Natural Hazards
https://doi.org/10.1007/s11069-019-03665-6

REVIEW ARTICLE

Drought forecasting through statistical models using standardised precipitation index: a systematic review and meta-regression analysis

Anshuka Anshuka[1] · **Floris F. van Ogtrop**[1] · **R. Willem Vervoort**[1]

Received: 9 October 2018 / Accepted: 9 July 2019

Abstract

Quality and reliable drought prediction is essential for mitigation strategies and planning in disaster-stricken regions globally. Prediction models such as empirical or data-driven models play a fundamental role in forecasting drought. However, selecting a suitable prediction model remains a challenge because of the lack of succinct information available on model performance. Therefore, this review evaluated the best model for drought forecasting and determined which differences if any were present in model performance using standardised precipitation index (SPI). In addition, the most effective combination of the SPI with its respective timescale and lead time was investigated. The effectiveness of data-driven models was analysed using meta-regression analysis by applying a linear mixed model to the coefficient of determination and the root mean square error of the validated model results. Wavelet-transformed neural networks had superior performance with the highest correlation and minimum error. Preprocessing data to eliminate non-stationarity performed substantially better than did the regular artificial neural network (ANN) model. Additionally, the best timescale to calculate the SPI was 24 and 12 months and a lead time of 1–3 months provided the most accurate forecasts. Studies from China and Sicily had the most variation based on geographical location as a random effect; while studies from India rendered consistent results overall. Variation in the result can be attributed to geographical differences, seasonal influence, incorporation of climate indices and author bias. Conclusively, this review recommends use of the wavelet-based ANN (WANN) model to forecast drought indices.

Keywords Drought forecasting · Meta-regression analysis · Standardised precipitation index · Model performance

---

✉ Anshuka Anshuka
aans0600@uni.sydney.edu.au

[1] School of Life and Environmental Science, Faculty of Science, University of Sydney, Sydney, Australia

1 Introduction

Drought is a complex and cyclical event characterised by a precipitation deficit, which has a ripple effect on agricultural and hydrological systems as well as on societies (Dai 2011; Delbiso et al. 2017). Albeit having subtle visual effects, the impacts of drought are severe without precautionary measures and tend to linger for a prolonged period even after termination (Wilhite 2002). An upsurge in drought occurrence and severity is highly probable in the future, leaving management bodies to turn to thorough risk-mitigation measures (Dai 2011; Mishra and Singh 2011). The slow progression and inception of drought can be advantageous to modellers to foresee these events in advance (Cancelliere et al. 2007; Rossi 2003). Reduction of risks has been an integral part of planning and policy developments. This, in recent times, has been made possible with the development of numerous modelling approaches giving insight into precipitation deficits and ultimately improving the ability to monitor droughts.

While there are many different published drought indices [e.g. standardised precipitation evapotranspiration index (SPEI), Palmer drought severity index (PDSI), standardised precipitation index (SPI), effective drought index (EDI) and reconnaissance drought index (RDI)], this study focuses on the SPI (McKee et al. 1993), which is the most commonly used drought index and has received recommendations from the World Meteorological Organisation (Hayes et al. 2011). It is computed on the basis of precipitation distribution, where monthly rainfall totals are standardised using a gamma or Pearson type III distribution (McKee et al. 1993). Investigation of the SPI has revealed many advantages—it is computed with minimum complexity; it provides a spatially consistent interpretation across various climates; and it is probabilistic in nature, therefore depicting ideal characteristics in forecasting and risk analyses (Guttman 1998; Zargar et al. 2011). Since the calculation is based exclusively on precipitation, it is highly beneficial in data-sparse regions where other parameters such as streamflow, evapotranspiration and soil moisture information may not be readily accessible (Hayes et al. 2005). The SPI can be calculated at different timescales to give insight into different types of drought; for instance, the short to medium timescale is suitable for meteorological and agricultural drought, while the longer timescales are suitable for hydrological and socio-economic drought (Gumus and Algin 2017). The SPI has multifaceted uses, and recently, it has been used to assess groundwater drought (Kumar et al. 2016), and to carry out spatio-temporal analysis of floods and droughts (Liu et al. 2018).

Records of drought extend beyond 1000 years in many parts globally (Dai 2011). Periodicity is an inherent condition among drought events (Deo et al. 2017a), characterised by a sudden peak in drought after a given number of years (Kane 1997), which can be captured in long records of historical data. For instance, in Virginia, long-term records of PDSI (96 years of data) were analysed by applying Markov chain model which gave information on long-term drought probabilities, duration of drought, expected periods of recovery and recurrence times (Lohani et al. 1998). Additionally, a European study by Ionita et al. (2012) illustrated quasi-periodic trends of self-calibrated PDSI and decadal Pacific Decadal Oscillation and Atlantic Multi-dimensional Oscillation using long-time series of data from 1900s. Legacy satellite data which consist of old records can be useful in assessing historical drought trends and characteristics and also to monitor the return episodes in a periodic manner (Sheffield and Wood 2007). A study in India reconstructed four time series of monthly meteorological, hydrological, soil moisture and vegetation droughts from 1981 to 2013 which gave insight on the most severe and widespread droughts and their spatial–temporal distributions (Zhang et al. 2017a). Information of drought severity, duration and extent from historical records can be used and extrapolated for future droughts (Hao et al. 2014). Considering the cyclical nature of droughts, assessing historical data makes an important contribution towards drought studies, particularly, the evolution of droughts under accelerated warming environment (Dai 2011).

Conditions which may be less severe than droughts but are characterised by prolonged period of dry days are known as dry spells (Mathugama and Peiris 2011). Importantly, prolonged dry spells have a significant impact on the agricultural production (Wetterhall et al. 2015). If a heavy rain eventuates once a year with the rest of the year receiving minimal to zero rainfall, is particularly harmful for crops, as opposed to receiving light rain regularly throughout the year (Usman and Reason 2004). The timing of dry spells, that is, the start and end date for a dry spell (the breaking point), can be more crucial in terms of crop growth and productivity rather than considering the accumulated annual seasonal rainfall. Frequency of dry spells and length of dry spells may form the basis for developing a forecasting tool useful for informing agricultural planning (Mathugama and Peiris 2011).

With the recognition of climate signals influencing precipitation patterns, climate indices form an integral part in assessing hydro-meteorological hazards. The Madden–Julian Oscillation, the El Niño Southern Oscillation (ENSO), the Pacific Decadal Oscillation, the North Atlantic Oscillation as well as the Intertropical Convergence Zone (ITCZ) position are common drivers which affect weather patterns (Hurrell 1995; McGree et al. 2016; Salinger et al. 2014; Van Der Wiel et al. 2015). These climate phenomena induce climatic variations, seasonality and anomalies in precipitation regimes (Brown et al. 2013; Xie 2009), therefore highlighting the importance of incorporating these in modelling studies.

In conjunction with drought and climate indices, drought modelling has opened avenues to investigate numerous drought parameters (Mishra and Singh 2011; Wilhite 2000). Modelling outputs may include initiation and termination of drought, nature of severity, probability of occurrence and lead time (Chen and Li 1998), which also creates understandings into characteristics such as severity and spatio-temporal extent (Khalili et al. 2011). Initiation and termination of a drought event are difficult to predict as droughts have a slow development and, until human activity becomes affected by visible impacts, their existence may remain unrecognised (Maybank et al. 1995). Numerous studies have examined drought class transitions from non-drought to mild, moderate and lastly extreme phases to establish early warning systems (Bonaccorso et al. 2015; Moreira et al. 2012). As such, various parameters of drought can be utilised to establish early warning systems; however, in this study, we focused on identifying empirical studies that forecast the SPI by using either the lagged relationship with climate indices and SPI or the autocorrelation of SPI. These studies forecast SPI at various lead times, where lead time refers to the early announcement of a likely drought event prior to the actual onset of the hazard. Forecasts are an essential component that can be incorporated into disaster management units as a preemptive approach to reduce risks (Paulo and Pereira 2007; Wilhite 2000).

In retrospect, some of the challenges associated with drought modelling have been highlighted previously. For instance, precipitation, which forms the primary data type in drought-related work, is non-stationary and seasonal by nature, which offers challenges to modellers in adopting a suitable model type (Zhang et al. 2017b). Difficulties also exist in determining relationships between climate drivers such as sea surface temperature anomalies and ENSO. Some papers are in favour of climate indices to improve forecasting (Ganguli and Reddy 2014), while others report no significant improvement in model performance with the incorporation of climate drivers (Morid et al. 2007). Inconsistent region characteristics in different geographical locations inhibit proper response rates of the models; therefore, thorough testing and validation of the approaches are required before it can be adopted to develop an early warning system (Deo et al. 2017a). Lastly, complexity exists in choosing a suitable index, bearing in mind factors such as the timescale and the type of drought to address (Mishra and Desai 2005; Wilhite 2000), which often leaves modellers perplexed. Over the years, researchers have yielded inconsistent results in drought forecasting, and inconsistencies exist in performance among different model types. This makes implementation of an early warning system a challenging task for governments, causing indecisiveness in which modelling system to adopt and lack of trust among potential forecast users. Minimal research has been performed to compare the effectiveness of different models collectively (Choubin et al. 2016). Therefore, the novelty of this systematic review involves application of an approach which to the best of our knowledge has not been previously applied in the field of hydrology. In this study, a meta-analysis was undertaken to determine the suitability of the data-driven models for forecasting the SPI. The study aimed (1) to determine differences if any in model performance and (2) to determine the most appropriate timescale to calculate SPI with respective lead time combination to ensure optimum model performance.

