# Measuring and restructuring the risk in forecasting drought classes: an application of weighted Markov chain based model for standardised precipitation evapotranspiration index (SPEI) at one-month time scale

*By* ZULFIQAR ALI[1,8], IJAZ HUSSAIN[1]*, AMNA NAZEER[2], MUHAMMAD FAISAL[3], MUHAMMAD ISMAIL[4], SADIA QAMAR[5], MARCO GRZEGORCZYK[6], FAISAL MAQBOOL ZAHID[7], and GUANGHENG NI[8], [1]*Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan;* [2]*Department of Mathematics, COMSATS University Islamabad, Islamabad, Pakistan;* [3]*Bradford Institute of Health Research, University of Bradford, UK;* [4]*Department of Statistics, COMSATS University Islamabad, Lahore Campus, Pakistan;* [5]*Department of Statistics, University of Sargodha, Sargodha, Pakistan;* [6]*Johann Bernoulli Institute, Groningen University, Groningen, 9747AG, The Netherlands;* [7]*Department of Statistics, Government College University Faisalabad, Faisalabad, Pakistan;* [8]*State Key Laboratory of Hydro-Science and Engineering and Dept. of Hydraulic Engineering, Tsinghua University, Beijing 100084, China*

## ABSTRACT

Drought monitoring and forecasting play a vital role in making drought mitigation policies. In previous research, several drought monitoring tools based on the probabilistic models have been developed for precise and accurate inferences of drought severity and its effects. However, the risk of inaccurate determination of drought classes always exists in probabilistic models. The aim of this paper is to reconnaissance the advantage of the weighted Markov chain (WMC) model to accommodate the erroneous drought classes in the monthly classifications of drought. It was assumed that to increase the precision in drought prediction, the role of standardised self-correlation coefficients as weight may incorporate to establish and restructure the accurate probabilities of risk for incoming expected drought classes in the WMC framework. Consequently, the current research is based on the experimental findings of seventeen meteorological stations located in the Northern Areas of Pakistan. In this study, the standardised precipitation evapotranspiration index (SPEI) at a 1-month time scale based drought monitoring approach is applied to quantify the historical classification of drought conditions. The exploratory analysis shows that the proportion of each drought class varies from zone to zone. However, a high proportion of near-normal drought classes has been observed in all the stations. For the prediction of future drought classes, transition probability matrices have been computed using R statistical software. Our findings show that the probability of occurrences of near-normal is very high. Overall, the results associated with this study show that the WMC method for drought forecasting is sufficiently flexible to incorporate the change of drought conditions; it may change both the transition probability matrix and the autocorrelation structure.

*Keywords: Drought, Markov chain, standardised precipitation evapotranspiration index (SPEI), autocorrelation*

## 1. Introduction

Drought, the highest-ranked natural hazard, is the primary source of severe destructive effects on the planet (White, 1974). Its sustained consequences lead to the sterilisation of agricultural land and the initiation of diseases. Factors associated with a higher risk of drought are the long duration of rainfall, high rate of evapotranspiration, low relative humidity, high temperature, and high wind

*Correspnding author. email:ijaz@qau.edu.pk

**1**

speed (Edwards et al., 2009). Moreover, many other environmental and ecological factors are also responsible for the recurrent occurrences of drought hazard. However, drought intensity, duration, and severity may vary from region to region. In recent decades, almost all the developing countries are facing water shortage due to continued expansion in agriculture, industrial, and energy sectors. Consequently, a perpetual increase in the difference between water demand and renewable freshwater resources will lead to significant social and economic issues.

However, to overcome the severe effect of the frequent occurrence of drought, forecasting plays a significant role in drought mitigation policies. In previous research, several forecasting and assessment tools for the characterisation of drought regions and the quantification of drought risk have been established. Several studies have been conducted in the field of hydrology and climatology for the assessment and modelling of drought classes for different regions in the world. Drought indices are one of the most used tools for the assessment and quantification of drought risk. A range of different drought indices involving different climatic parameters has been developed to detect dry and wet categories of a region for a specified time. Details on the list of drought indices corresponding with their variable requirement can be found in Svoboda et al. (2016).

