water

MDPI

Article

Accurate Multi-Scaler SPI Drought Prediction by Designing ORDWT-Based Signal Decomposition and Stacked LSTM

Abdulhaleem H. Labban [1], Mumtaz Ali [2,*] and Mohammed Diykh [2,3]

[1] Department of Meteorology, Faculty of Environmental Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia; alabban@kau.edu.sa
[2] School of Business, Law, Humanities and Pathways, University of Southern Queensland, Toowoomba, QLD 4305, Australia; mohammed.diykh@unisq.edu.au
[3] Technical Engineering College, Al-Ayen Iraqi University, Nasiriyah 64001, Iraq
* Correspondence: mumtaz.ali@unisq.edu.au

Abstract

Accurate drought prediction is essential for sustainable water resource management, agricultural planning, and climate resilience. However, drought signals are inherently nonlinear and non-stationary, making reliable predictions particularly challenging. Although current models have demonstrated promising prediction capabilities, their performance is often limited when learning directly from complex drought signals. To address this challenge, we proposed a novel drought prediction model that integrates the Overcomplete Rational Dilation Wavelet Transform (ORDWT) with a Stacked Long Short-Term Memory (SLSTM) network, referred to as the ORDWT-SLSTM. The proposed model first decomposes drought signals into multiple frequency components using ORDWT to preserve multi-scale characteristics while minimizing information loss. Subsequently, the decomposed components are modeled using SLSTM to capture complex temporal dependencies and long-term drought persistence patterns. The effectiveness of the proposed model was evaluated using the standardized precipitation index (SPI) at multiple prediction horizons (SPI-1, SPI-3, SPI-6, and SPI-12) across three climatic regions in Queensland, Australia, namely, Mackay, Gatton, and Darling Downs. The proposed model, the ORDWT-SLSTM, is compared with several benchmark models, including Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), and Bidirectional Long Short-Term Memory (BiLSTM). The Root-Mean-Square Error (RMSE), Mean Absolute Error (MAE) and Nash–Sutcliffe efficiency (NSE) are used to evaluate the proposed models. The proposed model, the ORDWT-SLSTM, obtained the best prediction performance compared with the state-of-the-art models: RMSE = 0.215, MAE = 0.125, NSE = 0.986; RMSE = 0.260, MAE = 0.147, NSE = 0.973; RMSE = 0.225, MAE = 0.130, NSE = 0.980 for Darling Downs, Gatton, and Mackay, respectively. The obtained results demonstrated the efficiency of the proposed ORDWT-SLSTM model in long-term and short-term drought prediction. The proposed ORDWT-SLSTM model offers significant practical value for operational drought management by providing reliable early-warning information that supports agricultural planning, water allocation decisions, and drought risk mitigation.

Keywords: ORDWT; LSTM; SPI; droughts; Queensland; climate change

Check for updates

Academic Editors: Chin H Wu and Pankaj Kumar

Received: 29 April 2026
Revised: 17 June 2026
Accepted: 7 July 2026
Published: 13 July 2026

1. Introduction

Drought occurs in many regions all over the world, and it requires intervention globally. Several regions in Queensland have been suffering severe drought for a long period of time, which has caused ecological destruction [1]. In the last ten years, drought at various levels has happened continually in Darling Downs [2]. Drought severity identification has become a challenge for researchers. Accurate drought prediction provides experts with the needed information to make the right decision to avoid disaster [3,4].

Currently, several indexes have been developed to monitor drought, including the Palmer Drought Severity Index (PDSI) and the standardized precipitation index (SPI) [5]. These indexes assess drought severity using several measures, including temperature and meteorological and hydrological data [6–10]. The SPI was first suggested by McKee et al. [11] to estimate the intensity and duration of drought.

Many statistical models have been designed to forecast droughts, such as ARIMA models and ARIMA-based neural networks [11]. Khan et al. [12] developed a hybrid model based on ARIMA and artificial neural networks to predict future drought values. Nury et al. [13] examined three models for drought prediction, including random forest (RF), ARIMA, and a long short-term memory recurrent neural network (LSTM-RNN). Recently, techniques such as Convolutional Neural Networks (CNNs), GRU, LSTM, and recurrent neural networks (RNNs) have been developed to predict SPI drought. Tang et al. [14] designed a hybrid model for flash drought prediction in China's Greater Bay Area. In that study, a CNN-LSTM hybrid coupled with causality awareness was developed. Xing et al. [15] suggested an SPI drought prediction model using an unsupervised Convolutional Autoencoder. In that study, the authors integrated an unsupervised Convolutional Autoencoder with Extreme Gradient Boosting (XGBoost). Mishra et al. [16] designed a hybrid deep learning model for SPI prediction.

Although these models obtained acceptable prediction results, they directly learn from raw or partially processed drought time series. However, drought data is nonlinear, non-stationary, and influenced by precipitation variability, temperature fluctuations, evapotranspiration dynamics, and delayed environmental responses. Consequently, standard deep learning approach networks often struggle to effectively extract both short-term fluctuations and long-term drought patterns from raw signals alone.

To address this issue, signal decomposition techniques have been increasingly integrated with prediction models. As a result, several studies have adopted Discrete Wavelet Transform (DWT), Empirical Mode Decomposition (EMD), and Multivariate Empirical Mode Decomposition (MEMD). For example, Kim and Valdes [17] combined a wavelet transform with neural networks. Ozger et al. [18] integrated wavelet transforms with a fuzzy logic model. Bacanli et al. [19] applied an adaptive neuro-fuzzy inference for drought prediction. Farokhnia et al. [20] used ANFIS to estimate SPI drought. Mishra and Singh [21] compared several methodologies for drought prediction. Abbes et al. [22] examined LSTM coupled with a Multi-Resolution Analysis Wavelet Transform for drought prediction. Their model was tested at different regions in Iran. Kadam et al. [23] compared several models, including RNN, GRU, and LSTM, for drought prediction. Bui et al. [24] investigated several models of random forest, including reduced error pruning tree (REPT), RT, and their integration with Bagging models, for water quality prediction. Kaur and Sood [25] applied support vector regression (SVR) and a deep neural network (DNN) for drought prediction in three regions in the Texas area, USA. They finally found that the DNN model surpassed other deep learning models.

Although deep learning-based decomposition techniques have improved drought prediction results, these techniques exhibit several inherent limitations. For example, DWT suffers from shift sensitivity and information loss due to down sampling operations, while EMD and MEMD are prone to mode mixing, end effects, and instability under noisy environmental conditions [26]. These limitations may distort the intrinsic drought signals and reduce the ability of prediction models to learn meaningful temporal dependencies across multiple drought scales. Furthermore, most existing decomposition-based drought prediction studies focus primarily on reducing prediction errors while paying limited attention to preserving multi-scale temporal information and maintaining the physical characteristics of drought evolution [27]. As a result, important drought-related features may be partially lost during decomposition, leading to a suboptimal prediction performance, particularly for longer SPI horizons where drought persistence and delayed climatic effects become more pronounced.

To overcome these limitations, this study proposes a novel ORDWT-SLSTM model that integrates the ORDWT with an SLSTM network. Unlike classic decomposition techniques, ORDWT avoids down sampling operations, thereby preserving signal continuity and reducing information loss while providing improved multi-resolution analysis. The rational dilation mechanism further enables more flexible frequency decomposition, allowing for the extraction of informative drought characteristics across different temporal scales. Subsequently, the proposed SLSTM model exploits the decomposed components to capture complex nonlinear dependencies and long-range temporal relationships more effectively than conventional shallow recurrent models.

The fundamental hypothesis of this work is that accurate drought prediction requires both effective multi-scale feature extraction and robust temporal dependency modeling. By combining the superior signal representation capability of the ORDWT with the hierarchical temporal learning capacity of SLSTM, the proposed model aims to provide a more reliable and physically meaningful representation of drought dynamics, ultimately improving forecasting accuracy and robustness across multiple SPI timescales and climatic conditions. The objectives of the current research can be summarized as follows:

- This study introduces a novel ORDWT-SLSTM model for multi-timescale drought prediction, combining the superior multi-resolution analysis capability of the ORDWT with the hierarchical temporal learning capacity of SLSTM.
- The proposed ORDWT-SLSTM model provides a systematic investigation of multi-scale drought forecasting across SPI-1, SPI-3, SPI-6, and SPI-12 timescales, allowing for the assessment of both short-term and long-term drought dynamics under diverse climatic conditions.
- Comprehensive experiments and statistical significance analysis across three stations in QLD, Australia, demonstrate that the proposed ORDWT-SLSTM model achieved superior forecasting accuracy, robustness, and generalization capability compared with state-of-the-art decomposition-based deep learning models across three Australian drought monitoring stations.

2. Materials and Methodology

2.1. Study Area and Data Description

The datasets span the period from January 1900 to December 2024. The collected variables include monthly rainfall, maximum temperature, and minimum temperature. Based on the precipitation records, drought conditions were characterized using the SPI at four temporal scales, namely, SPI-1, SPI-3, SPI-6, and SPI-12, representing short-term, medium-term, and long-term drought conditions, respectively. The Mackay station is recognized for mining, agriculture, and tourism, and it experiences a tropical climate. The dry season in Mackay carries high temperatures and humidity and erratic cyclones. The annual rainfall of approximately 1585 mm happens in December and March. The weather conditions are altering in Mackay, where average temperatures are already 1 °C above as compared with those in the last few decades, due to global warming. The climate change triggers abnormal drought conditions in the region, affecting crops and water resources and increasing wildfires. The soil moisture is reduced due to increased evaporation rates, which makes the drought-like conditions even worse. Gatton station is positioned in the Lockyer Valley agriculture region of Queensland, Australia. The climate is humid and subtropical with heavy summer rain and warm conditions and plays a key role in the agricultural output. The droughts and heatwaves are tough and threatening in this area. The major agricultural occupations in Gatton are intensive horticulture, cropping and grazing. Approximately 40% of the fresh vegetables are generated here. The Darling Downs station is an inland plateau in southern Queensland, where droughts have substantial socio-economic and environmental effects. The prolonged dry spells have severely decreased grain, cotton, and livestock production, causing financial obstructions and disrupting supply chains. Figure 1 shows a map of the study regions.

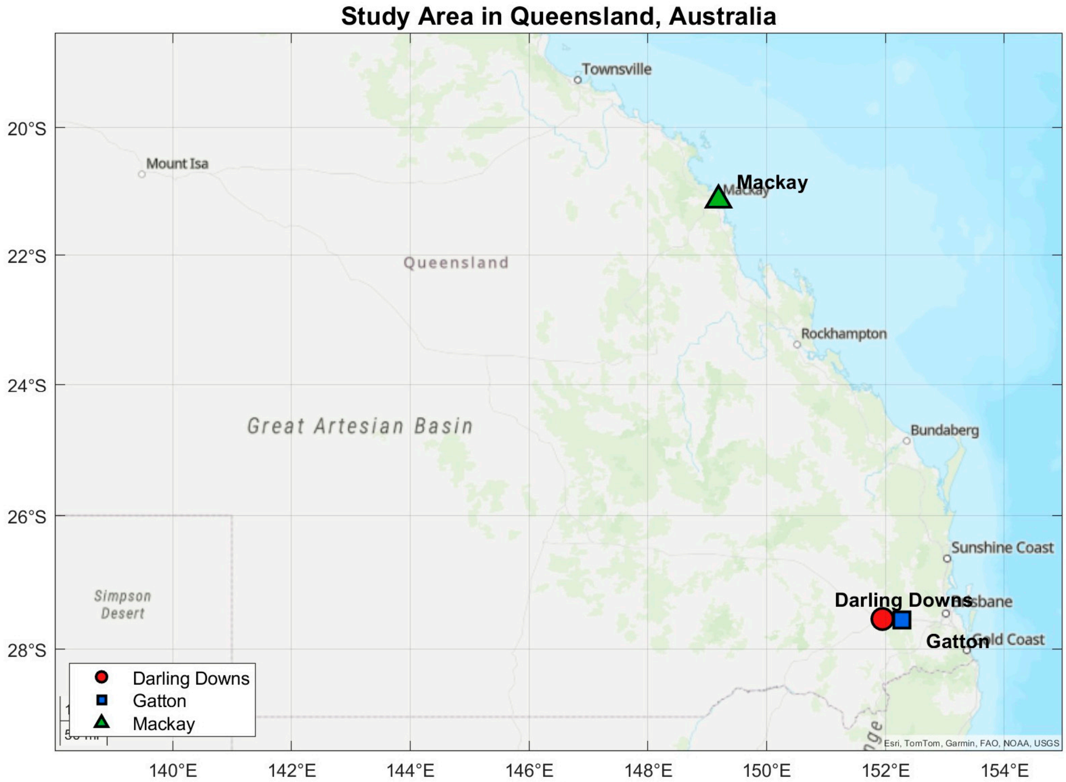

Figure 1. Map of study stations.

Table 1 summarizes the key statistical characteristics of the study datasets. The SPI distributions across all stations exhibit means close to zero and standard deviations near one, confirming the appropriate standardization of the drought indices. However, the minimum and maximum SPI values reveal considerable variability in the drought severity and wetness conditions among the stations. Darling Downs experienced the most extreme drought conditions, reaching a minimum SPI value of −4.542, while Mackay recorded the highest positive SPI value of 3.366, indicating periods of exceptionally wet conditions. The long temporal coverage, diverse climatic conditions, and multiple drought timescales make these datasets suitable for evaluating the robustness and generalization capability of the proposed drought forecasting framework.

Table 1. Statistical information of collected dataset.