2 Review of model types

2.1 Model selection

Much literature has shown that drought monitoring can be successfully achieved through the application of suitable index, climate signals and modelling tools (Mishra and Singh 2011). Forecasting models such as process-based and data-driven models are commonly utilised in hydrological applications. A physically based process model is underpinned by interpreting the physical processes of a system such as a river basin, while data-driven model identifies the best relationship based on the input data series (Solomatine and Ostfeld 2008). Information on parameters such as geology, soil, water abstractions and agricultural practices is not only difficult to access but also presents challenges to deduce sound scientific hydrological understanding for the different physical processes (Abrahart et al. 2008). This has led to the widespread use and development of data-driven models over process-based models in the field of hydrology which does not have such limitations. Data-driven models have showed great potential with improved skill to carry out drought forecasting, therefore, have been selected as the forecasting methodology for this study. Based on the review conducted by Mishra and Singh (2011), the following models were identified as the most commonly used data-driven models in the area of drought modelling: machine learning, time series models, probability models, regression models, and hybridised models such as wavelet-transformed neural networks; hence, these models names were used as search terms for this study.

2.1.1 Machine learning models

Techniques in artificial intelligence are becoming increasingly popular in the field of hydrology. One such approach is the artificial neural network (ANN) which originated through the idea of neurons based on the functional system in the human body (Hill et al. 1994). An ANN lacks a sound understanding of the physical relationship between the input and output variables, and yet produces reliable forecasts, which serves as a significant advantage where limited understanding exists between two variables (Hydrology 2000). An ANN is particularly useful because it can model nonlinear relationships in data, which is often the case in hydrological studies (Memarian et al. 2016). Notably, ANNs have been used to develop the satellite derived precipitation product, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Climate Data Record (Ashouri et al. 2015). Morid et al. (2007) used ANN to forecast SPI and EDI in Tehran, concluding that the EDI network model showed an impressive capacity to accurately predict the different drought classes and could potentially form a basis of operational forecasts. It has been also shown that short-term drought forecasts are better with a Recursive multi-step neural network approach and long-term forecasts have high accuracy with direct multi-step neural network approach (Mishra and Desai 2006). Similarly, Deo and Şahin (2015) demonstrated the strength of ANN to predict the drought severity, intensity and duration using SPEI in Australia.

A hybrid model, known as adaptive network-based fuzzy inference systems (ANFIS), emerged as a result of combining the learning techniques of ANN and neuro-fuzzy approach for adaptive network (Brown and Harris 1994; Nayak et al. 2004). For instance, it has been shown that fuzzy logic approach on its own has lower predictability; however a combined approach, that is, ANFIS, rendered better results to predict SPI (Keskin et al. 2009). Numerous studies have focused on the application of ANFIS in the area of drought modelling using SPI (El Ibrahimi and Baali 2018; Nguyen et al. 2017; Shirmohammadi et al. 2013). Bacanli et al. (2009) demonstrated better performance of the ANFIS model in comparison with feed-forward neural networks and multiple linear regression (MLR). The ANFIS model also had about 90% accuracy in predicting different drought classes of the effective drought index in Tehran (Farokhnia et al. 2011). High skill was also noted in the long-term forecast (6–12-month lead time) of the Palmer modified drought index by wavelet-transformed fuzzy logic over wavelet-transformed neural network (Özger et al. 2012).

Another popular machine learning model is the support vector machine (SVM) which was initially explored by Cortes and Vapnik (1995). SVMs are easy to train, have high efficiency and are able to handle noisy data well (Mokhtarzad et al. 2017; Raghavendra and Deka 2014). The least squares support vector machine (LSSVM) was used to forecast SPI, producing promising results (error value of approximately 0.1 and correlation of 0.9) (Deo et al. 2017a). Despite the promising results, LSSVM was not able to outperform multivariate adaptive regressions spline and M5 model tree. However, another comparative study between these machine learning models, that is, ANN, ANFIS and SVM, concluded SVM is superior to neural networks and ANFIS in predicting SPI (Mokhtarzad et al. 2017).

2.1.2 Time series models

A prevalent method of forecasting time series is by using stochastic approaches (Box et al. 2015). Autoregressive (AR), moving average (MA), autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA) models are all examples of time series models (Hyndman and Athanasopoulos 2018). The most commonly used among these models are ARIMA and SARIMA. Among many advantages, the models' proven superiority is attributable to characteristics such as moving average, exponential smoothing and most importantly, forecasting capability relative to time (Han et al. 2010). ARIMA/SARIMA models have been used for forecasting SPI, setting threshold levels for drought events and forecasting hydrological drought (Modarres 2007). Conversely, in a study in China, it was concluded that the ARIMA model had the lowest forecasting ability of SPI compared with ANN and wavelet-transformed ANN (WANN) (Zhang et al. 2017b). A noted disadvantage of the ARIMA method is its inability to handle nonlinearity and non-stationarity in the data. To overcome this problem, Mishra et al. (2007) demonstrated the use of a hybrid model comprising properties of ARIMA and ANN, which ultimately rendered results better than ANN and ARIMA models used independently.

2.1.3 Probability model

Markov chains are another type of model that uses stochastic processes (Modarres 2007). Markov models have been widely used in applications such as forecasting metrological droughts (Khadr 2016), probabilistic classification of drought states (Mallya et al. 2012), and identifying drought class transition probabilities (Paulo and Pereira 2007). A study using stochastic Markov model and SPI showed that a region has a higher probability of being in a state of drought (a month in the future) if the present conditions represent moderate or severe drought (Paulo et al. 2005). Similarly, Mishra et al. (2009) illustrated that for SPI the antecedent drought status will most likely persist in the following months, namely a ‘near normal’ or a ‘moderately dry’ period would follow a ‘moderately dry’ period. Markov models have also been used to illustrate the ‘persisting’ nature of drought in the UK with below average river discharge identified for as long as 6 years (Wilby et al. 2015).

2.1.4 Multiple linear regression models

MLR analysis was initially developed to observe the relationship between more than one predictor variable and a dependent or criterion variable (Pedhazur 1982). In MLR, the variable is forecasted using other variables, while in AR, the variable is forecasted using the lags of the variable itself (Hyndman and Athanasopoulos 2018). The popularity of MLR, as described in Chen and Li (1998), is based on straightforward implementation, robust statistical theory, availability of problem-solving tools to enhance the analysis and availability of standard deviations for the estimated parameters. Linear regression models in conjunction with weather variables, such as air pressure, air surface temperatures and SSTs, wind velocities and precipitation data, have been useful for hydrological extreme prediction. Examples of applications of MLR include using regression models to assess the relationship between the normalised difference vegetation index and SPI (Ji and Peters 2003). The former index was considered a useful tool to assess the response of vegetation to moisture in drought conditions. MLR has also been successfully applied to predict drought (Bacanli et al. 2009), barley yield in drought conditions (Odabas et al. 2014) and daily flood flow (Rezaeianzadeh et al. 2014). SPI and MLR had been used in Greece to carry out spatio-temporal analysis on drought and characterise the respective return periods (Loukas and Vasiliades 2004).

2.2 Model summary

Models have different operational principles, which differentiate them from one another. To better understand model performance, it is vital to understand the model operation and the advantages and disadvantages associated with the different types of model included in this review (Table 1).

