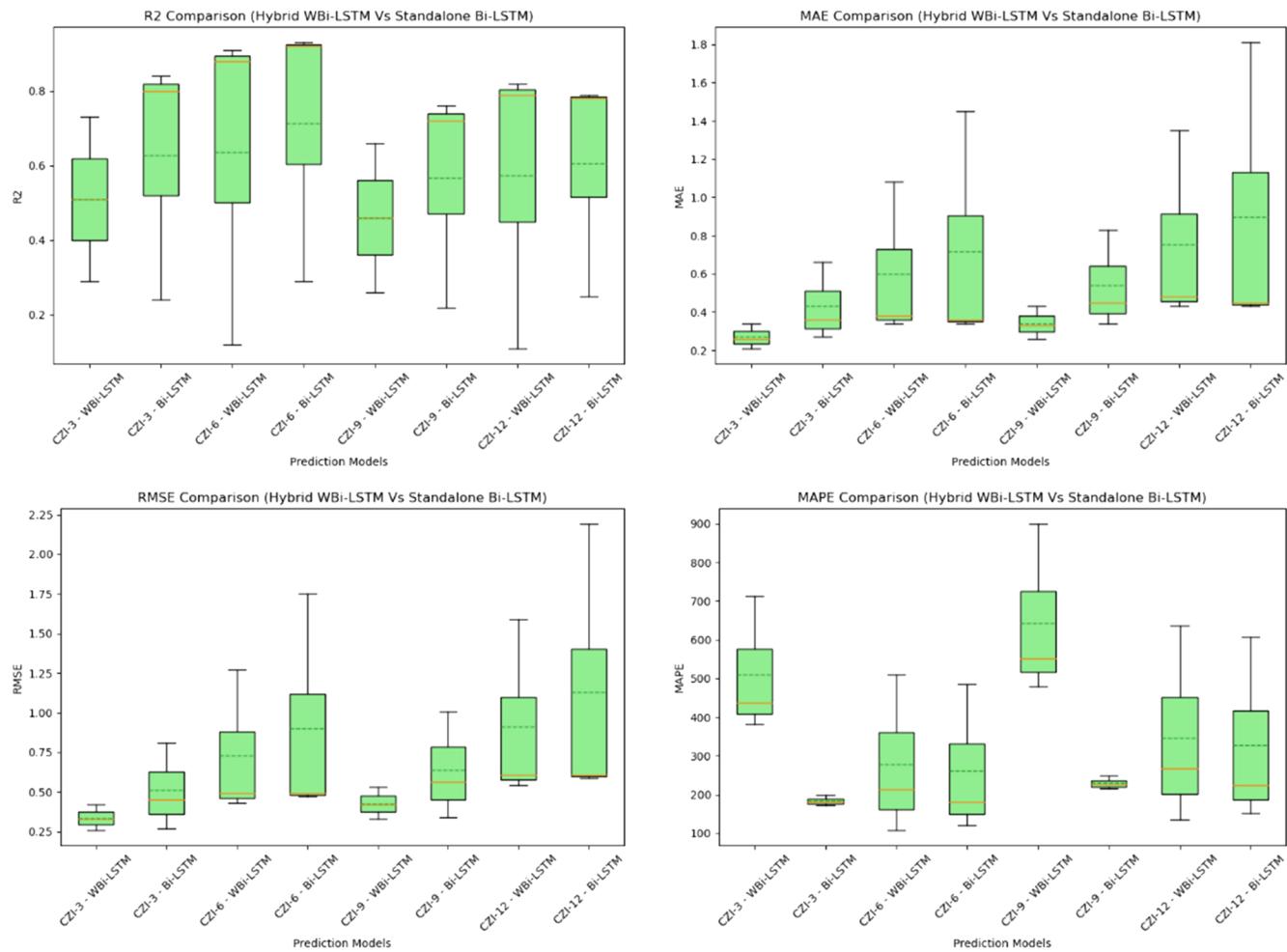


FIG. 17. Comparison of CZI prediction results for monthly time scales (CZI-3, CZI-6, CZI-9, and CZI-12) by wavelet decomposition (WBi-LSTM) and the standalone Bi-LSTM models.

standalone Bi-LSTM model achieves R^2 from 0.05 to 0.86, MAE values ranging from 0.19 to 0.68, RMSE values ranging from 0.24 to 0.90, and MAPE values ranging from 197.16 to 605.35.

The hybrid WBRF model consistently outperforms the standalone BRF model across all timespans for MCZI prediction (Figs. 18 and 19). For instance, in the case of MCZI-3, the hybrid WBRF model achieves R^2 of 0.42, 0.68, and 0.55 for M-1, M-2, and M-3, respectively, compared to R^2 of 0.336, 0.544, and 0.44 for the standalone BRF model. Similarly, the hybrid WBRF model consistently yields lower errors in terms of MAE and RMSE across different timespans compared to the standalone BRF model. Compared to the hybrid WBi-LSTM model, the hybrid WBRF model also demonstrates superior performance. For example, in the MCZI-6 timespan, the hybrid WBRF model achieves R^2 from 0.13 to 0.82, while the hybrid WBi-LSTM model achieves R^2 ranging from 0.12 to 0.74. Additionally, the hybrid WBRF model consistently exhibits lower MAE, RMSE, and MAPE values compared to the hybrid WBi-LSTM model across various timespans.

D. Hutchinson drought severity index (HDSI)

The results of the HDSI prediction model provide comprehension of their performance in forecasting HDSI, shown in Figs. 20 and 21. For the HDSI-3 timespan, the hybrid WBRF model achieves R^2 (0.34–0.45), suggesting a moderate correlation with the observed data (Fig. 20). Conversely, the standalone BRF model demonstrates a slightly lower R^2 (0.31–0.41). Additionally, while the hybrid WBRF model exhibits lower MAE and RMSE values compared to the standalone BRF model, the differences are not substantial. For instance, MAE values range from 1.38 to 1.73 for the hybrid WBRF model and from 1.91 to 2.08 for the standalone BRF model. RMSE (1.52–2.02) for the hybrid WBRF model and from 2.21 to 2.42 for the standalone BRF model. Transitioning to the HDSI-6 timespan, the hybrid WBRF model achieves R^2 (0.53–0.56), indicating a moderate to strong fit of the model to the data. Similarly, the standalone BRF model exhibits R^2 (0.45–0.48). In terms of MAE and RMSE, MAE (1.31–1.46) for the hybrid WBRF model and from 1.64 to 1.83 for the standalone BRF model. RMSE (1.52–1.67) for the hybrid WBRF