Station	Period	Records	Mean Monthly Rainfall (mm)	Annual Rainfall (mm/Year)	SPI Mean	SPI Std. Dev.	SPI Min	SPI Max
Mackay	1900–2024	1500	130.0	1560	0.005	0.973	−2.337	3.366
Gatton	1900–2024	1500	63.9	767	−0.002	1.014	−3.646	2.830
Darling Downs	1900–2024	1500	52.3	628	−0.003	1.020	−4.542	2.867

2.2. Methodology

2.2.1. Overcomplete Rational Dilation Wavelet Transform (ORDWT)

In this paper, drought data is decomposed using the ORDWT [28,29]. The ORDWT employs two filter banks. The dilation factor relies on the rational value of p/q, with the following condition: $q < p$. The wavelet function and scaling function are associated with the scaling relation and rational dilation, as shown in Equations (1) and (2):

$$\varphi(r) = \left(\frac{q}{p}\right)^{1/2} \sum_{i \in Z}^{1} h_0(i)\varphi\left(\frac{q}{p}r - i\right) \tag{1}$$

$$\psi(r) = \left(\frac{q}{p}\right)^{1/2} \sum_{i \in Z}^{1} h_1(i)\psi\left(\frac{q}{p}r - i\right) \tag{2}$$

where $\varphi(r)$ refers to the scaling function, $\psi(r)$ denotes the wavelet function, $h_0(i)$ refers to a scaling function filter, and $h_1(i)$ indicates the wavelet function filter. The ORDWT is employed to decompose the input time series into different bands according to the p, q and r parameters. At each decomposition level, the input time series is decomposed into a low passband and a stop band. The functions $H_1(\omega)$ and $H_0(\omega)$ are the frequency outcomes of $h_0(n)$ and $h_1(n)$, as shown in Figure 2. To compute the frequency outcomes, $H_0(\omega)$ and $H_1(\omega)$ are employed, as shown in Equations (3) and (4):

$$H_0(\omega) = \begin{cases} \sqrt{pq}, & \omega \in \left[0, \left(1 - \frac{1}{r}\right)\frac{1}{p}\pi\right] \\ \sqrt{pq}\,\theta_{H_0}\left(\frac{\omega - a}{b}\right), & \omega \in \left[\left(1 - \frac{1}{r}\right)\frac{1}{p}\pi, \pi\right] \\ 0, & \omega \in \left[\frac{1}{q}\pi, \pi\right] \end{cases} \tag{3}$$

$$H_1(\omega) = \begin{cases} 0, & \omega \in \left[0, \left(1 - \frac{1}{r}\right)\pi\right] \\ \sqrt{r}\theta_{H_1}\left(\frac{\omega - pa}{pb}\right), & \omega \in \left[\frac{r-1}{r}\pi, \frac{p}{q}\pi\right] \\ \sqrt{r}, & \omega \in \left[\frac{p}{q}\pi, \pi\right] \end{cases} \tag{4}$$

where a, b and θ are defined as:

$$a = \frac{\left(1 - \frac{1}{r}\right)\pi}{p} \tag{5}$$

$$b = \frac{1}{q} - \frac{\left(1 - \frac{1}{r}\right)1}{p} \tag{6}$$

$$\theta(\omega) = \frac{(1 + cos(\omega))}{2}\sqrt{2 - cos(\omega)},\ for\ \omega \in [0, \pi] \tag{7}$$

We employed the function $\theta(\omega)$ to obtain the down-pass filter ($H_0(\omega)$) and top passbands ($H_1(\omega)$). Daubechies wavelet filters are employed to derive $\theta(\omega)$. Orthonormal filters are important to designing any sharp edges; however, they can cause the slow decay rate of the $\psi(t)$ and $\varphi(t)$, which should be prevented.

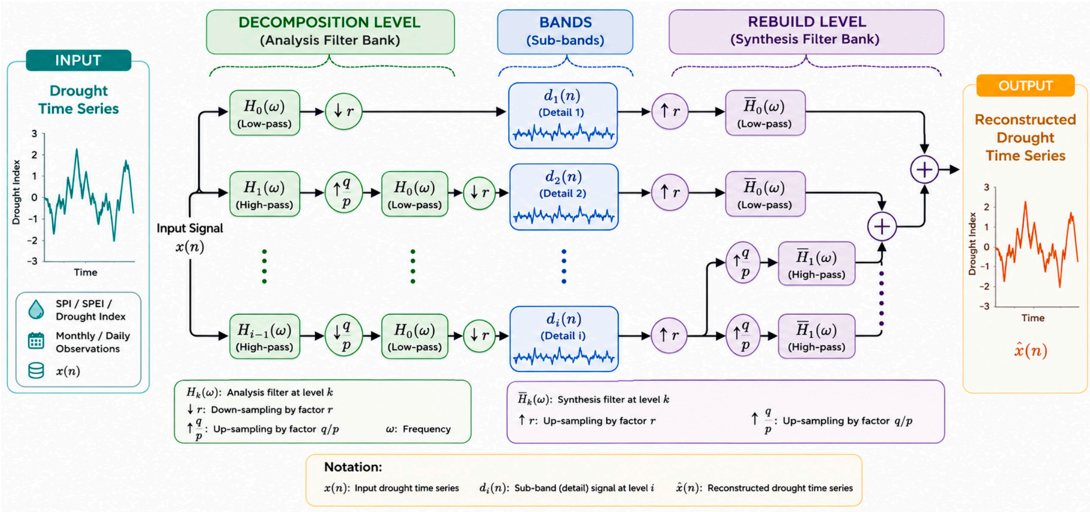

Figure 2. Decomposition process of ORDWT for analysis of drought time-series data.

To rebuild the decomposed time-series data, the conditional holds of the ORDWT ($|H_0(\omega/p)|^2/(pq) + |H_1(\omega)|^2/r = 1$) are checked. This indicates $x(n) = \overline{x}(n)$, as presented in Figure 2. The reconstruction process of the decomposed time series is defined as follows:

1. The length of the stop-band series is cut off as the decomposition proceeds. To recapture the length of the stop band, single-branch reconstruction (SBR) is employed to rebuild the sequences of the decomposition levels into equivalent wavelet sub-bands.
2. After decomposing the input time series to the jth level using SBR, the time series is rebuilt into the jth-1-level band:

$$\overline{x}(n) = \{A_j(n), D_j(n), D_{j-1}(n), \ldots, D_1(n)\} \tag{8}$$

where $A_j(n)$ denotes the rebuild approximation sub-bands of the time series, and $D_j(n)$ refers to the jth rebuild sub-bands of the time series.

1. Two functions, $\overline{H}_0(\omega)$ and $\overline{H}_1(\omega)$, are used to obtain the frequency responses of $\overline{h}_0(n)$ and $\overline{h}_1(n)$:

$$\overline{H}_0(\omega) = \frac{1}{p} h_0\left(\frac{\omega}{p}\right) \tag{9}$$

$$\overline{H}_1(\omega) = h_1(\omega) \tag{10}$$

where $\overline{h}_0(n)$ and $\overline{h}_1(n)$ are employed over the decomposition process.

In addition, for the jth decomposition level, the frequency responses, $\overline{H}_0(\omega)$ and $\overline{H}_1(\omega)$, are extended to:

$$\overline{H}_{0_j}(\omega) = \frac{1}{p_j} H_{0_j}\left(\frac{\omega}{p_j}\right),\ j \geq 1 \tag{11}$$

$$\overline{H}_{1_{j-1}}(\omega) = \begin{cases} H_1(\omega), & j = 0 \\ \frac{1}{p_j} H_{0j}\left(\frac{\omega}{p_j}\right) H_1\left(\left(\frac{q}{p}\right)^j \omega\right), & j \geq 1 \end{cases} \tag{12}$$

where

$$\overline{H}_{0_j}(\omega) = \begin{cases} \prod_{n=0}^{j-1} H(q_{j-1-n} p_n \omega), & \omega \in \left[0, \frac{\pi}{q_j}\right] \\ 0, & \omega \in \left[\frac{\pi}{q_j}, \pi\right] \end{cases},\ j \geq 1 \tag{13}$$

To rebuild the decomposed time series using the ORDWT, the p, q and r parameters should be satisfied with the following condition:

$$\left(1 - \frac{1}{r}\right)\frac{1}{p}\pi < \frac{\pi}{q} \Rightarrow \frac{p}{q} + \frac{1}{r} > 1 \quad (14)$$

This condition does not allow the passbands to overlap between bands. In this study, the proposed ORDWT method was applied to predict monthly rainfall. Also, p, q and r were chosen empirically as p = 5; q = 6; r = 2. In addition, the decomposition level (jth) was also selected empirically. Figures 3 and 4 present the ORDWT coefficients in terms of the decomposed signal and heatmap for Gatton station.

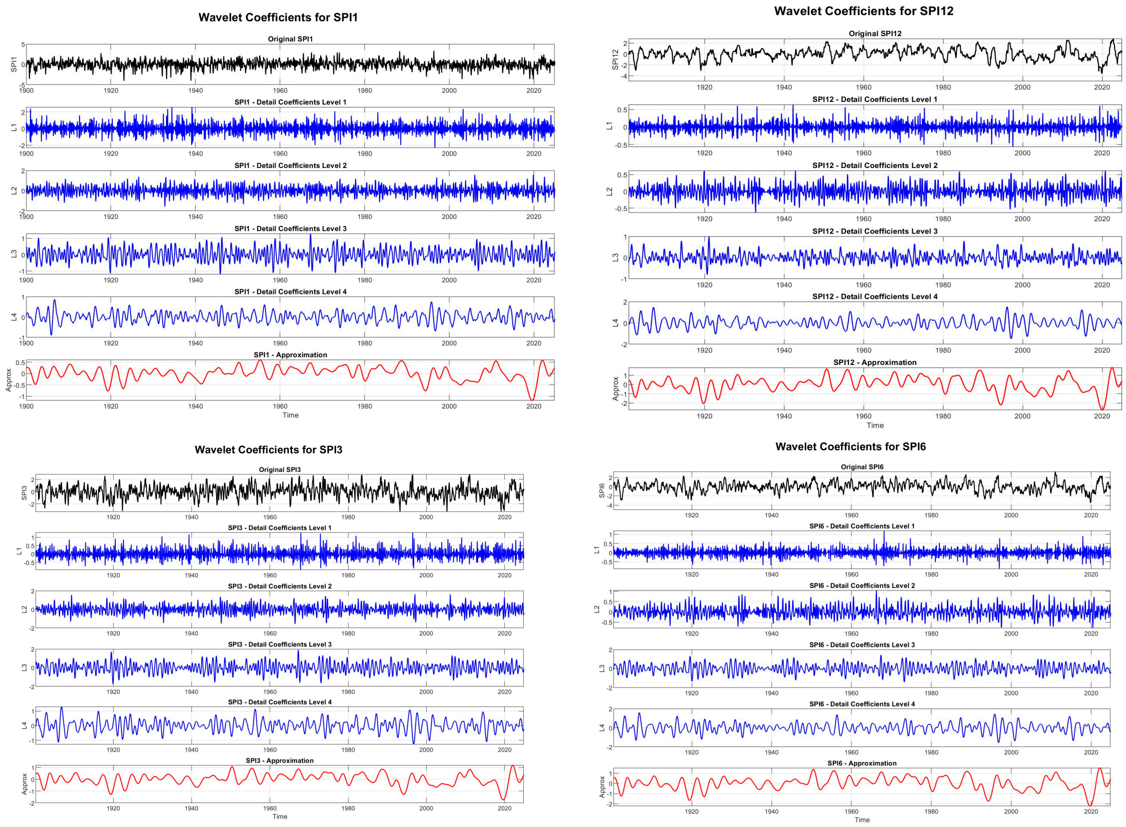

Figure 3. An example of ORDWT coefficients for SPI-1, SPI-3, SPI-6, and SPI-12.

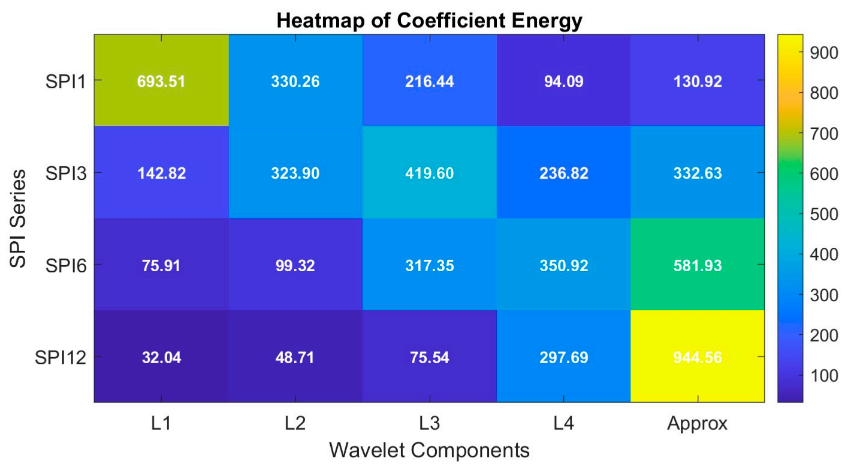

Figure 4. Heatmap of energy coefficients of ORDWT for SPI-1, SPI-3, SPI-6, and SPI-12.

2.2.2. Stacked Long Short-Term Memory (SLSTM)

The proposed prediction model integrates two stacked LSTM layers followed by a dense output layer for SPI forecasting. The stacked architecture enables the model to learn hierarchical temporal representations and capture both short-term and long-term drought dynamics more effectively than a single-layer LSTM. Each LSTM layer contains 64 hidden units, and the final dense layer generates the predicted SPI value. Figure 5 illustrates the architecture of the proposed SLSTM model. The input time-series sequence is processed through multiple SLSTM layers to hierarchically learn short-term and long-term temporal dependencies embedded within the drought data. The extracted high-level temporal features are subsequently passed to a fully connected dense layer, which maps the learned representations to the final prediction output. By employing multiple LSTM layers, the SLSTM architecture improved the model's ability to capture complex nonlinear and multi-scale temporal patterns, thereby improving the drought forecasting performance compared with classic deep learning models. The proposed SLSTM consists of multiple sequential LSTM layers with 64 hidden units per layer, followed by a dense output layer, thereby forming a deep recurrent architecture capable of learning hierarchical temporal representations from the decomposed drought signals. To improve the performance of the proposed model, we monitored the loss and metrics of the proposed model. The loss is adopted to select the hyperparameters, while the metrics prompt the call-back function in the case of failing to enhance the successive epoch.

