Long Lead Time Drought Forecasting Using Lagged Climate Variables and a Stacked Long Short-term Memory Model

Abhirup Dikshit[1], Biswajeet Pradhan[1,2,3,4*], Abdullah M. Alamri[5]

Long Lead Time Drought Forecasting Using Lagged Climate Variables and a Stacked Long Short-term Memory Model

[1]Centre for Advanced Modelling and Geospatial Information Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, New South Wales 2007, Australia

[2]Department of Energy and Mineral Resources Engineering, Sejong University, Choongmu-gwan, 209, Neungdongro Gwangjin-gu, Seoul 05006, Korea

[3]Centre of Excellence for Climate Change Research, King Abdulaziz University, P. O. Box 80234, Jeddah 21589, Saudi Arabia

[4]Earth Observation Centre, Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

[5]Department of Geology and Geophysics, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

*Corresponding Author: Biswajeet Pradhan, biswajeet.pradhan@uts.edu.au

Abstract

Drought forecasting with a long lead time is essential for early warning systems and risk management strategies. The use of machine learning algorithms has been proven to be beneficial in forecasting droughts. However, forecasting at long lead times remains a challenge due to the effects of climate change and the complexities involved in drought assessment. The rise of deep learning techniques can solve this issue, and the present work aims to use a stacked long short-term memory (LSTM) architecture to forecast a commonly used drought measure, namely, the Standard Precipitation Evaporation Index. The model was then applied to the New South Wales region of Australia, with hydrometeorological and climatic variables as predictors. The multivariate interpolated grid of the Climatic Research Unit was used to compute the index at monthly scales, with meteorological variables as predictors. The architecture was trained using data from the period of 1901–2000 and tested on data from the period of 2001–2018. The results were then forecasted at lead times ranging from 1 month to 12 months. The forecasted results were analysed in terms of drought characteristics, such as drought intensity, drought onset, spatial extent and number of drought months, to elucidate how these characteristics improve the understanding of drought forecasting. The drought intensity forecasting capability of the model used two statistical metrics, namely, the coefficient of determination (R^2) and root-mean-square error. The variation in the number of drought months was examined using the threat score technique. The results of this study showed that the stacked LSTM model can forecast effectively at short-term and long-term lead times. Such findings will be essential for government agencies and can be further tested to understand the forecasting capability of the presented architecture at shorter temporal scales, which can range from days to weeks.

Keywords: Drought forecasting; Deep learning; Lead time; Standard Precipitation Evaporation Index; New South Wales; Australia

Acronyms

LSTM – Long Short-term Memory

NSW – New South Wales

SPEI – Standardied Precipitation Evapotranspiration Index

SPI – Standardised Precipitation Index

CRU – Climate Research Unit

IOD – Indian Ocean Dipole

SAM – Southern Annular Mode

ENSO – El Niño–Southern Oscillation

PDO – Pacific Decadal Oscillation

SOI – Southern Oscillation Index

SST – Sea Surface Temperature

R^2 – Coefficient of Determination

RMSE – Root-mean-square Error

TS – Threat Score

ARIMA – Autoregressive Integrated Moving Average

ANN – Artificial Neural Network

1. Introduction

Droughts are amongst the most complex geohazards, and they have been recognised as the least understood among all ‘weather and climate extremes’ (Pulwarthy and Sivakumar, 2014). Droughts can last from a few weeks to decades and span from the local to the national scale (Pendergrass et al., 2020), causing significant damage to agriculture (Nasim et al., 2018), water resources (van Loon, 2015; Imad et al., 2019) and socioeconomic factors (Mishra and Singh, 2010). The impact of droughts is felt across different sectors, making establishing a universal drought definition practically impossible (Lloyd-Hughes, 2014). Therefore, defining different drought types on the basis of their impact on a specific sector is necessary (Vicente-Serrano et al., 2020). Historically, drought definitions have been classified into meteorological, agricultural, hydrological and socioeconomic (Mishra and Singh, 2010). However, some researchers have argued over expanding the definition to include other critical areas, such as groundwater (Mishra and Singh, 2010) and ecological (Crausbay et al., 2017) and environmental aspects (Vicente-Serrano et al., 2020). Such argument is well justified and will enable us to make a clearer distinction, and consequently, understand the propagation of drought. To date, however, such a consensus has yet to be reached in the drought community (Vicente-Serrano et al., 2020). The present study focuses on meteorological drought, which is a result of rainfall deficiency (Mishra and Singh, 2010).

To understand drought processes and effects, drought characteristics, such as intensity, duration and spatial extent, should be determined (van Loon, 2015; Parry et al., 2016). Such quantification can be performed on the basis of the truncation levels of a specific drought-affecting variable or by computing an indicator (Kallis, 2008). In general, indicators are used, in which a single (McKee et al., 1993) or a combination (Vicente-Serrano et al. 2011) of drought-affecting variables is utilised, conveying various drought characteristics. A deluge of information on the types of indices is available, and data should be used with advantages and limitations. Additional details regarding this can be found in Nagarajan (2009), Zargar et al. (2011) and Yihdego et al. (2019). One index that has demonstrated high capability in accurately assessing meteorological droughts under different climatic conditions is the Standardised Precipitation Evapotranspiration Index (SPEI) developed by Vicente-Serrano et al. (2010; 2012). SPEI can be considered superior to its predecessor, i.e. the Standardised Precipitation Index (SPI), which uses only rainfall to compute the index; by contrast, SPEI uses evapotranspiration and rainfall (Beguería et al., 2014). These indices are calculated at different time scales, representing short-term droughts (1–3 months) and long-term droughts (6–24 months).

Among different types of drought studies, forecasting at different lead times is particularly challenging (Hao et al., 2018). Historically, forecasting studies have revolved around the use of stochastic models, such as an autoregressive integrated moving average (ARIMA) model, which can understand the seasonality and lag in a time series (Han et al., 2010; Mishra and Singh, 2011). However, droughts are essentially nonlinear in nature, and thus, three types of models are used, namely, data-driven (Morid et al., 2007), physical (Wanders and Wood, 2016) and hybrid (Wang et al., 2012) models. The advantages and limitations of using these models have been discussed in the review articles of Mishra and Singh (2011) and Hao et al. (2018). Recently, focus on the use of data-driven models, which have been demonstrated to improve forecasting results compared with physical-based models, has increased (Abbott and Marohsay, 2014; Hao et al., 2018). Artificial neural networks (ANNs) are amongst the most popular and effective data-driven models, and they have been extensively used in the past decade and proven to be effective tools for forecasting at short and long lead times (Mishra and Desai, 2006; Barua et al., 2012; Özger et al., 2012; Dikshit et al., 2020a). Important references that highlight the advancements of neural networks in droughts or associated variables can be found in Rodrigues et al. (2018) and Fung et al. (2019). Despite obtaining satisfactory forecasting results, neural networks are incapable of dealing with non-stationarities in drought estimations and suffer from overfitting due to lag components involved in time series data (Alizadeh and Nikoo, 2018). Considering the aforementioned limitations, interest in the use of deep learning approaches, particularly LSTM, which is capable of retaining information for longer periods due to its recurrent and gate architecture, has been an increasing (Hochreiter and Schmidhuber, 1997). The use of LSTM in drought forecasting is still in its infancy, with the majority of studies focusing on forecasting drought variables, such as rainfall (Gao et al., 2020), sea surface temperature (SST) (Xiao et al., 2019), evaporation (Majhi et al., 2020) and El Niño–Southern Oscillation (ENSO) (Ham et al., 2019). In a recent paper published in Nature magazine, Reichestein et al. (2019) highlighted the various achievements in geosciences by using deep learning models and they provided several recommendations for future use. The present study is the first to forecast a drought index by using a stacked LSTM architecture at different lead times.

Drought occurrences are an amalgamation of a multitude of reasons, and thus, modellers are frequently befuddled when selecting variables that will be used as predictors in forecasting droughts (Deo et al., 2017). A significant milestone in drought forecasting is the discovery of atmospheric circulation patterns or teleconnections that affect drought events (Stahl and Demuth, 1999; Schubert et al., 2004). This finding has encouraged researchers to use climatic and SST indices as predictors for drought forecasting with long lead times (Woli et al., 2013; Seager and Hoerling, 2014; Schubert et al., 2016). Kirono et al. (2010) found that the relationship between climatic drivers and rainfall in Australia is one of the world's highest. Thus, the trend towards using large-scale climatic drivers as predictors for forecasting droughts has been increasing (Hao et al., 2018). Studies that utilise climatic drivers have been conducted and have achieved improvements in forecasting drought indices or variables.

Abbott and Marohasy (2014) adopted an ANN to forecast monthly rainfall at a lead time of 1 month using lagged climatic variables for the Queensland region of Australia; the ANN outperformed the dynamic models used by the Bureau of Meteorology. Similarly, Deo et al. (2017) and Feng et al. (2020) forecasted SPI at different lead times and achieved improvements in results when using lagged climate variables as predictors. Therefore, the current research also used climatic drivers as predictors and examined their implications for forecasting at different lead times.

The novel contribution of this work is to develop and validate the utility of a state-of-the-art stacked LSTM architecture for monthly SPEI forecasting at different lead times in the southeastern part of Australia. The methodology used in the present work is illustrated in Figure 1. The model adopts several hydrometeorological and climatic indices as input to the model. The major contribution of this work is the use of a global climatological dataset and a deep learning model to forecast droughts. This study is the first to use both aforementioned aspects, and it will help future research forecasts droughts at the country/global scale. The primary objectives of the present study are as follows: 1) to analyse the forecasting capabilities of deep learning models at longer lead times, 2) to understand variations in the forecasted results based on different drought characteristics (e.g. drought intensity, number of drought months and spatial extent) and 3) to use climatic variables as predictors for drought forecasting.