Besides drought monitoring and assessment tools, several probabilistic and deterministic forecasting models have been developed and used to predict and forecast drought classes for various climatological regions. In recent decades, the rapid increase in the development of theories associated with the stochastic process is found for modelling many real-life uncertain phenomena. An example includes, stock market and exchange rate fluctuations; signals such as speech; audio and video; medical data such as a patient's EKG, EEG, blood pressure or temperature; and random movements, such as Brownian motion or random walks. Among several other stochastic models, the theory of Markov chains is a promising approach to dynamic model activities that have a stochastic factor (Lange, 2010). To model uncertain events, especially in the field of engineering (Takahashi et al., 2007), economics (Lee and Chen, 2006), and physics (Crommelin and Vanden-Eijnden, 2006), Markov chain models play a significant role in the prediction and forecasting of the probabilities associated with such events. Markov chain models can be useful for forecasting future drought classes due to their multifaceted nature to enumerate uncertainties connected with hydro-meteorological variables causing droughts.

Sen (1990) derived stationary second-order Markov chains for finite sample lengths for exact probability distribution functions (PDF) for three representative annual flow series from various regions of the globe. Lohani and Loganathan (1997) used non-homogeneous Markov chain models to characterise the random behaviour of droughts using the Palmer Drought Severity Index. Paulo and Pereira (2007) used Markov Chain models on the drought classes determined by the SPI drought index to characterise the stochasticity of drought and predict three months of drought class. Bacanli et al. (2009) used the SPI drought index based on an Adaptive Neuro-Fuzzy Inference System (ANFIS) forecasting drought. Ali et al. (2017b) used a multilayer perceptron model and SPEI drought index for drought forecasting in the Northern Area and KPK.

However, it is difficult to adjust the transition probability matrix and the precision of the forecast that is affected by objective factors. To overcome this problem, in many applications Weighted Markov Chain (WMC) method have been employed in several disciplines, including hydrology and environmental sciences (Benoit, 2005; Le-Tian, 2005; De-di and Chen, 2006; Peng et al., 2010; Kaliakatsos-Papakostas et al., 2011; Zhou et al., 2011; Chen and Yang, 2012; Gong et al., 2014; Gui and Shao, 2017). Chen and Yang (2012) proposed a drought prediction model for SPI with different time scales under the weighted Markov chain framework for Anhui Province of Huaihe River in China. Outcomes associated with this model show that the weighted Markov Chain method is a useful approach for drought prediction, and it can be helpful for decision-making in regional drought management. In recent years, we have proposed a new weighting scheme of WMC for ordinal data (see Ali et al., 2018). In this article, the Standardised Precipitation Temperature Index (SPTI) (Ali et al., 2017) has been used to evaluate the proposed scheme.

However, analysing droughts by using a single variable is not enough to distinguish different regions because drought hazards relate to multiple variables. A comprehensive analysis of the characterisation of drought classes is required that make a joint analysis of rainfall, runoff, and soil moisture conditions (Vicente-Serrano et al., 2005). The SPI is the simplest and commonly used drought index that is based on accumulated precipitation (McKee et al., 1993). Vicente-Serrano et al. (2010) developed a new multi-scaler drought index, the standardised precipitation evapotranspiration index (SPEI). The methodological structure of SPEI is like SPI. However, in SPEI, instead of only precipitation, the mean monthly temperature is also used in the estimation of drought classes.