Table 1 Comparison of study models

Model	Overview	Advantages	Disadvantages
ANN	The ANN originated through the concept of neurons and operates on the basis of the functional system of neurons in the human body (Hill et al. 1994). This mathematical model entails a three layered architecture system (Wang 2003)	It automatically creates relationships between variables without being instructed, which proves advantageous to a modeller where there is limited understanding (Hydrology 2000)	The *black box* nature provides insufficient information about the internal mechanism to the user. Difficulty in determining proper architecture of the neural network and proneness to overfitting are other common downfalls (Tu 1996)
ARIMA/SARIMA	This is a method of forecasting time series using stochastic approaches (Box et al. 2015) that applies linear correlation among observations to forecast both stationary univariate data and non-stationary time series that exhibit trends such as seasonality	The model proves superior because of characteristics such as moving average, use of exogenous variable, exponential smoothing and, most importantly, its forecasting capability relative to time (Han et al. 2010)	The use of the traditional form of model selection techniques can be prone to bias and may need prior experience to be able to determine the best order for model development. The approach does not account for nonlinearity in the data
SVM	Cortes and Vapnik (1995) initially explored the SVM in an attempt to determine an estimate of limited samples by classifying data on a hyperplane in the best possible way	Raghavendra and Deka (2014) identified the advantages of SVM as follows: capability of handling linear and stationary data; the ability to predict unobserved data, and handle noisy conditions without overfitting	Assigning data to the hyperplane is a challenging task and may need identification of a suitable kernel function algorithm to assist in classifying the data therefore, a user with limited experience may experience difficulty. Moreover, it does not work well with large datasets and has a tedious training process (Sapankevych and Sankar 2009)
MM	A Markov model is a double stochastic process that is observable through another set of stochastic process (Modarres 2007). It has a robust statistical foundation underpinning an efficient learning algorithm	The flexibility of MM enables it to handle data of varying lengths. The model has demonstrated a diversity of uses in numerical alignment, data mining, classification, operational analysis and pattern discovery (Briggs and Sculpher 1998)	The limitations of HMMs are often associated with training, having restrictive conditional independence assumptions, and the absence of kernel-based methods due to explicit feature representations (Altun et al. 2003)
MLR	A modelling technique which predicts a response variable using more than one explanatory variable (Myers and Myers 1990)	The prediction is based on many independent criterion which could better explain the dependent variable in practical forecasts hence making it a better approach than simple linear regression (Wilks 2011)	The approach is not able to handle nonlinearity and non-stationarity in the data series (Boulanger et al. 2005; Mouatadid et al. 2015)

3 Materials and methods

The narrative or traditional form of literature review may be inadequate to establish explicit connections and a bigger picture of a collection of studies in a particular discipline (Tranfield et al. 2003). To synthesise information better, meta-analysis can be applied to studies where many variations in the results exist. It involves application of quantitative analysis such as regression analysis to procure objective results (Stanley and Jarrell 1989). Although much more common in the medical field and social science research, such as psychology, it is slowly being developed in other disciplines of applied sciences. For instance, Kroeker et al. (2010) used meta-analysis to explore the effect of ocean acidification on marine animals, and Rustad et al. (2001) investigated the response of carbon dioxide production from decomposition in soil, net nitrogen oxidation and aboveground plant biomass in an ecosystem warming scenario. The different data-driven model types as discussed have been shown to perform well; however, no synthesised information exists based on which the best-performing model can be determined, which was addressed in this study.

This systematic review was guided by the research principles, identified by Khan et al. (2003), that a systematic review should synthesise results according to an explicitly devised research question and methodology.

Step 1: framing the question

The first step entailed framing a relevant research question. The participant/population/problem–intervention–comparison–outcomes (PICO) framework by Schardt et al. (2007) was used to formulate the question for this analysis, as shown in Table 2.

On the basis of the above framework, the question was framed as follows: which data-driven models are most efficient in forecasting the SPI at varying lead times, and what, if any, significant differences are identified in the performance of the different models? For the purpose of this paper, a distinction was made between the words prediction and forecast. In statistics, a prediction occurs within the sample such that the Y value is predicted in an observation or within part of the sample, while a forecast is a subset of a prediction that provides projections out of the sample for future values by applying historical data and that can change according to external scenarios.

We also make a distinction between now-casting, forecasting and hindcasting. Forecasting refers to reproducing an aspect of a system ahead of time (Beven and Young 2013). Now-casting refers to a description of current parameters or near real-time parameters within 0–2 h, which is more applicable for weather. Conversely, hindcasting applies forecasting to historical data to simulate an event in the past (Soares and Cardoso 2018; Wandres et al. 2018). A forecast usually incorporates meteorological data and is initialised by the results of a now-cast. According to the World Meteorological Organisation, forecasts

Table 2 Components of the PICO framework used to determine the main elements of the study meta-analysis

Framework item	Explanation
Problem	Drought forecasting
Intervention	Through application of the SPI
Comparison	Compared for different data-driven models
Outcomes	Determine the effectiveness of models at certain lead times and timescales

can range from hours to 24 months. In this review, we mainly focus on forecasting of future occurrence of drought by using the SPI at a certain lead time. A short-term forecast is categorised by a lead time of 1–3 months, a medium-term forecast by a lead time of 4–6 months and a long-term forecast by a lead time of 9 months or more.

Step 2: search strategy

We conducted an exhaustive search of the electronic literature on the Web of Science Core Collection database by applying the search elements identified through the PICO framework. The Boolean technique was used to carry out the search, employing operators such as 'AND' and 'OR'. The specific search terms were 'drought forecasting' and 'standardised precipitation index' or 'SPI' and the specific model type (i.e. ANN, ARIMA, MM, MLR, ANFIS, SVM) as identified above. Searching for one concept at a time is claimed to give the best results (Tuttle et al. 2009); therefore, the search was conducted for each type of data-driven model individually, one at a time. The initial search of the Web of Science database yielded $n=74$ results (Table 3). Titles, abstracts and author information of these articles were extracted and kept for further review.

To broaden the search results, similar search terms were used in Google Scholar (GS) for each type of model with two additional search terms, which were 'lead time' and 'timescale'. Hence, the search phrase was as follows: 'drought forecasting' and 'standardised precipitation index' and '[model type]' and 'lead time' and 'timescale'. This provided more results and yielded a sample of $n=176$. Only the results of the first three pages of GS were considered to avoid grey literature, which would have compromised the quality of the study. This reduced our search from GS to $n=22$.

Step 3: specific inclusion and exclusion criteria

The abstracts were reviewed initially on the basis of three general criteria: (1) significant reviews of the literature that investigated drought forecasting by using the model of interest, (2) use of the SPI for forecasting and (3) evaluation of lead time as the drought parameter. In case a criterion was not met directly, but the study was in similar context, the abstract was set aside for further evaluation.

The PRISMA method (Liberati et al. 2009; Moher et al. 2009), as shown in Fig. 1, was employed to further select the studies for use in the meta-analysis. From the selected abstracts, full-text articles were retrieved for further review. Because the primary objective was to measure the performance of different models that forecast the SPI at different lead

Table 3 Search result for the Web of Science database based on specific model type

Model type	Number of results
ARIMA	11
ANN	10
SARIMA	4
SVM	7
ANFIS	7
MM	16
MLR	19

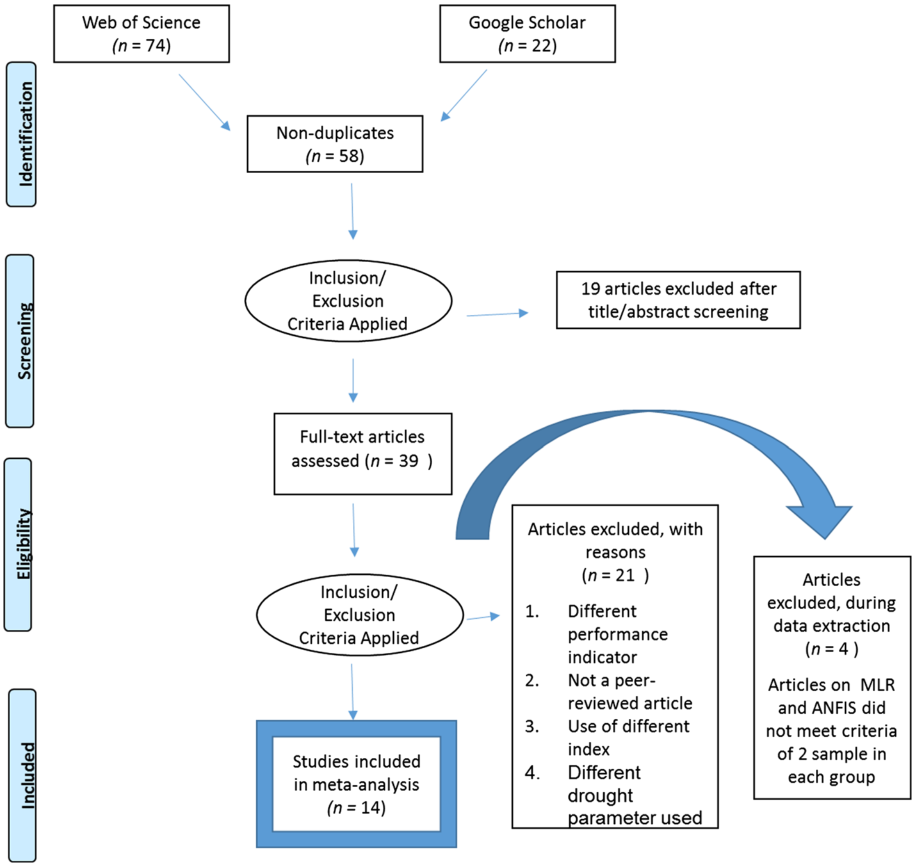

Fig. 1 PRISMA flowchart to aid in the selection of studies for the quantitative analysis

times, inclusion of studies was based on the following specific criterion: the study included performance measures such as NSE, RMSE, R^2, R, MSE, z test and corr. While going over the different studies, it was identified that R^2 and RMSE were the two most commonly used performance measures. Therefore, we only included those studies with the performance measures R^2 and RMSE or R and MSE, which could be converted to R^2 and RMSE. The coefficient of determination (R^2) represents the extent of association between two variables that could be, the observed and the predicted values. The root mean square error (RMSE) assesses the variance of errors and highlights the inconsistencies between the observed and the forecasted values (Adamowski et al. 2012). The performance measure values range from 0 to 1, where the higher R^2 value indicates a strong relationship between the data and 0 represents no statistical correlation. On the other hand, a lower RMSE value represents a better association.