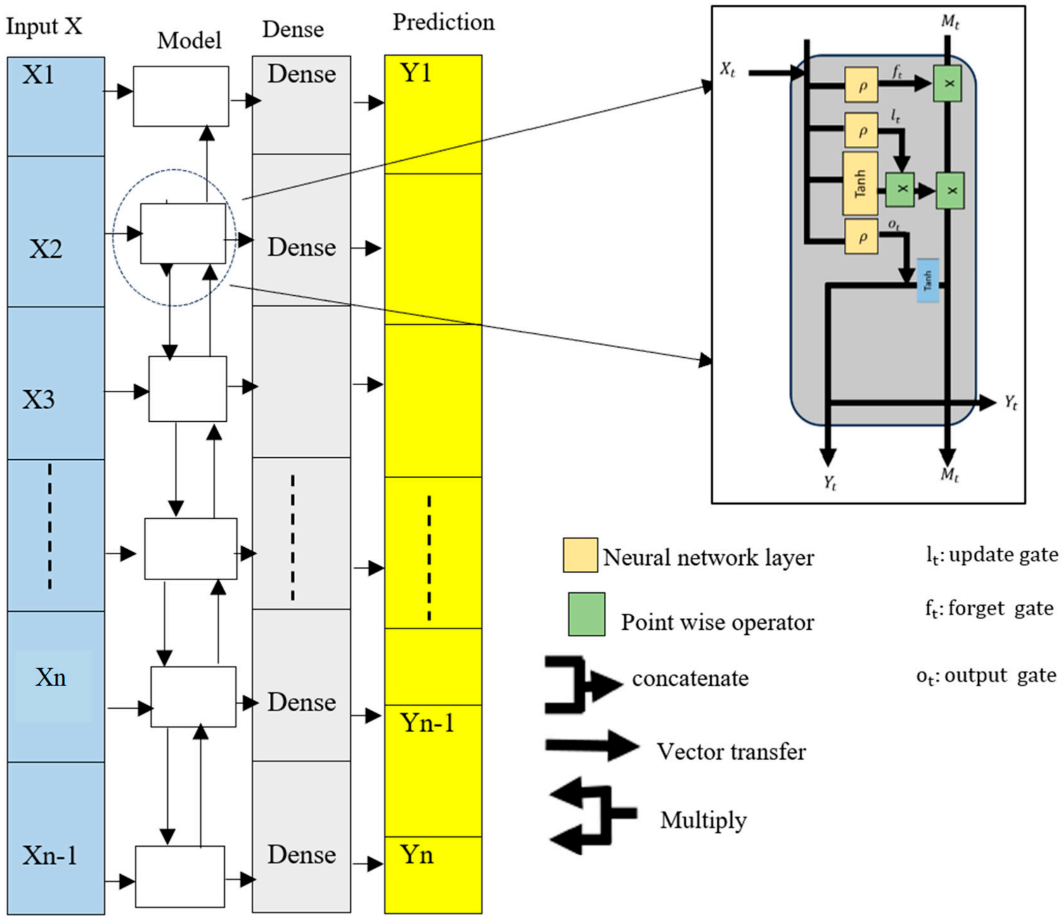

Figure 5. The proposed stacked LSTM for drought prediction.

The stacked architecture of LSTM is described in Figure 5:

$$ft = \sigma(Wf[ht - 1, xt] + bf) \tag{15}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{16}$$

$$\widetilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \tag{17}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \tag{18}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{19}$$

$$h_t = o_t \odot \tanh(C_t) \tag{20}$$

where f_t, i_t, and o_t denote the forget, input, and output gates, respectively; $\widetilde{C}_t$ is the candidate cell state; C_t is the updated memory cell; h_t is the hidden state; x_t is the input at time t; $\sigma(\cdot)$ is the sigmoid activation function; and $\odot$ denotes element-wise multiplication. The function ρ works as a filter to eliminate any irrelevant information. The memory state vector (M_t) is

$$M_t = ft \odot Mt - 1 + it \odot M \tag{21}$$

During the prediction phase, to maintain the functioning of long-term memory operation, self-state functions are adopted. The output of hidden states of each long short-term memory is expressed as

$$h_t = ot \odot tanh(M_t) \tag{22}$$

and the final prediction outcome (y_t) is represented as

$$y_t = W_d y_t + b_d \tag{23}$$

where y_t is the prediction output, W_d denotes the weight, and b_d refers to the bias of dense net.

A look-back window of 12 months was adopted in this study. Therefore, the previous 12 observations were used as input to predict the subsequent SPI value. This window length was selected to capture long-term and short-term dependencies that are important for drought forecasting. To preserve the temporal structure of the drought series, the dataset was divided chronologically into training (70%), validation (10%), and testing (20%) subsets. The earliest observations were assigned to the training set, followed by the validation and testing sets. Random shuffling was not applied, ensuring that future observations were never used during model training and preventing information leakage.

2.3. Materials and Method Development

The proposed ORDWT-SLSTM model was developed to address two fundamental challenges in drought prediction: non-stationary signal behavior and long-term temporal dependency learning. The ORDWT decomposes complex drought signals into simpler frequency components, reducing prediction difficulty and improving interpretability. Subsequently, the SLSTM network captures hierarchical temporal dependencies within each decomposed component. The integration of the ORDWT and SLSTM enables the model to effectively model both short-term fluctuations and long-term drought trends. This combination improves prediction robustness and provides more reliable drought predictions across diverse climatic regions, represented by the Mackay, Gatton, and Darling Downs stations. Figure 6 depicts the proposed model.

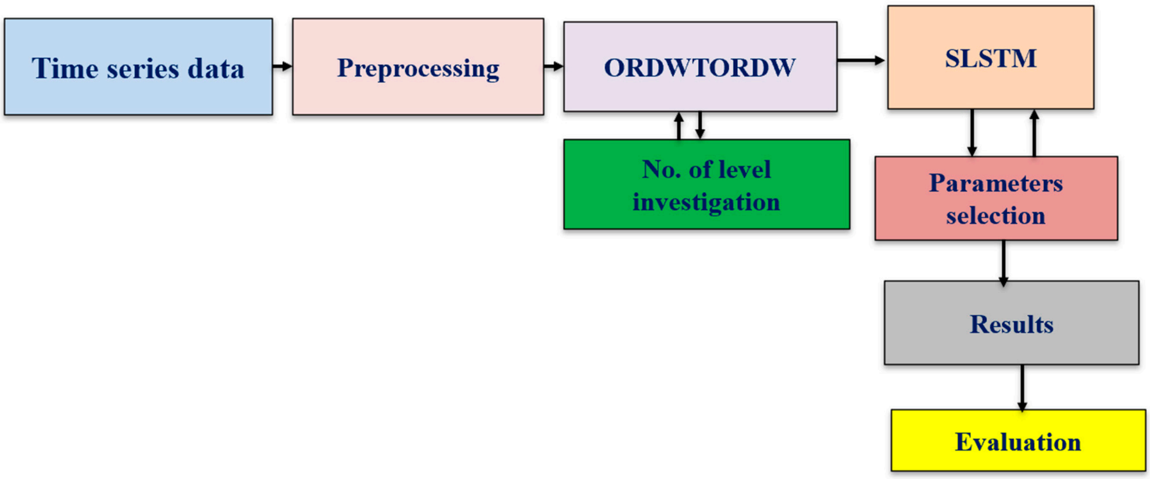

Figure 6. The proposed model for SPI drought detection.

Step 1: Data Collection: The time-series data were collected from three representative regions in Queensland, Australia, namely, Mackay, Gatton, and Darling Downs. These stations were selected because they represent distinct climatic conditions ranging from tropical coastal environments (Mackay) to subtropical inland agricultural regions (Gatton and Darling Downs). The data consisted of monthly climatic variables used to derive drought indicators and assess drought forecasting performance.

Step 2: Data Preprocessing: The collected time series were pre-processed and normalized to improve the data quality and enable efficient learning. Normalization prevents variables with large numerical ranges from dominating the learning process and improves network convergence. In this step, the missing and erroneous observations were removed using the Min-Max normalization formula:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where X represents the original observation, and X_{norm} is the normalized value.

Step 3: ORDWT: Drought time series were highly nonlinear and non-stationary due to the influence of climatic variability and seasonal fluctuations. Direct prediction of raw signals often results in reduced prediction accuracy. To address this issue, we applied the ORDWT to the collected time series. The ORDWT model decomposes the signal into multiple sub-components. The decomposition process separates low-frequency information from high-frequency components, thereby reducing signal complexity and improving prediction performance. The optimal decomposition level was determined experimentally by evaluating the prediction accuracy across multiple decomposition levels.

Step 4: SLSTM: The output of the ORDWT were sent to SLSTM network. Unlike a standard LSTM, SLSTM employs multiple hidden layers, allowing for the hierarchical extraction of temporal dependencies. The stacked architecture enables the model to capture both short-term and long-term drought evolution patterns. The memory mechanism of SLSTM effectively preserves drought-related information over extended periods, making it particularly suitable for drought forecasting applications.

Step 5: Evaluation: The performance is crucial to empirically proving the effectiveness of the proposed model. The evaluation phase involves employing statistical tests to evaluate predicted values of the proposed ORDWT-SLSTM model against actual values to determine the degree to which the proposed model aligns with the real output. The metrics listed in Table 2 are used to compare predicted values and actual values.

Table 2. Metrics used to evaluate proposed model.

Metrics	**Equations**
Mean Absolute Error (MAE)	$MAE = \frac{1}{m}\sum_{j=1}^{m}\left\|y_j - \ddot{y}_j\right\|$
Root-Mean-Square Error (RMSE)	$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(y_j - \ddot{y}_j)^2}$
Kling–Gupta Efficiency (KGE)	$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1\ )^2}$ Where $r = \frac{\sum_{i=1}^{n}(y_i-\overline{y})(\hat{y}_i-\overline{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i-\overline{\hat{y}})^2}}$, $\alpha = \frac{\sigma_{\hat{y}}}{\sigma_y}$, $\beta = \frac{\mu_{\hat{y}}}{\mu_y}$
Global Performance Index (GPI)	$GPI_i = \sum_{j=1}^{q}\alpha_j\left(\widetilde{M}_j \middle\| M_{ij}\right)$
Nash–Sutcliffe Efficiency (NSE)	$\mathrm{NSE} = 1 - \frac{\sum_{j=1}^{m}(y_j-\hat{y}_j)^2}{\sum_{j=1}^{m}(y_j-\ddot{y}_j)^2}$

Note(s): where y_j and $\hat{y}_j$ are the observed and predicted SPI values, respectively, $\overline{y}$ is the mean observed value, and m is the total number of observations. M_{ij} = normalized value of metric j for model i; $\widetilde{M}_j$ = median value of metric j; α_j = weighting coefficient; q = number of performance metrics; μ = mean; σ = standard deviation; r = correlation coefficient.

Extensive experiments were performed to verify the superior prediction performance of the proposed ORDWT-SLSTM. Several evaluation metrics were employed to validate the proposed model. All experiments were designed and conducted using MATLAB 2025a.

Step 6: Hyperparameter Selection: The parameters of the proposed model were selected carefully. Table 3 lists the parameters of the proposed model. Choosing the optimal parameters has significant effects on the accuracy of the prediction model. Firstly, we selected the parameters of the ORDWT carefully in this paper. Based on the simulation results, the number of decomposition levels (J) was set to 2, the rational dilation numerator (p) was set to $p = 5$, and the rational dilation denominator (q) was set to 6. Regarding the prediction model and baseline models, to obtain suitable parameters for each prediction model, we used the MSE as an indicator to evaluate the performance of each model. The loss function was set according to the MSE. The depth of the SLSTM was selected empirically. The epoch was set to 300.

Table 3. Parameters of proposed model.

Method	**Parameters**
ORDWT	p = 5; q = 6; r = 2
Stacked LSTM	Batch size = 64; learning rate = 0.001; dropout = 0.2; hidden units = 64; optimizer = Adam; number of LSTM layers = 2.
GRU	Hidden units = 128; dropout = 0.2; learning rate = 0.005; optimizer = Adam; batch size = 32; epochs = 150.
LSTM	Hidden units = 128; dropout = 0.2; learning rate = 0.001; optimizer = Adam; batch size = 32; epochs = 300.
BiLSTM	Hidden units = 128; dropout = 0.2; learning rate = 0.001; optimizer = Adam; batch size = 32; epochs = 300.

2.4. Benchmarks

Now, we should study some prediction models that operated as our baseline [28–33]. These models were compared with the proposed model to evaluate its performance. A short description of these models is given below:

- LSTM [28,29]: LSTM networks are a specialized type of RNN model. It includes specialized mechanisms named gates, namely, the forget gate, output gate, and input gate, as shown in Figure 7. These gates control how information flows through the network, allowing it to preserve important information over continued periods. The forget gate determines which past information to discard, the input gate manages the relevance of new information, and the output gate regulates what information the LSTM unit should pass to the next timestamp.
- BiLSTM [30]: BiLSTM is an extension of traditional LSTM. It consists of two independent LSTMs, as shown in Figure 8: one processing the input sequence forward and the other processing it backward. BiLSTM processes the input data in both the backward and forward directions. By examining input from past-to-future and future-to-past, it captures and contextualizes the information required for future prediction.
- GRU [31–33]: GRU is a type of RNN model. The GRU model was developed in 2014 to handle sequential data by resolving the vanishing gradient issue of traditional RNNs. GRUs are a simplified development of LSTM, where multiple gates are merged into update and reset gates. Figure 9 illustrates the GRU model.