The objective of this research is to handle all the difficulties in formulating a mathematical model for forecasting SPEI drought index at a one-month time scale under the WMC framework in various regions of Pakistan. We use autocorrelations from the historical series of SPEI
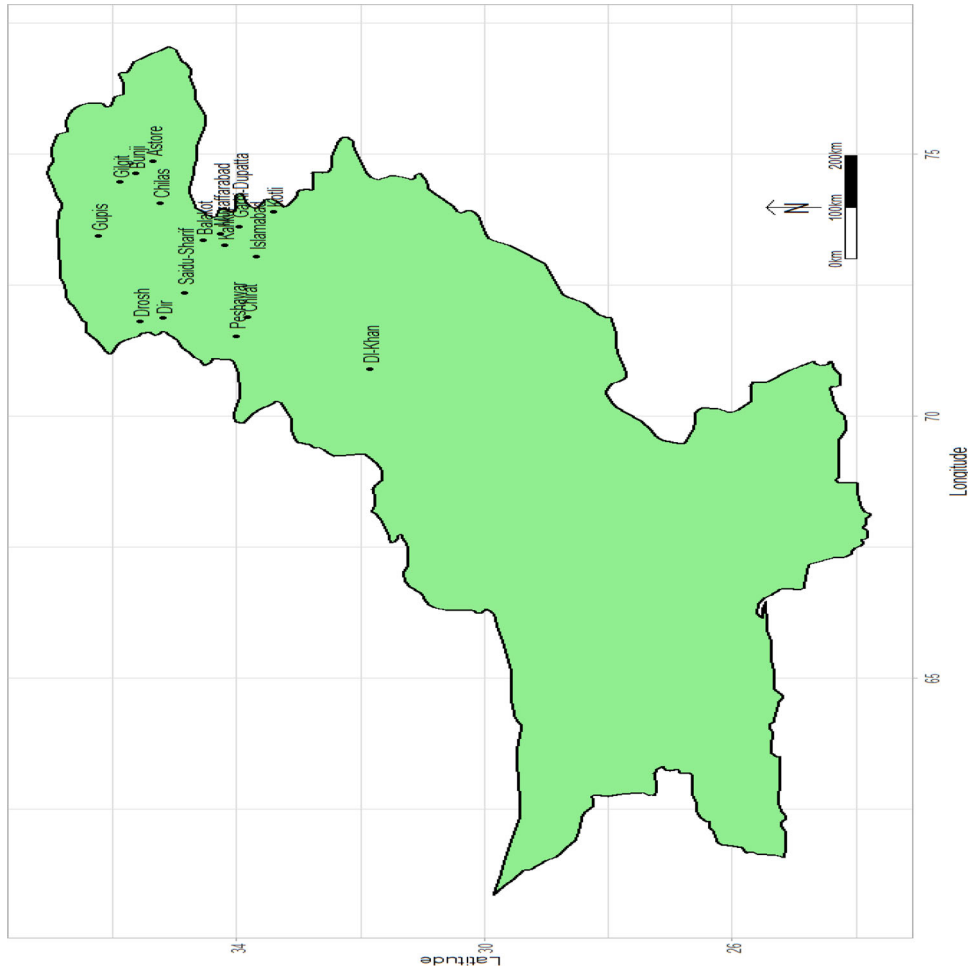
*Table 1.* Summary statistics of the selected stations.