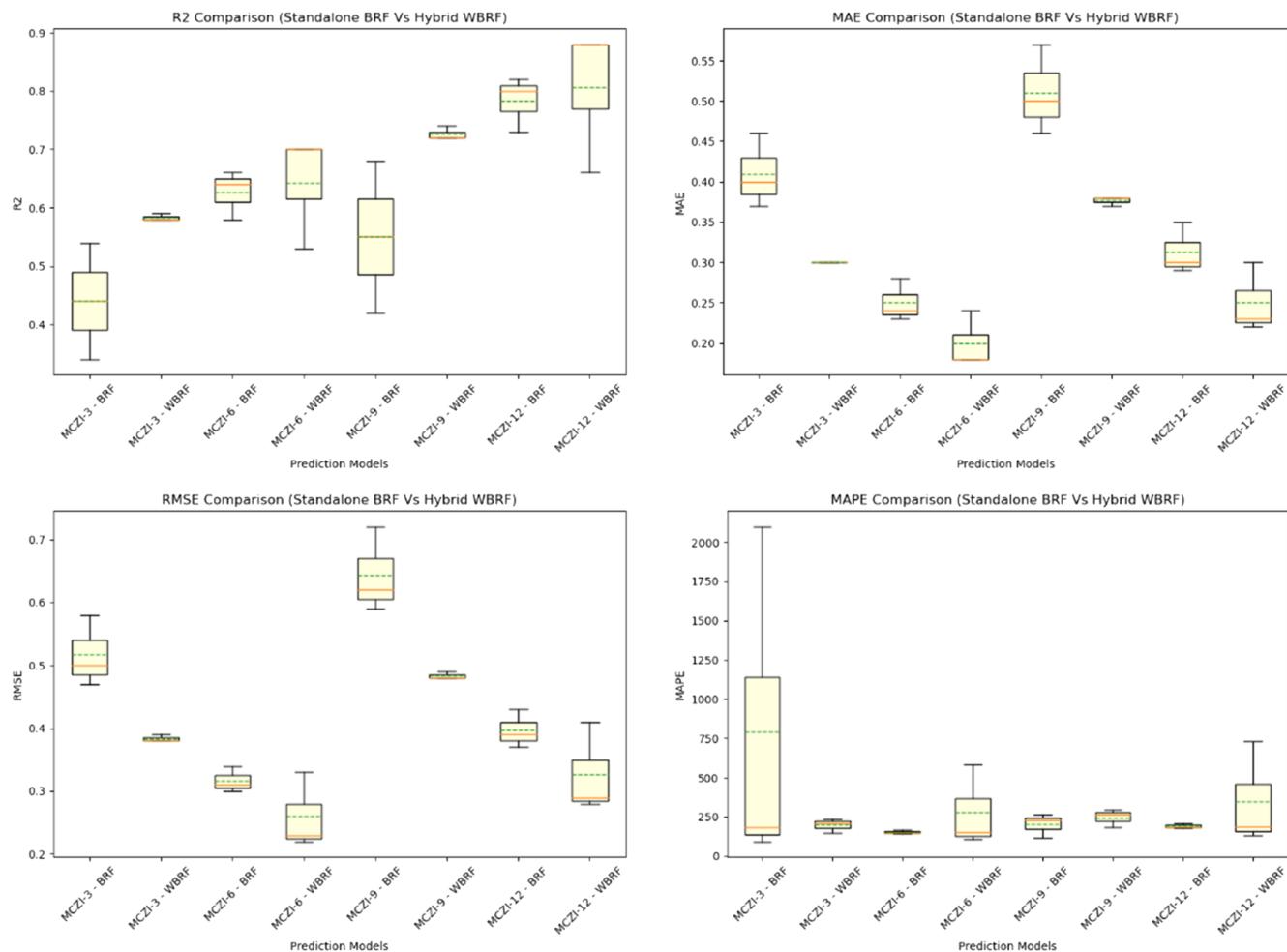


FIG. 18. Comparison of MCZI prediction results for monthly time scales (MCZI-3, MCZI-6, MCZI-9, and MCZI-12) by wavelet decomposition (WBRF) and the standalone BRF models.

model and from 1.90 to 2.09 for the standalone BRF model. For the HDSI-9 and HDSI-12 timespans, the hybrid WBRF model consistently outperforms the standalone BRF model across all metrics. The hybrid WBRF model achieves higher R^2 and lower MAE and RMSE in the HDSI-9 timespan, R^2 (0.55–0.60), while the standalone BRF model achieves R^2 (0.50–0.54). Similarly, in the HDSI-12 timespan, the hybrid WBRF model achieves R^2 (0.62–0.65), while the standalone BRF model achieves R^2 values ranging from 0.56 to 0.59.

In Fig. 21, the comparison of Bi-LSTM and hybrid model (WBi-LSTM) shows the results for different timespans. Starting with the HDSI-3 timespan, the hybrid WBi-LSTM model achieves R^2 (0.38–0.60), indicating a moderate correlation with observed data, whereas the standalone Bi-LSTM model exhibits R^2 (0.34–0.54). In terms of absolute errors, the MAE ranges from 0.54 to 1.50 compared to the MAE (0.68–1.88) for the standalone Bi-LSTM model. However, the standalone Bi-LSTM model achieves slightly lower RMSE values across all scenarios compared to the hybrid WBi-LSTM model. Transitioning to the HDSI-6 timespan, both models demonstrate

improved performance compared to the HDSI-3 timespan. The hybrid WBi-LSTM model exhibits higher R^2 (0.22–0.80), indicating a stronger fit of the model to the data, while the standalone Bi-LSTM model showcases R^2 (0.20–0.72). For the HDSI-9 and HDSI-12 timespans, the hybrid WBi-LSTM model consistently outperforms the standalone Bi-LSTM model across all metrics.

Based on the analysis of the HDSI prediction models across different timespans (HDSI-3, HDSI-6, HDSI-9, and HDSI-12), it is evident that the hybrid WBi-LSTM model consistently outperforms the standalone Bi-LSTM model in forecasting HDSI. Therefore, it can be concluded that the hybrid WBi-LSTM model is the preferred choice for forecasting HDSI, offering better predictive accuracy and reliability compared to the standalone Bi-LSTM model.

E. Rainfall anomaly index (RAI)

Beginning with the standalone BRF model, its performance, indicated by R^2 , ranges from 0.29 to 0.52 across the different