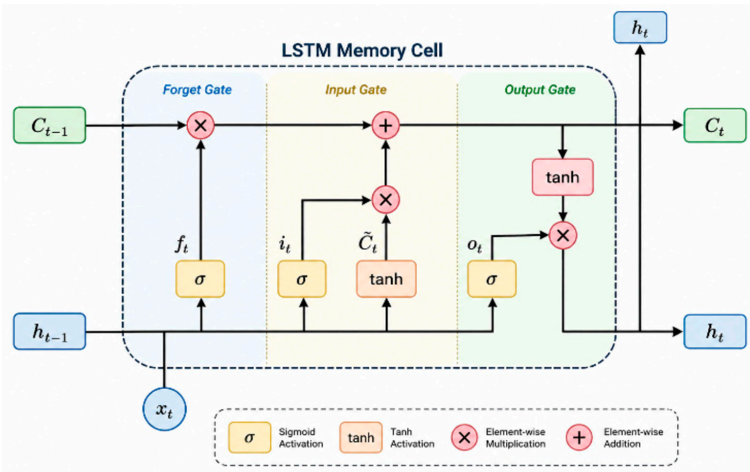

Figure 7. The main architecture of LSTM.

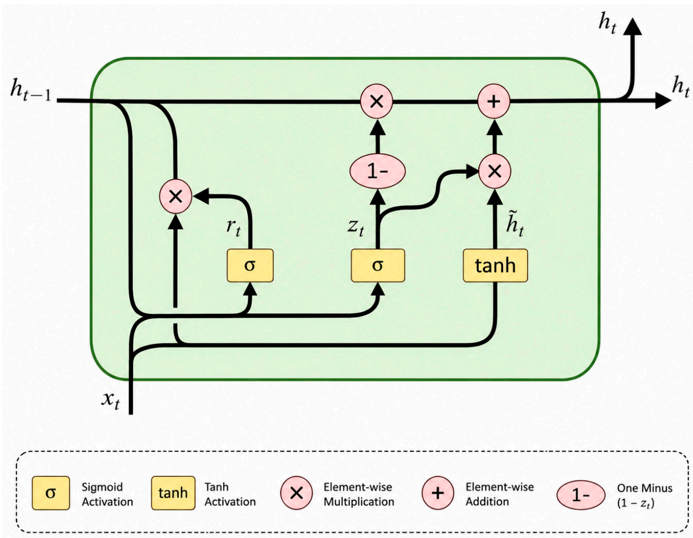

Figure 8. The architecture of the GRU model.

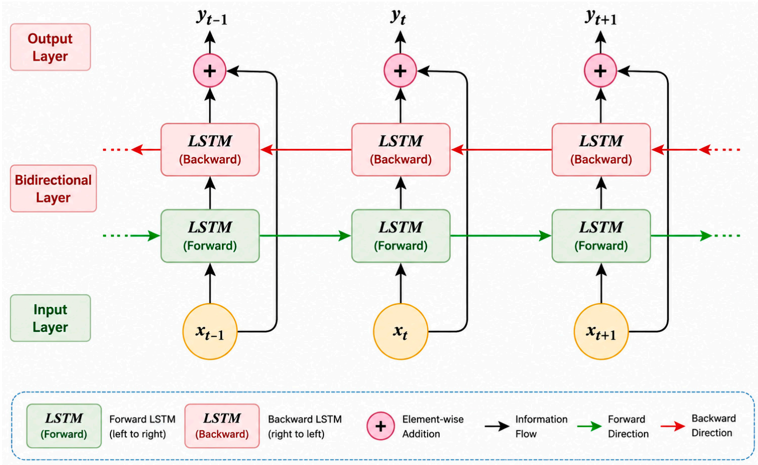

Figure 9. The main design of BiLSTM.

3. Results

In this paper, a robust SPI drought prediction model was developed. The ORDWT was integrated with SLSTM. A chronological sliding window strategy was used to construct the supervised learning dataset. The collected samples were divided chronologically into 70% training, 10% validation, and 20% testing subsets. The training set corresponded to the earliest portion of the time series, followed by the validation set, while the final 20% of observations were used exclusively for testing. This procedure avoids random shuffling and prevents temporal data leakage.

To determine the optimal decomposition level, ORDWT decomposition levels from 1 to 6 were tested as shown in Table 4. The selection criterion was based on the prediction performance using the RMSE and MAE. The fourth decomposition level consistently produced the lowest prediction errors across the stations and SPI timescales. Increasing the decomposition level beyond four did not yield additional performance gains and occasionally resulted in slight accuracy degradation due to excessive signal fragmentation. Therefore, decomposition level four was adopted throughout this study. Table 4 shows that the four levels of decomposition provided more accurate predictive results and improved the performance of the prediction SLSTM model.

Table 4. Prediction accuracy based on decomposition level.

Decomposition Level	RMSE	MAE	NSE
1	0.397	0.287	0.862
2	0.366	0.299	0.890
3	0.233	0.131	0.943
4	0.215	0.125	0.986
5	0.249	0.154	0.921
6	0.255	0.154	0.911

To quantitatively evaluate the decomposition quality and degree of mode mixing, the Orthogonality Index (OI) was employed. The OI measures the level of correlation between decomposed components, where values closer to zero designate better mode separation and reduced information leakage among modes. Consequently, lower OI reflects a superior decomposition performance and reduced mode mixing. Table 5 presents the QI values attained for different decomposition methods. The proposed ORDWT scored the lowest OI value (0.019), substantially lower than the DWT (0.088), EMD (0.069), and MEMD

(0.040). This result indicates that the ORDWT generates more independent and orthogonal sub-series with significantly reduced mode mixing. Improved mode separation assists the extraction of clearer multi-scale drought patterns and provides more informative inputs to the SLSTM forecasting model. Consequently, the superior decomposition quality of the ORDWT contributed directly to the improved prediction accuracy observed across the SPI timescales.

Table 5. Quantitative evaluation of decomposition quality using OI.

Method	Orthogonality Index (OI)	Average Inter-Mode Correlation
DWT	0.088	0.33
EMD	0.069	0.27
MEMD	0.040	0.19
ORDWT	0.019	0.08

3.1. Comparison with Other Decomposition Models

The proposed ORDWT-SLSTM was compared with DWT-SLSTM, EMD-SLSTM, and MVEMD-SLSTM. In this experiment, the proposed stacked LSTM was combined with DWT-SLSTM, EMD-SLSTM, and MVEMD-SLSTM to evaluate the proposed ORDWT-SLSTM model consistency in capturing temporal dependencies in the SPI-12, SPI-3, SPI-1, and SPI-6 time series. Tables 6–9 report the prediction rates in terms of the RMSE, MAE, and NSE across the three stations.

Table 6. Comparison with other decomposition models for SPI-1 prediction across three stations.

	Darling Downs			Gatton			MacKay		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.215	0.125	0.986	0.260	0.147	0.973	0.225	0.130	0.980
DWT-SLSTM	0.476	0.325	0.826	0.497	0.345	0.803	0.517	0.380	0.804
EMD-SLSTM	0.266	0.175	0.903	0.255	0.165	0.913	0.290	0.193	0.884
MEMD-SLSTM	0.247	0.157	0.915	0.236	0.155	0.903	0.240	0.141	0.890

Table 7. Comparison with other decomposition models for SPI-3 prediction across three stations.

	Darling Downs			Gatton			MacKay		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.223	0.144	0.951	0.270	0.153	0.960	0.241	0.151	0.961
DWT-SLSTM	0.534	0.404	0.811	0.530	0.410	0.790	0.552	0.421	0.791
EMD-SLSTM	0.273	0.183	0.899	0.260	0.172	0.88	0.311	0.211	0.870
MEMD-SLSTM	0.253	0.16	0.901	0.220	0.170	0.899	0.251	0.162	0.880

Table 8. Comparison with other decomposition models for SPI-6 prediction across three stations.

	Darling Downs			Gatton			MacKay		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.244	0.149	0.950	0.276	0.157	0.958	0.243	0.157	0.960
DWT-SLSTM	0.551	0.408	0.810	0.536	0.415	0.783	0.556	0.426	0.790
EMD-SLSTM	0.279	0.189	0.890	0.268	0.177	0.899	0.318	0.213	0.866
MEMD-SLSTM	0.259	0.169	0.900	0.289	0.176	0.895	0.259	0.168	0.878

Table 9. Comparison with other decomposition models for SPI-12 prediction across three stations.

	Darling Downs			Gatton			MacKay		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.288	0.152	0.945	0.289	0.159	0.952	0.290	0.161	0.955
DWT-SLSTM	0.576	0.424	0.800	0.566	0.488	0.745	0.599	0.499	0.777
EMD-SLSTM	0.288	0.199	0.886	0.298	0.209	0.865	0.376	0.298	0.854
MEMD -SLSTM	0.276	0.187	0.900	0.299	0.199	0.886	0.321	0.211	0.865

The results obtained for SPI-1 demonstrate, as shown in Table 6, that the proposed ORDWT-SLSTM model consistently outperformed the EMD-SLSTM, DWT-SLSTM, and MEMD-SLSTM across two stations in terms of the RMSE, MAE, and NSE. This improvement was significant for short-term drought prediction, where high-frequency variability and noise often degrade model performance. It was noticed that the DWT-SLSTM and EMD-SLSTM struggled to extract rapid fluctuations; however, the MVMD-SLSTM showed promising results because MVMD decomposes the signal into more stable intrinsic modes compared to VMD and DWT, which reduced noise and enhanced temporal learning. Consequently, the MVMD-SLSTM model scored the second lowest error values and higher NSE scores, indicating superior accuracy and reliability in capturing short-term drought dynamics. For SPI-3, the proposed ORDWT-SLSTM model showed clear superiority over all competing approaches, highlighting its ability to extract seasonal drought patterns. For SP-3, the signal exposes smoother variations compared to SPI-1 yet still contains nonlinear dependencies that challenge traditional models. The proposed ORDWT-SLSTM model leveraged multi-scale decomposition to isolate meaningful components, allowing the LSTM to learn temporal dependencies more effectively. As a result, the ORDWT-SLSTM model produced consistently lower RMSE and MAE values while achieving higher NSE, demonstrating its robustness in modeling seasonal drought variability across different locations.

Table 8 reports the case of SPI-6, which reflects medium-term drought conditions, where the performance gap between the proposed ORDWT-SLSTM model and baseline methods, MEMD-SLSTM, EMD-SLSTM, and MVEMD-SLSTM, become even more evident. Traditional decomposition methods such as DWT, EMD and DWT showed limitations in preserving signal consistency across longer temporal scales, leading to increased prediction errors. In contrast, the ORDWT provided better mode separation and reduced mode mixing, enabling a more accurate representation of underlying drought trends. This allowed the ORDWT-SLSTM model to capture long-range dependencies more effectively, resulting in improved predictive performance and higher stability across all stations. For SPI-12 long-term drought conditions, the superiority of the proposed ORDWT-SLSTM model was further proven, as shown in Table 9. At this scale, capturing persistent trends and low-frequency components is critical, and any decomposition inefficiency can significantly impact the prediction accuracy. The ORDWT model ensured the better decomposition and preservation of long-term signals, allowing the LSTM to model sustained drought patterns more accurately. Consequently, the proposed ORDWT-SLSTM model achieved the lowest RMSE and MAE values and the highest NSE across all stations, confirming its effectiveness in long-term drought forecasting and its robustness in handling complex temporal dependencies.

Overall, the consistent performance improvement in the ORDWT-SLSTM model across all SPI timescales showed its ability to effectively capture both short-term variability and long-term dependencies. By combining advanced signal decomposition with deep temporal learning, the proposed ORDWT-SLSTM model provides a more accurate and robust solution for drought prediction compared to existing methods.

For further evaluation, Taylor diagrams were employed. Figure 10 reports the results obtained. It can be observed that across SPI-1, SPI-3, SPI-6, and SPI-12, the proposed ORDWT-SLSTM model consistently demonstrated a superior performance compared to DWT-SLSTM, EMD-SLSTM, and MEMD-SLSTM. In each case, the ORDWT-SLSTM generated close values to the reference, indicating a standard deviation nearly matching the observed data, as well as a minimal centered RMSE and the highest correlation (approaching 0.96–0.98). This consistent proximity to the reference point across short-, medium-, and long-term drought indices showed that the ORDWT-SLSTM model effectively captured both high-frequency variability (SPI-1, SPI-3) and long-term trends (SPI-6, SPI-12). In contrast, the competing models were located farther from the reference, reflecting weaker correlation structures and larger deviation errors, particularly at longer timescales. The robustness of the ORDWT-SLSTM across all SPI horizons highlighted its strong generalization capability and its ability to preserve the intrinsic temporal dynamics of drought signals, making it the most reliable and accurate model for multi-timescale drought forecasting.

The regression plots in Figure 11 between the actual and predicted SPI-1, SPI-3, SPI-6, and SPI-12 provide powerful evidence of the ORDWT-SLSTM's superior predictive ability. Figure 10 displays the regression plots for Darling Downs, Gatton, and MacKay stations, where the ORDWT-SLSTM showed nearly perfect agreement with the actual values, recording an average of regression of 0.95 ± 1.6. This remarkable agreement implied that the proposed ORDWT-SLSTM produced close values to the actual values in all SPI cases. In addition, the MEMD-SLSTM model showed some promising results and scored the second highest regression values across the three stations. However, the DWT-SLSTM demonstrated an inconsistent performance and showed a poor performance across the three stations for SPI-1, SPI-3, SPI-6 and SPI-12. The Global Performance Indicator (GPI) was also employed to evaluate the proposed model against the hybrid models. Figure 12 displays the results obtained based on the GPIs for all three stations. It was observed that the proposed ORDWT-SLSTM model recorded the best values of the GPI across the three stations compared to the MEMD-SLSTM, EMD-SLSTM, and DWT-SLSTM. The results proved the ability of the proposed model in the prediction of short-term- and long-term-drought SPIs. The proposed ORDWT-SLSTM model recorded the highest GPI values across the three stations.