| Station | Jan | | | Feb | | | Mar | | | Apr | | | May | | | Jun | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp |
| Astore | 37.55 | 3.04 | −7.25 | 45.84 | 4.49 | −5.04 | 70.32 | 9.21 | −0.73 | 77.9 | 15.46 | 4.05 | 61.15 | 20.31 | 7.85 | 25.95 | 24.52 | 11.28 |
| Balakot | 93.14 | 14.31 | 2.16 | 149.63 | 15.73 | 4.21 | 173.24 | 20.03 | 8.04 | 122.97 | 25.79 | 12.66 | 77.52 | 31.28 | 17.4 | 100.7 | 34.8 | 20.61 |
| Kotli | 73.74 | 17.98 | 4.47 | 101.73 | 19.75 | 7.17 | 131.42 | 24.46 | 11.53 | 74.76 | 30.46 | 16.59 | 46.63 | 35.6 | 21.16 | 88.89 | 37.59 | 23.82 |
| Chirat | 39.73 | 9.39 | 3.33 | 76.42 | 10.97 | 4.21 | 98.79 | 15.73 | 7.75 | 62.7 | 22.53 | 13.23 | 27.27 | 28.69 | 17.81 | 26.69 | 32.28 | 20.88 |
| Chillas | 10.74 | 12.27 | 1.66 | 16.17 | 14.63 | 4 | 28.17 | 19.97 | 8.64 | 35.63 | 26.15 | 14.01 | 30.97 | 31.48 | 18.92 | 10.99 | 36.85 | 23.88 |
| Islamabad | 55.57 | 17.83 | 3.21 | 90.43 | 19.64 | 6.07 | 101.03 | 24.34 | 10.66 | 59.13 | 30.46 | 15.83 | 36.04 | 36.05 | 20.68 | 75.79 | 38.45 | 23.89 |
| Gupis | 7.6 | 4.75 | −5.81 | 13.11 | 7.09 | −3.34 | 16.82 | 12.63 | 1.7 | 43.09 | 19.04 | 6.93 | 27.08 | 23.9 | 10.85 | 18.74 | 28.7 | 14.67 |
| Peshawar | 37.55 | 18.47 | 4.35 | 61.32 | 19.99 | 6.95 | 82.91 | 24.3 | 11.72 | 59.95 | 30.58 | 16.89 | 22.55 | 36.73 | 21.98 | 19.81 | 40.11 | 25.57 |
| Saidu Sharif | 57.52 | 9.95 | −4.83 | 106.69 | 10.8 | −2.58 | 158.73 | 15.06 | 2.32 | 124.9 | 20.98 | 7.8 | 88.61 | 26.57 | 12.37 | 63.43 | 31 | 16.63 |
| Muzaffarabad | 95.54 | 16.39 | 3.11 | 144.23 | 18.04 | 5.43 | 165.95 | 22.61 | 9.51 | 106.18 | 28.59 | 14.05 | 77.79 | 33.85 | 18.3 | 120.76 | 37.14 | 21.51 |
| Bunji | 6.13 | 10.19 | 0.02 | 9.79 | 12.92 | 2.89 | 12.19 | 18.56 | 7.73 | 25.65 | 24.41 | 12.07 | 26.36 | 28.74 | 15.21 | 11.02 | 33.04 | 18.8 |
| DIK | 12 | 19.76 | 4.39 | 22.14 | 22.31 | 7.58 | 38.46 | 26.91 | 12.95 | 24.88 | 33.76 | 18.53 | 11.83 | 39.4 | 23.42 | 22.91 | 41.14 | 26 |
| Drosh | 48.03 | 9.5 | −0.18 | 77.71 | 11.11 | 0.92 | 110.9 | 16.47 | 4.77 | 91.13 | 23.66 | 10.25 | 57.86 | 29.89 | 15.14 | 20.5 | 35.58 | 19.97 |
| Garhi–Dupatta | 104.54 | 14.99 | 2.64 | 150.7 | 16.45 | 4.3 | 171.09 | 20.94 | 7.91 | 117.67 | 26.99 | 12.22 | 77.9 | 32.51 | 16.68 | 110.35 | 36.24 | 20.04 |
| Dir | 113.67 | 11.86 | −2.43 | 179.41 | 12.61 | −0.82 | 246.33 | 16.8 | 3 | 166.53 | 22.97 | 7.47 | 88 | 28.51 | 11.45 | 54.79 | 32.48 | 15.31 |
| Gilgit | 3.66 | 10.18 | −2.68 | 7.15 | 13.02 | 0.54 | 11.31 | 18.65 | 5.46 | 24.84 | 24.66 | 9.16 | 25.45 | 29.4 | 11.91 | 10.91 | 34.1 | 14.81 |
| Kakul | 66.12 | 12.96 | 0.62 | 107.97 | 13.95 | 2.23 | 145.55 | 18.09 | 6.07 | 113.96 | 23.67 | 10.42 | 70.76 | 28.92 | 14.61 | 97.42 | 32.1 | 18.24 |