The following exclusion criteria were applied in this study. (1) Studies that used performance measures other than R^2 and RMSE were excluded. (2) Authors were contacted via email if data were missing or if the entire analysis of the results was not published in the articles; if the authors did not provide the requested data, those studies were omitted. (3) Because the models of interest were data-driven models, other forms of model were omitted. (4) The SPI was found to be the most commonly used index, and other indices were not considered in this study. (5) Various parameters are utilised in the area of drought modelling; however, we concentrated on the lead time, eliminating studies that focused on class transitions, severity and duration. (6) A minimum of two studies is required in a meta-analysis to reach a conclusion (Valentine et al. 2010). Therefore, we applied this principle and ensured that under each model type there was a minimum of two study replicates to carry out the analysis. Owing to this criterion, studies using the adaptive network-based fuzzy inference system (ANFIS) and multiple linear regression (MLR) were omitted.

Step 4: data abstraction and appraisal of study quality

Selected studies were screened under more refined criteria to ensure quality assessment through the use of study design-based quality checklists. Only those studies that used a good quality and homogenous data of more than 30 years were used. Additionally, to ensure that the quality was maintained, the analysis was carried out on the model results which were validated. Evaluating forecast skill on the same dataset as that was used in training the model introduces bias and artificial skill in the results; therefore, it is essential to test the model on an independent set. Various methods of validation techniques were used in the selected studies. Although there is no fixed way of partitioning the data, the approach of two-way splitting of the data into training and validation is the easiest and most commonly used method (Mishra and Desai 2006; Mokhtarzad et al. 2017). For instance, 60% of the data is used to train the model, while 40% of the data is used to validate the model. Some studies have also incorporated a two stage of validating and testing the models with an additional independent set (Dehghani et al. 2014; Farokhnia et al. 2011), that is 90% of the data is used for training, 5% in validation and an additional 5% for testing. The error in the validation and test set often arises from the model overfitting the data or a poor fitting model. To overcome this issue, another popular method utilises cross-validation techniques (Deo et al. 2017b; Wong 2015), and this method was also noted in some studies.

Additionally, the data were solely from studies that were found in peer-reviewed journals; conference paper studies were omitted from this analysis. From the selected studies, a data matrix was created, comprising the first author, geographical location of the study, model of interest, SPI timescale and specified lead time (forecast range), and its respective performance measure indicator values of RMSE and R^2 were abstracted, which formed the primary component in the analysis. Each study was given a reference number ID. Where a study focused on more than one station in a region, average values of R^2 and RMSE were used.

Step 5: statistical analysis

In this study, the linear mixed model (LMM) analysis was applied from the following packages in R: the nonlinear mixed-effects (NLME) (Pinheiro 2009; Pinheiro et al. 2007); lme4 (Bates et al. 2014); lmerTest (Kuznetsova et al. 2017) and the lsmeans (Lenth 2016). The model type, SPI timescale and lead time months (i.e. forecast range) were the independent variables in the formula. The model type was used as the fixed effect, and the model categories were as follows: WANN, SVM, MM, ANN and ARIMA/SARIMA. The second factor evaluated was the timescale at which the SPI was computed; it was categorised as 3, 6, 9, 12 and 24, where SPI 3 referred to the SPI calculated for a 3-month timescale and SPI 24 referred to the SPI calculated for a 24-month timescale. The third factor was the lead time, which was divided into three categories: short-term forecast, referring to a lead time of 1–3 months; mid-term forecast, referring to a lead time of 4–6 months; and long-term forecast, referring to a lead time of 9 months or more. Last, the paper ID and authors nested in the geographical locations were assessed as a random effect to account for regional and researcher variation in the results of the selected studies. The LMM was applied to the weighted average of the R^2 and the RMSE on validated and forecasted results.

4 Results

The results from the linear mixed model demonstrate the wavelet-coupled ANN (WANN) model was the best-performing model with low error values and high correlation (Fig. 2). The average RMSE value of WANN is approximately 0.3; conversely, the RMSE value of other model types exceed 0.5. The compact letter display (a, b, c) in Fig. 2 shows whether the means for each of the RMSE and R^2 are significantly different from each other at a 95% significance level. For instance, in plot (a) LS Means RMSE, the SVM and MM had similar performance to each other indicated with alphabet (ab) for both, as did the time series ARIMA/SARIMA (labelled ARIMA on the graph) and ANN indicated with (b). However, while the SVM and MM show a similar performance to the WANN indicated with the letter (a), the performance of ARIMA and ANN is shown to be significantly different from WANN. In plot (b), LS Means R^2, MM shows association with all other model types indicated with (abc). ARIMA (a) and WANN (c) are distinctly different from all the other model types or shows no association with other model groups, while the two machine learning models, SVM and ANN, show similar performance. Overall, the performance of the ANN was not significantly different from those of the other model types, i.e. machine learning models do not necessarily perform better than other simpler models, such as time series. The ANN and ARIMA forecasts showed the least variations within the group. The WANN model explained 55–65% of the variation in the SPI, while the SVM was closer to 45–55%.

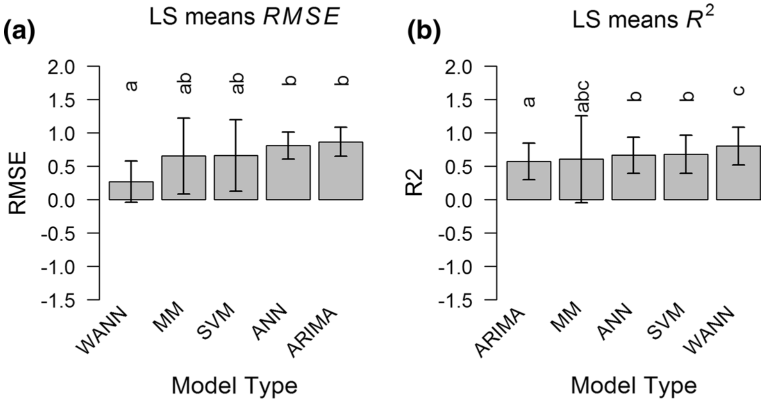

Fig. 2 Boxplot of estimated least square (LS) means of the linear mixed model (LMM) applied and the compact letter display for differences in means at the 95% significance level for **a** RMSE and **b** R^2 values, including 95% confidence intervals for wavelet artificial neural network (WANN), artificial neural network (ANN), Markov model (MM), support vector machine (SVM) and autoregressive integrated moving average (ARIMA) or seasonal ARIMA (SARIMA)

Generally, hydrological indices are calculated using monthly, annual or non-standardised timescales. In Fig. 3, LMM model applied to the SPI (as an independent factor) calculated at different timescales is shown. The best timescales to calculate the SPI were 12 and 24 months, which yielded the highest correlation values of R^2 and the smallest error (RMSE). The SPI at a timescale of 9 months yielded a reasonably high R^2 value with a high respective error value which could be due to outlier values in the data. The models forecasting SPI 12 and SPI 24 accounted for 30–50% of the variation, while the SPI 6 and 9 accounted for 60–80% of the variation. Additionally, the performance of SPI 12 and 24 was significantly different from the performance of SPI 3, 6 and 9. A number of studies have indicated improved predictability of SPI 12 compared to SPI 3 (Belayneh and Adamowski 2012; Djerbouai and Souag-Gamane 2016; Zhang et al. 2017b). In the case of SPI at the 3-month timescale, the variability in precipitation is high because monthly data for fewer months are used, that is, the calculated SPI value is for a accumulated precipitation for 3-month period, whereas in SPI 12, it is for 12-month period (Djerbouai and Souag-Gamane 2016; Guenang and Kamga 2014), resulting in a smoother time series for SPI at the 12-month timescale (Loukas and Vasiliades 2004); hence, it is likely that the models will generate better forecasts.