To further validate the performance of the proposed ORDWT-SLSTM model, a Wilcoxon signed-rank test was conducted using NSE prediction results from all three stations and all SPI forecasting horizons, resulting in 12 paired observations for each model comparison. The test was conducted to determine whether the observed improvements over competing decomposition-based models were statistically significant. The results of the Wilcoxon test are reported in Table 10. It can be observed that the proposed ORDWT-SLSTM consistently recorded higher NSE values than the DWT-SLSTM, EMD-SLSTM, and MEMD-SLSTM across all SPI horizons. The Wilcoxon signed-rank test confirmed that the performance improvements were statistically significant ($p < 0.05$), implying that the observed improvements were unlikely to have been caused by random variation. The largest performance gap was observed between the ORDWT-SLSTM and DWT-SLSTM, highlighting the effectiveness of the ORDWT in extracting multi-scale drought dynamics. Furthermore, statistically significant improvements over the EMD-SLSTM and MEMD-SLSTM demonstrated that the proposed DWT-SLSTM model provides more informative signal representations for subsequent temporal learning by the SLSTM model. Overall, the statistical analysis provides additional evidence supporting the strength, reliability, and generalization ability of the proposed ORDWT-SLSTM model for drought forecasting across multiple SPI timescales.

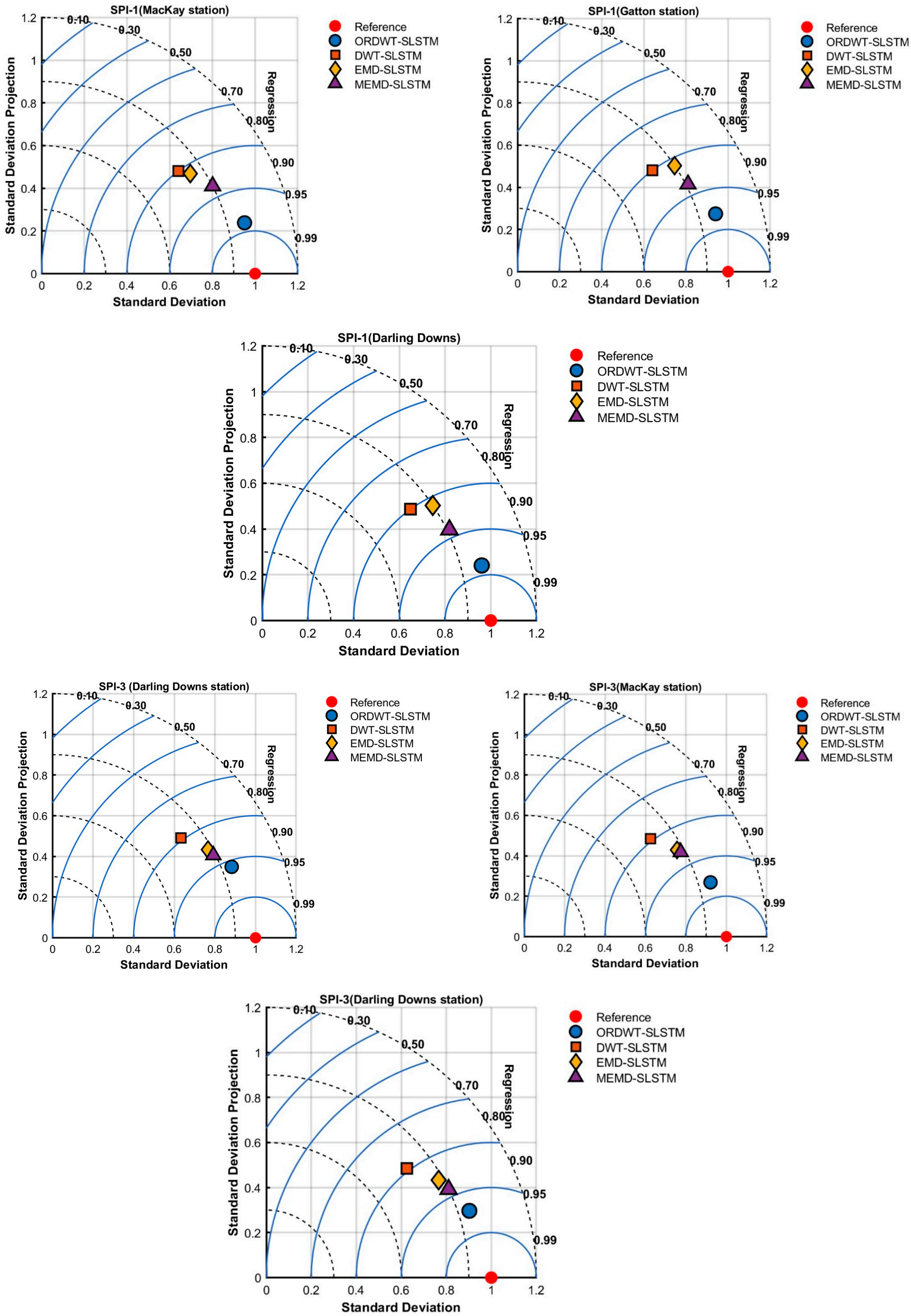

Figure 10. *Cont.*

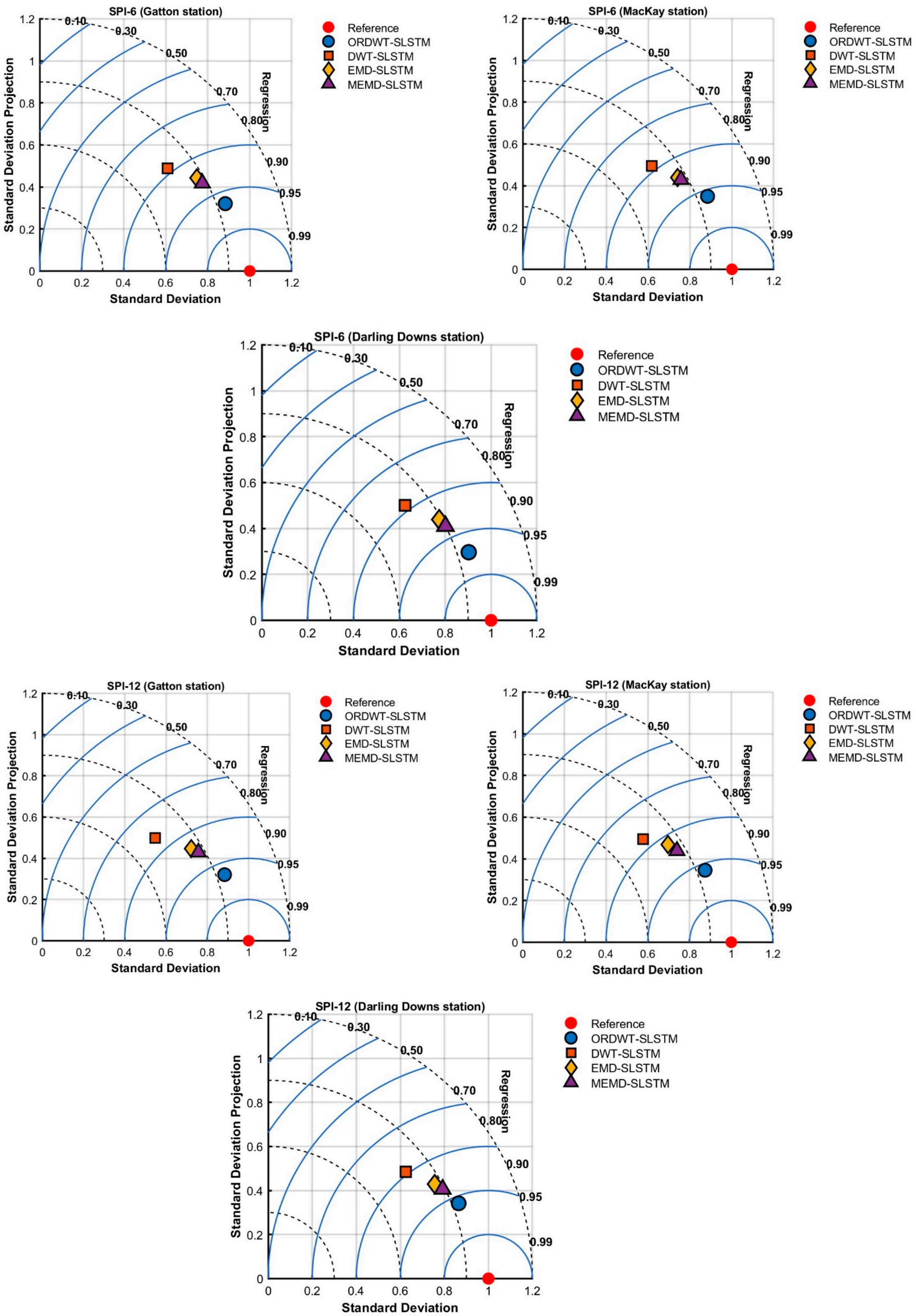

Figure 10. Taylor plots for proposed model and baseline approaches across three stations.