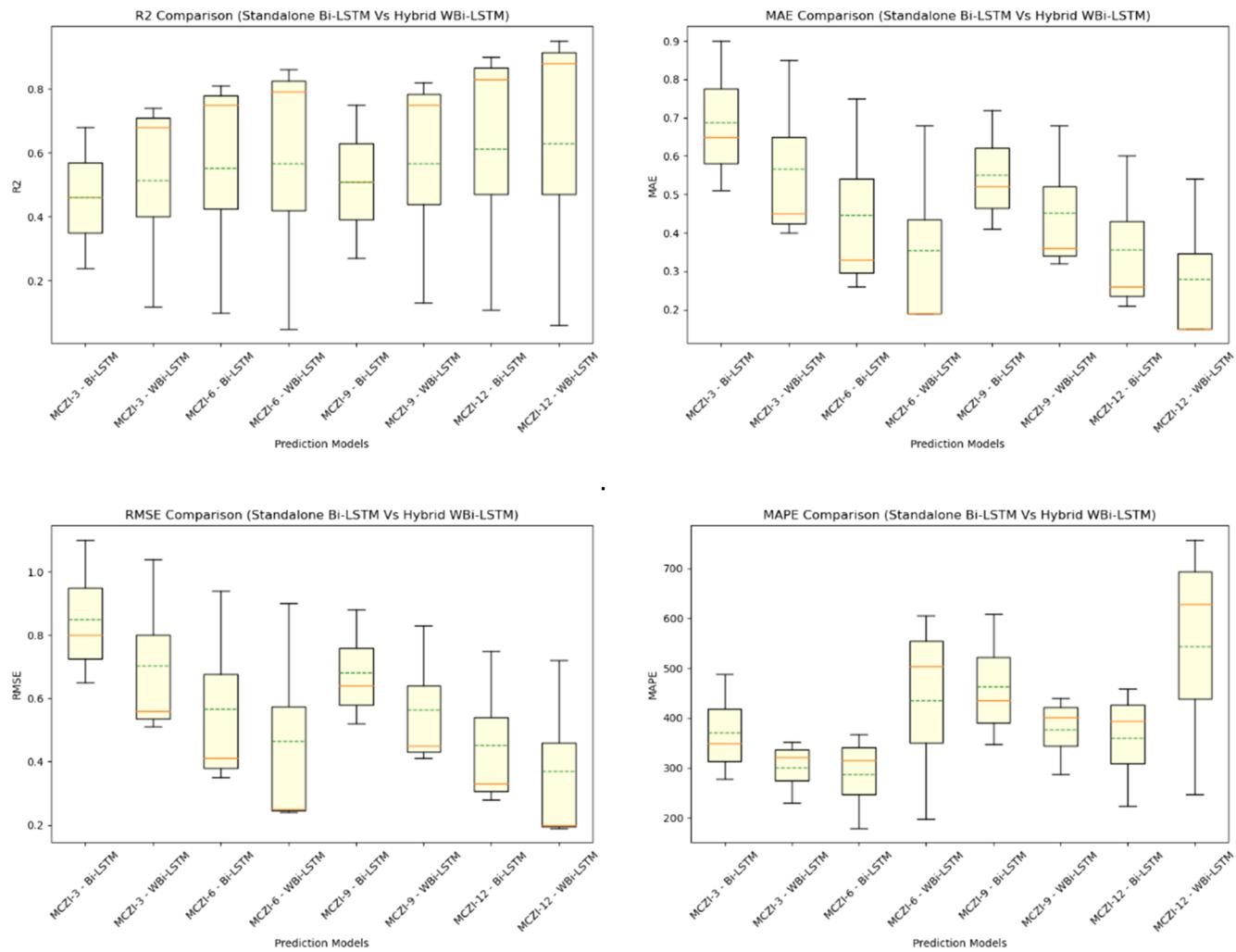


FIG. 19. Comparison of MCZI prediction results for monthly time scales (MCZI-3, MCZI-6, MCZI-9, and MCZI-12) by wavelet decomposition (WBi-LSTM) and the standalone Bi-LSTM models.

lead times (Fig. 22). These values reflect the proportion of variance in the RAI that the model captures. The standalone BRF model demonstrates moderate predictive power, albeit with certain limitations. The MAE and RMSE ranging from 1.39 to 2.95 and 1.69–2.95, respectively, signify the magnitude of prediction errors, with higher values indicating larger discrepancies between predicted and observed values. Despite its simplicity and ease of implementation, the standalone BRF model's performance may be constrained by its inability to capture temporal dependencies and frequency characteristics inherent in rainfall data. Conversely, the hybrid WBRF model exhibits superior predictive performance across most lead times and timespans, as evidenced by higher R^2 values ranging from 0.50 to 0.58. The incorporation of wavelet decomposition enhances the model's ability to capture both short-term fluctuations and long-term trends in rainfall patterns. This is reflected in the lower MAE and RMSE values ranging from 1.39 to 1.91 and

1.68–2.10, respectively, indicating reduced prediction errors compared to the standalone BRF model. By decomposing the time series into different scales using the Haar wavelet family, WBRF effectively captures the multiscale characteristics of rainfall data, thereby improving predictive accuracy. However, the hybrid WBRF model also presents certain limitations. The integration of wavelet decomposition adds computational complexity and may require careful parameter tuning, particularly in selecting the appropriate wavelet family and decomposition levels.

The standalone Bi-LSTM model demonstrates moderate to high predictive capability, with R^2 ranging from 0.11 to 0.80 across different lead times and timespans (Fig. 23). However, the model's performance appears to vary considerably, especially evident in higher MAPE values ranging from 184.82 to 676.50. These higher MAPE values indicate significant discrepancies between predicted and observed values, suggesting potential issues with model

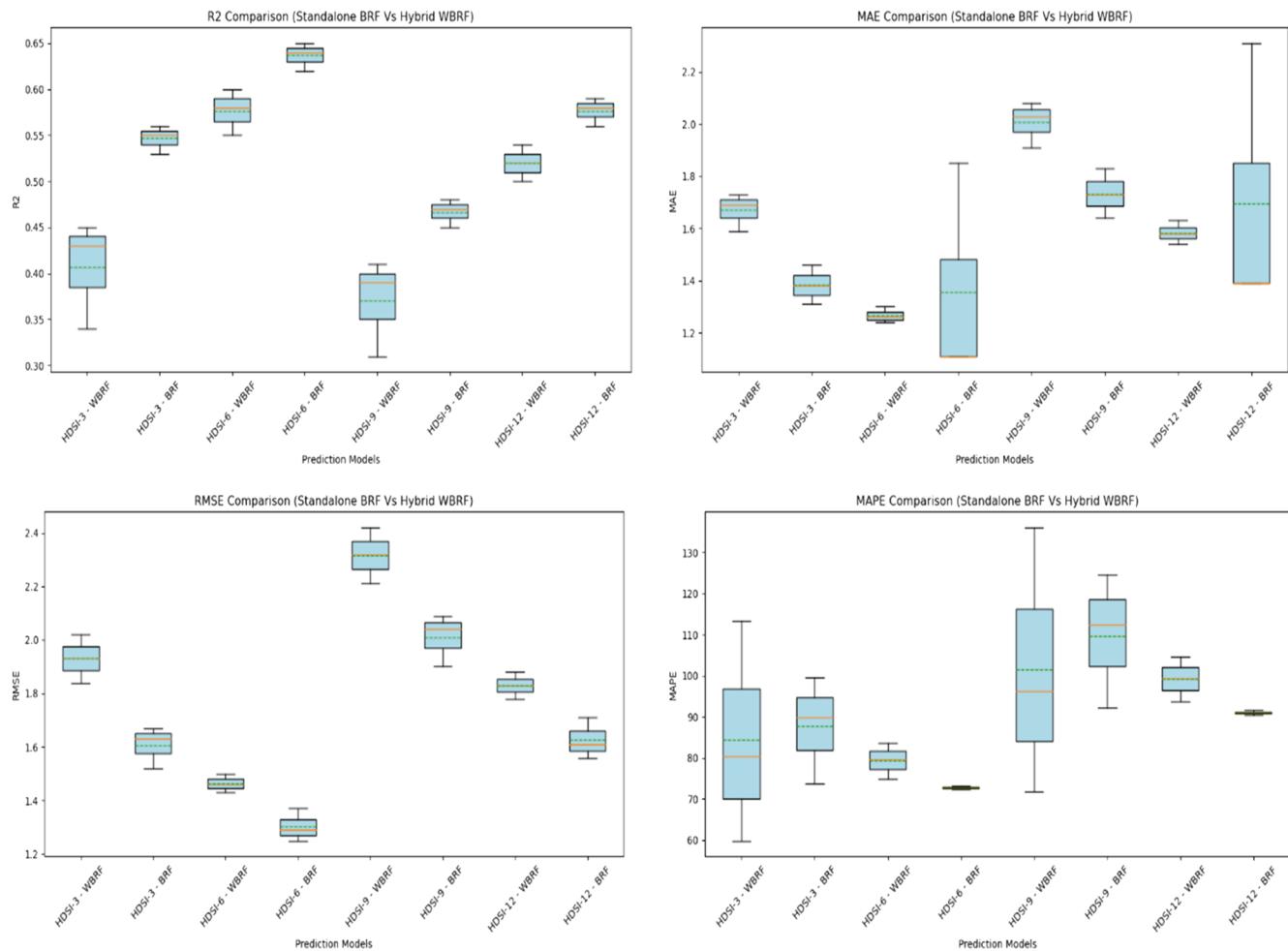


FIG. 20. Comparison of HDSI prediction results for monthly time scales (HDSI-3, HDSI-6, HDSI-9, and HDSI-12) by wavelet decomposition (WBRF) and the standalone BRF models.

accuracy, particularly for longer lead times and timespans. On the other hand, the hybrid WBi-LSTM model generally exhibits superior performance compared to the standalone Bi-LSTM model, particularly evident in higher R^2 and lower MAPE. For instance, in the SPI-12 timespan, the hybrid WBi-LSTM model achieves R^2 of 0.89, indicating better overall explanatory power compared to the standalone Bi-LSTM model. However, despite its enhanced performance, the hybrid WBi-LSTM model presents certain limitations. Notably, its performance varies across different lead times and timespans, with R^2 values fluctuating considerably. Moreover, the model's computational complexity, especially with the integration of wavelet decomposition, may hinder its practical applicability in real-time forecasting scenarios where computational efficiency is crucial.