Model assessment should be based on suitable lead time selection, such that a lead time provides sufficient time to implement proactive measures to reduce drought impacts (Nikbakht et al. 2012). Based on the performance measures (Fig. 4), the short-term forecasts or forecasts generated three months ahead of time, operated the best, with an accuracy of approximately 70%, and 50%, respectively, for a medium range forecast. The long-term forecast has accuracy of below 50%. Moreover, the accuracy of medium- and long-range forecasts is generally similar, while the accuracy of short-term forecast is significantly different as denoted by the compact letter display (a, b). A deterioration in the performance is noted as the forecast period increases, indicated by the overall low correlation and high error values.

The linear mixed model allows for the inclusion of random effects which can quantify variation between research groups or country to country variation (Bolker et al. 2009). Figure 5 shows that there is considerable variation between studies from different countries. The graph shows the variation between studies originating from a specific country as well as the variation within a study. For the analysis on RMSE, China had the most

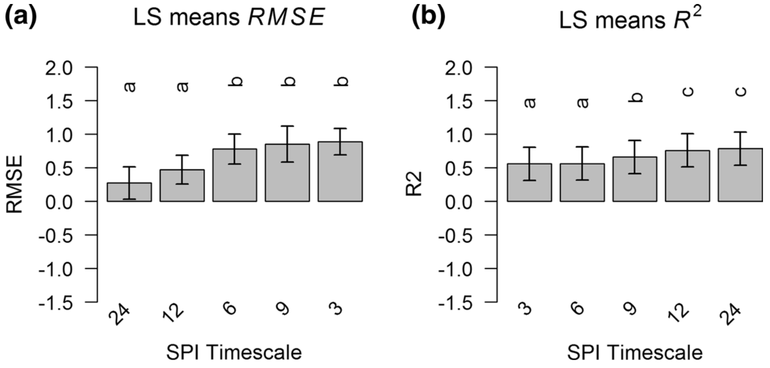

Fig. 3 Boxplot of estimated least square (LS) means of the linear mixed model (LMM) applied to the **a** RMSE and **b** R^2 values with SPI timescale as a factor, including 95% confidence intervals

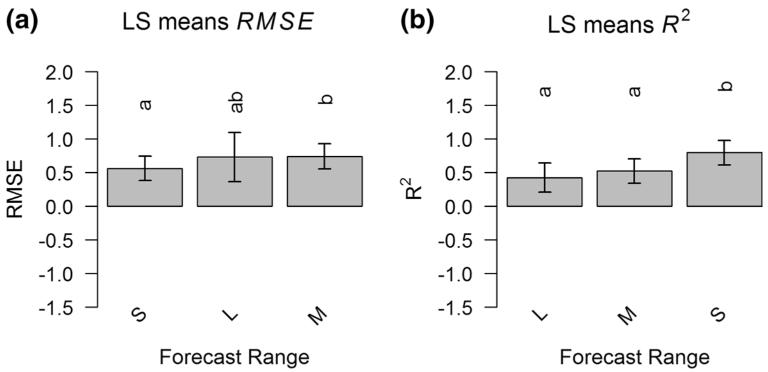

Fig. 4 Boxplot of estimated least square (LS) means with linear mixed model (LMM) applied to forecast accuracy according to prediction at various lead times measured in **a** RMSE and **b** R^2 values, including 95% confidence intervals. *S* short-range forecasts, *M* mid-range forecasts, *L* long-range forecasts

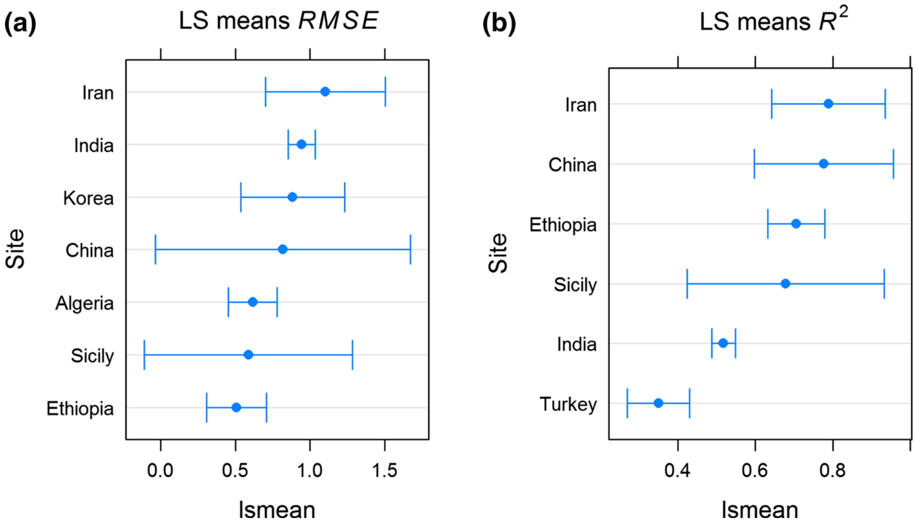

Fig. 5 A plot of random effects of linear mixed model applied to **a** RMSE and **b** R^2 including 95% confidence intervals. The manuscript authors were nested within study site to account for variations between model design, software, model validation and other unknown variations between research groups

variation in the results, while India was found to have the least variation. For the analysis on R^2, Sicily showed the highest variation in the results with consistent results found in studies from India. Additionally, the RMSE values for studies from India ranged from 0.5 to 1.0, while the RMSE values for studies from China ranged from 0 to 1.5. On the other hand, the R^2 values for studies in India were between 0.5 and 0.6, while the R^2 values for studies in China were 0.6–0.8. While studies from Iran showed a higher R^2 values as well as a higher error value. These could also highlight the quality of drought studies available in each area as well as the quality of work produced by the different authors.

5 Discussion

While most of the models yielded reasonably good results, the WANN was found the best-performing model in the validation models in this study (Fig. 2). Mouatadid et al. (2015), highlighted the advantages of using the ANN for drought prediction by comparing MLR and ANN in Australia, and showed that the ANN outperforms the linear regression models. A major advantage of ANNs is that they are able to learn patterns quickly even from complicated or noisy data (Belayneh et al. 2016). In contrast, a model such as time series assumes a linear relationship and hence is not able to capture non-linearity in the data, which is often the case for hydrological datasets (Djerbouai and Souag-Gamane 2016). However, according to our meta-analysis, the ANN alone is not superior to other types of machine learning models, probabilistic models such as the SVM, Markov chain and time series models comprising of ARIMA/SARIMA. The performance of the ANN is similar to other simpler models; however, the WANN proved superior in performance and the results of WANN were significantly better than other models. A number of studies have reported that the WANN is a better model as opposed to the regular ANN model (Adamowski and Sun 2010; Nourani et al. 2009; Wang and Ding 2003).

One crucial aspect of the WANN is that its performance is dependent on preprocessing of input data to account for potential non-stationarity in the data. Stationary data are those that have constant statistical properties such as mean and variance, are not time dependent and do not have abrupt changes in the series (Nason 2006). Ideally, stationary data are essential to design a model for future forecasting to accurately represent the phenomenon (Adhikari and Agrawal 2013; Huang et al. 1998). However, in the real world, data never comprise of constant statistical properties, giving rise to non-stationarities in the data. Wavelet transforms have been considered a potential solution to the problem of non-stationary data in hydrological forecasting and has been identified as one of the best methods to transform the data (Adamowski and Sun 2010; Huang et al. 1998). It operates by passing the time series to a function, which results in the decomposition of the original time series at multiple levels and scales, representing the same series with new values which are stationary in time (Djerbouai and Souag-Gamane 2016). This decomposition helps to de-noise the data, which is better handled by the models (Graps 1995). Additionally, this helps to yield more meaningful information from the series, which may have previously been masked by abrupt changes or seasonality in the data (Chaovalit et al. 2011). Wavelet analysis reduces the complexity of the data, causing the series to become more parsimonious (Djerbouai and Souag-Gamane 2016), resulting in a better-fitting model. However, this needs to be applied with caution, taking into account the filter type, length of series, type of decomposition and level of decomposition to be used, all of which affect the final result. For instance, decomposition into too many levels can result in the loss of meaningful information from the series (Guimarães Santos and Silva 2014). For model initialisation, the data should be partitioned into the respective training and validation stages before decomposing the series rather than splitting the data post decomposition as this can introduce bias in the forecasts due to the model being validated on data previously exposed during training stage (Deo et al. 2017b). An antecedent step in neural network models is to normalise the data before using it for training (Belayneh and Adamowski 2012). The validation of forecasts should be carried out using the observed SPI series (calculated or actual value) and not the normalised SPI data. Basing the validation on normalised series will potentially introduce artificial skill in the forecast and needs to be avoided.