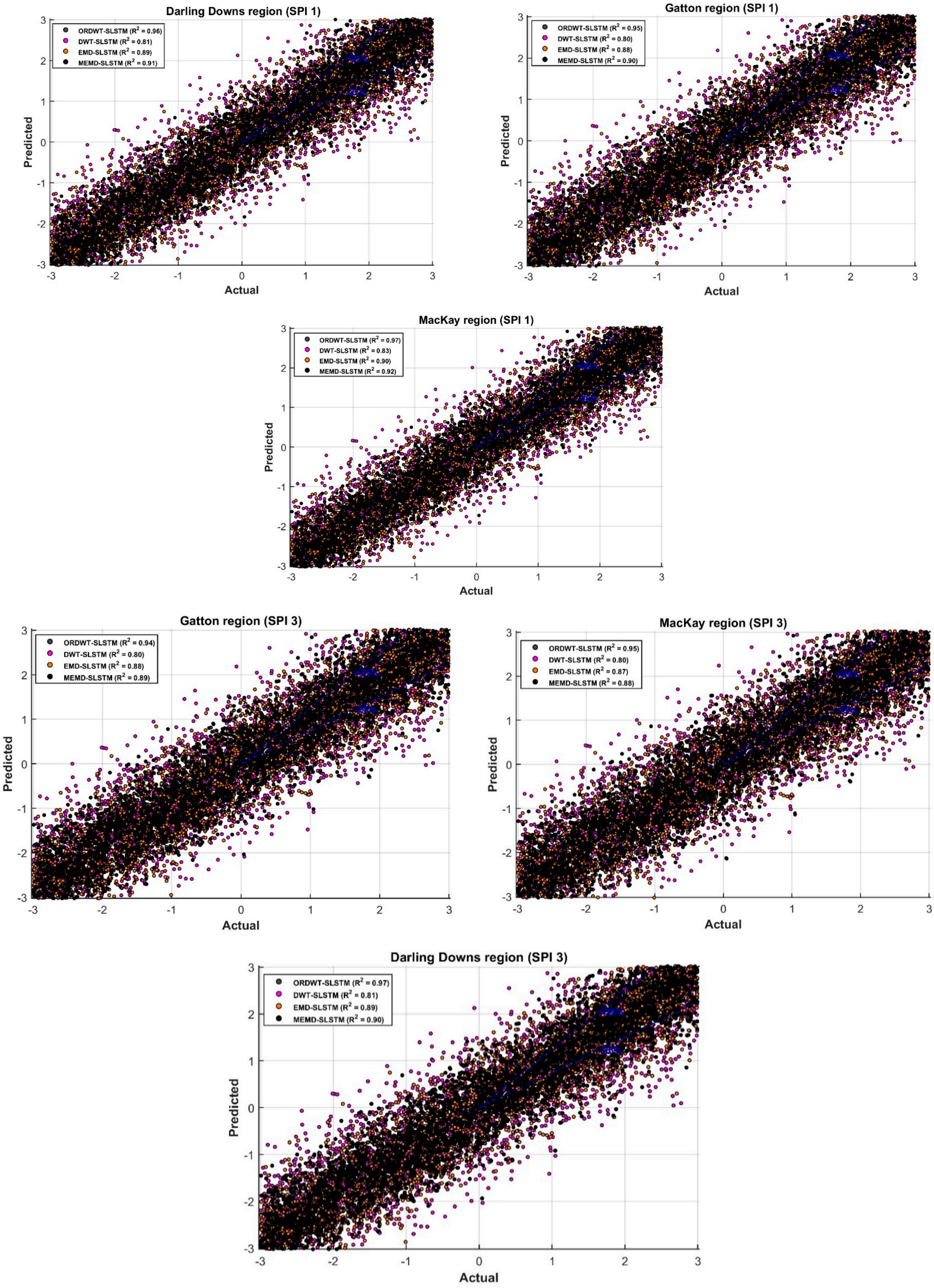

Figure 11. *Cont.*

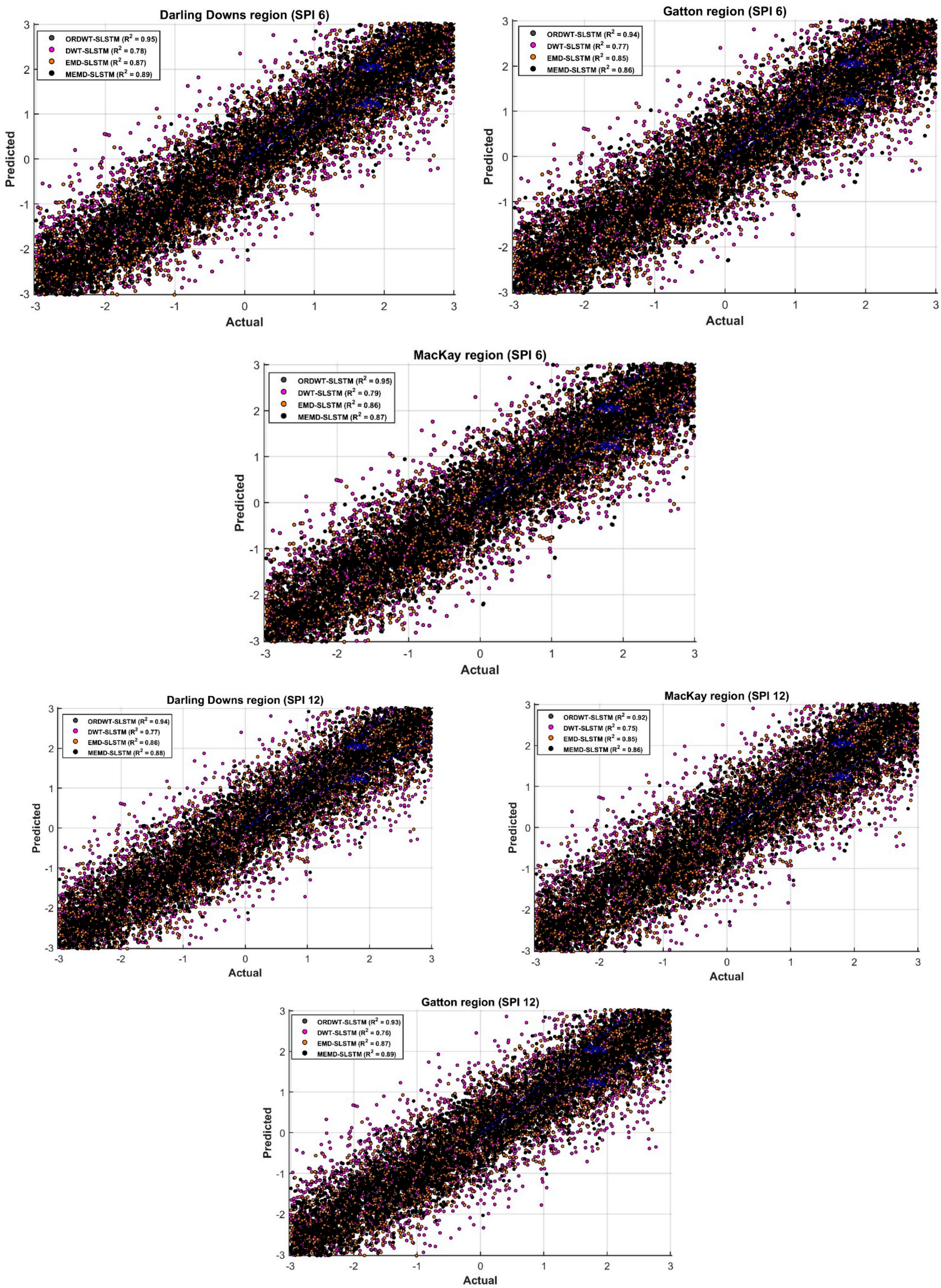

Figure 11. Regression plots for proposed model and compared approaches across three stations.

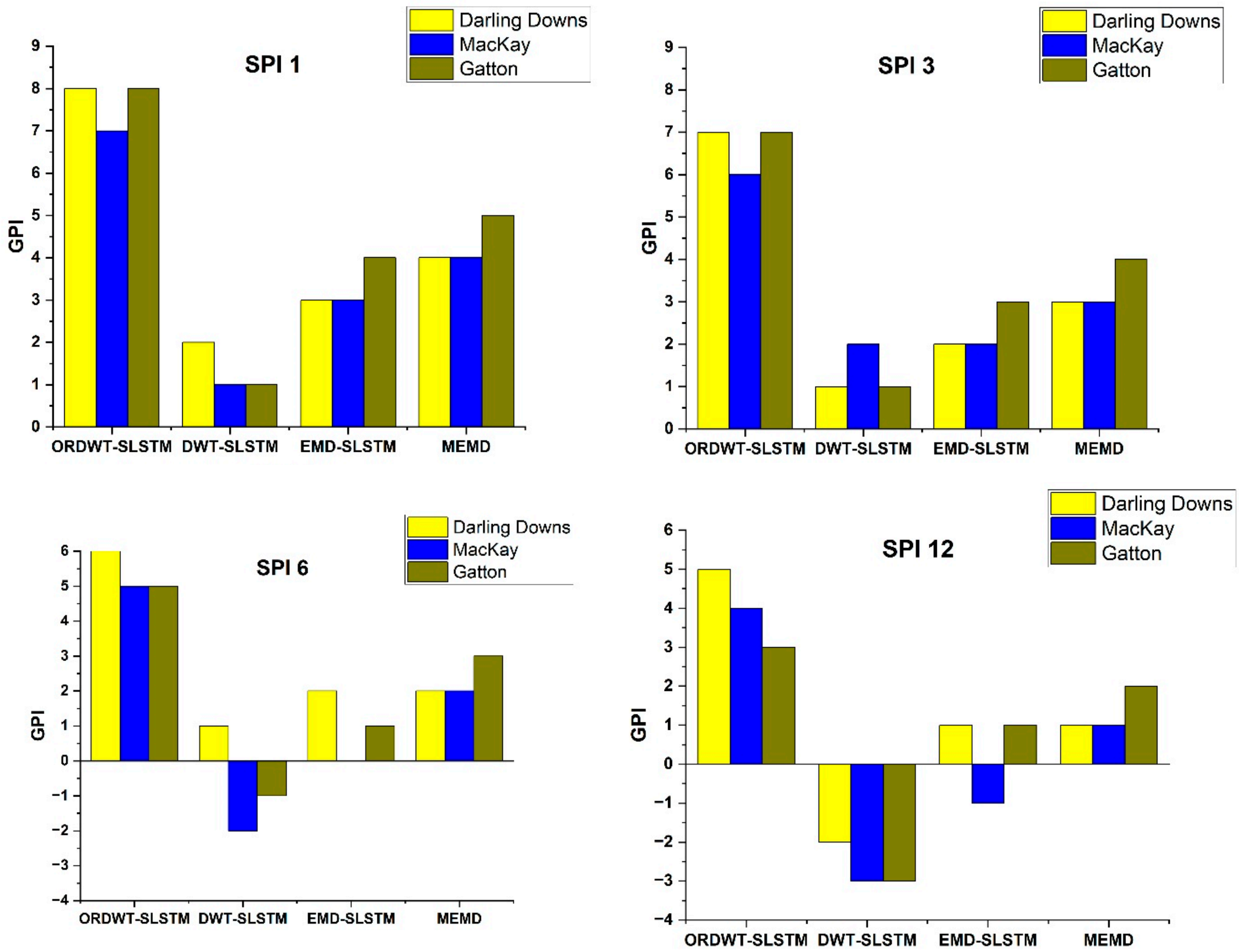

Figure 12. Bar graphs using GPI of the proposed model against compared approaches across three stations.

Table 10. Results of Wilcoxon test using NSE.

Comparison	Test Statistic (W)	*p*-Value	Significant (α = 0.05)
ORDWT-SLSTM vs. DWT-SLSTM	0.0	0.00002	Yes
ORDWT-SLSTM vs. EMD-SLSTM	4.0	0.00051	Yes
ORDWT-SLSTM vs. MEMD-SLSTM	7.0	0.00063	Yes

3.2. Effect of Integration of ORDWT with Other Deep Learning Models

The proposed ORDWT decomposition model was integrated with several deep learning models, including LSTM, BiLSTM, and GRU. The results across SPI-1, SPI-3, SPI-6, and SPI-12 consistently demonstrated the superior and stable performance of the proposed ORDWT-SLSTM model over all comparative approaches. Tables 11–14 show the results obtained. As shown in Table 11 for SPI-1, the proposed ORDWT-SLSTM model achieved the highest NSE alongside relatively low RMSEs and MAEs in some cases (SPI-1, SPI-3, and SPI-6), indicating strong predictive accuracy and robustness. The ORDWT-BiLSTM scored the second highest accuracy compared with the ORDWT-GRU. The ORDWT-SLSTM scored the lowest RMSE in one case (SPI-12). However, the ORDWT-GRU exhibited noticeable performance degradation, particularly in terms of higher errors and lower NSEs across Darling Downs, Gatton, and Mackay stations. Importantly, this consistency persists from short-term (SPI-1) to long-term (SPI-12) forecasting, confirming the ability of the ORDWT-SLSTM to effectively capture both rapid fluctuations and long-term drought dynamics. The stability of the ORDWT-SLSTM across multiple temporal scales highlights its strong generalization capability and its effectiveness at preserving the intrinsic structure of SPI signals, thereby establishing it as the most accurate and reliable framework for multi-timescale drought prediction. Figure 13 presents the Kling–Gupta Efficiency (KGE) values obtained by the four prediction models across different SPI horizons and stations. Overall, the proposed ORDWT-SLSTM consistently achieved the highest KGE values across most stations and prediction horizons, demonstrating its superior ability to reproduce the observed SPI drought. For Mackay station, the ORDWT-SLSTM attained KGE values ranging from 0.90 to 0.98, outperforming the ORDWT-LSTM, ORDWT-BiLSTM, and ORDWT-GRU across all SPI timescales. Similar trends were observed for Gatton and Darling Downs, where the proposed model maintained KGE values above 0.94 for most drought scales. The results indicate that integrating the ORDWT with the SLSTM architecture improved the representation of drought patterns by effectively capturing both short-term and long-term temporal dependencies. Compared with the ORDWT-LSTM, the stacked architecture provides additional hierarchical feature learning, resulting in improved correlation, variability representation, and bias reduction. Furthermore, the consistently lower KGE values obtained by the ORDWT-GRU suggest that the simplified gating mechanism may be less effective at modeling complex drought dynamics compared with the memory structure of LSTM-based models.

Table 11. Prediction results for SPI-1 across three stations.

	Darling Downs			**Gatton**			**MacKay**		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.215	0.125	0.986	0.260	0.147	0.973	0.225	0.130	0.980
ORDWT-LSTM	0.255	0.175	0.921	0.256	0.177	0.920	0.258	0.179	0.920
ORDWT-BiLSTM	0.244	0.154	0.934	0.245	0.157	0.932	0.248	0.159	0.935
ORDWT-GRU	0.288	0.217	0.899	0.289	0.213	0.893	0.290	0.233	0.891

Table 12. Prediction results for SPI-3 across three stations.

	Darling Downs			**Gatton**			**MacKay**		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.223	0.144	0.951	0.270	0.153	0.960	0.241	0.151	0.961
ORDWT-LSTM	0.257	0.178	0.920	0.259	0.178	0.918	0.266	0.182	0.917
ORDWT-BiLSTM	0.251	0.162	0.930	0.253	0.161	0.931	0.257	0.169	0.922
ORDWT-GRU	0.311	0.233	0.888	0.312	0.235	0.883	0.317	0.236	0.881

Table 13. Prediction results for SPI-6 across three stations.

	Darling Downs			**Gatton**			**MacKay**		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.244	0.149	0.950	0.276	0.157	0.958	0.243	0.157	0.960
ORDWT-LSTM	0.259	0.179	0.916	0.261	0.181	0.910	0.271	0.189	0.908
ORDWT-BiLSTM	0.258	0.169	0.925	0.262	0.171	0.925	0.269	0.173	0.920
ORDWT-GRU	0.318	0.237	0.882	0.322	0.245	0.867	0.333	0.256	0.863

Table 14. Prediction results for SPI-12 across three stations.

	Darling Downs			Gatton			MacKay		
	RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
ORDWT-SLSTM	0.288	0.152	0.945	0.289	0.159	0.952	0.290	0.161	0.955
ORDWT-LSTM	0.263	0.182	0.910	0.266	0.183	0.902	0.275	0.190	0.900
ORDWT-BiLSTM	0.264	0.173	0.912	0.268	0.174	0.920	0.269	0.177	0.914
ORDWT-GRU	0.324	0.241	0.881	0.327	0.248	0.861	0.327	0.248	0.861

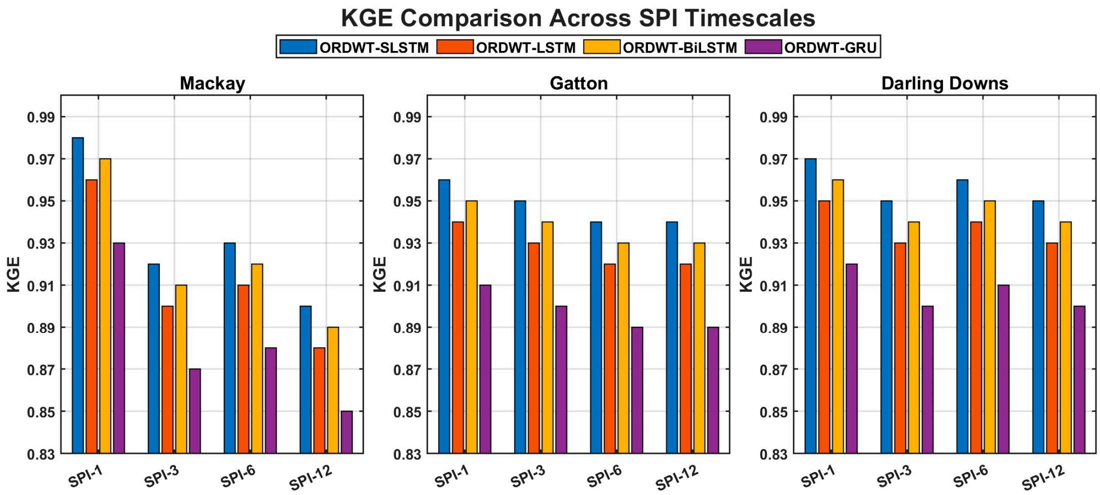

Figure 13. Performance evaluation using KGE metric.

