Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia

CrossMark

Ravinesh C. Deo [a,*], Mehmet Şahin [b]

[a] School of Agricultural Computational and Environmental Sciences, International Centre of Applied Climate Science (ICACS), University of Southern Queensland, Springfield 4300, Australia
[b] Department of Electrical and Electronics Engineering, Siirt University, 56100 Siirt, Turkey

ARTICLE INFO

ABSTRACT

The prediction of future drought is an effective mitigation tool for assessing adverse consequences of drought events on vital water resources, agriculture, ecosystems and hydrology. Data-driven model predictions using machine learning algorithms are promising tenets for these purposes as they require less developmental time, minimal inputs and are relatively less complex than the dynamic or physical model. This paper authenticates a computationally simple, fast and efficient non-linear algorithm known as extreme learning machine (ELM) for the prediction of Effective Drought Index (EDI) in eastern Australia using input data trained from 1957–2008 and the monthly EDI predicted over the period 2009–2011. The predictive variables for the ELM model were the rainfall and mean, minimum and maximum air temperatures, supplemented by the large-scale climate mode indices of interest as regression covariates, namely the Southern Oscillation Index, Pacific Decadal Oscillation, Southern Annular Mode and the Indian Ocean Dipole moment. To demonstrate the effectiveness of the proposed data-driven model a performance comparison in terms of the prediction capabilities and learning speeds was conducted between the proposed ELM algorithm and the conventional artificial neural network (ANN) algorithm trained with Levenberg–Marquardt back propagation. The prediction metrics certified an excellent performance of the ELM over the ANN model for the overall test sites, thus yielding Mean Absolute Errors, Root-Mean Square Errors, Coefficients of Determination and Willmott's Indices of Agreement of 0.277, 0.008, 0.892 and 0.93 (for ELM) and 0.602, 0.172, 0.578 and 0.92 (for ANN) models. Moreover, the ELM model was executed with learning speed 32 times faster and training speed 6.1 times faster than the ANN model. An improvement in the prediction capability of the drought duration and severity by the ELM model was achieved. Based on these results we aver that out of the two machine learning algorithms tested, the ELM was the more expeditious tool for prediction of drought and its related properties.

© 2014 Elsevier B.V. All rights reserved.

Acronyms

ANN	artificial neural network
BOM	Bureau of Meteorology
BP	back propagation
d	Willmott's Index of Agreement
DEP	deviation of effective precipitation
EDI	Effective Drought Index
EDI_o	observed Effective Drought Index
EDI_p	predicted Effective Drought Index
ELM	extreme learning machine
EP	effective precipitation
GCM	global circulation model
Hardlim	hard limit

* Corresponding author: Tel.: +61 7 3470 4430.
   E-mail addresses: ravinesh.deo@usq.edu.au, physrcd@yahoo.com (R.C. Deo).

IOD	Indian Ocean Dipole
JISAO	Joint Institute of the Study of the Atmosphere and Ocean
LM	Levenberg–Marquardt
Logsig	logarithmic sigmoid
MAE	mean absolute error
MEP	mean effective precipitation
MLP	multi-layer perceptron
POAMA	Predictive Ocean Atmosphere Model of Australia
R^2	Coefficient of Determination
Radbas	radial bias
RDDI	Rainfall Decile Drought Index
RMSE	root mean square error
SAM	Southern Annular Mode
SLFM	Single Layer Feedforward Neural Network
SOI	Southern Oscillation Index
PDO	Pacific Decadal Oscillation
SPI	Standardized Precipitation Index
SPOTA	Seasonal Pacific Ocean Temperature Analysis
SST	sea surface temperature
ST	standard deviation
SVD	singular value decomposition
SVM	support vector machine
tansig	hyperbolic-tangent sigmoid
trainbfg	BFGS quasi-Newton
trainbr	Bayesian regulation
trainlm	Levenberg–Marquardt
trainoss	one-step secant
trainscg	scaled conjugate gradient
tribas	triangular basis

1. Introduction

To be prepared for the detrimental consequences of drought on freshwater planning and water resource environments, the forecasting of future drought is a priori knowledge. For this purpose, basically two types of models are considered in literature: physical models which predict coupled effects of the ocean and the atmosphere, known as Global Circulation Model (GCM) and statistical models that assimilate observed values of hydro-meteorological properties (e.g. temperature or rainfall) to forecast future drought events. In Australia, the GCM framework implemented as the official model used by the Australian Bureau of Meteorology (BOM) is the Predictive Ocean Atmosphere Model for Australia (POAMA) (Hudson et al., 2011; Zhao and Hendon, 2009) and that by the Queensland Department of Environment and Resource Management is the statistical analysis of climate indices by the Seasonal Pacific Ocean Temperature Analysis (SPOTA-1) (Day et al., 2010). However the predictions of rainfall by GCMs on some occasions have failed to predict very wet or very dry conditions that produced significant economic consequences (van den Honert and McAneney, 2011). For example the floods between November 2010 and January 2011 that left three-quarters of Queensland, Australia a disaster zone (Hurst, 2011) were not predicted well in advance (Abbot and Marohasy, 2014; Inquiry, 2011; Seqwater, 2011). Despite improvements in the performance of numerical weather models, they do not provide quantitative precipitation forecasts at enough spatial and temporal scales (Kuligowski and Barros, 1998). Consequently there is a gap in rainfall prediction capability by the GCM especially beyond 1 week or shorter than a season (Hudson et al., 2011). A study by Fawcett and Stone (2010) found that the prediction skill level of the GCM models was "only moderate" although better than the climatological persistence models or randomly guessed forecasts. A review of 27 GCM simulations used by the Intergovernmental Panel on Climate Change Coupled Model Intercomparison Project Phase 5 found significant disparities between the various physical models (Irving et al., 2012). Hence there is a growing need for evaluation of the modeling frameworks and computational approaches used for prediction of rainfall, and obviously for the future drought and flood events.

One alternative to the physical model is the machine learning (ML) (or statistical model) that is now being experimented in a wide variety of climate applications. The ML models utilize, assimilate and 'learn' from the evidence of past climate trends using observational dataset to predict the future. Many types of ML algorithms have recently been proposed in literature, including the co-integration methods that analyze relationships between stationary and non-stationary data (Kaufmann et al., 2011; Kaufmann and Stern, 2002), regression approaches for evaluating time-series properties of air temperature (Douglass et al., 2004; Stone and Allen, 2005), neural networks for predicting rainfall (Abbot and Marohasy, 2012, 2014), wavelet or vector-regression for hydro-meteorological forecasting (Belayneh and Adamowski, 2012) and rainfall events prediction using rule-based fuzzy inference systems (Asklany et al., 2011). The practical advantages of the ML algorithm over the GCM are the explanation of the externally driven climate without the need for complex physical models, easiness of experimentation, validation and evaluation, low computational burden, much more simple and fast in the training and the testing phases, the applicability to the data for a specific point of measurement (a specific area, for example) and the competitive performance compared to physical models (Ortiz-García et al., 2014). Consequently many studies are using different ML algorithms to demonstrate nearly coincident or in some cases, even better prediction yields than the GCM models. In fact, recent studies that compared rainfall predictions using ML models with the physical models demonstrated dramatic improvements in the prediction capability of the former models (Abbot and Marohasy, 2012, 2014; Luk et al., 2000; Mekanik et al., 2013; Nasseri et al., 2008). In particular, the work of Abbot and Marohasy (2012, 2014) that compared rainfall prediction from an ML algorithm with the POAMA used in Australia over geographically distinct regions in Queensland found that the former approach was superior as evidenced by the lower root mean square errors, lower mean absolute errors and higher correlation coefficients between the observed and modeled rainfall values.

A frequently used ML algorithm used in climate sciences is the artificial neural network (ANN). ANN is a powerful and versatile data-driven algorithm for capturing and representing complex input and output relationships (Abbot and Marohasy, 2012, 2014; Govindaraju, 2000; Şahin et al., 2013). This model has been tested for rainfall and temperature predictions in many parts of the world including Australia (Abbot and Marohasy, 2012, 2014; Masinde, 2013; Nastos et al., 2014; Ortiz-García et al., 2014, 2012; Shukla et al., 2011). However a major

challenge encountered by the ANN is the requirement of iterative tuning of model parameters, slow response of the gradient based learning algorithm used and the relatively low prediction accuracy compared to the more advanced ML algorithms (e.g. Acharya et al., 2013; Şahin et al., 2014). Therefore in this study we have adopted a much improved class of ML algorithm, known as extreme learning machine (ELM) as a statistical model in a problem of predicting the monthly Effective Drought Index (EDI) (Byun and Wilhite, 1999).