Considering the performance of both models across different lead times and timespans (Figs. 22 and 23), the hybrid WBi-LSTM model emerges as the superior choice for predicting RAI for 3-, 6-, 9-, and 12-month time spans. The hybrid WBi-LSTM and WBRF model consistently demonstrates higher R^2 and lower MAE and

RMSE compared to the standalone Bi-LSTM model and BRF. For instance, in the SPI-12 timespan, the hybrid WBi-LSTM model achieves an impressive R^2 of 0.89, signifying strong explanatory power, while maintaining relatively lower MAPE values. Overall, the hybrid WBi-LSTM model stands out as the preferred choice for accurate and reliable RAI prediction.

IV. DISCUSSION

This study contributes to the formulation of a comprehensive drought management program by investigating optimal prediction models for meteorological droughts in the Mun River Basin, Thailand. Utilizing data from ten meteorological stations, standalone and hybrid models based on SPI, CZI, MCZI, HDSI, and RAI predictions at different time intervals were employed. Consistent with prior research, short-term rainfall deficits and declines in river flow and groundwater levels suggest long-term droughts in Thailand. Additionally, meteorological characteristics such as temperature, precipitation, relative humidity, and duration of sunshine were identified as

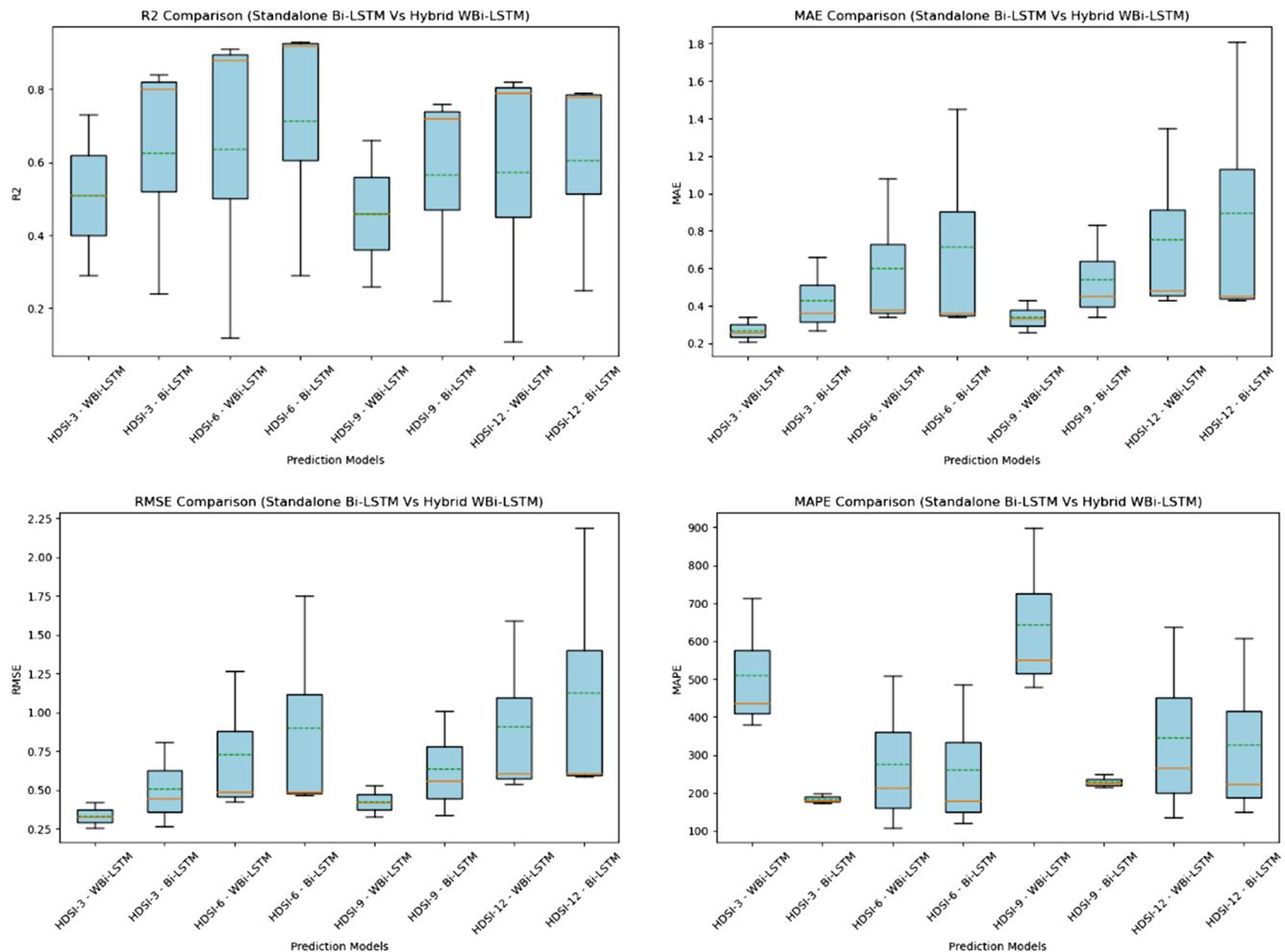


FIG. 21. Comparison of HDSI prediction results for monthly time scales (HDSI-3, HDSI-6, HDSI-9, and HDSI-12) by wavelet decomposition (WBi-LSTM) and the standalone Bi-LSTM models.

critical elements for drought prediction, aligning with findings from studies in other countries like Pakistan and China.

Thailand, especially the Mun River basin, characterized by a predominantly tropical climate with high temperatures and humidity, faces droughts as a predominant climate stressor in the country.³⁹ The occurrence, severity, and duration of droughts vary based on meteorological, hydrological, and agricultural factors.¹⁵ This study focused on meteorological-based drought, and in an endeavor to develop optimal prediction models for predicting droughts in Thailand, this study utilized data from ten meteorological stations from the Mun River basin and employed standalone and hybrid models based on the (SPI, CZI, MCZI, HDSI, and RAI) predictions at different time intervals (3-, 6-, 9-, and 12-months), contributing to the formulation of a drought management program. Consistent with prior research,^{39,73,74} these findings showed that short-term rainfall deficits in droughts (three and six months) and declines in river flow and groundwater levels (12 months) suggest long-term droughts

in Thailand. The use of PPT, Tmax, Tmin, Tave, and RH were recognized as critical elements for drought prediction. Similarly, in Pakistan, meteorological characteristics such as relative humidity, temperature, and wind speed were deemed most significant for accurately predicting droughts.⁷⁵ ML models highlighted the substantial influence of temperature, precipitation, relative humidity, and duration of sunshine on drought in China.⁷⁶ This study also found that the model's performance and efficacy varied depending on the period and incorporating other variables. As a result, hybrid models that employ wavelet decomposition are preferred over a single unified model. This proposal is supported by an examination of climatic variables that influence the SPI and the identification of time-based best models. Similar patterns in drought predictions have been recorded in other Asian nations, emphasizing the importance of robust and accurate methods for drought prediction and management.^{25,34,74,77,78} In wavelet decomposition, for example, in a three-month ahead forecast employing the Haar