Our results indicate that the SPI calculated for a 12-month or 24-month timescale produces the optimum drought forecasts for a region (Fig. 3). A possible explanation for this may be that the SPI for longer timescales is smoother and hence easier to forecast as opposed to the SPI for shorter timescales (Loukas and Vasiliades 2004). The application of the SPI at shorter timescales is not recommended and needs to be performed with caution to prevent overestimation or underestimation of hydro-meteorological events (Fluixá-Sanmartín et al. 2018; Wu et al. 2007). Calculation of SPI with short timescales results in a non-normally distributed SPI series owing to the high incidence of zero precipitation days in the precipitation data; consequently, the index fails to characterise a drought event effectively (Wu et al. 2007). Our results are consistent with previous work reporting that longer timescales are the best timescale to calculate the SPI and carry out drought modelling (Jain et al. 2015). Additionally, drought events with longer duration are not easily identifiable with short timescales (Fluixá-Sanmartín et al. 2018). However, we cannot entirely dismiss the use of shorter SPI timescales in other aspects of drought management. For instance, it has been shown for a river basin in Spain that the longer timescales (i.e. SPI 12 and SPI 24) are not ideal for drought quantification within surface flows, that correlate well to short SPI timescales, whereas reservoir storages are more sensitive to longer timescales (7–10 months) (Vicente-Serrano and López-Moreno 2005). Furthermore, a study that used the SPI at short and long timescales to interpolate historic drought events revealed that the short timescales (3 and 6) captured more drought events than did the long timescales (12 and 24) (Buttafuoco and Caloiero 2014). Therefore, SPI timescale selection should be dependent on the type of study undertaken, the type of catchment or area, the regional features and the type of drought examined (Vasiliades et al. 2011). This study emphasises the need to test the drought monitoring indices and the suitable timescale for sensitivity in a specific region before integrating it in management plans.

From the meta-analysis, the best results are achieved when forecasts are made 1–3 months prior, representing short-range forecasts (Fig. 4). From a planning and disaster management perspective, risk-reduction strategies should allow for sufficient time for contingency planning prior to the onset of a drought event (Karavitis et al. 2011). Governments end up spending large amounts of money in drought relief programmes after the disaster has materialised (Wilhite et al. 1986). A sufficient time frame to release the warning before the actual onset of drought will be beneficial for public awareness and also in reducing the associated costs in the aftermath of the disaster (Steinemann 2006). As water is critical in the early growth stages of many crops, a short-term forecast (1–3 month) in advance will also assist growers to make informed agronomic decisions. Forecasts produced 1–3 months have the highest accuracy and can be used with confidence in early warning systems and would be reliable for the forecast users. Therefore, this study recommends a lead time for forecasts ranging from 1 to 3 months, enabling the government to implement mitigation strategies ahead of time, which may result in a better action plan and financial security.

This review revealed a general paucity of studies representing different geographical locations of the world. A number of the studies were situated in the Middle East and Asia region, which may be linked to the type of indices commonly used in these areas. The variation in the results was the highest for studies undertaken in China and Sicily, and lowest for studies undertaken in India (Fig. 5). This indicates that the performance gap in studies in India was not significantly different, and therefore, the methodology can be replicated for other regions with similar climate. This review focused primarily on the SPI, and the SPI is a widely used index in these localities, which is not surprising given the simple calculation of the SPI, particularly in data-sparse areas. Other areas, for example, the continental USA, may be more reliant on indices such as the SPEI and the PDSI. Because studies on drought forecasting to establish early warning system are still scarce, there remain opportunities for more research in the areas of meteorology and hydrology.

The variability in the results may be due to various factors, which will be discussed. Studies that use climate indices in conjunction with drought indices to forecast may report different results than others that utilise the drought index on its own. An example of this is shown in the works of Memarian et al. (2016) who forecasted SPI by incorporating lagged global climatic signals and precipitation. Sensitivity analysis of models indicated that models with ENSO and precipitation indices produce better results. However, there is still uncertainty in this regard as other studies report no improvement in result with the use of climate signals. The ocean–atmospheric interactions driven by sea surface temperature have been useful in improving rainfall forecasts in general (Enfield and Alfaro 1999; Luo et al. 2005), but only to a certain limit (Westra and Sharma 2010). Ultimately, it depends on the extent of influence by the climate drivers on precipitation regimes in an area which needs to be determined for that specific location. In our meta-analysis, global climate signals were not used as a factor to determine the model performance because of the lower number of samples and inconsistencies in the data. Thus, this may be investigated further for regions where climate indices can be incorporated to determine the accuracy and improvability of forecasts.

Lastly, although the minimum number of study requirements was met, a limitation for this review was the small sample size for each model type. The ANN and ARIMA had the largest sample sizes, further highlighting their usage for drought modelling and forecasting. This meta-analysis can be replicated as more study samples become available in future for other model types with different drought indices.

6 Conclusion

Our meta-analysis demonstrated that the performances of the different machine learning, probabilistic and time series models are generally reasonable and similar to each other. However, preprocessing the data by wavelet analysis to remove non-stationarities renders the best results with minimum error and highest correlation overall in every aspect of the analysis. Additionally, the most reliable timescales to calculate the SPI are 12 and 24 months as this results in a smoother series which is better handled by the different models. However, a suitable timescale is ultimately dependent on the area of study and the type of drought to be addressed. Therefore, it is ideal to test the timescale for sensitivity in a region before implementing it in operational early warning system. A short-term forecast has a high accuracy and therefore can be used as a suitable lead time to inform the government and other agencies to reduce drought-related risks. The general trend observed from this review is a paucity of studies representative of different geographical locations. In hindsight, a large majority of the studies have used SPI for drought modelling; therefore, it was the index of interest for this study. However, the future recommendation is to carry out a similar analysis with indices other than the SPI (e.g. the PDSI, EDI or SPEI) to determine which index works best with a model type. Additionally, future studies on climate drivers as a factor may be performed to determine the connection with and influence on overall model performance and sensitivity when more studies become available.

Acknowledgements We would like to convey our thanks to the Department of Foreign Affairs and Trade, Australia, and Asia Pacific Network for Global Change Research for providing funding for this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abrahart RJ, See LM, Solomatine DP (2008) Practical hydroinformatics: computational intelligence and technological developments in water applications, vol 68. Springer, Berlin

Adamowski J, Sun K (2010) Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. J Hydrol 390:85–91

Adamowski J, Fung Chan H, Prasher SO, Ozga-Zielinski B, Sliusarieva A (2012) Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. Water Resour Res 48:1–14

Adhikari R, Agrawal R (2013) An introductory study on time series modeling and forecasting. arXiv:13026613

Altun Y, Tsochantaridis I, Hofmann T (2003) Hidden markov support vector machines. In: Proceedings of the 20th international conference on machine learning (ICML-03), pp 3–10

Ashouri H et al (2015) PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull Am Meteorol Soc 96:69–83

Bacanli UG, Firat M, Dikbas F (2009) Adaptive neuro-fuzzy inference system for drought forecasting. Stoch Environ Res Risk Assess 23:1143–1154

Bates D, Maechler M, Bolker B, Walker S (2014) lme4: linear mixed-effects models using Eigen and S4. R Package Version 1:1–23

Belayneh A, Adamowski J (2012) Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. Appl Comput Intell Soft Comput 2012:6

Belayneh A, Adamowski J, Khalil B, Quilty J (2016) Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. Atmos Res 172:37–47

Beven K, Young P (2013) A guide to good practice in modeling semantics for authors and referees. Water Resour Res 49:5092–5098

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White J-SS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24:127–135

Bonaccorso B, Cancelliere A, Rossi G (2015) Probabilistic forecasting of drought class transitions in Sicily (Italy) using standardized precipitation index and North Atlantic oscillation index. J Hydrol 526:136–150

Boulanger J-P, Leloup J, Penalba O, Rusticucci M, Lafon F, Vargas W (2005) Observed precipitation in the Paraná-Plata hydrological basin: long-term trends, extreme conditions and ENSO teleconnections. Clim Dyn 24:393–413

Box GE, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control. Wiley, Hoboken

Briggs A, Sculpher M (1998) An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 13:397–409

Brown M, Harris CJ (1994) Neurofuzzy adaptive modelling and control. Prentice Hall, Upper Saddle River

Brown JR, Moise AF, Colman RA (2013) The South Pacific Convergence Zone in CMIP5 simulations of historical and future climate. Clim Dyn 41:2179–2197

Buttafuoco G, Caloiero T (2014) Drought events at different timescales in southern Italy (Calabria). J Maps 10:529–537. https://doi.org/10.1080/17445647.2014.891267

Cancelliere A, Di Mauro G, Bonaccorso B, Rossi G (2007) Drought forecasting using the standardized precipitation index. Water Resour Manag 21:801–819