| Station | Jul | | | Aug | | | Sep | | | Oct | | | Nov | | | Dec | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp | Rainfall | Min temp | Max temp |
| Astore | 25.7 | 27.22 | 14.49 | 26.55 | 26.73 | 14.37 | 27.19 | 23.28 | 10.12 | 21.04 | 17.86 | 4.61 | 19.48 | 11.64 | −0.22 | 29.79 | 5.82 | −4.38 |
| Balakot | 332.96 | 32.37 | 21.4 | 269.38 | 31.29 | 20.71 | 119.69 | 30.6 | 17.11 | 53.85 | 27.28 | 11.51 | 43.66 | 21.97 | 6.46 | 60.82 | 16.8 | 3.29 |
| Kotli | 273.2 | 33.75 | 23.8 | 261.43 | 32.68 | 23.26 | 101.61 | 32.49 | 20.87 | 33.79 | 30.3 | 15.77 | 22.75 | 25.29 | 9.68 | 44.54 | 20.28 | 5.23 |
| Chirat | 95.96 | 29.14 | 19.64 | 104.77 | 26.88 | 18.77 | 35.55 | 26.29 | 17.87 | 19.05 | 22.6 | 14.49 | 15.2 | 17.01 | 10.07 | 23.41 | 12.22 | 5.97 |
| Chillas | 12.65 | 39.11 | 27.06 | 17.69 | 38.14 | 26.12 | 10.47 | 34.65 | 22.15 | 6.7 | 28.51 | 15 | 6.18 | 20.87 | 7.67 | 10.61 | 14.06 | 3.02 |
| Islamabad | 310.68 | 35.03 | 24.52 | 336.87 | 33.63 | 23.77 | 126.93 | 33.47 | 21.19 | 36.08 | 30.86 | 14.84 | 15.39 | 25.8 | 8.58 | 30.97 | 20.38 | 4.27 |
| Gupis | 16.18 | 31.8 | 17.51 | 26.4 | 30.44 | 16.47 | 13.55 | 26.53 | 12.27 | 8.82 | 20.51 | 6.38 | 2.8 | 14.2 | 0.9 | 4.77 | 6.99 | −3.56 |
| Peshawar | 60.64 | 37.44 | 26.65 | 81.44 | 35.79 | 25.82 | 28.61 | 34.95 | 22.95 | 22.21 | 31.28 | 16.64 | 14.21 | 25.85 | 10.18 | 17.08 | 20.72 | 5.37 |
| Saidu Sharif | 107.4 | 30.41 | 17.81 | 132.44 | 28.63 | 16.54 | 74.24 | 26.68 | 13.66 | 51.88 | 22.66 | 8.39 | 31.49 | 17.24 | 3.05 | 30.06 | 12.39 | −1.65 |
| Muzaffarabad | 337.29 | 34.83 | 22.64 | 218.79 | 33.91 | 22.52 | 108.34 | 33.24 | 19.36 | 43.13 | 30 | 13.5 | 36.94 | 23.93 | 7.8 | 62.62 | 18.2 | 4.09 |
| Bunji | 17.56 | 35.3 | 22.49 | 22.87 | 34.66 | 21.41 | 15.26 | 31.33 | 16.28 | 6.91 | 25.54 | 10.31 | 3.26 | 18.69 | 4.54 | 5.51 | 12.26 | 1.09 |
| DIK | 71.61 | 38.43 | 26.43 | 69.45 | 37.25 | 25.82 | 32.36 | 36.43 | 23.53 | 8.28 | 33.36 | 17.56 | 3.07 | 27.91 | 10.69 | 7.04 | 22.43 | 5.67 |
| Drosh | 21.49 | 36.97 | 22.54 | 19.63 | 35.85 | 21.48 | 19.66 | 32.96 | 17.4 | 31.23 | 27.04 | 11.27 | 30.93 | 19.24 | 5.94 | 34.07 | 12.46 | 1.94 |
| Garhi–Dupatta | 250.6 | 33.97 | 21.96 | 211.34 | 32.84 | 21.72 | 104.83 | 32.09 | 18.12 | 45.04 | 28.33 | 12.16 | 41.76 | 22.44 | 6.6 | 68.96 | 17.02 | 3.48 |
| Dir | 150.96 | 31.81 | 18.92 | 149.85 | 30.72 | 18.16 | 81.05 | 29.4 | 13.55 | 69.59 | 25.39 | 7.27 | 62.07 | 20.03 | 2.39 | 69.57 | 14.7 | −0.98 |
| Gilgit | 14.95 | 36.18 | 18 | 17.21 | 34.95 | 17.35 | 11.25 | 31.65 | 12.38 | 6.35 | 25.95 | 6.38 | 3.31 | 18.83 | 0.66 | 5.12 | 12.26 | −2.22 |
| Kakul | 257.65 | 29.56 | 19.45 | 240.79 | 28.49 | 18.87 | 105.49 | 27.85 | 15.86 | 54.86 | 25.16 | 10.51 | 30.21 | 20.59 | 5.82 | 51.03 | 15.81 | 2.3 |

*Figure 1.*    Study area: meteorological stations of Northern Area and KPK (Pakistan).

drought index with a one-month time scale as a weight in first-order Markov chain transition probability matrices to forecast the next incidence of drought categories for seventeen meteorological stations located in the Northern Areas and KPK (Pakistan). The drought has become a recurrent phenomenon in the country. In the recent decade, due to severe drought hazards, the economic system of the country was severely disturbed. In recent decades, several authors had been working to explore the geographical and hydrological importance of this region. Awan (2002) and Archer and Fowler (2004) explored and inferred different climatic variables in terms of regression, spatial correlation, and temporal variation. Ahmad et al. (2012) evaluated the significance of these mountainous areas that have substantial potential in hydro-power production and water resources. Ali et al. (2017a) have compared the performance of SPTI with SPI and SPEI using time series data on precipitation and temperature of these stations.