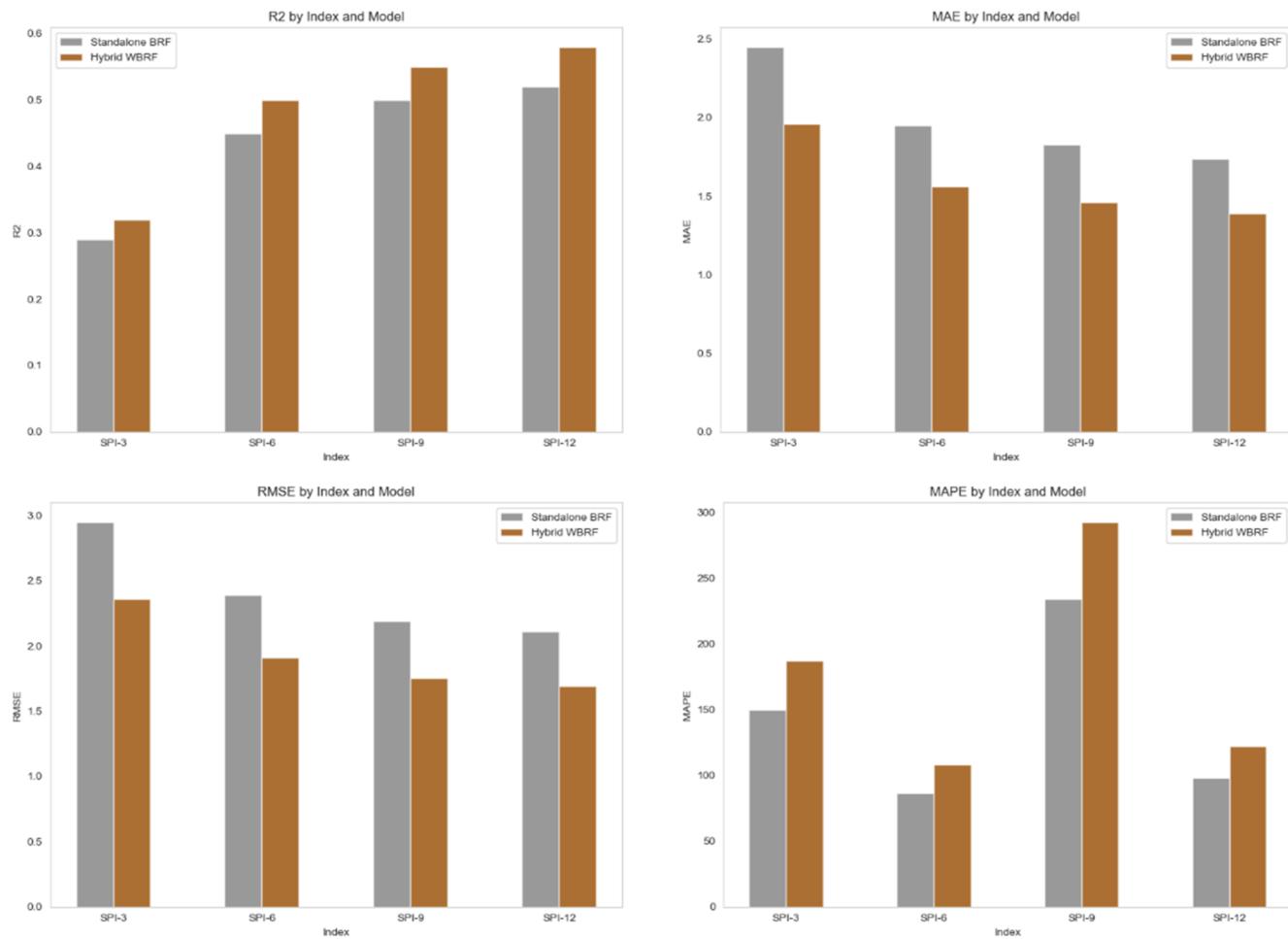


FIG. 22. Comparison of RAI prediction results for monthly time scales (RAI-3, RAI-6, RAI-9, and RAI-12) by wavelet decomposition (WBRF) and the standalone BRF models.

wavelet function, the second decomposition level adequately reflects the volatility in the 12-month series. However, for the three-month ahead forecast, a breakdown up to the third level is required. These findings verify the limits indicated in Nourani *et al.*⁷⁹ Furthermore, the prediction performance of standalone models is projected to improve significantly after input data decomposition and preprocessing. The results of the current study justified the results proposed by recent studies^{80,81} that hybrid models using the “decomposition-prediction” technique improve prediction accuracy. Economically, severe droughts can result in significant losses, which vary based on factors such as drought severity, economic resiliency, and reliance on agricultural and water supplies.⁸² Addressing early drought prediction and mitigation is crucial for countries dealing with water scarcity, needing comprehensive programs that include readiness, response, and recovery.^{10,83} Prioritizing drought mitigation in policy agendas is critical for increasing resilience and building effective response mechanisms.

In the evaluation of drought prediction models, different statistical metrics were employed. R^2 measures the proportion of the

variance in the dependent variable that is predictable from the independent variables, ranging from 0 to 1, with higher values indicating better model performance.²⁷ An R^2 value close to 1 signifies that the model explains most of the variability in the observed data, indicating high accuracy and reliability. For instance, an R^2 of 0.97 for SPI-3 means that 97% of the variance in the SPI-3 observations is captured by the model, demonstrating excellent predictive capability. MAE measures the average magnitude of the errors between predicted and observed values without considering their direction. It provides a straightforward interpretation of prediction accuracy, with lower values indicating better performance.²¹ For example, an MAE of 0.144 for SPI-6 suggests that the model’s predictions deviate from the actual values by an average of 0.144 units, which implies high accuracy and reliability, particularly in practical applications where precise predictions are critical. RMSE is the square root of the average of squared differences between predicted and observed values, penalizing larger errors more than MAE, making it sensitive to outliers. Lower RMSE values indicate better model performance. For example, an RMSE of 0.3 for SPI-12 indicates that the model