Figure 14 reports the PBIAS results for the different horizons across the three stations. The proposed ORDWT-SLSTM consistently produced PBIAS values closest to zero, indicating minimal underestimation of drought conditions. In contrast, the DWT-SLSTM exhibited the largest positive and negative bias values across most SPI horizons, suggesting less reliable prediction behavior. The EMD-SLSTM and MEMD-SLSTM achieved moderate bias reduction but remained inferior to the proposed model. The superior bias characteristics of the ORDWT-SLSTM showed that the ORDWT technique effectively preserves the underlying drought signal while reducing prediction distortion, resulting in more accurate and unbiased SPI forecasts across diverse climatic conditions.

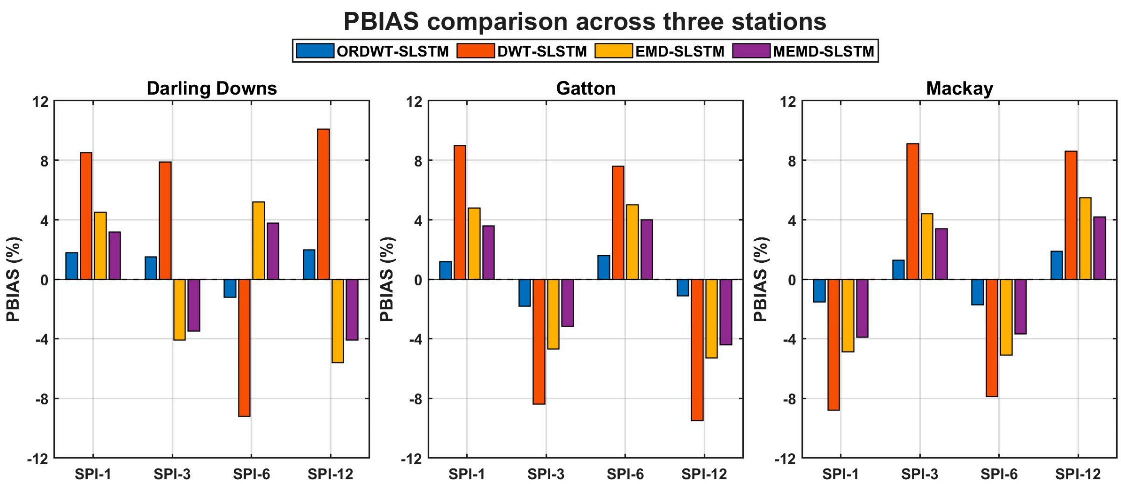

Figure 14. Comparison of prediction biases across forecasting models using PBIAS metric.

To further evaluate the model robustness and predictive uncertainty, each prediction model was independently executed five times using different random initialization seeds. The mean RMSE, standard deviation, and 95% confidence intervals were calculated across the repeated experiments. As shown in Table 15, the proposed ORDWT-SLSTM achieved the lowest average RMSE (0.223 ± 0.002) and the narrowest confidence interval [0.222, 0.224], indicating superior prediction accuracy and highly stable performance. The competing ORDWT-LSTM, ORDWT-BiLSTM, and ORDWT-GRU models exhibited larger RMSE values and wider confidence intervals. The small variability observed across repeated runs confirms that the proposed model was robust to random initialization and provides a reliable SPI forecasting performance. In addition, training and validation losses were monitored during model training, and no significant divergence was observed, indicating limited overfitting. These findings further support the effectiveness of integrating orthogonal wavelet decomposition with the stacked LSTM architecture for drought prediction.

Table 15. Uncertainty analysis of forecasting models based on five independent runs.

Model	**Average RMSE ± SD**	**95% CI**
ORDWT-SLSTM	0.236 ± 0.004	[0.233, 0.239]
ORDWT-BiLSTM	0.254 ± 0.005	[0.250, 0.258]
ORDWT-LSTM	0.261 ± 0.006	[0.256, 0.266]
ORDWT-GRU	0.315 ± 0.007	[0.309, 0.321]

In addition, uncertainty analysis of the forecasting models using 95% confidence was conducted. Figure 15 reports the prediction uncertainty analysis of the ORDWT-SLSTM, ORDWT-BiLSTM, ORDWT-LSTM, and ORDWT-GRU for SPI-1, SPI-3, SPI-6, and SPI-12 prediction. The black line refers to the actual SPI-12 values, while the red line represents the model predictions. The blue shaded region denotes the 95% confidence interval, reflecting the uncertainty of the mean prediction, while the yellow shaded region represents the 95% prediction interval, showing the expected variability in future observations. The narrower uncertainty bands observed for the ORDWT-SLSTM model show greater prediction stability and robustness compared with the ORDWT-LSTM, and ORDWT-GRU models. The findings in Figure 15 support our results in Tables 11–14. The proposed ORDWT-SLSTM model consistently demonstrated superior stability and reliability for long-term and short-term drought forecasting.

3.3. Ablation Study

To assess the contribution of the proposed ORDWT- SLSTM model, an ablation study was conducted by comparing the standalone SLSTM model with the proposed ORDWT-SLSTM, ORDWT-LSTM, and LSTM. The SLSTM and LSTM models were trained directly on the raw drought time series, whereas the ORDWT-SLSTM and ORDWT-LSTM first decomposed the input signal into multiple frequency components before temporal modeling. Table 16 presents comparative results. The standalone SLSTM model achieved a competitive prediction performance due to its ability to capture nonlinear temporal dependencies. However, the proposed ORDWT-SLSTM consistently achieved lower RMSE and MAE values together with higher NSE values across all evaluated SPI timescales. The improvements observed demonstrate that the improvements achieved by the proposed ORDWT-SLSTM model are not solely attributable to the SLSTM model. Instead, the ORDWT decomposition plays a critical role by reducing signal complexity, preserving multi-scale drought characteristics, and providing more informative representations for subsequent temporal learning. By separating the drought signal into multiple frequency components, the ORDWT enables the SLSTM and LSTM models to learn both short-term fluctuations and long-term drought persistence patterns more effectively than when trained directly on raw observations. Overall, the ablation study confirms that the integration of the ORDWT contributes substantially to the predictive accuracy and validates the effectiveness of the proposed decomposition learning strategy for drought prediction.

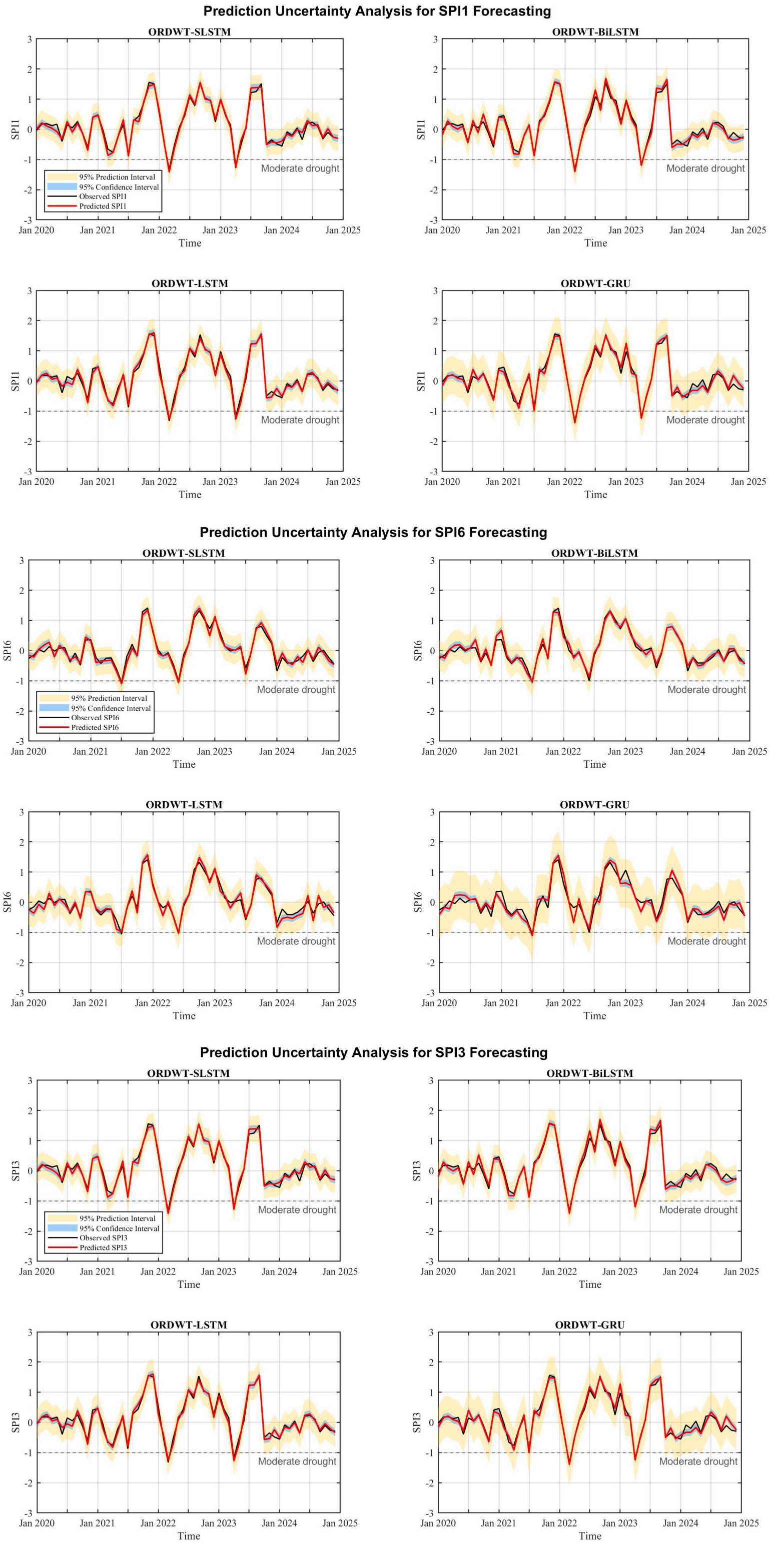

Figure 15. *Cont.*

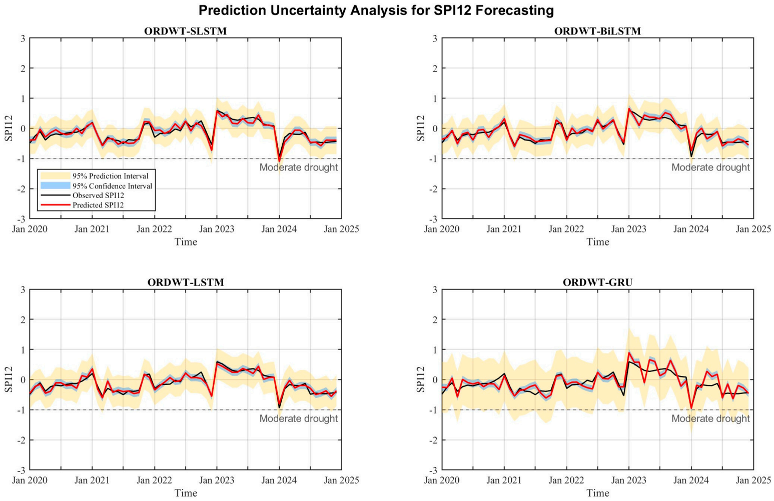

Figure 15. Uncertainty analysis of forecasting models using 95% confidence and prediction intervals.

Table 16. Comparing the standalone SLSTM model with the proposed ORDWT-SLSTM, ORDWT-LSTM, and LSTM.

Prediction Horizon	Model	Darling Downs			Gatton			MacKay		
		RMSE	MAE	NSE	RMSE	MAE	NSE	RMSE	MAE	NSE
SPI-12	ORDWT-SLSTM	0.288	0.152	0.945	0.289	0.159	0.952	0.290	0.161	0.955
	ORDWT-LSTM	0.263	0.182	0.910	0.266	0.183	0.902	0.275	0.190	0.900
	SLSTM	0.272	0.190	0.900	0.293	0.187	0.899	0.295	0.188	0.900
	LSTM	0.374	0.282	0.831	0.381	0.289	0.822	0.390	0.291	0.820
SPI-6	ORDWT-SLSTM	0.244	0.149	0.950	0.276	0.157	0.958	0.243	0.157	0.960
	ORDWT-LSTM	0.259	0.179	0.916	0.261	0.181	0.910	0.271	0.189	0.908
	SLSTM	0.266	0.187	0.910	0.267	0.189	0.911	0.269	0.190	0.910
	LSTM	0.372	0.275	0.841	3.744	0.277	0.839	0.375	0.280	0.833
SPI-3	ORDWT-SLSTM	0.223	0.144	0.951	0.270	0.153	0.960	0.241	0.151	0.961
	ORDWT-LSTM	0.257	0.178	0.920	0.259	0.178	0.918	0.266	0.182	0.917
	SLSTM	0.255	0.173	0.912	0.254	0.172	0.916	0.242	0.171	0.918
	LSTM	0.376	0.278	0.837	0.376	0.279	0.839	0.376	0.277	0.840
SPI-1	ORDWT-SLSTM	0.215	0.125	0.986	0.260	0.147	0.973	0.225	0.130	0.980
	ORDWT-LSTM	0.255	0.175	0.921	0.256	0.177	0.920	0.258	0.179	0.920
	SLSTM	0.211	0.165	0.922	0.210	0.164	0.923	0.200	0.154	0.925
	LSTM	3.755	0.266	0.841	0.375	0.261	0.843	0.372	0.257	0.851

3.4. Comparison with Present Time-Series Prediction Models

To further evaluate the effectiveness of the proposed ORDWT-SLSTM model, additional experiments were conducted using recent drought prediction architectures, including transformers and informer models. These models were selected because they have demonstrated promising performances in drought prediction. Table 17 summarizes the prediction performances of all evaluated models across the SPI-1, SPI-3, SPI-6, and SPI-12 prediction horizons. The results show that informer models provided a competitive prediction performance due to their ability to capture temporal dependencies within drought time series. However, informer-based models struggled with long-term drought prediction (SPI-12, SPI-6). In addition, informer models showed weak performance compared with transformer models on all horizons. The proposed ORDWT-SLSTM model consistently achieved the best overall performance. This finding suggests that accurate drought forecasting depends not only on the prediction architecture itself but also on the quality of the input signal representation. By employing the ORDWT, the model effectively decomposes complex drought signals into informative multi-scale components while preserving important temporal characteristics. Consequently, the subsequent SLSTM network can learn more discriminative temporal patterns from the decomposed signals than models trained directly on raw drought time series.