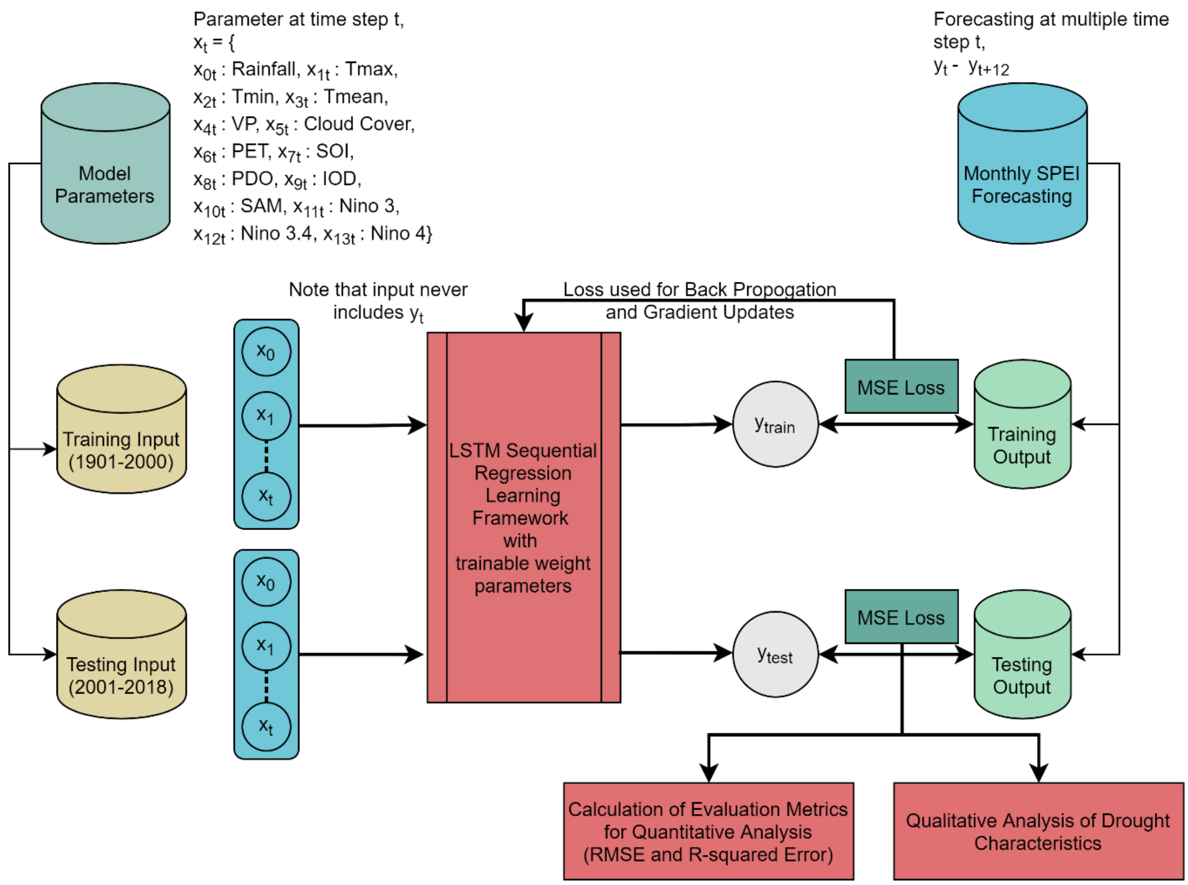


Figure 1. Flowchart of the methodology used in the present study

The remainder of this paper is organised as follows. The ‘Study area’ section describes the history of droughts in the region and the effects of various climatic indices on drought occurrence. It also discusses the datasets and the drought indices used in the study. Section 2, i.e. ‘Model development’, describes the architecture used in the present study. The ‘Results’ section presents the findings in terms of metrics and drought characteristics. The ‘Discussion’ section compares the major findings in the present study and the available literature along with the limitations. Lastly, the ‘Conclusion’ section concludes this study and summarises its results.

2. Study Area

The study area in the present work is New South Wales (NSW), which is located in the southeast of Australia. The study area was selected because of its long history of droughts. It is also one of Australia’s major agricultural belts. The country frequently experiences drought conditions and is the driest inhabited continent in the world (Ummenhofer et al., 2009). The economic impact of droughts from 2017 to 2019 was estimated to be US$8.1 billion (Wittwer, 2020). The region suffered three major droughts and several minor droughts since 1900 (Dikshit et al., 2020a). The Millennium Drought (2001–2009) is regarded as the country’s worst drought. The effect of climate change has worsened the drought situation in NSW, with increased intensity and frequency during hot days (Cai et al., 2014). The recent bushfires in this region were aggravated by drought conditions, dry vegetation and temperature increase (Steffen et al., 2019).

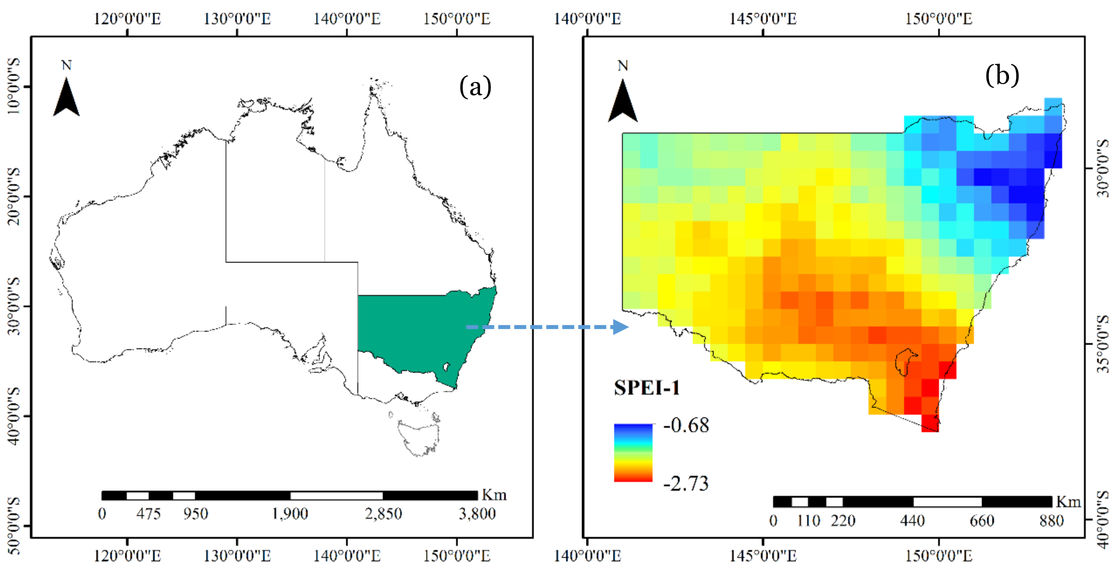


Figure 2. Location of the study area and the lowest SPEI (1) value during the Millennium Drought (2001–2009)

2.1 Dataset and Variables

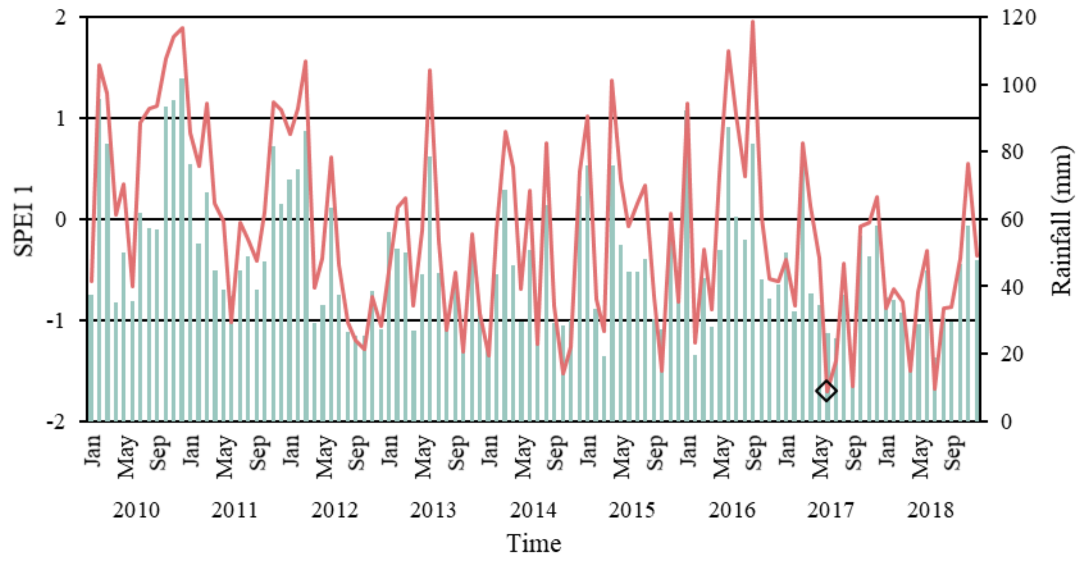
The dataset used for determining SPEI and the potential meteorological predictors can be either ground-based or interpolated grids (AghaKouchak et al., 2015). The challenge in ground-based datasets is that they are prone to manual errors and typically lack long time series data. Interpolated grids are a viable solution to such problems, and they have been extensively used in various geohazard studies. Sun et al. (2018) examined 30 different interpolated datasets for rainfall values and found that the Climate Research Unit (CRU) provided observational values. However, the researchers also suggested that dataset choice should be based on a study's objective and research area. The CRU dataset (with a spatial resolution of 0.5° × 0.5°) was developed by the University of East Anglia and established using several stations conducting quality control and homogeneity check (Harris et al., 2020). In the present study, the CRU TS 4.03 dataset, which spans the years 1901 to 2018, was used.