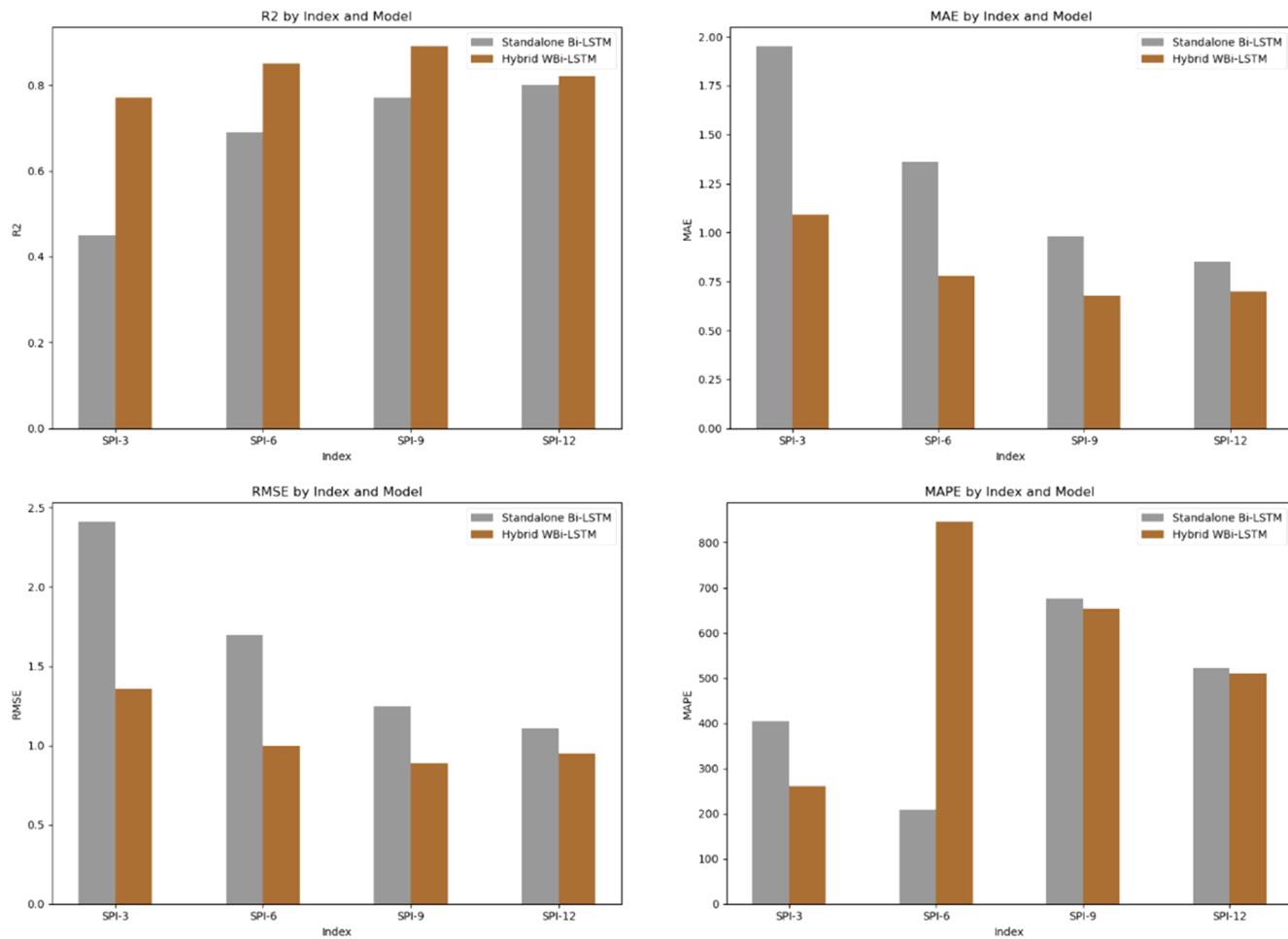


FIG. 23. Comparison of SPI prediction results for monthly time scales (RAI-3, RAI-6, RAI-9, and RAI-12) by wavelet decomposition (WBi-LSTM) and the standalone Bi-LSTM models.

predictions have a small average error, thus reflecting high reliability. RMSE is particularly useful in highlighting models that not only make accurate predictions but also manage to minimize large errors. The combination of these metrics provides a comprehensive evaluation of model performance: high R^2 values suggest that the models are effective in capturing the overall trend and variance in the data, demonstrating strong predictive power; low MAE values indicate that the models have high accuracy, consistently making predictions that are close to the actual observations; and low RMSE values reflect that the models are reliable, with fewer large deviations between predicted and observed values. In summary, these statistical measures collectively show that the hybrid models (WBi-LSTM and WBRF) not only capture the underlying patterns in the data (high R^2) but also provide accurate (low MAE) and reliable (low RMSE) predictions. This detailed discussion underscores the robustness of the proposed models in predicting meteorological droughts in the Mun River Basin, highlighting their

practical applicability and superior performance over standalone models.

This study's findings on meteorological drought prediction provide important support for reducing extreme situations. By emphasizing preventative measures, authorities can reduce the impact of droughts on communities while also promoting sustainability. The WBi-LSTM and WBRF models are particularly useful for modeling drought time series, highlighting local changes, and efficiently handling complicated datasets. The study encountered limitations in the manual selection of decomposition levels and hyperparameter tuning for models like WBRF and WBi-LSTM. Computational complexity and variability in model performance based on input combinations and timespans were also acknowledged. Ensuring quality data preprocessing and considering geographical variability are essential for enhancing the effectiveness and generalizability of predictive models for meteorological droughts.

V. CONCLUSION

The study utilized the wavelet transform, an effective method for analyzing drought signals and time series, to predict meteorological drought in the Mun River Basin. Investigating a critical research gap, it assessed the performance of standalone bootstrapped random forest (BRF) and bi-directional long short-term memory (Bi-LSTM) models, alongside their hybrid counterparts incorporating wavelet decomposition, for predicting various meteorological drought indices, including SPI, CZI, MCZI, HDSI, and RAI, over different timespans (3, 6, 9, and 12 months).

This novel approach addresses a significant research gap by combining advanced machine learning techniques with wavelet decomposition to enhance drought prediction accuracy. The evaluation of predictive accuracy using statistical metrics such as R^2 , MAE, RMSE, and MAPE highlighted the superiority of hybrid models incorporating wavelet decomposition over standalone models across all evaluated drought indices and timespans. For instance, both wavelet decomposed BRF (WBRF) and wavelet decomposed Bi-LSTM (WBi-LSTM) models demonstrated superior performance in SPI prediction compared to their standalone counterparts, indicating higher R^2 and lower MAE, RMSE, and MAPE values. Similarly, hybrid models consistently outperformed standalone models in predicting CZI, MCZI, HDSI, and RAI, with the WBi-LSTM model emerging as the preferred choice for RAI prediction. This integration of wavelet decomposition with machine learning models not only enhances predictive accuracy but also addresses the challenge of capturing multi-scale temporal patterns and complex dependencies present in precipitation data.

The findings underscore the importance of employing advanced modeling techniques to improve the accuracy of drought prediction, which is crucial for effective water resource management and decision-making. While hybrid models demonstrated superior performance, future research could explore further enhancements to these models, such as incorporating additional features or ensemble techniques, to improve predictive accuracy and robustness. Investigating the applicability of these models in real-time forecasting scenarios is also warranted, considering computational efficiency and practical implementation challenges.

ACKNOWLEDGMENTS

We would like to acknowledge the Department of Mathematics, King Mongkut's University of Technology Thonburi (KMUTT), Thailand, Science Research and Innovation (TSRI), and the National Science, Research and Innovation Fund (NSRF) for their technical and financial support in this work. This research project is supported by King Mongkut's University of Technology Thonburi (KMUTT), Thailand Science Research and Innovation (TSRI), and the National Science, Research and Innovation Fund (NSRF) fiscal year 2024.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

U.W.H. and M.W. contributed equally to this work.