The development of better predictive models for forecasting rainfall and drought is an appealing problem as drought events in eastern Australia are known to significantly impact sustainable economic growth, infrastructure, daily living, agricultural industry and natural ecosystems (e.g. Deo, 2011; Deo et al., 2009; Dijk et al., 2013; McAlpine et al., 2007; McAlpine et al., 2009). However, in predicting future drought the application of the EDI for drought assessment has been considered superior to the many other indices used in literature (Dogan et al., 2012; Kim et al., 2009; Morid et al., 2006; Pandey et al., 2008). In particular, the merits of the EDI were emphasized in a study by Pandey et al. (2008) that investigated that drought in Orissa (India) showed the greater capability of the EDI compared to the other drought indices (e.g., Standardized Precipitation Index, SPI and the Rainfall Decile Drought Index, RDDI) in quantifying water resources in relation to drought. Another study by Morid et al. (2006) in Tehran (Iran) showed the significantly better performance of EDI, and its better response in detecting the start of drought when compared to other approaches (e.g., percent of normal, SPI, China-Z index, Z-Score). Moreover, Dogan et al. (2012) found the EDI to be very sensitive to subtle changes in rainfall, and Kim et al. (2009) showed the superiority of the EDI for continuous monitoring of short and long-term drought. The study of Byun et al. (2008) also showed the ability of the EDI for prediction of future drought in Korea using periodicity analysis of long-term observed dataset and the study of Kim et al. (2011) developed a spatio-temporal map using wavelet transformations for the Korean region. Despite a plethora of studies using the EDI for drought assessment in East Asia, to our best knowledge, no study on the prediction of the EDI has been conducted in Australia. In this context, the prediction of the EDI is a novel and an enlightening problem of interest, as this index can assist with the detection of actual drought onset and termination dates, and drought severity or intensity based on various water scarcity conditions (Byun et al., 2008; Byun and Wilhite, 1999; Dogan et al., 2012; Kim and Byun, 2009; Kim et al., 2009, 2011; Masinde, 2013; Morid et al., 2006; Pandey et al., 2008). The prediction of the EDI can also facilitate the measurement of short-term dry spells and long-term (or ongoing) drought events and more precisely, the detection of current usable water resources and precipitation return to normal for assessment of drought severity conditions (Dogan et al., 2012; Kim et al., 2009; Morid et al., 2006; Pandey et al., 2008).

The study described in this paper used the ELM algorithm for predicting the monthly EDI, and compared the output with the ANN algorithm. The prediction capabilities of both models were assessed by performance metrics like the Mean Absolute Error, Root-Mean Square Error, Coefficient of Determination and Willmott's Index of Agreement (Acharya et al., 2013; Rajesh and Prakash, 2011; Willmott, 1982). The application of these two models is well-established in computational science

although they are not so familiar to the climate science or hydrologic engineering community. The ELM algorithm (Huang et al., 2006) is a fast three step method designed for better convergence of the output using the Single Layer Feedforward Neural Network (SLFNs) with N hidden neurons and randomly chosen input weights. The hidden layer bias can exactly learn N distinct observations, therefore enlightening the use of observed meteorological dataset in predicting the future. A suite of 13 inputs for each model was used for training, where the primary meteorological inputs are the observed monthly rainfall and air temperature data at 4 observational sites in the region. Together, the primary large-scale climate mode indices known to impact drought events in eastern Australia, namely the Southern Oscillation Index (SOI), Indian Ocean Dipole (IOD), Southern Annular Mode (SAM) and Pacific Decadal Osculation (PDO) were used, which have recently been proven useful in rainfall prediction problems (Abbot and Marohasy, 2012, 2014; Mekanik et al., 2013). Our study was also motivated by many research works that showed greater potential of ELM algorithm over the simple arithmetic means approach and the singular value decomposition in predicting rainfall and drought events (e.g. Acharya et al., 2013) and also the superiority of the ELM over the ANN model for predicting solar radiation and air temperatures (Şahin, 2012; Şahin et al., 2013, 2014).

The purpose of this investigation is then threefold: A first objective consists of evaluating the capability of the ELM model considered in the problem of predicting monthly Effective Drought Index using meteorological datasets and climate mode indices as input variables. A second objective consists of a deep statistical analysis of the results disseminated by the ELM model in relation to the predictions by ANN model. A third objective is to deduce the optimum network architecture of the ELM and the ANN models, computational speeds and performance parameters. The conclusion is that the ELM model performs significantly better than the ANN model and therefore, is potentially useful for research in areas of effective management of water environments, agriculture and infrastructure. Moreover, accurate forecasts of drought using the proposed data-driven model can complement the different types of climate forecasting programs already undertaken, and can thus be an alternative modeling framework for rainfall and drought studies.

The rest of the paper is structured as follows: the next section presents a theoretical overview of the machine learning approaches used with the ELM as the primary, and ANN as the secondary (comparative) model in a prediction problem of the Effective Drought Index in eastern Australia. The design of the basic network architecture of the two learning algorithms has been given. Section 3 discusses the meteorological data utilized, computational procedures of the Effective Drought Index and the model development and evaluation parameters used in the experimental component together with statistical analysis for validating the model fidelity considered in this paper. Section 4 presents and discusses the results of our work and finally, concluding remarks for closing the paper are made in Section 5.

2. Basic theory of machine learning algorithm

2.1. Extreme learning machine (ELM)

Extreme learning machine (ELM) developed by Huang et al. (2006) is the state-of-art novel machine learning

algorithm for Single Layer Feedforward Neural Network (SLFNs). Consequently the ELM model has been widely used for the solution of estimation problems in many different fields and is now gaining attention within the climate research and applied engineering community (Acharya et al., 2013; Belayneh and Adamowski, 2012; Şahin et al., 2014). These investigations and others have demonstrated important advantages of the ELM model over the traditional neural network or vector classification schemes. The ELM model is easy to use and no parameters need to be tuned except the predefined network architecture, thus avoiding many complications faced by the gradient-based algorithms such as learning rate, learning epochs, and local minima. Importantly the ELM model has also been proven to be a faster algorithm compared with other conventional learning algorithms such as back propagation (BP) or support vector machines (SVM) (Rajesh and Prakash, 2011).

In the ELM approach most of the training is accomplished in time span of seconds or at least in minutes in large complex applications which are not easily achieved by using the traditional neural network models (Acharya et al., 2013; Sánchez-Monedero et al., 2014). The ELM model possesses similar generalization performance to the back propagation, the SVM and the singular value decomposition (SVD) algorithms in data classification and prediction problems. Therefore, the ELM model has been considered as an ideal computational algorithm for forecasting atmospheric and meteorological variables including solar energy, air temperature and rainfall that generally have large and complex datasets to be dealt with (Leu and Adi, 2011; Şahin, 2012; Şahin et al., 2013, 2014; Sánchez-Monedero et al., 2014; Wu and Chau, 2010).

In Fig. 1 the basic schematic topological structure of an ELM network has been illustrated. Briefly, the basic theory of the ELM model states that for N arbitrary distinct input samples $(x_k, y_k) \in R^n \times R^n$, the standard SLFNs with M hidden nodes and an activation $g(.)$ function are mathematically described as

$$\sum_{i=1}^{M} \beta_i g(x_k, ; c_i, a_i) = y_k \quad k = 1, 2, 3, \dots, N \quad (1)$$

where $c_i \in R$ is the randomly assigned bias of the ith hidden node and $w_i \in R$ is the randomly assigned input weight vector connecting the ith hidden node and the input nodes. β_i is the weight vector connection the ith hidden node to the output node. $g(x_k; c_i, w_i)$ is the output of the ith hidden node with respect to the input sample x_k. Each input is randomly assigned to the hidden nodes in ELM network. Then, Eq. (1) can be written as

$$H \beta = Y \quad (2)$$

where

$$H = \begin{bmatrix} g(x_1, ; c_1, , w_1) & \dots & g(x_1, ; c_M, , w_M) \\ g(x_N, ; c_1, , w_1) & \dots & g(x_N, ; c_M, , w_M) \end{bmatrix}_{N \times M} \quad (3)$$

$$H \beta = \left(\beta_1^T \beta_2^T, \dots, \beta_L^T \right)_{m \times M}^T \quad (4)$$

and the output (Y)

$$Y = \left(t_1^T t_2^T, \dots, t_L^T \right)_{m \times M}^T. \quad (5)$$

The output weights are derived by finding the least square solutions to the aforementioned linear system, which is given by
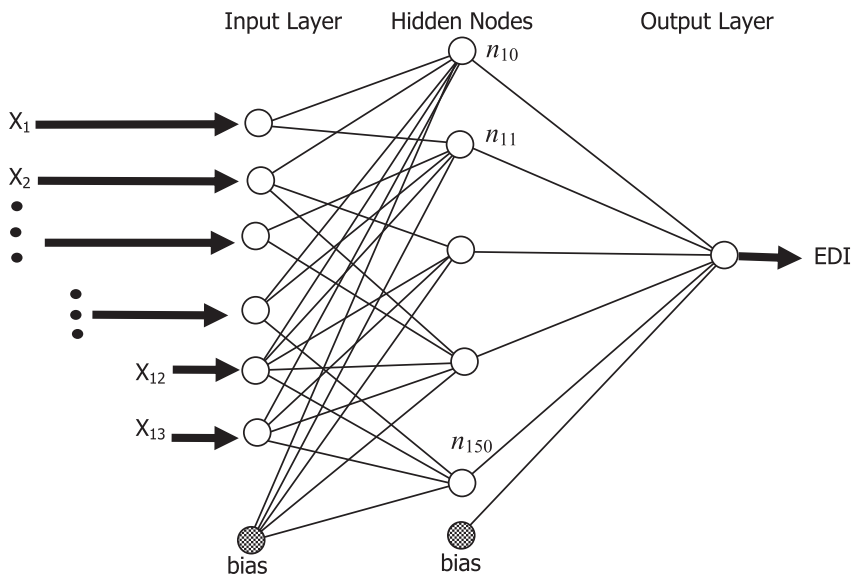
$$\beta = H^+ Y \quad (6)$$