The climatic drivers that affect droughts in the region arise from atmospheric circulation patterns emerging from the Pacific, Indian, and Southern Oceans (Duc et al., 2017). The atmospheric ocean phenomenon arising from the Pacific Ocean is known as ENSO and the Pacific Decadal Oscillation (PDO). The phase and strength of an ENSO event are defined using an index, which can be based either on SST or surface atmospheric pressure (Hanley et al., 2003). SST-based indices that affect NSW include Niño 3.0, Niño 3.4 and Niño 4.0; meanwhile, the surface atmospheric pressure index is the Southern Oscillation Index (SOI)

(Duc et al., 2017). Similarly, SST anomalies arising from the Indian Ocean are depicted using Indian Ocean Dipole (IOD), which is known as the Southern Annular Mode (SAM) in the Southern Ocean. These climatic drivers are known to influence rainfall during different seasonal periods (Hendon et al., 2007; Risbey et al., 2009; Duc et al., 2017). Power et al. (1998) found that droughts in the region mostly result from the El Niño phase of the ENSO cycle. Ummenhofer et al. (2009) determined that the Millennium Drought was caused by a combination of SAM and the negative IOD phase. Their study also suggested that anomalies in the Indian Ocean lead to severe drought conditions. The wet periods after the Millennium Drought were due to a strong La Niña event and a positive SAM event (Gergis et al., 2012). The aforementioned studies highlight that large-scale climatic indices are intertwined with drought occurrences, and thus, considering these variables as predictors is essential.

2.2 Drought Index and Characteristics

Amongst available drought indices, SPEI has been found to be a useful index for encapsulating drought characteristics. SPEI has been tested extensively in different parts of the world, highlighting various climatic regions (Vicente-Serrano et al., 2012). The determination of SPEI involves the use of climatic water balance (CWB), which is the difference between rainfall and potential evaporation(PET). CWB is computed at different time scales (1 month for the present study), and the calculated values are fit to a log-logistic probability distribution, which transforms the original values to standardised units (Beguería et al., 2014). A detailed explanation for the calculation was provided by Vicente-Serrano et al. (2010) and Beguería et al. (2014). The global SPEI database at different monthly scales using the CRU dataset can be accessed from https://spei.csic.es/database.html. Once the values are computed, they can be used to understand different drought characteristics. Figure 3(a) describes the monthly SPEI variation, and Figure 3(b) depicts the spatial precipitation regression map for the NSW region from January 2010 to December 2018 based on the mean gridded value of the CRU dataset. The marker highlights the lowest SPEI value (i.e. drought intensity) from 2011 to 2018. Drought onset is initiated when SPEI values become negative, and it ends when the values become positive. The period between onset and end is called drought duration, which can range from a few months to several years (Deo and Şahin, 2015). Drought severity is the cumulative deficit of SPEI values during a drought event (Zhang et al., 2015). The spatial extent of droughts is illustrated in Figure 1(b). The values are representative of different drought conditions, as highlighted in Table 1.



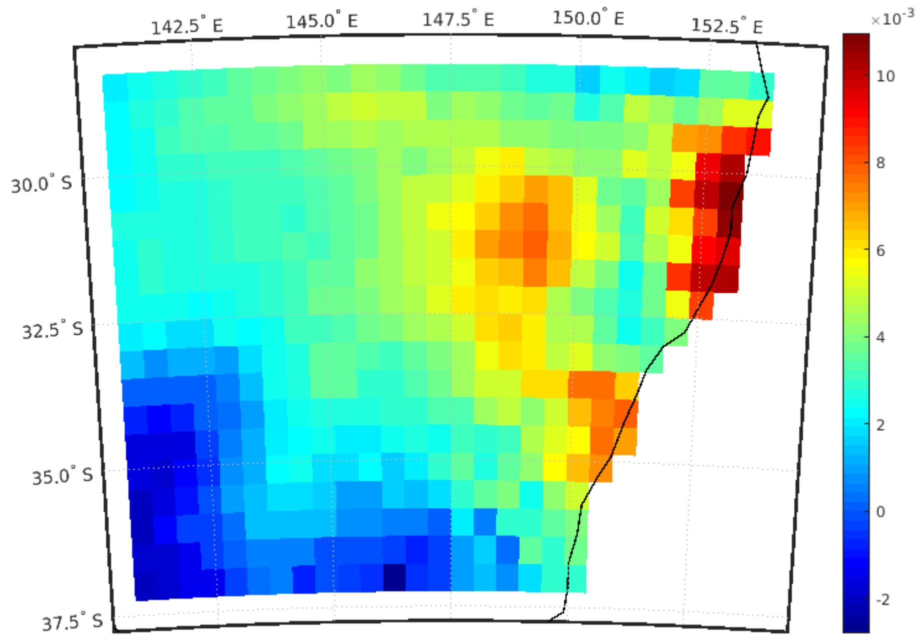


Figure 3. (a) Temporal SPEI 1 variation and (b) spatial regression rainfall map of the region from 2010 to 2018 based on the CRU TS dataset

Table 1: Drought categories per SPEI values (Rhee and Im, 2017)

SPEI classifications	Categories
≤−2.0	Extremely dry
−1.99 to −1.5	Severely dry
−1.49 to −1.0	Moderately dry
−0.99 to 0.99	Near normal
1.0 to 1.49	Moderately wet

1.5 to 1.99	Severely wet
≥2.0	Extremely wet

3. Model Development

After collecting SPEI data, the values of all the predictors were collected from relevant sources. The predictors used in the present study can be categorised into the following: a) hydrometeorological, which includes variables, such as temperature (minimum, maximum and mean), PET, rainfall and cloud cover and b) climatic indices (SOI, PDO, SAM, IOD and Niño indices 3, 3.4 and 4). Hydrometeorological variables were collected from the CRU dataset, and climatic indices were collected from the Earth System Research Laboratory (2020). The objective of this study is to provide maximum input datasets for training and a sufficient testing dataset. Therefore, data were trained from 1901 to 2000 and tested from 2001 to 2018. The next step was to determine the lag periods of large-scale climate predictors. The number of lag months to be used is not definite, and different studies have used varying lag periods, with each period fulfilling their respective objective. For example, Mekanik et al. (2016) used a lag period of 3 months to forecast rainfall, and Feng et al. (2020) used a lag period of 12 months to forecast SPI.

In the present study, instead of testing different lag periods, the lag of climatic indices based on the mean gridded SPEI value was analysed using cross-correlation. Cross-correlation was applied between the SPEI values and the predictor variables during the training period (Deo et al., 2017). Table 2 provides the optimum lag and the corresponding correlation coefficients of various climatic indices with a maximum lag period of 12 months. The results show that the SSTs exhibit a lag period of 2 months. Meanwhile, amongst the climatic indices, PDO has a lag period of 8 months. In the case of meteorological variables, rainfall and cloud cover are highly correlated with no lag period and their coefficients are 0.94 and 0.78, respectively. This finding is significant because most studies have disregarded cloud cover, primarily due to the use of ground-based data sources, which lack such datasets. From the obtained results, the proposed architecture considered climatic variables and meteorological variables as input and forecasted monthly SPEI at different lead times. The influence of lag periods on forecasting was also assessed.

Table 2. Correlation coefficients between climatic indices and monthly SPEI

Predictor	Optimum lag time (in months)	Correlation coefficient
SOI	0	0.31
PDO	8	0.15

SAM	0	0.16
IOD	0	0.12
Niño 3	2	0.23
Niño 3.4	2	0.26
Niño 4	2	0.25

3.1 LSTM Architecture

A detailed explanation of the LSTM model was provided in Olah (2015) and Goodfellow et al. (2016). A brief summary of this architecture is presented in this paper. The structure of LSTM is similar to a chain, as shown in Figure 4, wherein the basic building block is a cell and its state is the key to the mode. Gates that determine cell state have three types: input, forget and output gates. The gates analyse and control the amount of information that can pass through them, and they consist of a sigmoid neural layer and point-wise multiplication operation (Olah, 2015). The working mechanism of the gates and information flow can be expressed using the following equations:

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f), \quad (1)$$

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i), \quad (2)$$

$$C'_t = tanh(W_C . [h_{t-1}, x_t] + b_c), \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * C'_t, \quad (4)$$

$$o_t = \sigma(W_o . [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h_t = \sigma_t * tanh(C_t), \quad (6)$$

where x_t is the input vector at time t and σ is the activation function similar to *Sigmoid* or *ReLU*. W_f, W_i, W_C and W_o are the applied weights to the concatenation of the new input x_t and output h_{t-1} from the previous cell, with b_f, b_i, b_c and b_o as the corresponding biases (Xiao et al., 2019). f_t, i_t and o_t are the outputs of three sigmoid functions, σ, and their values range from 0 to 1. They control the information that is forgotten in the old cell state C_{t-1} and passed to the new cell C_t, with the new information being C'_t and h_t being the output information from the cell.