Usa Wannasingha Humphries: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal). **Muhammad Waqas:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Software (equal); Writing – original draft (equal). **Phyo Thandar Hliang:** Validation (equal); Visualization (equal). **Porntip Dechpichai:** Formal analysis (equal); Methodology (equal); Resources (equal); Visualization (equal); Writing – review & editing (equal). **Angkool Wangwongchai:** Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹Y.-S. Ham *et al.*, “Comparison of LSTM network, neural network and support vector regression coupled with wavelet decomposition for drought forecasting in the western area of the DPRK,” *Nat. Hazards* **116**(2), 2619–2643 (2023).
- ²A. Dikshit, B. Pradhan, and A. M. Alamri, “Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model,” *Sci. Total Environ.* **755**, 142638 (2021).
- ³B. Lloyd-Hughes, “The impracticality of a universal drought definition,” *Theor. Appl. Climatol.* **117**, 607–611 (2014).
- ⁴S. M. Vicente-Serrano *et al.*, “A review of environmental droughts: Increased risk under global warming?,” *Earth-Sci. Rev.* **201**, 102953 (2020).
- ⁵A. Zargar *et al.*, “A review of drought indices,” *Environ. Rev.* **19**(NA), 333–349 (2011).
- ⁶E. Olukayode Oladipo, “A comparative performance analysis of three meteorological drought indices,” *J. Climatol.* **5**(6), 655–664 (1985).
- ⁷M. Peña-Gallardo *et al.*, “Complex influences of meteorological drought timescales on hydrological droughts in natural basins of the contiguous United States,” *J. Hydrol.* **568**, 611–625 (2019).
- ⁸X. Liu *et al.*, “Agricultural drought monitoring: Progress, challenges, and prospects,” *J. Geogr. Sci.* **26**, 750–767 (2016).
- ⁹A. F. Van Loon, “Hydrological drought explained,” *WIREs Water* **2**(4), 359–392 (2015).
- ¹⁰D. A. Wilhite, *Droughts: A Global Assessment* (Routledge, 2016).
- ¹¹J. D. Petersen-Perlman, I. Aguilar-Barajas, and S. B. Megdal, “Drought and groundwater management: Interconnections, challenges, and policy responses,” *Curr. Opin. Environ. Sci. Health* **28**, 100364 (2022).
- ¹²S. Mukherjee, A. Mishra, and K. E. Trenberth, “Climate change and drought: A perspective on drought indices,” *Current Clim. Change Rep.* **4**, 145–163 (2018).
- ¹³A. K. Mishra and V. P. Singh, “Drought modeling – A review,” *J. Hydrol.* **403**(1–2), 157–175 (2011).
- ¹⁴A. K. Mishra and V. P. Singh, “A review of drought concepts,” *J. Hydrol.* **391**(1–2), 202–216 (2010).
- ¹⁵M. Pedro-Monzón *et al.*, “A review of water scarcity and drought indexes in water resources planning and management,” *J. Hydrol.* **527**, 482–493 (2015).
- ¹⁶J. Dong *et al.*, “Standardized precipitation evapotranspiration index (SPEI) estimated using variant long short-term memory network at four climatic zones of China,” *Comput. Electron. Agric.* **213**, 108253 (2023).

- ¹⁷T. B. McKee, N. J. Doesken, and J. Kleist, "The relationship of drought frequency and duration to time scales," in Proceedings of the 8th Conference on Applied Climatology, 1993.
- ¹⁸J. Xiaosheng *et al.*, "Research on determination of station indexes and division of regional flood/drought grades in China," *J. Appl. Meteor. Sci.* **8**(1), 26–33 (1997).
- ¹⁹M. Van Rooy, "A rainfall anomaly index independent of time and space," *Notos* **14**, 43–48 (1965).
- ²⁰D. Khadka *et al.*, "Multivariate and multi-temporal analysis of meteorological drought in the northeast of Thailand," *Weather Clim. Extremes* **34**, 100399 (2021).
- ²¹M. Waqas *et al.*, "Advancements in daily precipitation forecasting: A deep dive into daily precipitation forecasting hybrid methods in the tropical climate of Thailand," *MethodsX* **12**, 102757 (2024).
- ²²U. W. Humphries *et al.*, "Determination of crop water requirements and potential evapotranspiration for sustainable coffee farming in response to future climate change scenarios," *Smart Agric. Technol.* **8**, 100435 (2024).
- ²³S. Pimonsree *et al.*, "Urbanization-induced changes in extreme climate indices in Thailand during 1970–2019," *Atmos. Res.* **265**, 105882 (2022).
- ²⁴S. Kornkosa *et al.*, "Analysis and prediction of meteorological drought area by using standardized precipitation index in northeast, Thailand," *Int. J. Environ. Sci. Dev.* **12**(12), 372 (2021).
- ²⁵M. A. Al Mamun *et al.*, "Identification of influential weather parameters and seasonal drought prediction in Bangladesh using machine learning algorithm," *Sci. Rep.* **14**(1), 566 (2024).
- ²⁶M. Waqas *et al.*, "Potential of artificial intelligence-based techniques for rainfall forecasting in Thailand: A comprehensive review," *Water* **15**(16), 2979 (2023).
- ²⁷M. Waqas *et al.*, "Incorporating novel input variable selection method for in the different water basins of Thailand," *Alexandria Eng. J.* **86**, 557–576 (2024).
- ²⁸X. He *et al.*, "Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks," *Water Resour. Manage.* **33**, 1571–1590 (2019).
- ²⁹A. Dikshit and B. Pradhan, "Explainable AI in drought forecasting," *Mach. Learn. Appl.* **6**, 100192 (2021).
- ³⁰M. Achite *et al.*, "Modeling of meteorological, agricultural, and hydrological droughts in semi-arid environments with various machine learning and discrete wavelet transform," *Theor. Appl. Climatol.* **154**(1), 413–451 (2023).
- ³¹D. Salim *et al.*, "Comparative study of different discrete wavelet based neural network models for long term drought forecasting," *Water Resour. Manage.* **37**(3), 1401–1420 (2023).
- ³²S. Bachmair *et al.*, "Developing drought impact functions for drought risk management," *Nat. Hazards Earth Syst. Sci.* **17**(11), 1947–1960 (2017).
- ³³K. Fung *et al.*, "Drought forecasting: A review of modelling approaches 2007–2017," *J. Water Clim. Change* **11**(3), 771–799 (2020).
- ³⁴S. Prabnakorn, "Integrated flood and drought mitigation measures and strategies," in *Case Study: The Mun River Basin, Thailand* (CRC Press, 2020).
- ³⁵A. Kikon and P. C. Deka, "Artificial intelligence application in drought assessment, monitoring and forecasting: A review," *Stochastic Environ. Res. Risk Assess.* **36**(5), 1197–1214 (2022).
- ³⁶J. Liu *et al.*, "Distributive characteristics of riverine nutrients in the Mun River, Northeast Thailand: Implications for anthropogenic inputs," *Water* **11**(5), 954 (2019).
- ³⁷J. Liu *et al.*, "Impacts of anthropogenic changes on the Mun River water: Insight from spatio-distributions and relationship of C and N species in northeast Thailand," *Int. J. Environ. Res. Public Health* **16**(4), 659 (2019).
- ³⁸S. Yadav *et al.*, "Land use impact on the water quality of large tropical river: Mun River Basin, Thailand," *Environ. Monit. Assess.* **191**(10), 614 (2019).
- ³⁹S. Li and H. Pi, "Deconstruction of dryness and wetness patterns with drought condition assessment over the Mun River Basin, Thailand," *Land* **11**(12), 2244 (2022).
- ⁴⁰A. Wangwongchai *et al.*, "Imputation of missing daily rainfall data: A comparison between artificial intelligence and statistical techniques," *MethodsX* **11**, 102459 (2023).
- ⁴¹V. Costa, W. Fernandes, and M. Naghettini, "A Bayesian model for stochastic generation of daily precipitation using an upper-bounded distribution function," *Stochastic Environ. Res. Risk Assess.* **29**, 563–576 (2015).
- ⁴²F. E. Grubbs and G. Beck, "Extension of sample sizes and percentage points for significance tests of outlying observations," *Technometrics* **14**(4), 847–854 (1972).
- ⁴³S. Li *et al.*, "Comparison of hybrid machine learning models to predict short-term meteorological drought in Guanzhong region, China," *Water Sci. Technol.* **87**(11), 2756–2775 (2023).
- ⁴⁴N. M. Alcazar *et al.*, "Estudo cefalométrico comparativo dos espaços naso e bucofaríngeo nas más oclusões Classe I e Classe II, Divisão 1, sem tratamento ortodôntico, com diferentes padrões de crescimento," *Rev. Dent. Press Ortod. Ortop. Facial* **9**, 68–76 (2004).
- ⁴⁵J. A. Costa and G. P. Rodrigues, "Space-time distribution of rainfall anomaly index (RAI) for the Salgado Basin, Ceará State - Brazil," *Cienc. Nat.* **39**(3), 627–634 (2017).
- ⁴⁶T. Raziee, "Revisiting the rainfall anomaly index to serve as a simplified standardized precipitation index," *J. Hydrol.* **602**, 126761 (2021).
- ⁴⁷E. B. Wilson and M. M. Hilferty, "The distribution of chi-square," *Proc. Natl. Acad. Sci. U. S. A.* **17**(12), 684–688 (1931).
- ⁴⁸M. Kendall and A. Stuart, *The Advanced Theory of Statistics. Vol. 1: Distribution Theory* (Griffin, London, 1977).
- ⁴⁹D. I. Smith, M. F. Hutchinson, and R. McArthur, *Climatic and Agricultural Drought: Payments and Policy* (Australian National University, Centre for Resource and Environmental Studies, 1992).
- ⁵⁰J. H. F. Flores, P. M. Engel, and R. C. Pinto, "Autocorrelation and partial autocorrelation functions to improve neural networks models on univariate time series forecasting," in *The 2012 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2012).
- ⁵¹A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.* **15**(4), 723–736 (1984).
- ⁵²M. Shoaib *et al.*, "A comparative study of various hybrid wavelet feedforward neural network models for runoff forecasting," *Water Resour. Manage.* **32**, 83–103 (2018).
- ⁵³I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **41**(7), 909–996 (1988).
- ⁵⁴I. Daubechies, *Ten Lectures on Wavelets* (SIAM, 1992).
- ⁵⁵J. Estévez *et al.*, "Monthly precipitation forecasts using wavelet neural networks models in a semiarid environment," *Water* **12**(7), 1909 (2020).
- ⁵⁶S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989).
- ⁵⁷T. Partal, H. K. Cigizoglu, and E. Kahya, "Daily precipitation predictions using three different wavelet neural network algorithms by meteorological data," *Stochastic Environ. Res. Risk Assess.* **29**, 1317–1329 (2015).
- ⁵⁸W. Phusakulkajorn, C. Lursinsap, and J. Asavanant, "Wavelet-transform based artificial neural network for daily rainfall prediction in southern Thailand," in *2009 9th International Symposium on Communications and Information Technology* (IEEE, 2009).
- ⁵⁹F. Soares and M. J. Anzanello, "Support vector regression coupled with wavelength selection as a robust analytical method," *Chemom. Intell. Lab. Syst.* **172**, 167–173 (2018).
- ⁶⁰R. X. Gao and R. Yan, *Wavelets: Theory and Applications for Manufacturing* (Springer Science & Business Media, 2010).
- ⁶¹A. Haar, *Zur Theorie der Orthogonalen Funktionensysteme* (Georg-August-Universität, Gottingen, 1909).
- ⁶²Y. Meyer, "Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs," *Sém. Bourbaki* **662**, 1985–1986 (1985).
- ⁶³M. S. Chavan *et al.*, "Implementation of SYMLET wavelets to removal of Gaussian additive noise from speech signal," in *Proceedings of the Recent Researches in Communications, Automation, Signal Processing, Nanotechnology, Astronomy and Nuclear Physics: 10th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications (EHAC'11)*, 2011.
- ⁶⁴S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.* **9**(8), 1735–1780 (1997).
- ⁶⁵N. Salaeh *et al.*, "Long-short term memory technique for monthly rainfall prediction in thale sap Songkhla River Basin, Thailand," *Symmetry* **14**(8), 1599 (2022).