Fig. 1. The topological structure of the extreme learning machine network used in this study. Input layer is denoted as X_1, X_2, …, X_{13}, hidden nodes as n_{10}, n_{11}, …, n_{150} and the output layer generates the predicted values of the Effective Drought Index (EDI).

where H^+ is the Moore–Penrose generalized inverse of the hidden layer matrix H and the input weights and hidden biases are randomly chosen, and the output weights are determined analytically. The principle which distinguishes ELM from the traditional neural network methodology is that all the parameters of the feed-forward networks (input weights and hidden layer biases) are not required to be tuned in the former. The studies of Tamura and Tateishi (1997) and Huang (2003) showed in their work that the SLFNs with randomly chosen input weights efficiently learn distinct training examples with minimum error. After randomly choosing input weights and the hidden layer biases, SLFNs can be simply considered as a linear system. The output weights which link the hidden layer to the output layer of this linear system can now be analytically determined through simple generalized inverse operation of the hidden layer output matrices. This simplified approach makes ELM model many times faster than that of traditional feedforward learning algorithms (Acharya et al., 2013; Şahin et al., 2013, 2014).

2.2. Artificial neural network (ANN)

In literature the radial basis function networks and multi-layer perceptron (MLP) are the examples of feed-forward networks used in artificial neural network models. However, MLP is the simplest and most commonly used ANN architecture (Sözen and Ali Akçayol, 2004). Consequently MLP is often adopted in hydrologic forecasting areas due to the simplicity of model design and its success in forecasting rainfall and its properties in many real-life applications (Belayneh and Adamowski, 2012; Şahin et al., 2013). In this study we adopted the MLP procedure in our ANN model trained with the Levenberg–Marquardt (LM) back propagation for comparison with results of the ELM model. The schematic topological structure of an ANN model has been illustrated in Fig. 2.

In basic terms, the MLP algorithm used in this study consisted of an input layer with 13 inputs and a hidden layer with 64 nodes. An output layer was used to generate the monthly EDI values as the predicted output using the ANN approach (Kim and Valdés, 2003) viz

$$y_k(t) = f_o \left[\sum_{j=1}^{m} w_{kj} . f_n \left(\sum_{i=1}^{N} w_{ji} x_i(t) + w_{j0} + w_{k0} \right) \right] \qquad (7)$$

where N is the number of samples, m ($=64$) is the number of hidden neurons, $x_i(t)$ is the ith output variable at the time-step used, w_{ji} is the weight that connects the ith neuron in the input layer and the jth neuron in the hidden layer, w_{j0} is the bias for the hidden jth hidden neuron, f_n is the activation function of the hidden neuron, w_{kj} is the weight that connects the jth neuron in the hidden layer and kth neuron in the output layer, w_{k0} is the bias for the kth neuron, f_o is the activation function for the output neuron and $y_k(t)$ is the predicted kth output at time-step t. In this study the LM back propagation algorithm was chosen because of its efficiency in the performance and the reduced computational time in training and testing models (Adamowski and Chan, 2011).

3. Materials and methods

3.1. Climate data

For prediction of the drought index, we relied on a suite of 13 input predictor variables where 5 of them were the site-specific parameters (year, month, latitude, longitude and elevation) and 8 were the meteorological variables that were closely related to drought events (Table 2). Our choice of the meteorological inputs was decided by considering the primary climate drivers that describe attributes of rainfall variability and consequently, drought in eastern Australia. Recent works have also used similar inputs for the prediction of hydrological variables. For example, Abbot and Marohasy (2012, 2014) used a combination of large-scale climate indices with air and sea surface temperatures to show good capability of these variables for predicting rainfall in Queensland. For the purpose of our study we capitalized the high quality meteorological data from
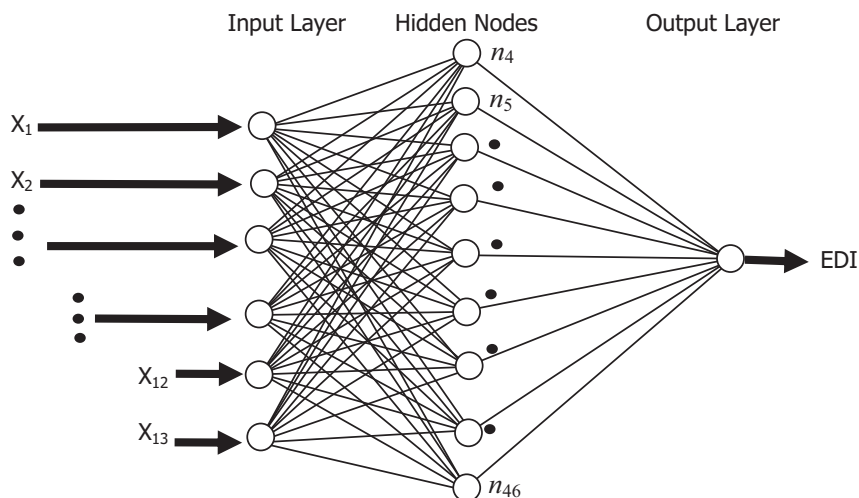

Fig. 2. The topological structure of the artificial neural network model used for comparison purpose. Input layer is denoted as X_1, X_2, ..., X_{13}, hidden nodes as n_1, n_2, ..., n_{46} and the output layer generates the predicted values of the Effective Drought Index (EDI).

the Australian Bureau of Meteorology (BOM) for 4 sites in eastern Australia in the period 1957–2011 (Fig. 3; Table 1). The inputs were monthly rainfall and air temperature (minimum, maximum and mean) values (Table 2) originally collated from on-site, historical observations (Jones et al., 2009; Lavery et al., 1997).

Quality controls tests were already done to adjust the raw dataset for inhomogeneities caused by station relocations and adverse exposures using objective statistical methods (Torok and Nicholls, 1996) including detection and removal of gross single-day errors and metadata inhomogeneities. Rather than making the homogeneity adjustments in the mean values, the daily records were adjusted for discontinuities at the 5, 10… 90, 95 percentile levels. Missing data were deduced missing data by generating artificial values based on cumulative distributions (Haylock and Nicholls, 2000). Consequently these records have since been used extensively for climate change studies (Alexander et al., 2006; Suppiah and Hennessy, 1998).

As the supplementary training dataset, we used the following large-scale climate mode indices as regression covariates for predicting the monthly EDI: the Southern Oscillation Index (SOI) and the Indian Ocean Dipole (IOD) data from Australian Bureau of Meteorology (Trenberth, 1984), the Pacific Decadal Oscillation (PDO) index from the Joint Institute of the Study of the Atmosphere and Ocean (JISAO) (Mantua et al., 1997; Zhang et al., 1997) and the Southern Annular Model (SAM) from the
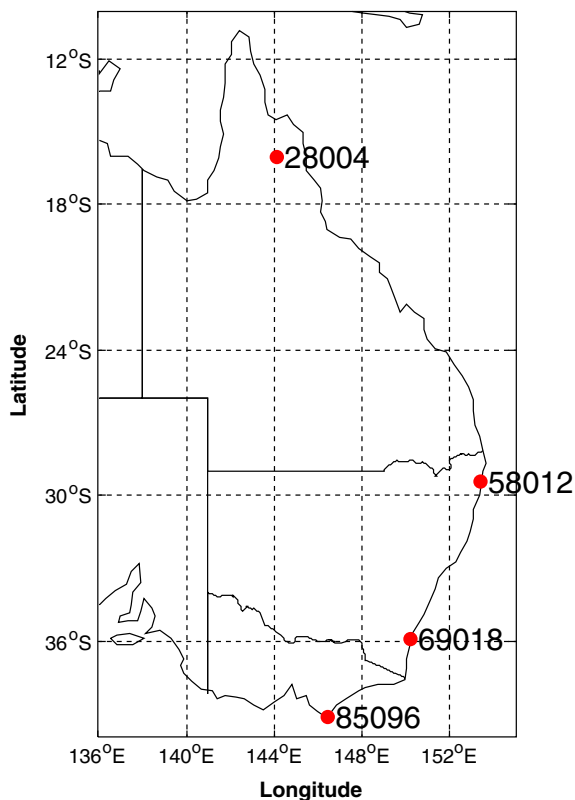
Table 1
The geographical description of the locations in this study.