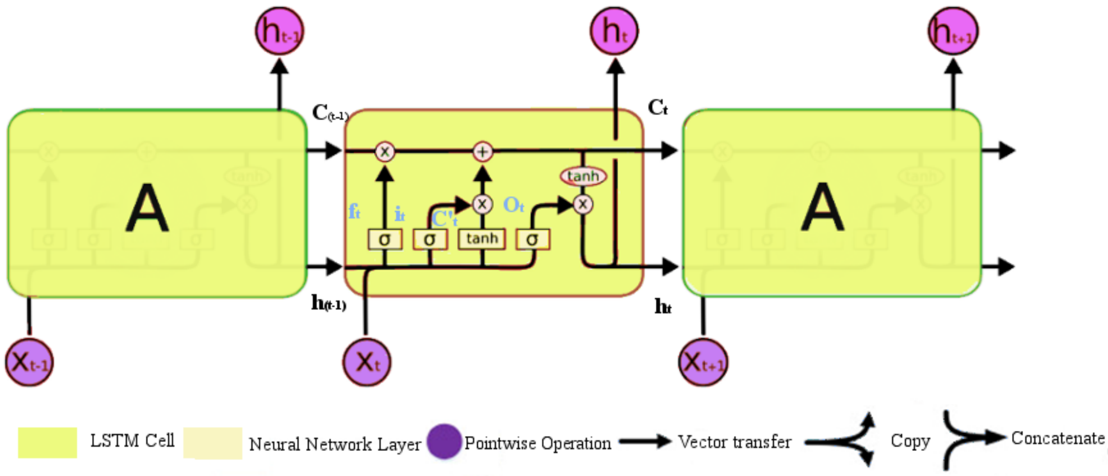


Figure 4. Structure of the LSTM network (modified from Olah, 2015)

In the present study, the LSTM architecture has 5 layers, which consists of 14 input layers, 2 LSTM layers and 2 dense layers. This network provided reliable results based on several experiments. In the LSTM network, the network is a 3D tensor, which represents the batch size for training (sample_size), the time window used to forecast SPEI (time_step) and the number of features (feature_n). In the case of the time step window, the sequence length was set as 20 based on the trial-and-error approach, indicating that the last 20 months of SPEI sequence were used to forecast the 21st month. In terms of features, the input features were the hydrometeorological variables and climatic indices. The number of cells in the first and second LSTM layers was 50 and 25, respectively. Meanwhile, the dense layers were set as 12 and 1, respectively. A dropout mechanism was applied after the LSTM layer to prevent overfitting, which was set as 0.25. The LSTM deep neural network was applied with Keras (Francois, 2015) using TensorFlow as the back end. The architecture of the network used is shown in Figure 5.

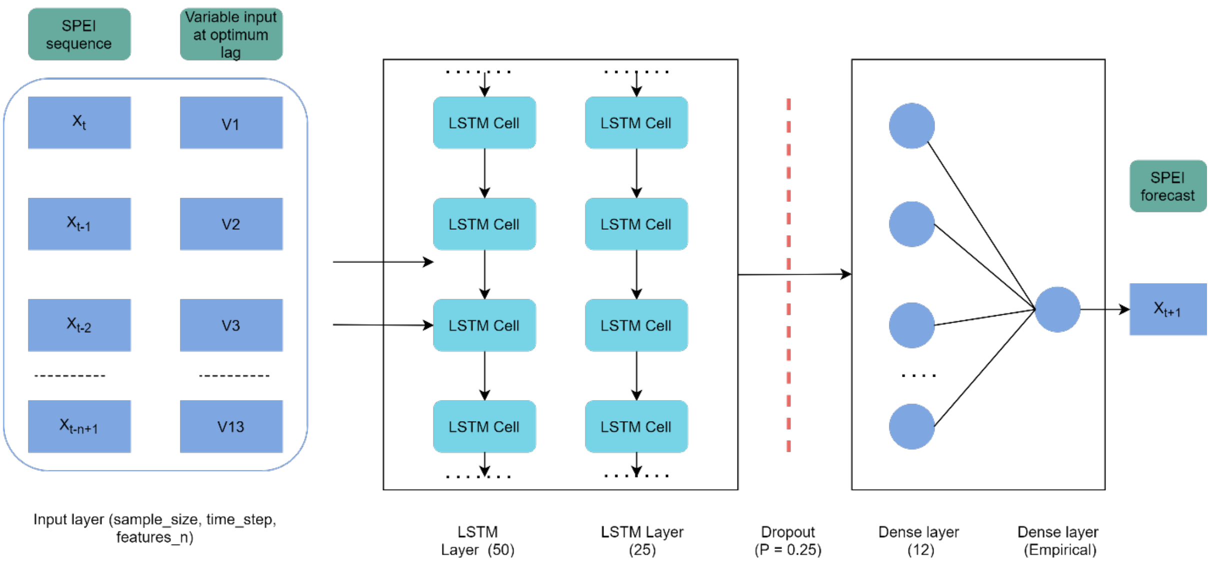


Figure 5. LSTM architecture used in the present study

Forecasting at multiple time steps was reviewed by Taieb et al. (2012), and they described five multiple time step forecasting techniques. In the present study, the sequence-to-sequence (seq2seq) forecasting approach or sliding window technique was used, wherein the forecasted value at time t_i is shifted towards the forecast values at time t_{i+1}. Moreover, the second dense layer increased as the number of forecasted lead months increased. Similarly, if the lead time is 6 months, then the second dense layer will be six instead of one.

The statistical metrics used to analyse the performance of the model were implemented using the coefficient of determination (R^2) and the root-mean-square error (RMSE) method. RMSE is frequently used as a metric because it penalises large errors and is suitable for forecasting purposes. R^2 represents the extent of association between the observed and forecasted values. The value ranges from 0 to 1, where 1 indicates an exact match and 0 denotes no association. By contrast, a lower RMSE value depicts better performance. The history of performance metrics used in forecasting with machine learning (ML) models was highlighted in Botchkarev (2018).

4. Results

The statistical metrics at different lead times during the test period are depicted in Figures 6(a) and 6(b). The results signify that the forecasting capability of the model diminishes as lead time increases.

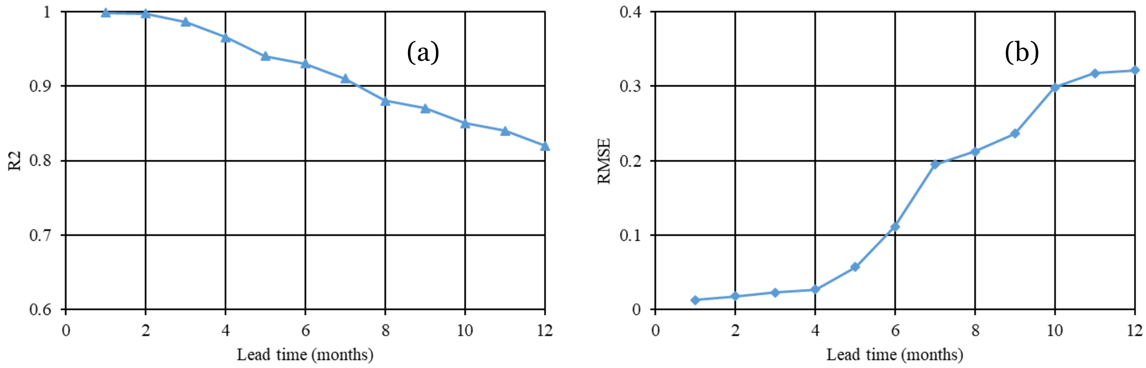


Figure 6. Statistical metrics (a) R^2 and (b) RMSE for monthly SPEI forecasting at different lead times

Analysis of the forecasted results was conducted in terms of different drought characteristics at four different lead times (1 month, 3 months, 6 months and 12 months), as shown in Figure 7. Depicting all the forecasted results during the test period is infeasible. Thus, the first instance of the forecasted SPEI sequence was presented. That is, for a lead time of 1 month, the comparison between the observed and forecasted values is shown for January 2001. Similarly, for a lead time of 3 months, the comparison is made for March 2001; for a lead time of 6 months, the comparison is made for June 2001 and for a lead time of 12 months, the comparison is made for December 2001. The number of grids in the region was 310. Amongst these grids, the percentage of pixels under the influence of drought (SPEI <-0.99) was 22.9% in January 2001 and 21.9% in June 2001. No pixels were observed during the drought periods of March and December 2001. The forecasting results showed that the percentage of pixels under drought periods was 27.7%, 26.1% and 4.5% for lead times of 1 month, 6 months and 12 months, respectively. No pixel was found during drought periods at a lead time of 3 months. As the results indicate, although no pixel was found during the drought periods at a lead time of 12 months, the forecasted result depicted specific pixels during drought periods. Such variation is expected because the lower limit of near-normal drought class (SPEI <-0.99) is within the vicinity of the mild drought class (Table 1). Therefore, examining drought intensity values in terms of different drought characteristics and not relying solely on statistical metrics are essential.