- ⁶⁶A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," in *International Conference on artificial Neural Networks* (Springer, 2005).
- ⁶⁷Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," [arXiv:1508.01991](https://arxiv.org/abs/1508.01991) (2015).
- ⁶⁸A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks* **18**(5–6), 602–610 (2005).
- ⁶⁹L. Breiman, "Random forests," *Mach. Learn.* **45**, 5–32 (2001).
- ⁷⁰M. Waqas *et al.*, "Assessment of advanced artificial intelligence techniques for streamflow forecasting in Jhelum River Basin," *Pak. J. Agric. Res.* **34**, 580 (2021).
- ⁷¹O. E. Obisesan, "Machine learning models for prediction of meteorological variables for weather forecasting," *Int. J. Environ. Clim. Change* **14**(1), 234–252 (2024).
- ⁷²M. W. M. Saifullah *et al.*, "Modelling of rainfall-runoff process by GEP, RBF-SVM and M5 model tree in Jhelum River Basin, Pakistan," in *International Conference on Hydrology and Water Resources (ICHWR'21)*, Lahore, Pakistan, 25 March 2021 (Centre of Excellence in Water Resources Engineering, UET Lahore-Pakistan, 2021), p. 108.
- ⁷³P. Ganguli and M. J. Reddy, "Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach," *Hydrol. Processes* **28**(19), 4989–5009 (2014).
- ⁷⁴A. Nohegar, S. Mahmoodabadi, and A. Norouzi, "Comparison the suitability of SPI, PNI and DI drought index in Kahurestan watershed (Hormozgan Province/South of Iran)," *J. Environ Earth Sci.* **5**(8), 71–77 (2015).
- ⁷⁵N. Khan *et al.*, "Prediction of droughts over Pakistan using machine learning algorithms," *Adv. Water Resour.* **139**, 103562 (2020).
- ⁷⁶R. Zhang *et al.*, "Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China," *Sci. Total Environ.* **665**, 338–346 (2019).
- ⁷⁷Z. Wang *et al.*, "Does drought in China show a significant decreasing trend from 1961 to 2009?," *Sci. Total Environ.* **579**, 314 (2017).
- ⁷⁸Y. Ge *et al.*, "Drought frequency change: An assessment in northern India plains," *Agric. Water Manage.* **176**, 111 (2016).
- ⁷⁹V. Nourani, Ö. Kisli, and M. Komasi, "Two hybrid artificial intelligence approaches for modeling rainfall-runoff process," *J. Hydrol.* **402**(1–2), 41–59 (2011).
- ⁸⁰E. E. Başakın, Ö. Ekmekcioğlu, and M. Özger, "Drought prediction using hybrid soft-computing methods for semi-arid region," *Model. Earth Syst. Environ.* **7**, 2363–2371 (2021).
- ⁸¹X. Wang *et al.*, "An adaptive daily runoff forecast model using VMD-LSTM-PSO hybrid approach," *Hydrol. Sci. J.* **66**(9), 1488–1502 (2021).
- ⁸²J. Freire-González, C. Decker, and J. W. Hall, "The economic impacts of droughts: A framework for analysis," *Ecol. Econ.* **132**, 196–204 (2017).
- ⁸³D. A. Wilhite, M. J. Hayes, and C. L. Knutson, "Drought preparedness planning: Building institutional capacity," in *Drought and Water Crises: Science, Technology, and Management Issues*, 107 (Taylor and Francis Group, 2005), pp. 93–136.