Station name	BOM ID	Lon (°E)	Lat (°S)	Elevation (m)
Moruya Heads Pilot Station	69018	150.15	35.91	150.2
Palmerville	28004	144.08	16.00	144.1
Wilsons Promontory Lighthouse	85096	146.42	39.13	146.4
Yamba Pilot Station	58012	153.36	29.43	153.4

British Antarctic Survey database (Marshall, 2003) (Table 2). In its original form, the SOI is typically calculated using Troup's method from the differences in mean sea level pressure between Tahiti and Darwin. The PDO is created using the UKMO Historical SST dataset for 1900–81, Reynolds Optimally Interpolated SST (V1) (Morid et al., 2007) for January 1982–Dec 2001 and OI.v2 SST fields from January 2002 onwards and the IOD is a coupled ocean and atmosphere phenomenon in the equatorial Indian Ocean that together with the SOI impacts the Australian climate (Saji et al., 2005) and the SAM describes that the changing position of the westerly wind belt influences the strength and position of cold fronts and mid-latitude storm systems, and is an important driver of rainfall variability in southern Australia (Hendon et al., 2007).

The inclusion of large-scale climate drivers in our proposed ML algorithm for predicting future drought trends manifested in historical changes of these indices was consistent with previous approaches. The choice of air temperatures and climate indices followed the approach of related studies on rainfall prediction problems in Australia (e.g. Abbot and Marohasy, 2012, 2014; Mekanik et al., 2013) and the prediction of the Effective Drought Index elsewhere (e.g. Iran and South Africa, Masinde, 2013; Morid et al., 2007). Hence the best combination of most relevant meteorological variables and large-scale climate drivers was used to improve model performance in the present study.

3.2. Calculation of the Effective Drought Index

The daily rainfall dataset was assimilated by the FORTRAN code for calculating the EDI following the method of Byun and Wilhite (1999) (http://atmos.pknu.ac.kr/~intra3/). For reasons of space, a concise description of the primary mathematical

Table 2
The input parameters (meteorological variables) used in the ELM and ANN models.

Meteorological input parameters	
Monthly mean precipitation (mm)	PRCP
Monthly mean air temperature (°C)	Tmean
Monthly maximum air temperature (°C)	Tmax
Monthly minimum air temperature (°C)	Tmin
Large-scale climate mode indices	
Southern Oscillation Index	SOI
Pacific Decadal Oscillation	PDO
Southern Annular Mode	SAM
Indian Ocean Dipole	IOD

Fig. 3. Map of study locations with Bureau of Meteorology (BOM) station ID.

details of the EDI has been documented but the readers are referred to the original work of Byun and Wilhite (1999) for more details. In this study, the EDI was calculated from the daily effective precipitation (EP), which was the summed value with time-dependent reduction function following Eq. (1). EP was compared with the climatological mean EP (MEP) (Eq. (2)) and the results were normalized (Eq. (3)). If P_m was rainfall and N was the duration of the preceding period, then EP for current day was

$$EP = \sum_{N=1}^{D} \left(\left(\sum_{m=1}^{N} P_m \right) / N \right) = P_1 + \frac{(P_1 + P_2)}{2} + \frac{(P_1 + P_2 + P_3)}{3} + \cdots$$
$$+ \frac{(P_1 + P_2 + \cdots + P_{365})}{365} = P_1 \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{365} \right) \quad (8)$$
$$+ P_2 \left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{365} \right) + \cdots + P_{365} \left(\frac{1}{365} \right)$$

$$DEP = EP - MEP \quad (9)$$

$$EDI = \frac{DEP}{ST(DEP)}. \quad (10)$$

In their investigation, the work of Byun and Jung (1998) showed that the rainfall–runoff model exhibits a similar effect to Eq. (1). In our study the D was the duration of summation at 365, the most common precipitation cycle worldwide. Thus, for calculating the daily values of the EDI, the precipitation values accumulated over the previous 365 days with weighted values were used. The MEP and ST(DEP) represented the climatological mean of EP and standard deviation of DEP for each calendar day, respectively. The base period for calculating MEP and ST(DEP) needs to be at least 30 years (Kim et al., 2009) so we used 1971–2000, which is a common period used for analysis of Australian climate (Smith et al., 2008). During the computation process, if DEP continued to be negative for more than 2 days, the duration of summation D was increased by the number of days for which the DEP was negative which had no upper limit. Owing to this function, EDI was able to consider continuity of drought without limitations on the timescale. In this study the daily values of the EDI were converted to their corresponding monthly (observed) EDI values in order to match monthly meteorological input data and climate indices, and

then compared with the predictand values of the monthly EDI from the ELM and the ANN models.

3.3. Network architecture and optimum ELM and ANN model

All the model simulations using the ELM and the ANN algorithms were conducted in the MATLAB environment running under the Pentium 4, 2.93 GHz CPU system. Table 3 shows the parameters of the ELM and the ANN model. The 55 years of available data (1957–2011) were portioned into two parts, viz the training (1958–2008) and the testing (2009–2011) phases. The training dataset was used for designing both network models. After training the proposed network, a weight matrix was obtained and applied to the independent inputs in the "test" set. Then the final outcomes were compared with the observed (actual) values of the Effective Drought Index.

For designing the ELM model three layers were used to build the architecture (see Table 3) for predicting monthly EDI trained with data from 1957 to 2008, and tested over 2009 to 2011 with observed Effective Drought Index (EDI_o). The number of neurons was 13 (input) with eight as meteorological properties and 4 as the climate mode indices of interest (Table 2) accompanied by the year, month, station elevation, latitude and longitude as invariants for each given site. The ELM output layer had one neuron representing the predicted monthly Effective Drought Index (EDI_p) but in hidden layers a maximum of 150 neurons are tested (Table 1). A taxonomy of activation functions was tried one by one, which included sigmoid, log-sigmoid, hyperbolic-tangent sigmoid, radial bias, triangular bias, hyperbolic-tangent sigmoid and hard-limit. In each trial the numbers of nodes in hidden layer were increased gradually by an interval of five. Then, the nearly optimal node for ELM was selected as 50 with the hard-limit activation function and 13–50–1 neurons in the architecture of the ELM model (Fig. 1).

In addition, to show the potential of the proposed ELM model for predicting the Effective Drought Index a performance comparison in terms of the estimation capability was made between the ELM and the conventional feedforward ANN model run with the BP algorithm. Because it is a well-known universal estimator, the ANN model can rather be considered as standard benchmark (Gencoglu and Uyar, 2009; Nasr et al., 2002). ANN architecture used is illustrated in Fig. 2. In accordance with Maier and Dandy (2000), all data prior to its

Table 3
The ELM and ANN modeling frameworks employed in the present study. The acronyms are as follows: Activation functions [hardlim (hard-limit), tribas (triangular basis), radbas (radial basis), logsig (log-sigmoid), tansig (hyperbolic tangent sigmoid)] and ANN back propagation training algorithms [trainscg (scaled conjugate gradient), trainoss (one-step secant), trainbfg (BFGS quasi-Newton), trainbr (Bayesian regulation), trainlm (Levenberg–Marquardt)]. The functions in boldface were the optimum ones used in prediction of the EDI.

ELM		ANN	
Number of layers	3	Number of layers	3
Neurons	Input: 13	Neurons	Input: 13
	Hidden: 10…150		Hidden: 4…46
	Output: 1 (EDI)		Output: 1 (EDI)
Activation function	sig; sin; **hardlim**; tribas; radbas; logsig; tansig	Training algorithm	transcg; trainoss; trainbfg; trainbr; **trainlm**
		Hidden transfer function	**tansig**; logsig
		Output transfer function	**linear**; logsig
Learning rule	ELM for SLFNs	Learning rule	Back propagation
Model architecture	13–60–1	Model architecture	13–26–1

inclusion into the ANN model were scaled appropriately. In the present investigation, the input neurons were scaled in the range of [−1, 1] and a transfer function was implemented to explain the nonlinear relationship between input and output neurons (Chattopadhyay, 2007). For determining the optimum ANN model to be used in this work, the set of five back propagation training algorithms used were as follows: scaled conjugate gradient (trainscg), one-step secant (trainoss), BFGS quasi-Newton (trainbfg), Bayesian regulation (trainbr) and Levenberg–Marquardt (trainlm). Additionally, the set of two commonly used family of hidden transfer functions (hyperbolic-tangent sigmoid and log-sigmoid) and the three output functions (linear, hyperbolic-tangent sigmoid & log-sigmoid) were all tried on the testing datasets one at a time in order to seek the optimum model for the final experiments. Like in the case of the ELM model, the number of neurons in the hidden layer was varied gradually but this time over the range of 4 to 46 neurons. Consequently, the optimum training algorithm was the Levenberg–Marquardt, the hidden transfer function was the hyperbolic-tangent sigmoid and the output transfer function was linear with an ANN architecture of 13–26–1 (see Fig. 2).

The estimation capability of the monthly EDI from the two machine learning algorithms was statistically evaluated using the following score metrics or prediction error indicators: Root-Mean Square Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2) (Paulescu et al., 2011; Ulgen and Hepbasli, 2002) and the Willmott's Index of Agreement (d) (Acharya et al., 2013; Willmott, 1982) viz

$$MAE = \frac{1}{N} \sum_{i=1}^{n} \left| \left(EDI_{p_i} - EDI_{o_i} \right)_t \right| \qquad (11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{n} \left(EDI_{p_i} - EDI_{o_i} \right)_t^2} \qquad (12)$$

$$R^2 = \left(EDI_{p_i} - EDI_{o_i} \right)_t^2 \qquad (13)$$

$$d = 1 - \left[\frac{\sum_{i=1}^{N} \left(EDI_{pi} - EDI_{oi} \right)^2}{\sum_{i=1}^{N} \left(\left| EDI'_{pi} \right| - \left| EDI'_{oi} \right| \right)^2} \right], 0 \le d \le 1 \qquad (14)$$

where EDI_{p_i} and EDI_{oi} are the predicted and the observed Effective Drought Index in period t (testing slice) respectively, i

Table 4
Skill metric in terms of the overall Mean Absolute Error (MAE), Root-Mean Square Error (RMSE), Square-root of Coefficient of Determination (R^2) and Willmott's Index of Agreement (d).