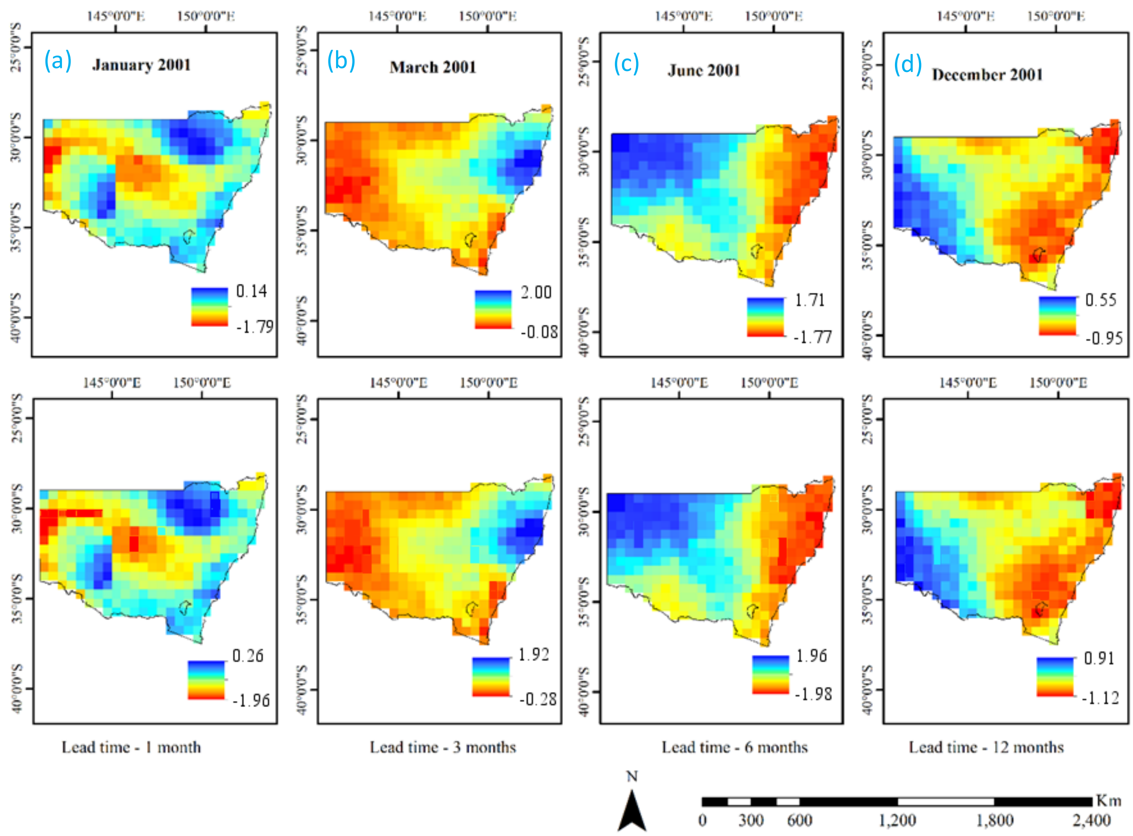


Figure 7. Comparison of the observed and forecasted SPEI 1 values at the first instance with lead times of a) 1 month, b) 3 months, c) 6 months and d) 12 months. The top row depicts the observed values, and the bottom row depicts the forecasted values.

To examine the importance of the LSTM architecture, annual rainfall and mean temperature anomaly maps are shown in Figures 8(a) and 8(b). The black line indicates the splitting of the dataset into training and testing sets. For forecasting, ML models learn uniform weightage across time steps. As the figure suggests, a significant variation in rainfall and temperature anomalies is observed during the entire study period. This phenomenon necessitates the use of decay over weights across periods. Hence, the use of LSTM is encouraged to learn decayed weights. The forget gate in LSTM ensures that the model can effectively capture the decay-weighted lag–lead sequence relationship without the vanishing gradient problem. This condition is also reflected in the spatial anomaly maps of rainfall and mean temperature (Figure 9). Temporal and spatial anomaly maps were prepared with the baseline period of 1961–1990. The studied months in Figure 9 depicts the initiation of the Millennium Drought. Ummenhofer et al. (2009) found that the droughts during this period resulted from high temperatures and climatic anomalies.

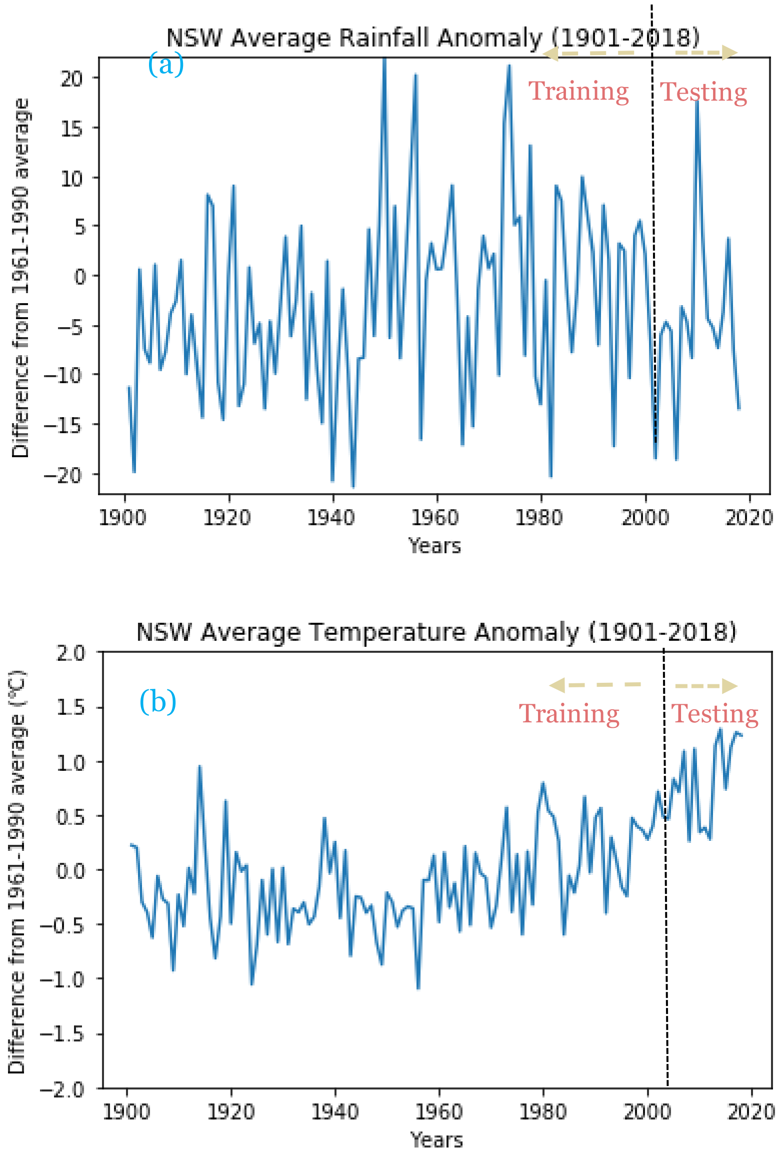


Figure 8. Temporal anomaly graphs of (a) rainfall and (b) mean temperature from 1901 to 2018. The black line indicates the splitting of data into training and testing sets.