Table 17. Comparison with more recent models across three stations.

Prediction Horizon	Model	Darling Downs			Gatton			MacKay		
		RMSE	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**	**RMSE**	**MAE**	**NSE**
SPI-12	ORDWT-SLSTM	0.288	0.152	0.945	0.289	0.159	0.952	0.290	0.161	0.955
	SLSTM	0.272	0.190	0.900	0.293	0.187	0.899	0.295	0.188	0.900
	Transformer	0.353	0.221	0.803	0.356	0.233	0.800	0.357	0.235	0.802
	Informer	0.311	0.211	0.856	0.316	0.212	0.854	0.317	0.215	0.856
SPI-6	ORDWT-SLSTM	0.244	0.149	0.950	0.276	0.157	0.958	0.243	0.157	0.960
	SLSTM	0.266	0.187	0.910	0.267	0.189	0.911	0.269	0.190	0.910
	Transformer	0.367	0.234	0.799	0.368	0.240	0.798	0.369	0.242	0.799
	Informer	0.321	0.215	0.851	0.339	0.217	0.850	0.341	0.239	0.849
SPI-3	ORDWT-SLSTM	0.223	0.144	0.951	0.270	0.153	0.960	0.241	0.151	0.961
	SLSTM	0.255	0.173	0.912	0.254	0.172	0.916	0.242	0.171	0.918
	Transformer	0.311	0.200	0.821	0.312	0.201	0.820	0.310	0.198	0.821
	Informer	0.291	0.200	0.865	0.295	0.203	0.862	0.291	0.188	0.871
SPI-1	ORDWT-SLSTM	0.215	0.125	0.986	0.260	0.147	0.973	0.225	0.130	0.980
	SLSTM	0.211	0.165	0.922	0.210	0.164	0.923	0.200	0.154	0.925
	Transformer	0.300	0.189	0.830	0.308	0.192	0.811	0.311	0.194	0.821
	Informer	0.276	0.180	0.891	0.279	0.179	0.888	0.281	0.186	0.891

3.5. Time Complexity Analysis

Table 18 presents the complexity time analysis of the proposed SLSTM against the standard prediction models. As expected, the GRU model exhibited the lowest computational time, requiring only 107 K trainable parameters and 171 s of training time. The LSTM and BiLSTM models demonstrated moderate computational requirements, while the proposed SLSTM model increased both the model capacity and training time due to its deeper stacked architecture. The proposed ORDWT-SLSTM model scored the highest training time because of the additional multi-decomposition stage. Importantly, the number of trainable parameters remains comparable to that of the standalone SLSTM model, indicating that the performance improvement was primarily attributed to the enhanced feature representation provided by the ORDWT rather than increased model complexity. The increase in computational cost was justified by the substantial improvement in the prediction accuracy observed across all SPI timescales and stations. Furthermore, the testing time remained low, indicating that the proposed ORDWT-SLSTM model was suitable for practical drought prediction applications once model training was completed. These results demonstrate that the proposed SLSTM model achieved an effective balance between predictive performance and computational time, making it a viable solution for long-term drought monitoring and early-warning systems.

https://doi.org/10.3390/w18141686

Table 18. Complexity time analysis.

Model	Trainable Parameters (K)	Training Time	Testing Time
LSTM	140	220 s	0.76 s
GRU	107	171 s	0.65 s
BiLSTM	230	270 s	0.83 s
SLSTM	280	350 s	0.96 s
ORDWT-SLSTM	280	800 s	1.12 s

4. Conclusions

A novel ORDWT-SLSTM model for long-term and short-term SPI drought prediction was developed. SLSTM and the ORDWT were combined and compared against state-of-the-art models. Three regions in Queensland, named Mackay, Darling Downs, and Gatton, were selected to evaluate the proposed model. The experimental results showed that the proposed ORDWT-SLSTM model outperforms traditional decomposition models, including DWT, EMD, and MEMD, and DWT. In addition, the proposed ORDWT model was integrated with LSTM, BiLSTM, and GRU, and the results showed that the SLSTM outperformed the classic deep learning models in terms of the RMSE, MAE, and NSE metrics for certain station-specific cases. Although the proposed ORDWT-SLSTM model demonstrated promising drought prediction results, several opportunities remain for further improvement. First, the proposed ORDWT-SLSTM model was tested using data from three stations in Queensland. This may limit its generalizability to other climatic regions. Second, future research will investigate the integration of Artificial Intelligence (XAI) techniques, such as SHAP, Local Interpretable Model-Agnostic Explanations, and attention-based interpretation mechanisms, to improve the transparency and interpretability of the prediction process. Such approaches would provide valuable insights into how the model utilizes multi-scale drought information when generating predictions, thereby improving user trust and supporting decision making in operational drought management. Moreover, optimization strategies will be applied to reduce computational time and enable real-time applications.

Author Contributions: Conceptualization, A.H.L., M.A. and M.D.; methodology, A.H.L., M.A. and M.D.; validation, A.H.L.; formal analysis, M.A. and M.D.; investigation, A.H.L. and M.D.; resources, M.A. and M.D.; data curation, A.H.L. and M.D.; writing—original draft preparation, M.D. and M.A.; writing—review and editing; visualization, M.D.; supervision, M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia under grant no. (IPP: 1227-155-2025). The authors, therefore, acknowledge with thanks DSR for technical and financial support.

Data Availability Statement: The datasets presented in this article can only be requested from the corresponding author due to data confidentiality. The authors do not have permission to share the data publicly.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Oyounalsoud, M.S.; Yilmaz, A.G.; Abdallah, M.; Abdeljaber, A. Drought prediction using artificial intelligence models based on climate data and soil moisture. *Sci. Rep.* **2024**, *14*, 19700. [CrossRef] [PubMed]
2. Dayal, K.S.; Deo, R.C.; Apan, A.A. Spatio-temporal drought risk mapping approach and its application in the drought-prone region of south-east Queensland, Australia. *Nat. Hazards* **2018**, *93*, 823–847. [CrossRef]
3. McKee, T.B.; Doesken, N.J.; Kleist, J. The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology, Anaheim, CA, USA, 17–22 January 1993; Volume 17, pp. 179–183.
4. de Oliveira Luz, A.E.; Negri, R.G.; Santos, L.B.L. Deep learning approaches for drought forecasting, monitoring and impact assessment: A comprehensive systematic review. *Appl. Soft Comput.* **2026**, *197*, 115196. [CrossRef]
5. Liu, Q.; Zhang, S.; Zhang, H.; Bai, Y.; Zhang, J. Monitoring drought using composite drought indices based on remote sensing. *Sci. Total Environ.* **2020**, *711*, 134585. [CrossRef] [PubMed]
6. Dai, H.; Xiong, L.; Ma, Q.; Duan, Z. Deep learning model for drought prediction based on large-scale spatial causal network in the Yangtze River Basin. *J. Hydrol.* **2025**, *654*, 132808. [CrossRef]
7. Liu, Q.; Du, M. Global assessment of drought risk to expanded urban land from 2020 to 2100. *Land Use Policy* **2026**, *167*, 108040. [CrossRef]
8. Ashouri, H.; Hsu, K.L.; Sorooshian, S.; Braithwaite, D.K.; Knapp, K.R.; Cecil, L.D.; Nelson, B.R.; Prat, O.P. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 69–83. [CrossRef]
9. Zarikos, I.; Politi, N.; Gounaris, N.; Vlachogiannis, D.; Sfetsos, A. Assessing Climate Change Impacts on Precipitation Volume and Drought Characteristics Across Basin and Sub-Basin Scales in Greece. *Water* **2026**, *18*, 872. [CrossRef]
10. Zheng, X.; Chen, Y.; Liu, C.; Bento, V.A.; Wu, X.; Zhang, R.; Qi, J.; Wang, Q. Comprehensive Analysis of Spatial–Temporal Patterns and Trends of Compound Drought and High Temperature Events from 1982 to 2023 Across China. *Water* **2026**, *18*, 943. [CrossRef]
11. Kebede, A.; Warrach-Sagi, K.; Schwitalla, T.; Wulfmeyer, V.; Abebe, T.; Tadesse, T. A multivariate approach to drought monitoring: Improving robustness and accuracy through a new drought index in regions with high climate variability, applied to the drought-prone region of Ethiopia. *J. Hydrol. Reg. Stud.* **2026**, *64*, 103332. [CrossRef]
12. Khan, M.M.H.; Muhammad, N.S.; El-Shafie, A. Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *J. Hydrol.* **2020**, *590*, 125380. [CrossRef]
13. Nury, A.H.; Taher, A.; Alam, S.; Afroz, R.; Anti, S.D.; Majumdar, S.N.; Munna, G.M. Assessment of groundwater variability using ARIMA, random forest, and LSTM-RNN in the northeastern region of Bangladesh. *H2Open J.* **2025**, *8*, 336–360. [CrossRef]
14. Tang, S.; Wang, S.; Jiang, J.; Zheng, Y. Improved flash drought forecasting and attribution: A spatial-temporal causality-aware deep learning approach. *J. Hydrol.* **2026**, *667*, 134945. [CrossRef]
15. Xing, X.; Wei, S.; Chen, X.; Qian, J.; Peng, S.; Sun, J.; Sun, B.; Chen, C. A deep learning-based composite agricultural drought index for monitoring and impact assessment in Central Asia. *Agric. Water Manag.* **2026**, *323*, 110043. [CrossRef]
16. Mishra, A.K.; Desai, V.R.; Singh, V.P. Drought forecasting using a hybrid stochastic and neural network model. *J. Hydrol. Eng.* **2007**, *12*, 626–638. [CrossRef]
17. Kim, T.W.; Valdés, J.B. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J. Hydrol. Eng.* **2003**, *8*, 319–328. [CrossRef]
18. Özger, M.; Başakın, E.E.; Ekmekcioğlu, Ö.; Hacısüleyman, V. Comparison of wavelet and empirical mode decomposition hybrid models in drought prediction. *Comput. Electron. Agric.* **2020**, *179*, 105851. [CrossRef]
19. Bacanli, U.G.; Firat, M.; Dikbas, F. Adaptive neuro-fuzzy inference system for drought forecasting. *Stoch. Environ. Res. Risk Assess.* **2009**, *23*, 1143–1154.
20. Farokhnia, A.; Morid, S.; Byun, H.R. Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. *Theor. Appl. Climatol.* **2011**, *104*, 71–81.
21. Mishra, A.K.; Singh, V.P. Analysis of drought severity-area-frequency curves using a general circulation model and scenario uncertainty. *J. Geophys. Res. Atmos.* **2009**, *114*, D06120. [CrossRef]
22. Abbes, A.B.; Inoubli, R.; Rhif, M.; Farah, I.R. Combining deep learning methods and multi-resolution analysis for drought forecasting modeling. *Earth Sci. Inform.* **2023**, *16*, 1811–1820. [CrossRef]
23. Kadam, C.M.; Bhosle, U.V.; Holambe, R.S. Deep learning-driven regional drought assessment: An optimized perspective. *Earth Sci. Inform.* **2024**, *17*, 1523–1537. [CrossRef]
24. Bui, D.T.; Khosravi, K.; Tiefenbacher, J.; Nguyen, H.; Kazakis, N. Improving prediction of water quality indices using novel hybrid machine-learning algorithms. *Sci. Total Environ.* **2020**, *721*, 137612. [CrossRef] [PubMed]
25. Kaur, A.; Sood, S.K. Deep learning based drought assessment and prediction framework. *Ecol. Inform.* **2020**, *57*, 101067. [CrossRef]
26. Al Ghayab, H.R.; Li, Y.; Diykh, M.; Sahi, A.; Abdulla, S.; Alkhuwaylidee, A.R. EEG based over-complete rational dilation wavelet transform coupled with autoregressive for motor imagery classification. *Expert Syst. Appl.* **2025**, *269*, 126433. [CrossRef]

27. Bayram, I.; Selesnick, I.W. Overcomplete discrete wavelet transforms with rational dilation factors. *IEEE Trans. Signal Process.* **2008**, *57*, 131–145. [CrossRef]
28. Yu, J.; Pu, H.; Sun, D.W. Stacked long and short-term memory (SLSTM)-assisted terahertz spectroscopy combined with permutation importance for rapid red wine varietal identification. *Talanta* **2025**, *291*, 127650. [CrossRef] [PubMed]
29. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of LSTM and BiLSTM in forecasting time series. In *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data)*; IEEE: New York, NY, USA, 2019; pp. 3285–3292.
30. Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. In *Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*; IEEE: New York, NY, USA, 2017; pp. 1597–1600.
31. Diykh, M.; Ali, M.; Jamei, M.; Abdulla, S.; Uddin, M.P.; Farooque, A.A.; Labban, A.H.; Alabdally, H. Empirical curvelet transform based deep DenseNet model to predict NDVI using RGB drone imagery data. *Comput. Electron. Agric.* **2024**, *221*, 108964. [CrossRef]
32. Diykh, M.; Ali, M.; Farooque, A.A.; Aldhafeeri, A.A.; Jamei, M.; Labban, A. A robust artificial intelligence informed over complete rational dilation wavelet transform technique coupled with deep learning for long-term rainfall prediction. *Eng. Appl. Artif. Intell.* **2026**, *165*, 113426. [CrossRef]
33. Diykh, M.; Ali, M.; Labban, A.H.; Prasad, R.; Jamei, M.; Abdulla, S.; Farooque, A.A. Designing empirical fourier decomposition reinforced with multiscale increment entropy and deep learning to forecast dry bulb air temperature. *Results Eng.* **2025**, *26*, 104597. [CrossRef]