Skill metric	ELM	ANN
RMSE	0.331	0.335
MAE	0.004	0.125
R^2	0.944	0.759
d	0.930	0.920

is month of tested data and N (=36) is length (number of samples in the test set) in period t (2009 to 2011).

The prediction metrics in Eqs. (11)–(14) represented the machine learning model's ability to simulate the monthly drought indices deduced from actual rainfall measurements. Table 4 shows the performance capability of the ELM and the ANN models used in this study. The optimum ELM model obtained by the hard-limit activation function with the learning rule "ELM for Single Layer Feedforward Neural Network" (SLFNs) yielded prediction metrics as follows: MAE (0.004), RMSE (0.331) and r (0.944). Contrarily, the baseline comparison by the ANN model yielded MAE (0.125), RMSE (0.335) and r (0.759). Finally the Willmott's Index of Agreement was slightly better (0.93) for the ELM compared with the ANN model (0.92).

4. Results and discussion

For all of the individual stations tested in this study a scatter plot of the monthly EDI_p versus the monthly EDI_o was prepared to assess the statistical correlation of the observed and the modeled values of the drought index (Fig. 4). A much larger degree of scatter is visible for the ANN model between the monthly EDI_p and EDI_o. This was especially pronounced for Palmerville, Wilsons Promontory and Yamba, as also evidenced by the greater deviation of linear regression model in each subplot. A performance assessment based on linear regression equation

$$EDI_p = mEDI_o + C \qquad (15)$$

is listed in Table 5, together with the best-fit slopes (m), square of correlation coefficient (r^2), maximum deviations of the predictions from linear model (maxDev) and the y-intercepts. The ELM model performed significantly better in predicting the monthly EDI than the ANN model as exemplified by the values of m closer to 1, relatively higher r^2 and lower maxDev for all stations. Importantly the slopes and correlation for the ELM model showed best performance for Moruya and Palmerville (1.01, 0.95 and 0.908, 0.975 respectively) and slightly worse for Wilsons (0.92, 0.816). Despite closer match of regression slope with a value of 1 for Moruya, the maxDev was significantly higher than that of Palmerville and Wilsons primarily due to an outlier in the EDI_p that deviated from the linear regression of the ELM model. Overall, a better prediction skill of the ELM model was exhibited for all sites in this investigation.

A comparison of the performance based on statistical analysis of errors of the predicted output with the observed values for each model for the four sites is shown in Table 6. The ELM outperformed the ANN model by all means of performance indicators between the predicted and observed Effective Drought Index. This was typified by the RMSE being dramatically lower by 50% (Moruya and Yamba), 70% (Palmerville) and 40% (Wilsons). Correspondingly the MAE was also lower for the ELM compared to the ANN model. If the comparison of coefficient of determination is made, then the ELM model exhibited the best prediction performance with r ≈ 0.904 to 0.953, relative to r ≈ 0.804 to 0.704 for the ANN model. Also, the location Palmerville yielded the best values of 0.156 (RMSE) and 0.988 (r) and Wilsons yields the worst, 0.357 (RMSE) and 0.904 (r). Only a slight improvement of the ELM
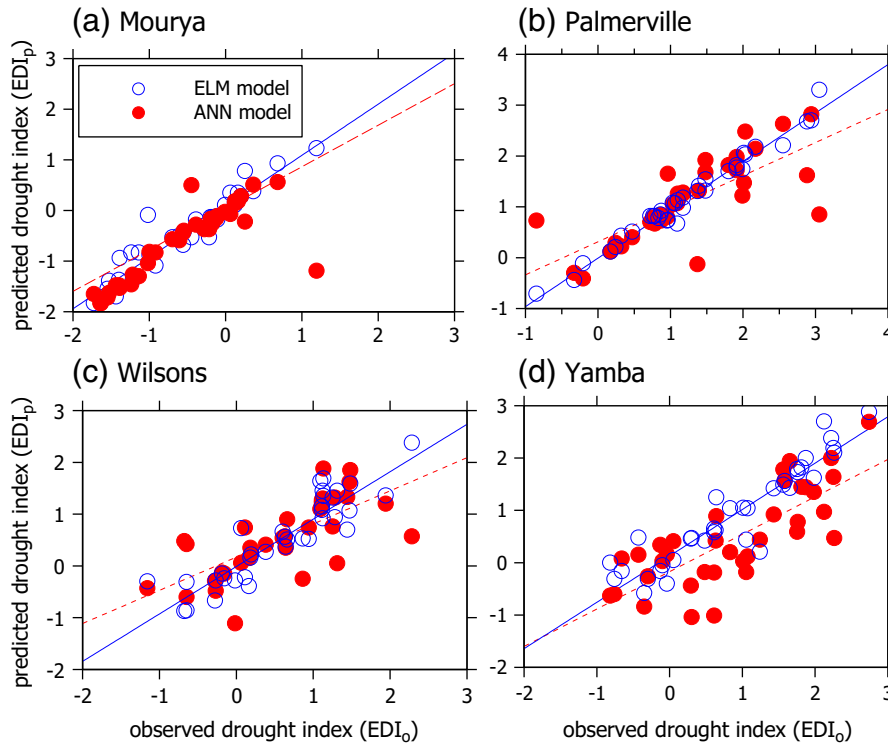
Fig. 4. Scatter plot of the predicted (EDI_p) and observed Effective Drought Index (EDI_o) based on extreme learning machine and artificial neural network for the testing period 2009–2011. A linear regression fit is also displayed for each model.

over the ANN model was shown by the index of agreement for Palmerville and Wilsons. However, in terms of the overall performance for all study locations the ANN model yielded twice the magnitude of RMSE and 22 times the MAE over the ELM model.

In studies of climate properties where relatively large datasets are often dealt with, the speed of the forecasting model is a paramount property for determining the fidelity of the model that has been used. Therefore, we compared the performance of the machine learning algorithms in terms of their learning and training speeds (Table 6). As shown the ELM model executed almost 32 times faster than the ANN model trained with BP. It was evident that the testing time for the ANN was 6 times longer than the testing time for the ELM model. Clearly the ELM obtained the fastest learning speed for estimating the monthly EDI, therefore was considered as having an edge over the conventional ANN model.

The spread of the predicted and the observed Effective Drought index has been illustrated in Fig. 5 using a Boxplot, which represents the degree of spread for the predicted data

using its respective quartiles. The lower end lies between the lower quartile Q1 (25th percentile) and upper quartile Q3 (75th percentile), with the second quartile Q2 (50th percentile) as the median of the data is represented by a vertical line. Two horizontal lines (known as whisker) are extended from the top and bottom of the box. The bottom whisker extends from Q1 to the smallest non-outlier in the dataset, whereas the other one goes from Q3 to the largest non-outlier. It is noticeable that the medians of predicted and observed EDI for ELM model were nearly identical for all four stations, whereas that for Yamba was significantly different for the ANN model. Additionally the discernment of the upper and lower quartiles of the ELM model with the observed data was the smallest. However, the lower quartile of ELM model for Wilsons appeared to be over-estimated and that for Moruya was under-estimated. The upper whisker of the ANN model for Moruya and Wilsons was significantly shorter than the ELM model. Moreover, the distribution of the upper and lower quartile and median for the ANN model for Yamba was shifted downwards. In all cases except Palmerville, spread of the ELM model was less with

Table 5
Model performances based on linear regression ($EDI_p = m\ EDI_o + C$) of the observed Effective Drought Index (EDI_o) with predicted (EDI_p) from 2009–2011 for each site. The square of correlation coefficient (r^2) and the maximum deviation (maxDev) of the predicted index from the modeled index are also shown.

Locations	ELM				ANN			
	m	r^2	maxDev	C	m	r^2	maxDev	C
Moruya Heads Pilot	1.01	0.908	0.842	0.079	0.80	0.065	1.970	−0.189
Palmerville	0.95	0.975	0.385	−0.001	0.65	0.563	1.460	0.312
Wilsons Promontory	0.92	0.816	0.755	−0.012	0.64	0.496	1.280	0.169
Yamba Pilot Station	0.87	0.867	1.040	0.130	0.72	0.601	1.290	−0.171

Table 6

Model performances based on score metrics for the ELM and ANN models for each site in terms of the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2) for the observed (EDI_o) and predicted Effective Drought Index (EDI_p). Willmott's Index of Agreement (d) is also shown.

| Locations	ELM				ANN			
	RMSE	MAE	R^2	d	RMSE	MAE	R^2	d
Moruya Heads Pilot Station	0.240	0.074	0.908	0.94	0.449	0.089	0.651	0.94
Palmerville	0.156	0.068	0.976	0.92	0.609	0.119	0.563	0.91
Wilsons Promontory Lighthouse	0.357	0.064	0.817	0.95	0.585	0.051	0.496	0.93
Yamba Pilot Station	0.354	0.028	0.868	0.90	0.766	0.429	0.600	0.90
Overall (all-station)	0.277	0.008	0.892	0.93	0.602	0.172	0.578	0.92
Learning speeds								
Overall (all-station) training time (s)				1.078				6.560
Overall (all-station) testing time (s)				0.002				0.063

respect to observed Effective Drought Index, and therefore, indicated the relatively better performance of this model.