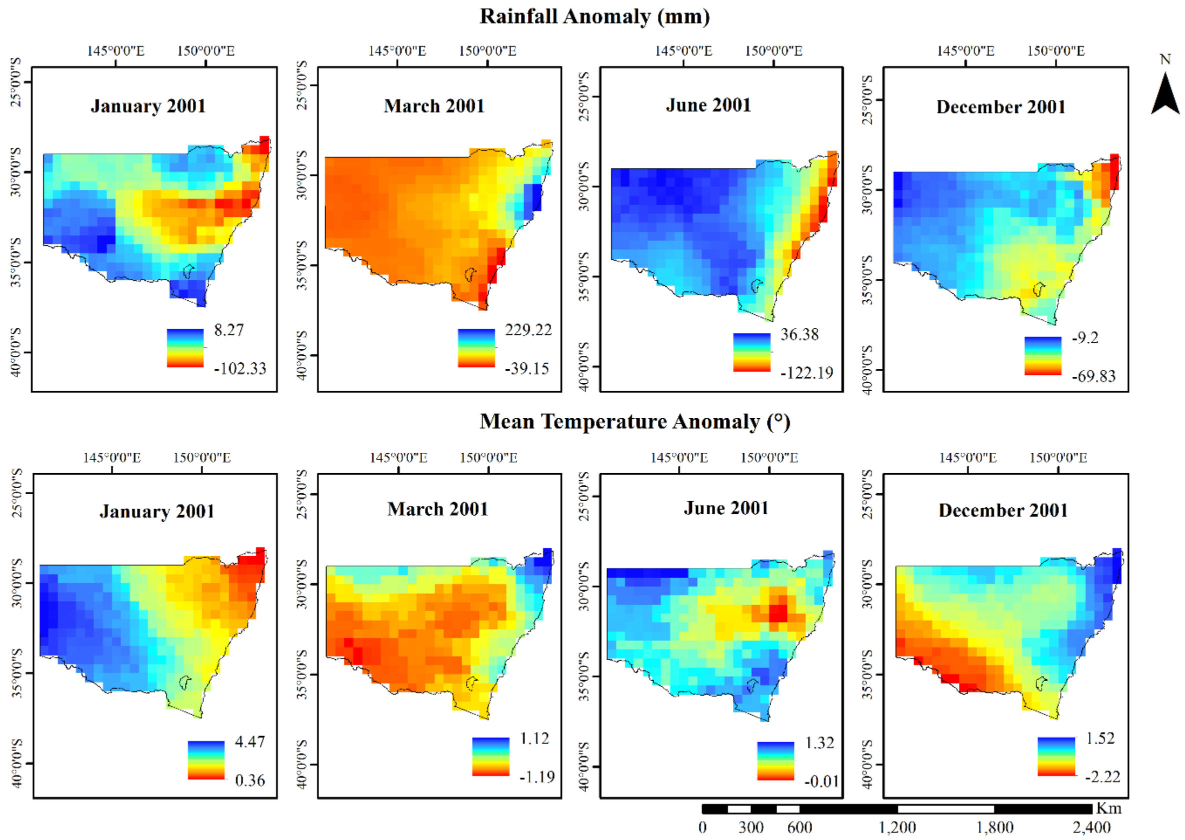
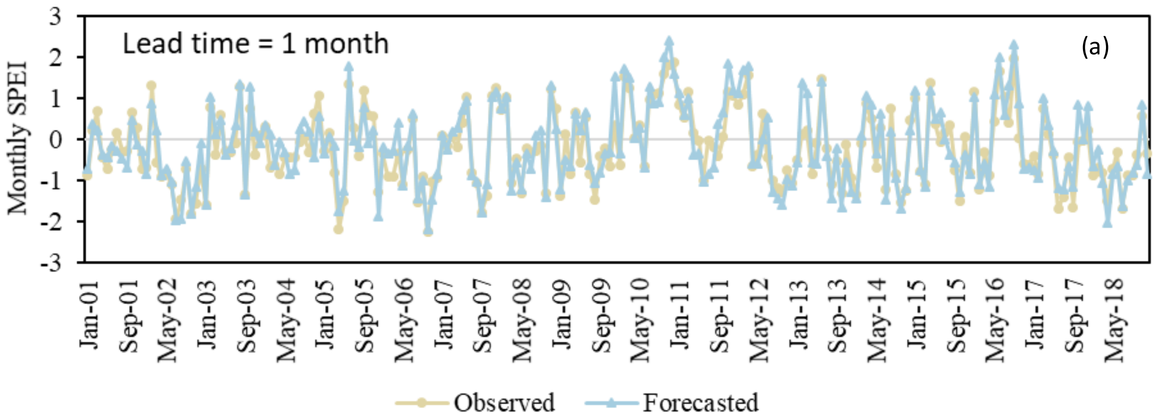
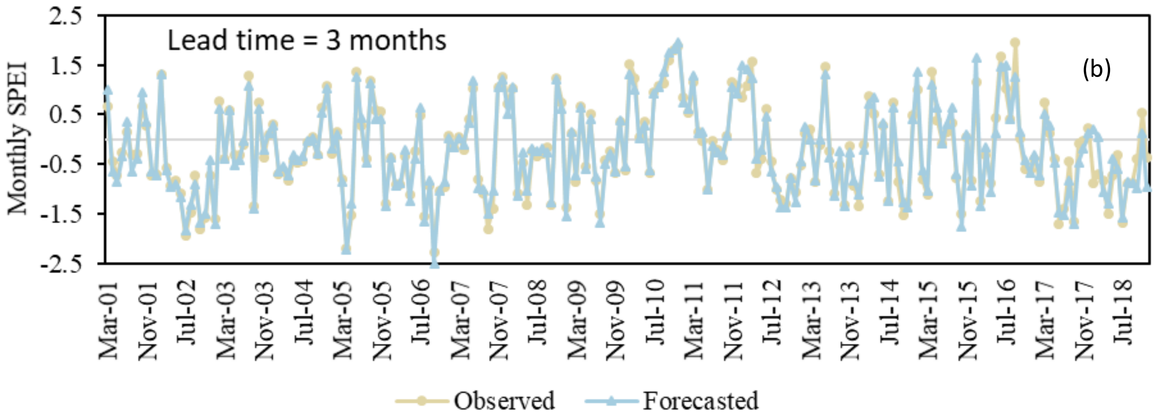
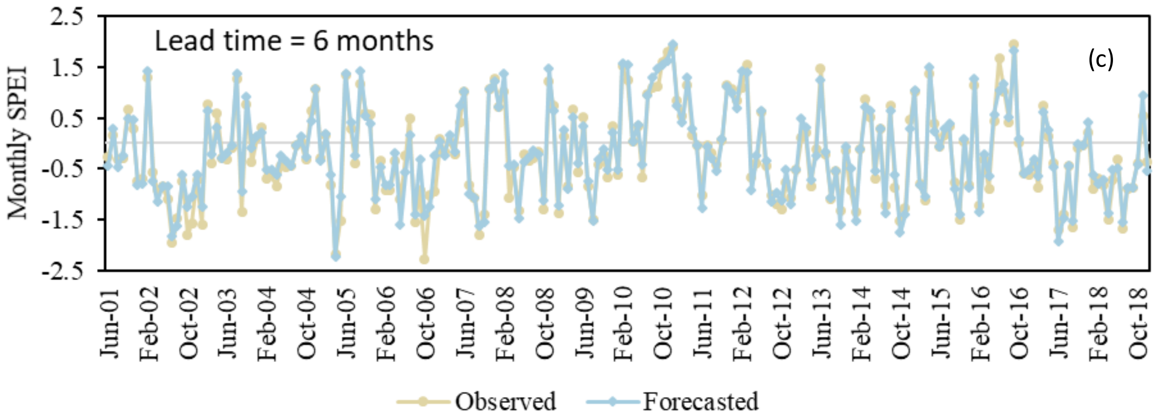


Figure 9. Spatial anomaly maps of precipitation and mean temperature

Thereafter, a comparison between the mean gridded observed and forecasted SPEI values was conducted. Subsequently, variations in terms of drought intensity, drought duration and number of drought months were analysed. The variation in drought intensity values at different lead times is illustrated in Figures 10(a) and 10(d). A period was considered a drought month when the mean SPEI grid value was less than −0.99. On the basis of this assumption, the number of observed drought months during the test period was 41 months (~19%), with the first drought onset in January 2001. To understand the forecasted results with the observed values, a useful statistical metric, namely, threat score (TS), was used. TS measures the fraction of correctly predicted forecasted results corresponding to the observed values. The mathematical formula is $TS = \frac{hits}{hits+misses+false\ alarms}$ (Jollifee and Stephenson, 2003). The value of TS ranges from 0 to 1, with 1 being the perfect score and 0 indicating no skill. The results indicated that TS was 0.93 in the case with a lead time of 1 month, 0.91 in the case with a lead time of 3 months, 0.86 in the case with a lead time of 6 months and 0.78 in the case with a lead time of 12 months. These results show that the model is capable of adequately forecasting monthly SPEI values. In addition, the forecasted results were analysed in terms of other drought characteristics, such as onset, end and duration. A closer look at the results showed that drought duration and end are generally correctly forecasted at lead times of 1 month and 3 months. However, onset is frequently erratic, i.e. either predicted earlier or later than observed. Similarly, for lead times of 6 and 12 months, onset and end are inconsistent with the observed values. Therefore, understanding the objective of the study and using the model adequately are of utmost importance. The results of the present study suggest that the model can determine drought onset at shorter lead times; however, caution is necessary when specifically determining onset at longer lead times.







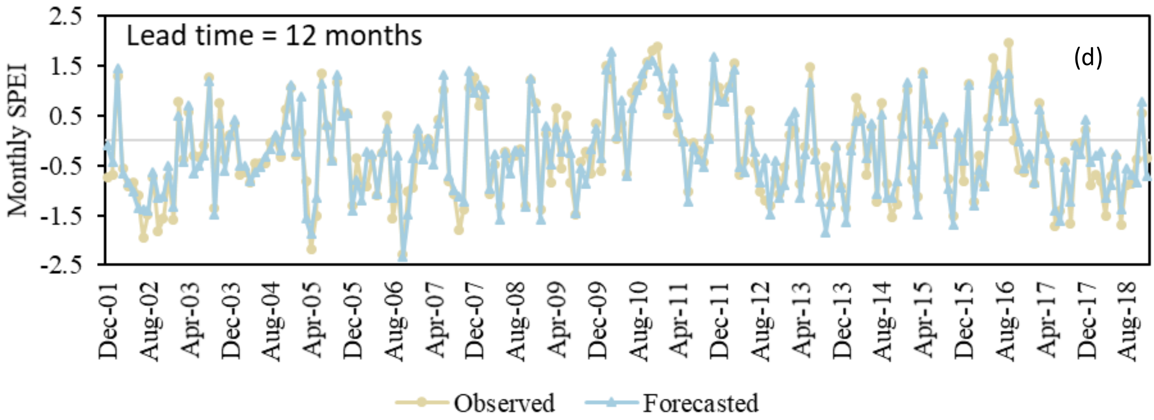


Figure 10. Variation between observed and forecasted monthly SPEI values at lead times of (a) 1 month, (b) 3 months, (c) 6 months and (d) 12 months

5. Discussion

Recent studies have aptly highlighted the superiority of deep learning models over traditional ML models in forecasting various drought aspects (Xiao et al., 2019; Reichstein et al., 2019; Ham et al., 2019; Dikshit et al. 2020c). Poornima and Pushpalatha (2019) forecasted SPEI for the city of Hyderabad, India using a single LSTM layer at a lead time of 12 months with hydrometeorological variables. Their study, which used rainfall and relative humidity as predictors, achieved an RMSE value of 0.2 and an accuracy of 97.05%. Their findings indicated that the LSTM model performed better than other stochastic and ML models. Agana and Homaifar (2017) used a deep belief network to forecast the Standardized Streamflow Index (SSI) for the Colorado River basin by using lagged values of SSI as input. They found that this network performed better than ML models. Earlier studies in the NSW region have used different ML models to predict droughts. For example, Deo and Sahin (2015) used ANN to predict the monthly SPEI at five different sites and achieved an R^2 value of 0.992. Similarly, Feng et al. (2019) predicted the SPEI for 3 months along agricultural belts by using three ML models. They found different statistical metrics for each specific study site. However, when considering the entire state and using a global climatological dataset, Dikshit et al. (2020a) found the R^2 value to be 0.72. This result showed that more emphasis should be given on studying larger areas and using a global climatological dataset, which will be helpful in regional drought management planning.

The critical findings of the present study can be grouped into two parts. The first is understanding the variables that affect drought forecasting at different lead times, and the second includes the findings regarding drought characteristics. In terms of variables, specific predictors, such as minimum and mean temperatures, did not considerably affect statistical metrics and can be eliminated if computational resources are scarce. The present study complements the previous findings of using lagged climatic variables as predictors, which improves drought forecasting results at longer lead times. Moreover, the impact of climatic driver changes over time should be noted (Ummenhofer et al., 2009). The architecture of the LSTM model helps capture this phenomenon. Although SOI was correlated with SPEI, it did not significantly affect forecasting. By contrast, the Niño indices were identified as crucial factors that affect forecasting. One possible reason for this finding can be the nonlinear impact between the indices captured by the LSTM model. Previous studies have shown that a nonlinear relationship exists between precipitation anomalies and ENSO events (Power et al., 2017; Fung et al., 2020).