Chaovalit P, Gangopadhyay A, Karabatis G, Chen Z (2011) Discrete wavelet transform-based time series analysis and mining. ACM Comput Surv (CSUR) 43:6

Chen C-H, Li K-C (1998) Can SIR be as popular as multiple linear regression? Stat Sinica 8(2):289–316

Choubin B, Khalighi-Sigaroodi S, Malekian A, Kişi Ö (2016) Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. Hydrol Sci J 61:1001–1009. https://doi.org/10.1080/02626667.2014.966721

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/bf00994018

Dai A (2011) Drought under global warming: a review. Wiley Interdiscip Rev Clim Change 2:45–65. https://doi.org/10.1002/wcc.81

Dehghani M, Saghafian B, Nasiri Saleh F, Farokhnia A, Noori R (2014) Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation. Int J Climatol 34:1169–1180

Delbiso TD, Altare C, Rodriguez-Llanes JM, Doocy S, Guha-Sapir D (2017) Drought and child mortality: a meta-analysis of small-scale surveys from Ethiopia. Sci Rep 7:2212

Deo RC, 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 161:65–81

Deo RC, Kisi O, Singh VP (2017a) Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. Atmos Res 184:149–175

Deo RC, Tiwari MK, Adamowski JF, Quilty JM (2017b) Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. Stoch Environ Res Risk Assess 31:1211–1240

Djerbouai S, Souag-Gamane D (2016) Drought forecasting using neural networks, wavelet neural networks, and stochastic models: case of the Algerois Basin in North Algeria. Water Resour Manag 30:2445–2464

El Ibrahimi A, Baali A (2018) Application of several artificial intelligence models for forecasting meteorological drought using the standardized precipitation index in the Saïss Plain (Northern Morocco). Int J Intell Eng Syst 11(1):267–275

Enfield DB, Alfaro EJ (1999) The dependence of Caribbean rainfall on the interaction of the tropical Atlantic and Pacific Oceans. J Clim 12:2093–2103

Farokhnia A, Morid S, Byun H-R (2011) Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. Theor Appl Climatol 104:71–81

Fluixá-Sanmartín J et al (2018) Searching for the optimal drought index and timescale combination to detect drought: a case study from the lower Jinsha River basin, China. Hydrol Earth Syst Sci 22:889–910

Ganguli P, Reddy MJ (2014) Ensemble prediction of regional droughts using climate inputs and the SVM–copula approach. Hydrol Process 28:4989–5009

Graps A (1995) An introduction to wavelets. IEEE Comput Sci Eng 2:50–61

Guenang GM, Kamga FM (2014) Computation of the standardized precipitation index (SPI) and its use to assess drought occurrences in Cameroon over recent decades. J Appl Meteorol Climatol 53:2310–2324

Guimarães Santos CA, Silva GBLD (2014) Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. Hydrol Sci J 59:312–324

Gumus V, Algin HM (2017) Meteorological and hydrological drought analysis of the Seyhan–Ceyhan River Basins. Turk Meteorol Appl 24:62–73

Guttman NB (1998) Comparing the Palmer drought index and the standardized precipitation index. J Am Water Resour As 34:113–121

Han P, Wang PX, Zhang SY, Zhu DH (2010) Drought forecasting based on the remote sensing data using ARIMA models. Math Comput Model 51:1398–1403. https://doi.org/10.1016/j.mcm.2009.10.031

Hao Z, AghaKouchak A, Nakhjiri N, Farahmand A (2014) Global integrated drought monitoring and prediction system. Sci Data 1:140001

Hayes M, Svoboda M, Le Comte D, Redmond KT, Pasteris P (2005) Drought monitoring: new tools for the 21st century. Taylor and Francis, Routledge

Hayes M, Svoboda M, Wall N, Widhalm M (2011) The Lincoln declaration on drought indices: universal meteorological drought index recommended. Bull Am Meteorol Soc 92:485–488

Hill T, Marquez L, O'Connor M, Remus W (1994) Artificial neural network models for forecasting and decision making. Int J Forecast 10:5–15

Huang NE et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond A Math Phys Eng Sci 454(1971):903–995

Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. Science 269:676–679

Hydrology ATCoAoANNi (2000) Artificial neural networks in hydrology. II: hydrologic applications. J Hydrol Eng 5:124–137

Hyndman RJ, Athanasopoulos G (2018) Forecasting: principles and practice. OTexts[Online]. Available at: https://otexts.com/fpp2/intro.html

Ionita M, Lohmann G, Rimbu N, Chelcea S, Dima M (2012) Interannual to decadal summer drought variability over Europe and its relationship to global sea surface temperature. Clim Dyn 38:363–377

Jain VK, Pandey RP, Jain MK, Byun H-R (2015) Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin Weather and Climate. Extremes 8:1–11. https://doi.org/10.1016/j.wace.2015.05.002

Ji L, Peters AJ (2003) Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. Remote Sens Environ 87:85–98. https://doi.org/10.1016/S0034-4257(03)00174-3

Kane R (1997) Prediction of droughts in north-east Brazil: role of ENSO and use of periodicities. Int J Climatol J R Meteorol Soc 17:655–665

Karavitis CA, Alexandris S, Tsesmelis DE, Athanasopoulos G (2011) Application of the standardized precipitation index (SPI) in Greece. Water 3:787–805

Keskin ME, Terzi O, Taylan ED, Küçükyaman D (2009) Meteorological drought analysis using data-driven models for the Lakes District. Turk Hydrol Sci J 54:1114–1124

Khadr M (2016) Forecasting of meteorological drought using hidden Markov model (case study: the upper Blue Nile river basin, Ethiopia). Ain Shams Eng J 7:47–56

Khalili D, Farnoud T, Jamshidi H, Kamgar-Haghighi AA, Zand-Parsa S (2011) Comparability analyses of the SPI and RDI meteorological drought indices in different climatic zones. Water Resour Manag 25:1737–1757

Khan KS, Kunz R, Kleijnen J, Antes G (2003) Five steps to conducting a systematic review. J R Soc Med 96:118–121

Kroeker KJ, Kordas RL, Crim RN, Singh GG (2010) Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. Ecol Lett 13:1419–1434

Kumar R et al (2016) Multiscale evaluation of the standardized precipitation index as a groundwater drought indicator. Hydrol Earth Syst Sci 20:1117–1131

Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: tests in linear mixed effects models. J Stat Softw 82(13):1–26. https://doi.org/10.18637/jss.v082.i13

Lenth RV (2016) Least-squares means: the R package lsmeans. J Stat Softw 69:1–33

Liberati A et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 6:e1000100

Liu D, You J, Xie Q, Huang Y, Tong H (2018) Spatial and temporal characteristics of drought and flood in Quanzhou based on standardized precipitation index (SPI) in recent 55 years. J Geosci Environ Prot 6:25

Lohani V, Loganathan G, Mostaghimi S (1998) Long-term analysis and short-term forecasting of dry spells by Palmer Drought Severity Index. Hydrol Res 29:21–40

Loukas A, Vasiliades L (2004) Probabilistic analysis of drought spatiotemporal characteristics in Thessaly region, Greece. Nat Hazards Earth Syst Sci 4:719–731

Luo J-J, Masson S, Behera S, Shingu S, Yamagata T (2005) Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. J Clim 18:4474–4497

Mallya G, Tripathi S, Kirshner S, Govindaraju RS (2012) Probabilistic assessment of drought characteristics using hidden Markov model. J Hydrol Eng 18:834–845

Mathugama S, Peiris T (2011) Critical evaluation of dry spell research. Int J Basic Appl Sci 11:153–160

Maybank J, Bonsai B, Jones K, Lawford R, O'brien E, Ripley E, Wheaton E (1995) Drought as a natural disaster. Atmos Ocean 33:195–222

McGree S, Schreider S, Kuleshov Y (2016) Trends and variability in droughts in the Pacific Islands and Northeast Australia. J Clim 29:8377–8397

McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th conference on applied climatology, American Meteorological Society Boston, MA, vol 22, pp 179–183

Memarian H, Bilondi MP, Rezaei M (2016) Drought prediction using co-active neuro-fuzzy inference system, validation, and uncertainty analysis (case study: Birjand, Iran). Theor Appl Climatol 125:541–554

Mishra A, Desai V (2005) Drought forecasting using stochastic models. Stoch Environ Res Risk Assess 19:326–339

Mishra A, Desai V (2006) Drought forecasting using feed-forward recursive neural network. Ecol Model 198:127–138

Mishra AK, Singh VP (2011) Drought modeling—a review. J Hydrol 403:157–175

Mishra A, Desai V, Singh V (2007) Drought forecasting using a hybrid stochastic and neural network model. J Hydrol Eng 12:626–638

Mishra AK, Singh VP, Desai VR (2009) Drought characterization: a probabilistic approach. Stoch Environ Res Risk Assess 23:41–55. https://doi.org/10.1007/s00477-007-0194-2