The functional relationship between the observed and the predicted Effective Drought Index using machine learning algorithms is illustrated in Fig. 6 for test samples starting from January 2009 to December 2011. A significantly smaller degree of scatter in EDI for the ELM model was exhibited if comparisons are made with the observed values of the index. In fact, for the ANN model in Yamba, the departure from observed datasets became larger between the 24 and 36 month timescale in addition to similar discrepancy between 0 and 12 months. In general, the departure of the predicted EDI from observed values was significantly larger for the ANN model, which was verified emphatically by the prediction error yield per month (Fig. 7). This base error for each month's prediction model was especially pronounced for the ANN model when the data for Palmerville, Wilsons and Yamba are studied and reflected consistently the results in Table 6 with the lower MAE and RMSE by the ELM over the ANN model.

A conclusive argument can be made by a baseline study of the drought properties for the monthly predictions of the EDI

for the period January 2009 to December 2011. Based on the predicted values of the EDI, the severity and the duration of drought at the four sites were deduced following the approach of Kim et al. (2011) where a drought month was identified when the monthly value of the EDI was negative (i.e. rainfall conditions were lower rainfall than the normal period). The severity of the drought was then the accumulated value of the negative EDI and the duration as the sum of all months when this drought status was sustained.
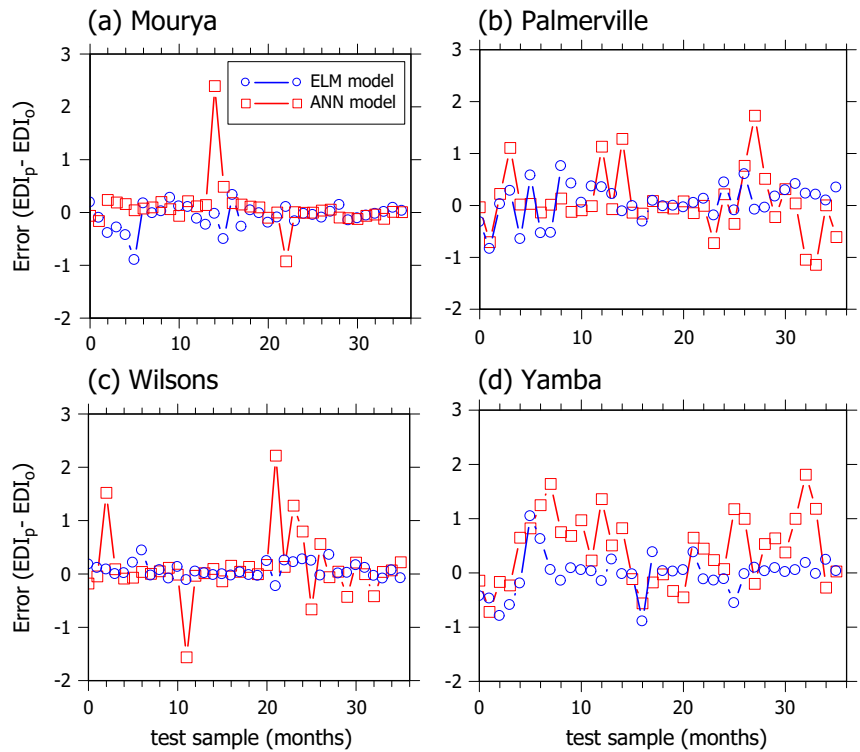
Fig. 8 presents the performance of the ELM and the ANN models for quantifying drought severity based on the accumulated negative EDI and drought duration in tested period from 2009–2011. Here each bar represented the difference in the drought property between the ELM and observed, and the ANN and observed dataset. Apart from the site (Wilsons) where the predictions of the drought properties by the ANN model were statistically better, those for the Moruya, Palmerville and Yamba were dramatically better by the ELM model. In fact the best performance of the ELM model was obtained for the first two stations with the absolute differences in the prediction skill of drought severity as −6.05% and −5.12%, respectively.
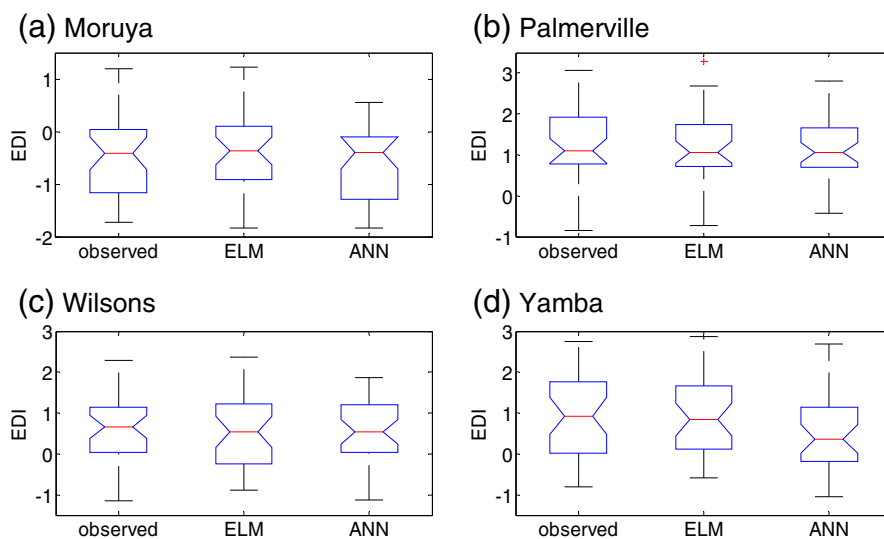

Fig. 5. Boxplot of Effective Drought Index (EDI) from observed data and predictions from machine learning models viz., extreme learning machine (ELM) and artificial neural network (ANN) for testing period 2009–2011.
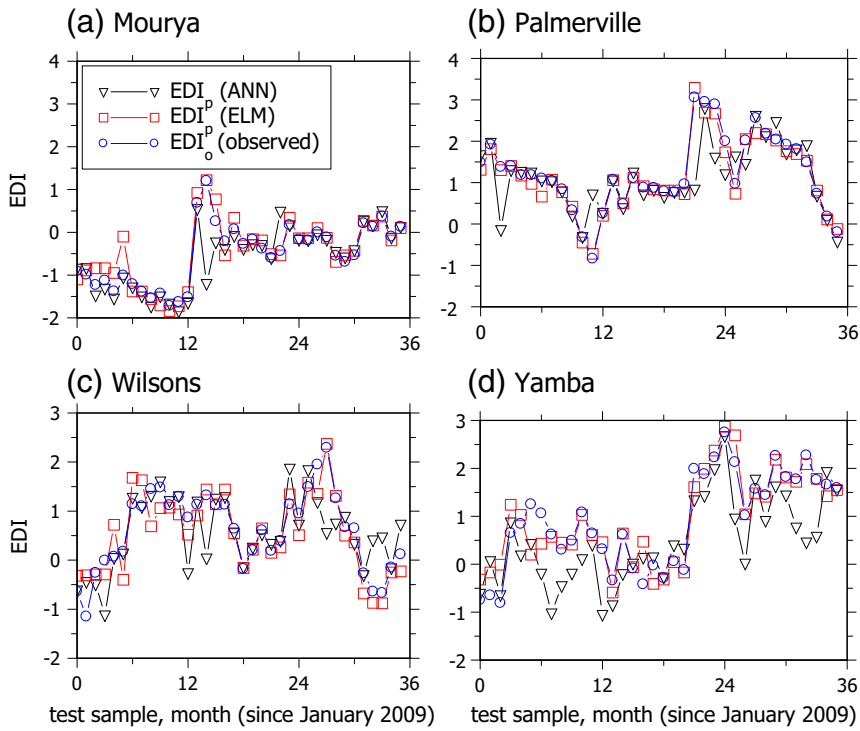
Fig. 6. Time-series of the Effective Drought Index (EDI) from observed (EDI$_o$) and predicted data (EDI$_p$) using machine learning models viz., extreme learning machine (ELM) and artificial neural network (ANN) for testing period 2009–2011.