Furthermore, a comparison between the observed and forecasted results based on different drought characteristics was conducted to understand the forecasting results. The drought characteristics used were intensity, duration, onset, termination and number of drought months. The analysis based on drought characteristics shows that drought intensity variation increases as lead time becomes longer. However, considering that only a few pixels were located on the borderline between drought and non-drought, the results may have overpredicted or underpredicted the values, leading to different drought classes. This result is not a reflection of the model's limitation, but instead, of the manner in which drought indices have been categorised. Given that droughts involve a multitude of characteristics, forecasting studies for any lead time should analyse results in terms of different characteristics and not focus solely on drought intensity. This suggestion was highlighted while analysing the variation between the observed and forecasted values within the spatial context (Figure 7). When examining disparity in terms of drought characteristics, TS also decreased as lead time increased. However, the interesting finding was that the accurate forecasting of drought onset and end diminished as lead time became longer. In fact, onset was typically determined later at longer lead times, suggesting that the architecture can fulfil some of this study's objectives and can provide a general understanding of future drought scenarios. Understanding that an index value is not an absolute reflection of ground reality but a possible drought scenario is also essential. Considering the manner in which drought categories have been designed, the index value frequently reflects different drought conditions, particularly at a pixel level. When combining all the pixel values, errors add up and can often depict contrasting results. However, the architecture presented in this work can be used to fulfil specific drought forecasting objectives in terms of different drought characteristics. Moreover, this study's results will be more helpful when a regional drought management plan is being considered instead of a localised management plan.

One limitation of the present work is understanding the spatial variation or autocorrelation of SSTs and climatic indices, which is a useful direction that should be considered in the future

(Legendre, 1993). The relationship between predictors and the forecasted monthly SPEI at the spatial scale requires further examination. Such an examination can be conducted using convolutional neural network LSTM architecture, wherein a certain grid size is used for feature extraction. In a recent study on forecasting ENSO at longer lead times of up to 18 months, Ham et al. (2019) used such an architecture to identify the hot spots of predictors. This study is the first to use a stacked LSTM architecture in drought forecasting. On the basis of the findings, deep learning approaches can outperform traditional ML models.

6. Conclusions

Droughts are amongst the most complex natural hazards due to the multitude of variables that affect their occurrences. One of the most challenging tasks towards effective drought management is forecasting droughts at long lead times. Accordingly, a deep learning model was used in the present study to forecast droughts at different lead times by using meteorological and climatic variables as predictors in the NSW region of Australia. A stacked LSTM model was developed to forecast monthly SPEI using the 1901–2018 dataset. This dataset was divided into the training period (1901–2000) and the testing period (2001–2018). Thereafter, a sliding window technique was used to forecast SPEI at different lead times during the testing period. The findings of this study indicated that the lagged climatic variables improve forecasting capabilities at longer lead times but do not have a significant effect at shorter lead times (1–3 months). Assessment of the forecasted results was performed on the basis of statistical metrics and by examining different drought characteristics. With regard to statistical metrics, the results showed that the LSTM model outperforms traditional data-driven models. Future work should explore deep learning models by improving the proposed architecture and experimenting with different models, such as ensemble models. In terms of drought characteristics (intensity, onset and end), the results varied across different lead times. The findings of this study demonstrate that statistical metrics do not provide sufficient assurance as researchers delves into understanding drought forecasting in terms of drought characteristics. This study can be highly useful to regional drought management planners, helping them prepare for future drought scenarios.

Credit authorship contribution statement

Abhirup Dikshit: Conceptualisation, Methodology, Modelling, Writing of Original Draft.

Biswajeet Pradhan: Validation, Visualisation, Review and Editing, Supervision, Conceptualisation, Funding.

Abdullah M. Alamri: Review and Editing.

Acknowledgements

This research was supported by the Centre for Advanced Modelling and Geospatial Information Systems, Faculty of Engineering and Information Technology, University of Technology Sydney. It was also supported by Researchers Supporting (Project number: RSP-2020/14), King Saud University, Riyadh, Saudi Arabia.

References

1) Abbot, J., Marohasy, J. (2014) Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. Atmos. Res., 138, 166-178, https://doi.org/10.1016/j.atmosres.2013.11.002
2) Alizadeh, M.R., Nikoo, M.R. (2018) A fusion-based methodology for meteorological drought estimation using remote sensing data. Remote Sens. Environ., 211, 229-247, https://doi.org/10.1016/j.rse.2018.04.001
3) AghaKouchak, A., Farahmand, A., Melton, F.S. et al. (2015) Remote sensing of drought: Progress, challenges and opportunities. Rev. Geophys., 53(2), 452-480, https://doi.org/10.1002/2014RG000456
4) Barua, S., Ng, A., Perera, C. (2012) Artificial neural network–based drought forecasting using a nonlinear aggregated drought index. J. Hydrol. Eng., 17, 1408–1413, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000574
5) Beguería, S., Vicente-Serrano, S., Reig, F., Latorre, B. (2014) Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. Int. J. Clim. 34(10):3001–3023, https://doi.org/10.1002/joc.3887
6) Ben Taieb, S., Bontempi, G., Atiya, A.F., Sorjamaa, A. (2012) A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert Syst. Appl. 39(8), 7067–7083, https://doi.org/10.1016/j.eswa.2012.01.039
7) Botchkarev, A. (2018) Performance metrics (error measures) in machine learning regression, forecasting and prognostics: properties and typology. Article 03006, ArXiv abs/1809
8) Cai, W., Purich, A., Cowan, T., van Rensch, P., Weller, E., (2014) Did climate change–induced rainfall trends contribute to the Australian Millennium Drought? J. Climate, 27, 3145-3168, https://doi.org/10.1175/JCLI-D-13-00322.1
9) Crausbay, S.D., Ramirez, A.R., Carter, S.L. et al (2017) Defining Ecological Drought for the Twenty-First Century. Bull. Amer. Meteor. Soc. 98(12):2543-2550, https://doi.org/10.1175/BAMS-D-16-0292.1
10) Deo, R.C., Şahin, M. (2015) Application of the Artificial Neural Network model for prediction of monthly Standardized Precipitation and Evapotranspiration Index using

hydrometeorological parameters and climate indices in eastern Australia. Atmos. Res., 161-162:65-81, https://doi.org/10.1016/j.atmosres.2015.03.018

11) Deo, R.C., Kisi, K., Singh, V.P. (2017) Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. Atmos. Res., 184, 149-175, https://doi.org/10.1016/j.atmosres.2016.10.004

12) Dikshit, A., Pradhan, B., Alamri, A.M. (2020a) Temporal hydrological drought index forecasting for New South Wales, Australia using machine learning approaches. Atmosphere-Basel 11, 585, https://doi.org/10.3390/atmos11060585

13) Dikshit, A., Pradhan, B., Alamri, A.M. (2020b) Short-term spatio-temporal drought forecasting using random forests model at New South Wales, Australia. Appl. Sci-Basel 10, 4254, https://doi.org/10.3390/app10124254

14) Dikshit, A., Pradhan, B., Alamri, A.M. (2020c) Pathways and challenges of the application of artificial intelligence to geohazards modelling. Gondwana Res., https://doi.org/10.1016/j.gr.2020.08.007

15) Duc, H.N., Rivett, K., MacSween, K. et al. (2017) Association of climate drivers with rainfall in New South Wales, Australia, using Bayesian Model Averaging. Theor. Appl. Climatol., 127, 169–185, https://doi.org/10.1007/s00704-015-1622-8

16) ESRL Earth System Research Laboratory, National Oceanic and Atmospheric Administration (2020) Available at: https://www.esrl.noaa.gov/, Accessed 28th Jan 2020

17) Feng, P., Wang, B., Luo, J.J. et al. (2020) Using large-scale climate drivers to forecast meteorological drought condition in growing season across the Australian wheat belt. Sci. Total Environ., 724, 138162, https://doi.org/10.1016/j.scitotenv.2020.138162

18) Feng, P., Wang, B., Liu, D.L., Yu, Q. (2019) Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. Agr. Syst., 173, 303-316, https://doi.org/10.1016/j.scitotenv.2020.138162

19) Francois, C. (2015) Keras. Github, https://github.com/keras-team/keras

20) Fung, K., Huang, Y., Koo, C., Soh, Y. (2019) Drought forecasting: A review of modelling approaches 2007–2017. J. Water Clim. Chang., 1–29, https://doi.org/10.2166/wcc.2019.236

21) Gergis, J., Gallant, A.J.E., Braganza, K. et al. (2012) On the long-term context of the 1997–2009 'Big Dry' in South-Eastern Australia: insights from a 206-year multi-proxy rainfall reconstruction. Clim. Change 111, 923–944, https://doi.org/10.1007/s10584-011-0263-x

22) Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y. (2016) Deep learning, MIT Press

23) Ham, Y., Kim, J., Luo, J. (2019) Deep learning for multi-year ENSO forecasts. Nature 573, 568–572, https://doi.org/10.1038/s41586-019-1559-7

24) Hao, Z., Singh, V.P., Xia, Y. (2018) Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. Rev. Geophy., 56(1):108-141, https://doi.org/10.1002/2016RG000549

25) Hanley, D.E., Bourassa, M., O'Brien, J.J. et al. (2003) A Quantitative Evaluation of ENSO Indices. J. Clim. 2003, 16, 1249–1258, https://doi.org/10.1175/1520-0442(2003)16<1249:AQEOEI>2.0.CO;2