Modarres R (2007) Streamflow drought time series forecasting. Stoch Environ Res Risk Assess 21:223–233

Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097

Mokhtarzad M, Eskandari F, Vanjani NJ, Arabasadi A (2017) Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. Environ Earth Sci 76:729

Moreira EE, Mexia JT, Pereira LS (2012) Are drought occurrence and severity aggravating? A study on SPI drought class transitions using log-linear models and ANOVA-like inference. Hydrol Earth Syst Sci 16:3011–3028

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

Mouatadid S, Deo RC, Adamowski JF (2015) Prediction of SPEI using MLR and ANN: a case study for Wilsons Promontory Station in Victoria. In: 2015 IEEE 14th international conference on machine learning and applications (ICMLA), IEEE, pp 318–324

Myers RH, Myers RH (1990) Classical and modern regression with applications, vol 2. Duxbury Press, Belmont

Nason GP (2006) Stationary and non-stationary time series. In: Mader HM, Coles SG, Connor CB, Connor LJ (eds) Statistics in volcanology, vol 1. Special publications of IAVCEI. Geological Society of London, London

Nayak PC, Sudheer K, Rangan D, Ramasastri K (2004) A neuro-fuzzy computing technique for modeling hydrological time series. J Hydrol 291:52–66

Nguyen V, Li Q, Nguyen L (2017) Drought forecasting using ANFIS-a case study in drought prone area of Vietnam. Paddy Water Environ 15:605–616

Nikbakht SA, Zahraie B, Nasseri M (2012) Seasonal meteorological drought prediction using support vector machine. Water Wastewater 23:72–84

Nourani V, Komasi M, Mano A (2009) A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resour Manag 23:2877

Odabas MS, Leelaruban N, Simsek H, Padmanabhan G (2014) Quantifying impact of droughts on barley yield in North Dakota, USA using multiple linear regression and artificial neural network. Neural Netw World 24:343

Özger M, Mishra AK, Singh VP (2012) Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. J Hydrometeorol 13:284–297

Paulo AA, Pereira LS (2007) Prediction of SPI drought class transitions using Markov chains. Water Resour Manag 21:1813

Paulo A, Ferreira E, Coelho C, Pereira L (2005) Drought class transition analysis through Markov and Loglinear models, an approach to early warning. Agric Water Manag 77:59–81

Pedhazur EJ (1982) Multiple regression in behavioral research: explanation and prediction. Harcourt Brace Jovanovish College Publishers, San Diego

Pinheiro J (2009) nlme: linear and nonlinear mixed effects models. http://cran r-project org/web/packages/nlme/index html. Accessed 25 Dec 2017

Pinheiro J, Bates D, DebRoy S, Sarkar D, Team RC (2007) Linear and nonlinear mixed effects models. R Package Version 3:1–89

Raghavendra NS, Deka PC (2014) Support vector machine applications in the field of hydrology: a review. Appl Soft Comput 19:372–386. https://doi.org/10.1016/j.asoc.2014.02.002

Rezaeianzadeh M, Tabari H, Yazdi AA, Isik S, Kalin L (2014) Flood flow forecasting using ANN, ANFIS and regression models. Neural Comput Appl 25:25–37

Rossi G (2003) Requisites for a drought watch system. In: Rossi G, Cancelliere A, Pereira LS, Oweis T, Shatanawi M, Zairi A (eds) Tools for drought mitigation in Mediterranean regions. Springer, Dordrecht, pp 147–157

Rustad L et al (2001) A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. Oecologia 126:543–562

Salinger M, McGree S, Beucher F, Power SB, Delage F (2014) A new index for variations in the position of the South Pacific convergence zone 1910/11–2011/2012. Clim Dyn 43:881–892

Sapankevych NI, Sankar R (2009) Time series prediction using support vector machines: a survey. IEEE Comput Intell Mag 4:24–38

Schardt C, Adams MB, Owens T, Keitz S, Fontelo P (2007) Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 7:16

Sheffield J, Wood EF (2007) Characteristics of global and regional drought, 1950–2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J Geophys Res Atmos 112(D17):1–21. https://doi.org/10.1029/2006JD008288

Shirmohammadi B, Moradi H, Moosavi V, Semiromi MT, Zeinali A (2013) Forecasting of meteorological drought using Wavelet-ANFIS hybrid model for different time steps (case study: southeastern part of east Azerbaijan province, Iran). Nat Hazards 69:389–402

Soares PM, Cardoso RM (2018) A simple method to assess the added value using high-resolution climate distributions: application to the EURO-CORDEX daily precipitation. Int J Climatol 38:1484–1498

Solomatine DP, Ostfeld A (2008) Data-driven modelling: some past experiences and new approaches. J Hydroinform 10:3–22

Stanley TD, Jarrell SB (1989) Meta-Regression analysis: a quantitative method of literature surveys. J Econ Surv 3:161–170

Steinemann AC (2006) Using climate forecasts for drought management. J Appl Meteorol Climatol 45:1353–1361

Tranfield D, Denyer D, Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br J Manag 14:207–222

Tu JV (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 49:1225–1231. https://doi.org/10.1016/S0895-4356(96)00002-9

Tuttle BD, Isenburg MV, Schardt C, Powers A (2009) PubMed instruction for medical students: searching for a better way. Med Ref Serv Q 28:199–210

Usman MT, Reason C (2004) Dry spell frequencies and their variability over southern Africa. Clim Res 26:199–211

Valentine JC, Pigott TD, Rothstein HR (2010) How many studies do you need? A primer on statistical power for meta-analysis. J Educ Behav Stat 35:215–247

Van Der Wiel K, Matthews AJ, Joshi MM, Stevens DP (2015) The influence of diabatic heating in the South Pacific Convergence Zone on Rossby wave propagation and the mean flow. Q J R Meteorol Soc 142:901–910

Vasiliades L, Loukas A, Liberis N (2011) A water balance derived drought index for Pinios River Basin, Greece. Water Resour Manag 25:1087–1101

Vicente-Serrano SM, López-Moreno J (2005) Hydrological response to different time scales of climatological drought: an evaluation of the standardized precipitation index in a mountainous Mediterranean basin. Hydrol Earth Syst Sci Dis 9:523–533

Wandres M, Pattiaratchi C, Hetzel Y, Wijeratne E (2018) The response of the southwest Western Australian wave climate to Indian Ocean climate variability. Clim Dyn 50:1533–1557

Wang S-C (2003) Artificial neural network. Interdisciplinary computing in java programming. Springer, Boston, pp 81–100

Wang W, Ding J (2003) Wavelet network model and its application to the prediction of hydrology. Nat Sci 1:67–71

Westra S, Sharma A (2010) An upper limit to seasonal rainfall predictability? J Clim 23:3332–3351

Wetterhall F, Winsemius H, Dutra E, Werner M, Pappenberger E (2015) Seasonal predictions of agro-meteorological drought indicators for the Limpopo basin. Hydrol Earth Syst Sci 19:2577–2586

Wilby RL, Prudhomme C, Parry S, Muchan K (2015) Persistence of hydrometeorological droughts in the United Kingdom: a regional analysis of multi-season rainfall and river flow anomalies. J Extreme Events 2:1550006

Wilhite DA (2000) Drought as a natural hazard: concepts and definitions. DigitalCommons@University of Nebraska-Lincoln. Chap. 1, pp 1–18. https://pdfs.semanticscholar.org/978b/179885ad9cd08da8ef466ca717425a4eb82c.pdf

Wilhite DA (2002) Combating drought through preparedness. Natural resources forum, vol 4. Wiley, Hoboken, pp 275–285

Wilhite DA, Rosenberg NJ, Glantz MH (1986) Improving federal response to drought. J Clim Appl Meteorol 25:332–342

Wilks DS (2011) Statistical methods in the atmospheric sciences, vol 100. Academic Press, Cambridge

Wong T-T (2015) Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit 48:2839–2846

Wu H, Svoboda MD, Hayes MJ, Wilhite DA, Wen F (2007) Appropriate application of the standardized precipitation index in arid locations and dry seasons. Int J Climatol 27:65–79

Xie S-P (2009) Ocean-Atmosphere Interaction and Tropical Climate. The encyclopedia of life support systems (EOLSS). Trop Meteorol 1:1–13

Zargar A, Sadiq R, Naser B, Khan FI (2011) A review of drought indices. Environ Rev 19:333–349

Zhang X, Obringer R, Wei C, Chen N, Niyogi D (2017a) Droughts in India from 1981 to 2013 and implications to wheat production. Sci Rep 7:44552

Zhang Y, Li W, Chen Q, Pu X, Xiang L (2017b) Multi-models for SPI drought forecasting in the north of Haihe River Basin, China. Stoch Environ Res Risk Assess 31:2471–2481