Fig. 7. Prediction error yield for the modeled (EDI$_p$) and observed Effective Drought Index (EDI$_o$) using extreme learning machine and artificial neural network for the testing period 2009–2011.
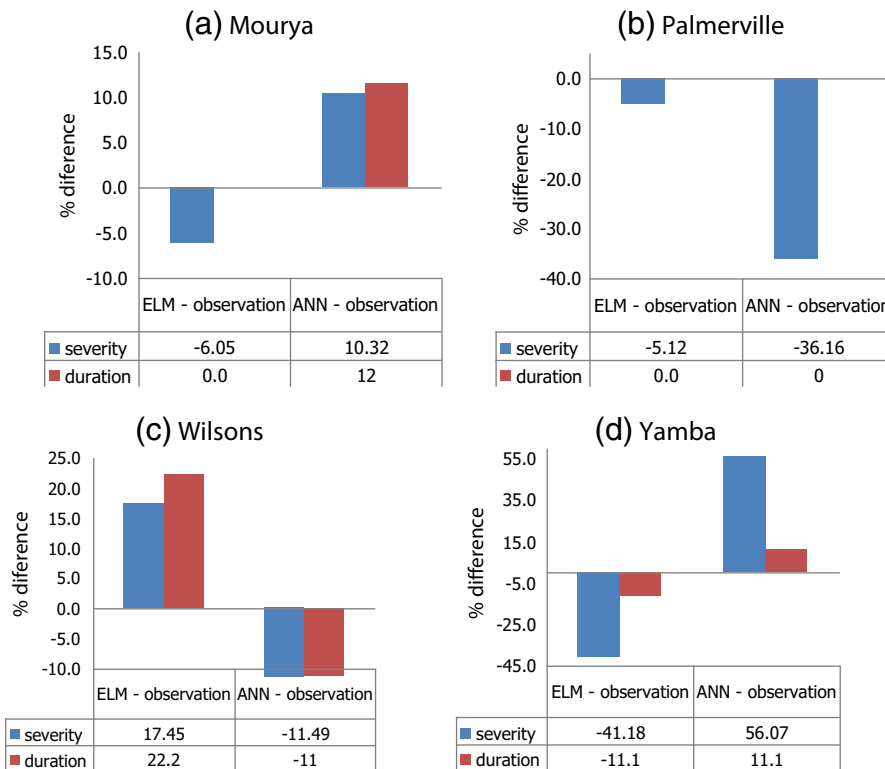
Fig. 8. Performance of the ELM and ANN models for quantifying drought severity based on accumulated negative EDI and the drought duration following the approach of Kim et al. (2011) for the testing period 2009–2011. The table shows the actual percentage difference between predicted and observed values of the EDI.

While the ELM model underestimated the severity of drought by about 41.18% for Yamba, the ANN model tended to overestimate it by almost 56%. Interestingly the prediction skill of drought duration was significantly better for Moruya by the ELM model and was also in parity for those for the stations, Palmerville and Yamba. Overall the ELM model showed better prediction skill of drought properties than the ANN model and therefore appears to be a more useful algorithm for modeling of drought and assessment of climate risk.

5. Summary and conclusion

The prediction of drought events is a topic of significant interest for the management of water resources agriculture, facilities maintenance, control and infrastructural (floodgates, airports, motor-roads, etc.). Our study attempted to determine an effective data-driven machine learning model for predicting the monthly Effective Drought Index (Byun and Wilhite, 1999) using meteorological datasets from eastern Australia for the first time. A new machine learning model (ELM), which was an improved version of the SLFN architecture, was investigated and the prediction skills were compared with the conventional ANN model with back propagation algorithm. The monthly variables used as inputs to both models were the mean rainfall and mean, maximum and minimum temperatures and the climate mode indices (Southern Oscillation Index, Pacific Decadal Oscillation, Indian Ocean Dipole and Southern Annular Mode). The models were trained using data for the period 1957–2008 and tested over 2009–2011.

The basis of this study was to investigate the feasibility of the ELM model for predicting nonlinear relationship between input variables and the monthly values of the Effective Drought Index. In results we assessed the performance of the primary ELM model compared with the traditional ANN model using prediction skill metrics like spread distribution, Mean Absolute Error, Root-Mean Square Error, Coefficient of Determination and the Willmott's Index of Agreement. The prediction skills were examined for the monthly Effective Drought Index tested between 2009 and 2011 and compared with ANN.

In summary, the following findings have been enumerated briefly:

1. After trials and errors on different activation functions to be used, the optimum ELM model was designed by using the hard-limit function with the 13–60–1 neuron architecture. For comparison with the ANN model, the Levenberg–Marquardt training algorithm with back propagation was used with the transfer function as the hyperbolic-tangent sigmoid and linear output function and the ANN model architecture was 13–26–1.
2. To show the applicability of the ELM model, a performance comparison in terms of the estimation capability and learning speed was made between the ELM and the conventional ANN models. The learning speed of the ELM was 32 times and testing speed was 6 times faster than the ANN model. Thus, the ELM model reduces significantly the computational time over the ANN model as required in many meteorological and hydrology areas with typically large datasets.

3. In terms of the match between the predicted and the observed values of the EDI, there was a significant improvement by the ELM model for capturing the variability of monthly EDI over the conventional ANN model. The monthly variability of the EDI by the ELM model was very close to the observed values calculated directly from rainfall data using Eq. (1). This was verified by the ELM model yielding significantly smaller magnitudes of Mean Absolute Error, Root-Mean Square Error, Coefficient of Determination and the Willmott's Index of Agreement for individual and overall stations.

4. The spread of the predicted monthly EDI by the ELM model was substantially smaller for majority of the sites and the median is closer to the observed values of the drought index. This indicated the prediction yield of the ELM model returned individual monthly EDI values closer to the observed, and therefore was superior to the ANN model.

5. To examine the performance of the ELM over the ANN model, the severity and duration of drought based on predicted monthly EDI were assessed. The performance of the ELM model in more accurate representation of predicted properties of drought for at least three out of the four stations studied.

Based on our results, the ELM model is seen to enhance the prediction skill of the monthly Effective Drought Index over the ANN model, and therefore, can overcome deficiencies in prediction when applied to climate analysis that typically requires thousands of training data points and time efficacy of the modeling framework. Moreover the success of using monthly rainfall and temperature and the climate mode indices as inputs for the preferred ELM model is a promising approach for the future or evaluation of the likelihood of future drought and great assistance in the design of hydrologic systems in engineering problems and water resource management.

Acknowledgments

References

Abbot, J., Marohasy, J., 2012. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. Adv. Atmos. Sci. 29 (4), 717–730.

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.

Acharya, N., Shrivastava, N.A., Panigrahi, B., Mohanty, U., 2013. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. Clim. Dyn. 1–8.

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. 407 (1), 28–40.

Alexander, L., et al., 2006. Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res. Atmos. (1984–2012) 111 (D5).

Asklany, S.A., Elhelow, K., Youssef, I., Abd El-wahab, M., 2011. Rainfall events prediction using rule-based fuzzy inference system. Atmos. Res. 101 (1), 228–236.

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.

Byun, H., Jung, J., 1998. Quantified diagnosis of flood possibility by using effective precipitation index. J. Korean Water Res. Assoc. 31 (6), 657–665.

Byun, H.-R., Wilhite, D.A., 1999. Objective quantification of drought severity and duration. J. Clim. 12 (9), 2747–2756.

Byun, H.-R., et al., 2008. Study on the periodicities of droughts in Korea. Asia-Pac. J. Atmos. Sci. 44 (4), 417–441.

Chattopadhyay, S., 2007. Feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India. Acta Geophys. 55 (3), 369–382.

Day, K.A., Ahrens, D.G., Peacock, A., 2010. Seasonal Pacific Ocean Temperature Analysis-1 (SPOTA-1) as at November 1, 2010. Report Issued by the Queensland Climate Change Centre of Excellence, Queensland Government, Brisbane, Australia (2 pp.).

Deo, R.C., 2011. Links between native forest and climate in Australia. Weather 66 (3), 64–69.

Deo, R.C., et al., 2009. Impact of historical land cover change on daily indices of climate extremes including droughts in eastern Australia. Geophys. Res. Lett. 36 (8).

Dijk, A.I., et al., 2013. The Millennium Drought in southeast Australia (2001–2009): natural and human causes and implications for water resources, ecosystems, economy, and society. Water Resour. Res. 49 (2), 1040–1057.

Dogan, S., Berktay, A., Singh, V.P., 2012. Comparison of multi-monthly rainfall-based drought severity indices, with application to semi-arid Konya closed basin, Turkey. J. Hydrol. 470, 255–268.

Douglass, D.H., Blackman, E.G., Knox, R.S., 2004. Temperature response of Earth to the annual solar irradiance cycle. Phys. Lett. A 323 (3), 315–322.

Fawcett, R., Stone, R., 2010. A comparison of two seasonal rainfall forecasting systems for Australia. Aust. Meteorol. Oceanogr. J. 60 (1), 15.

Gencoglu, M.T., Uyar, M., 2009. Prediction of flashover voltage of insulators using least squares support vector machines. Expert Syst. Appl. 36 (7), 10789–10798.

Govindaraju, R.S., 2000. Artificial neural networks in hydrology. II: hydrologic applications. J. Hydrol. Eng. 5 (2), 124–137.

Haylock, M., Nicholls, N., 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. Int. J. Climatol. 20 (13), 1533–1541.

Hendon, H.H., Thompson, D.W., Wheeler, M.C., 2007. Australian rainfall and surface temperature variations associated with the Southern Hemisphere annual mode. J. Clim. 20 (11), 2452–2467.