26) Harris, I., Osborn, T.J., Jones, P., Lister, D. (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 2020, 7, 1–18, https://doi.org/10.1038/s41597-020-0453-3

27) Hendon, H.H., Thompson, D.W.J., Wheeler, M. (2007) Australian Rainfall and Surface Temperature Variations Associated with the Southern Hemisphere Annular Mode. J. Clim., 20, 2452–2467, https://doi.org/10.1175/JCLI4134.1

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

29) Imad H. U., Akhund M., Ali M., Pathan A. A., Ahmed A. (2019) Non-volumetric pricing is a threat to water reserves. Civil Engineering Journal 5 (2), 422–428, https://doi.org/10.28991/cej-2019-03091256

30) Jollifee, I.T., D.B. Stephenson, D.B. (2003) Forecast verification: A Practitioner's Guide in Atmospheric Science. John Wiley & Sons, Chichester, West Sussex, England

31) Kallis, G. (2008) Droughts. Annu. Rev. Env. Resour., 33, 85–118, https://doi.org/10.1146/annurev.environ.33.081307.123117

32) Khan, N., Sachindra, D.A., Shahid, S. et al. (2020) Prediction of droughts over Pakistan using machine learning algorithms. Adv. Water Resour., 139, 103562, https://doi.org/10.1016/j.advwatres.2020.103562

33) Kirono, D.G., Chiew, F.H., Kent, D.M. (2010) Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. Hydrol. Process., 24 (10):1237-1247, https://doi.org/10.1002/hyp.7585

34) Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? Ecology, 74(6), 1659-1673, https://doi.org/10.2307/1939924

35) Lloyd-Hughes, B. (2014) The impracticality of a universal drought definition. Theor. Appl. Climatol., 117 (3-4), 607-611, https://doi.org/10.1007/s00704-013-1025-7

36) Majhi, B., Naidu, D., Mishra, A.P. et al. (2020) Improved prediction of daily pan evaporation using Deep-LSTM model. Neural Comput & Applic., 32, 7823–7838, https://doi.org/10.1007/s00521-019-04127-7

37) Mishra, A.K.; Desai, V. (2006) Drought forecasting using feed-forward recursive neural network. Ecol. Model., 198, 127–138, https://doi.org/10.1016/j.ecolmodel.2006.04.017

38) Mishra, A.K., Singh, V.P. (2011) Drought modeling–A review. J. Hydrol., 403, 157–175, https://doi.org/10.1016/j.jhydrol.2011.03.049

39) Mekanik, F., Imteaz, M., Talei, A. (2016) Seasonal rainfall forecasting by adaptive network-based fuzzy inference system (ANFIS) using large scale climate signals. Clim. Dynam., 46 (9–10), 3097-3111, https://doi.org/10.1007/s00382-015-2755-2

40) Morid, S., Smakhtin, V., & Bagherzadeh, K. (2007) Drought forecasting using artificial neural networks and time series of drought indices. Int. J. Climatol., 27, 2103–2111. https://doi.org/10.1002/joc.1498

41) Nagarajan R. (2009) Drought Indices. In: Drought Assessment. Springer, Dordrecht, https://doi.org/10.1007/978-90-481-2500-5_5

42) Nasim, W., Ahmad, A., Amin, A. et al. (2018) Radiation efficiency and nitrogen fertilizer impacts on sunflower crop in contrasting environments of Punjab, Pakistan. Environ. Sci. Pollut. Res., 25, 1822–1836. https://doi.org/10.1007/s11356-017-0592-z

43) Olah, C. (2015) Understanding LSTM Networks, http://colah.github.io/posts/2015-08-Understanding-LSTMs/

44) Özger, M., Mishra, A.K., Singh, V.P. (2012) Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. J. Hydrometeorol., 13(1):284-297, https://doi.org/10.1175/JHM-D-10-05007.1

45) Pendergrass, A.G., Meehl, G.A., Pulwarty, R. et al. (2020) Flash droughts present a new challenge for subseasonal-to-seasonal prediction. Nat. Clim. Change. 10, 191–199, https://doi.org/10.1038/s41558-020-0709-0

46) Poornima, S., Pushpalatha, M. (2019) Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. Soft Comput., 23, 8399–8412, https://doi.org/10.1007/s00500-019-04120-1

47) Power, S., Tseitkin, F., Torok, S. et al. (1998) Australian temperature, Australian rainfall and the Southern Oscillation, 1910–1992: coherent variability and recent changes. Aust. Meteorol. Mag., 47(2):85–101

48) Power, S., Delage, F.P., Chung, C.T., Ye, H., Murphy, B.F. (2017) Humans have already increased the risk of major disruptions to Pacific rainfall. Nat. Commun., 8, 14368, https://doi.org/10.1038/ncomms14368

49) Pulwarty, R. S., Sivakumar, M. V. K. (2014) Information systems in a changing climate: early warnings and drought risk management. Weather Clim. Extrem., 3, 14–21, https://doi.org/10.1016/j.wace.2014.03.005

50) Reichstein, M. et al. (2019) Deep learning and process understanding for data-driven Earth system science. Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1

51) Risbey, J., Pook, M., McIntosh, P., Wheeler, M., Hendon, H. (2009) On the remote drivers of rainfall variability in Australia. Mon. Weather Rev., 137:3233–3253, https://doi.org/10.1175/2009MWR2861.1

52) Rodrigues, E., Gomes, Á., Gaspar, A.R., Henggeler Antunes, C. (2018) Estimation of renewable energy and built environment-related variables using neural networks—A review. Renew. Sustain. Energ. Rev., 94, 959–988, https://doi.org/10.1016/j.rser.2018.05.060

53) Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., & Bacmeister, J. T. (2004) Causes of long-term drought in the US Great Plains. J. Clim., 17(3), 485–503, https://doi.org/10.1175/1520-0442(2004)017<0485:COLDIT>2.0.CO;2

54) Seager, R., Hoerling, M. (2014) Atmosphere and ocean origins of north American droughts. J. Clim., 27(12), 4581–4606, https://doi.org/10.1175/JCLI-D-13-00329.1

55) Stahl, K., Demuth, S. (1999) Methods for regional classification of stream flow drought series: Cluster analysis. University of Freiburg Assessment of the Regional Impact of Droughts in Europe Tech. Rep. 1, 8 pp

56) Steffen, W., Hughes, L., Mulling, G. et al. (2019) Dangerous Summer: Escalating Bushfire, Heat and Drought Risk; Climate Council of Australia: Potts Point, Australia.

57) Sun, Q., Miao, C., Duan, Q. et al. (2018) A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev. Geophys., 2018, 56, 79–107, https://doi.org/10.1002/2017RG000574

58) Ummenhofer, C.C., England, M.H., McIntosh, P.C. et al. (2009) What causes southeast Australia's worst droughts? Geophys. Res. Lett., 36(4), 1-5, https://doi.org/10.1029/2008GL036801

59) Van Loon, A.F. (2015) Hydrological drought explained. Wiley Interdiscip. Rev.-Water, 2, 359–392, https://doi.org/10.1002/wat2.1085

60) Vicente-Serrano, S., Beguería, S., I López-Moreno, J. (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J. Clim., 23, 1696–1718, https://doi.org/10.1175/2009JCLI2909.1

61) Vicente-Serrano, S., Beguería, S., Lorenzo-Lacruz, J., et al. (2012) Performance of drought indices for ecological, agricultural, and hydrological applications. Earth Interact., 16, 1–27, https://doi.org/10.1175/2012EI000434.1

62) Vicente-Serrano, S.M., Quiring, S.M., Peña-Gallardo, M., Yuan, S., Domínguez-Castro, F., (2020). A review of environmental droughts: Increased risk under global warming? Earth-Sci. Rev., 201, 102953, https://doi.org/10.1016/j.earscirev.2019.102953

63) Wanders, N., & Wood, E. F. (2016) Improved sub-seasonal meteorological forecast skill using weighted multi-model ensemble simulations. Environ. Res. Lett., 11(9), 094007. https://doi.org/10.1088/1748-9326/11/9/094007
64) Wang, Q. J., Schepen, A., & Robertson, D. E. (2012). Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging. J. Clim., 25, 5524–5537. https://doi.org/10.1175/JCLI-D-11-00386.1
65) Wittwer, G. (2020) Estimating the regional economic impacts of the 2017 to 2019 drought on NSW and the rest of Australia; Victoria University, Centre of Policy Studies/IMPACT Centre: Budapest, Hungary.
66) Woli, P., Jones, J., Ingram, K., Paz, J. (2013) Forecasting drought using the agricultural reference index for drought (ARID): a case study. Weather Forecast., 28(2):427-443, https://doi.org/10.1175/WAF-D-12-00036.1
67) Xiao, C., Chen, N., Hu, C. et al. (2019) Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. Remote Sens. Environ., 233, 111358, https://doi.org/10.1016/j.rse.2019.111358
68) Yihdego, Y., Vaheddoost, B., Al-Weshah, R.A. (2019) Drought indices and indicators revisited. Arab J. Geosci., 12, 69, https://doi.org/10.1007/s12517-019-4237-z
69) Zargar, A., Sadiq, R., Naser, B., Khan, F.I. (2011) A review of drought indices. Environ. Rev., 19(NA):333–349, https://doi.org/10.1139/a11-013
70) Zhang, Q., Qi, T., Singh, V.P. et al. (2015) Regional frequency analysis of droughts in China: a multivariate perspective. Water Resour. Manage., 29, 1767–1787, https://doi.org/10.1007/s11269-014-0910-x