Huang, G.-B., 2003. Learning capability and storage capacity of two-hidden-layer feedforward networks. IEEE Trans. Neural. Netw. 14 (2), 274–281.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: theory and applications. Neurocomputing 70 (1), 489–501.

Hudson, D., Alves, O., Hendon, H.H., Marshall, A.G., 2011. Bridging the gap between weather and seasonal forecasting: intraseasonal forecasting for Australia. Q. J. R. Meteorol. Soc. 137 (656), 673–689.

Hurst, D., 2011. Three-quarters of Queensland a disaster zone. Brisbane Times, Fairfax MediaAvailable online at http://www.brisbanetimes.com.au/environment/weather/threequarters-of-queensland-a-disaster-zone-20110111-19mf8.html.

Inquiry, Q.F.C.o, 2011. Queensland Foods Commission of Inquiry: Interim Report, (1180 pp., Brisbane, Queensland, Australia).

Irving, D.B., Whetton, P., Moise, A.F., 2012. Climate projections for Australia: a first glance at CMIP5. Aust. Meteorol. Oceanogr. J. 62 (4), 211–225.

Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. Aust. Meteorol. Oceanogr. J. 58 (4), 233.

Kaufmann, R.K., Stern, D.I., 2002. Cointegration analysis of hemispheric temperature relations. J. Geophys. Res. Atmos. (1984–2012) 107 (D2), 8–10 (ACL 8-1-ACL).

Kaufmann, R.K., Kauppi, H., Mann, M.L., Stock, J.H., 2011. Reconciling anthropogenic climate change with observed temperature 1998–2008. Proc. Natl. Acad. Sci. 108 (29), 11790–11793.

Kim, D.-W., Byun, H.-R., 2009. Future pattern of Asian drought under global warming scenario. Theor. Appl. Climatol. 98 (1–2), 137–150.

Kim, T.-W., Valdés, J.B., 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. J. Hydrol. Eng. 8 (6), 319–328.

Kim, D.-W., Byun, H.-R., Choi, K.-S., 2009. Evaluation, modification, and application of the Effective Drought Index to 200-Year drought climatology of Seoul, Korea. J. Hydrol. 378 (1), 1–12.

Kim, D.-W., Byun, H.-R., Choi, K.-S., Oh, S.-B., 2011. A spatiotemporal analysis of historical droughts in Korea. J. Appl. Meteorol. Climatol. 50 (9), 1895–1912.

Kuligowski, R.J., Barros, A.P., 1998. Experiments in short-term precipitation forecasting using artificial neural networks. Mon. Weather Rev. 126 (2), 470–482.

Lavery, B., Joung, G., Nicholls, N., 1997. An extended high-quality historical rainfall dataset for Australia. Aust. Meteorol. Mag. 46 (1), 27–38.

Leu, S.-S., Adi, T.J.W., 2011. Probabilistic prediction of tunnel geology using a Hybrid Neural-HMM. Eng. Appl. Artif. Intell. 24 (4), 658–665.

Luk, K., Ball, J., Sharma, A., 2000. A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. J. Hydrol. 227 (1), 56–65.

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model Softw. 15 (1), 101–124.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C., 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc. 78 (6), 1069–1079.

Marshall, G.J., 2003. Trends in the Southern Annular Mode from observations and reanalyses. J. Clim. 16 (24), 4134–4143.

Masinde, M., 2013. Artificial neural networks models for predicting effective drought index: factoring effects of rainfall variability. Mitig. Adapt. Strateg. Glob. Chang. 1–24.

McAlpine, C., et al., 2007. Modeling the impact of historical land cover change on Australia's regional climate. Geophys. Res. Lett. 34 (22).

McAlpine, C., et al., 2009. A continent under stress: interactions, feedbacks and risks associated with impact of modified land cover on Australia's climate. Glob. Chang. Biol. 15 (9), 2206–2223.

Mekanik, F., Imteaz, M., Gato-Trinidad, S., Elmahdi, A., 2013. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. J. Hydrol. 503, 11–21.

Morid, S., Smakhtin, V., Moghaddasi, M., 2006. Comparison of seven meteorological indices for drought monitoring in Iran. Int. J. Climatol. 26 (7), 971–985.

Morid, S., Smakhtin, V., Bagherzadeh, K., 2007. Drought forecasting using artificial neural networks and time series of drought indices. Int. J. Climatol. 27 (15), 2103–2111.

Nasr, G., Badr, E., Younes, M., 2002. Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches. Int. J. Energy Res. 26 (1), 67–78.

Nasseri, M., Asghari, K., Abedini, M., 2008. Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. Expert Syst. Appl. 35 (3), 1415–1421.

Nastos, P., Paliatsos, A., Koukouletsos, K., Larissi, I., Moustris, K., 2014. Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece. Atmos. Res. 144, 141–150.

Ortiz-García, E., Salcedo-Sanz, S., Casanova-Mateo, C., Paniagua-Tineo, A., Portilla-Figueras, J., 2012. Accurate local very short-term temperature prediction based on synoptic situation Support Vector Regression banks. Atmos. Res. 107, 1–8.

Ortiz-García, E., Salcedo-Sanz, S., Casanova-Mateo, C., 2014. Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. Atmos. Res. 139, 128–136.

Pandey, R., Dash, B., Mishra, S., Singh, R., 2008. Study of indices for drought characterization in KBK districts in Orissa (India). Hydrol. Process. 22 (12), 1895–1907.

Paulescu, M., Tulcan-Paulescu, E., Stefu, N., 2011. A temperature-based model for global solar irradiance and its application to estimate daily irradiation values. Int. J. Energy Res. 35 (6), 520–529.

Rajesh, R., Prakash, J.S., 2011. Extreme learning machines—a review and state-of-the-art. Int. J. Wisdom Based Comput. 1 (1), 35–49.

Şahin, M., 2012. Modelling of air temperature using remote sensing and artificial neural network in Turkey. Adv. Space Res. 50 (7), 973–985.

Şahin, M., Kaya, Y., Uyar, M., 2013. Comparison of ANN and MLR models for estimating solar radiation in Turkey using NOAA/AVHRR data. Adv. Space Res. 51 (5), 891–904.

Şahin, M., Kaya, Y., Uyar, M., Yıldırım, S., 2014. Application of extreme learning machine for estimating solar radiation from satellite data. Int. J. Energy Res. 38 (2), 205–212.

Saji, N., Ambrizzi, T., Ferraz, S., 2005. Indian Ocean Dipole mode events and austral surface air temperature anomalies. Dyn. Atmos. Oceans 39 (1), 87–101.

Sánchez-Monedero, J., Salcedo-Sanz, S., Gutiérrez, P., Casanova-Mateo, C., Hervás-Martínez, C., 2014. Simultaneous modelling of rainfall occurrence and amount using a hierarchical nominal–ordinal support vector classifier. Eng. Appl. Artif. Intell. 34, 199–207.

Seqwater, 2011. January 2011 Flood Event: Report on the Operation of Somerset Dam and Wivenhoe Dam REVIEW OF HYDROLOGICAL ISSUES Queensland Government, Melbourne, Australia (77 pp.).

Shukla, R.P., Tripathi, K.C., Pandey, A.C., Das, I., 2011. Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach. Atmos. Res. 102 (1), 99–109.

Smith, I., Wilson, L., Suppiah, R., 2008. Characteristics of the northern Australian rainy season. J. Clim. 21 (17), 4298–4311.

Sözen, A., Ali Akçayol, M., 2004. Modelling (using artificial neural-networks) the performance parameters of a solar-driven ejector-absorption cycle. Appl. Energy 79 (3), 309–325.

Stone, D.A., Allen, M., 2005. Attribution of global surface warming without dynamical models. Geophys. Res. Lett. 32 (18).

Suppiah, R., Hennessy, K.J., 1998. Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990. Int. J. Climatol. 18 (10), 1141–1164.

Tamura, S.i, Tateishi, M., 1997. Capabilities of a four-layered feedforward neural network: four layers versus three. IEEE Trans. Neural. Netw. 8 (2), 251–255.

Torok, S., Nicholls, N., 1996. A historical annual temperature dataset. Aust. Meteorol. Mag. 45 (4).

Trenberth, K.E., 1984. Signal versus noise in the Southern Oscillation. Mon. Weather Rev. 112 (2), 326–332.

Ulgen, K., Hepbasli, A., 2002. Comparison of solar radiation correlations for Izmir, Turkey. Int. J. Energy Res. 26 (5), 413–430.

van den Honert, R.C., McAneney, J., 2011. The 2011 Brisbane floods: causes, impacts and implications. Water 3 (4), 1149–1173.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bull. Am. Meteorol. Soc. 63 (11), 1309–1313.

Wu, C., Chau, K., 2010. Data-driven models for monthly streamflow time series prediction. Eng. Appl. Artif. Intell. 23 (8), 1350–1367.

Zhang, Y., Wallace, J.M., Battisti, D.S., 1997. ENSO-like interdecadal variability: 1900–93. J. Clim. 10 (5), 1004–1020.

Zhao, M., Hendon, H.H., 2009. Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. Q. J. R. Meteorol. Soc. 135, 337–352.