In this research considered seventeen meteorological stations having different climatology and estimated

historical time series data on SPEI drought index for a one-month time scale. Time series data on precipitation and temperature are used to estimate SPEI values for these stations.

The organisation of this paper is as follows. A brief description of the study area, the estimation method of SPEI, and the mathematical formulation of WMC method is presented in Section 1. Section 2 consists of temporal representations of SPEI drought classes and results associated with WMC based forecasting. Section 3 is based on the results of this paper. In contrast, discussion and conclusion have been presented in Sections 4 and 5, respectively.

## 2. Materials and methods

### 2.1. Study area and data

We consider seventeen meteorological stations located in different climatic regions of Northern Areas of Pakistan.

These stations have high variability in rainfall throughout the season. Concerning the climatological statistics (see Table 1), the behaviour of all stations varies from zone to zone.

In each season, some of the stations are continuing to bear extremely vulnerable drought conditions. In the current research, twelve meteorological stations exhibiting cold and humid climate and five meteorological stations having mild cold and arid climate are included to check the efficiency of SPEI drought index from the global warming perspective. Primary data on monthly total rainfall, mean minimum temperature, and mean maximum temperature of each station for the period 1976–2017 were collected from the Karachi data processing center through Pakistan Meteorological Department (PMD), Islamabad. Figure 1 shows the geographical distribution of the study area located in different climatological zones.

## 2.2. Standardised precipitation evapotranspiration Index-SPEI

There are several procedures to report drought severity using a multiscalar drought index (Ali et al., 2017a). McKee et al. (1993) developed an SPI drought index, which is based on long-term precipitation records to quantify precipitation scarcity for different time scales. One of the significant advantage of using the SPI index is that it can be used to monitor drought for various time scales. Vicente-Serrano et al. (2010) developed the standardised precipitation evapotranspiration index (SPEI): a multiscalar drought index. In SPEI, the water balance model based on the difference between precipitation and potential evapotranspiration (PET) is used with a similar estimation procedure of SPI. One significant advantage of SPEI over SPI is to include the effect of evaporation in rainfall data to characterise the regions under study. The method employed in SPEI drought index for the determination of drought classes is a little bit different from SPI. SPEI uses both temperature and precipitation for the characterisation of drought classes, whereas, in the SPI drought index, only monthly cumulative precipitation data is used. In SPEI drought index, monthly time series data on drought classes is obtained by standardising the distribution of the difference between precipitation and Potential Evapotranspiration (PET). Several methods are available in the literature for the estimation of PET. The choice of the method for the estimation of PET depends on the availability of data and the sensitivity of the PET values. In their original paper, Vicente-Serrano et al. (2010) had chosen the simplest approach to calculate PET by using Thornthwait (TH) equation (Thornthwaite 1948).

However, this method underestimates PET values at arid regions in cold climatic regions, whereas overestimates PET values in humid regions (Ali et al., 2017a). The Hargreaves method for the estimation of PET overcomes this issue. However, it needs additional climatic parameters (i.e. mean minimum and mean maximum temperature etc.) (Hargreaves, 1994). In this research, we use the Hargreaves method for the estimation of PET based on minimum mean and minimum-maximum temperature using SPEI package in R. Further detailed procedures on computation, temporal behaviour, and normality test on SPEI time series data can be found in Ali et al. (2019).

## 2.3. Markov chain

A stochastic process, or random process $\{Z = Z(t), t \ \varepsilon \ T\}$ is a set of random variables indexed by time. That is, for all $t$ in the index set $T$, $Z(t)$ is a random variable. If the index set $T$ is a countable set, we call $Z(t)$ a discrete-time stochastic process. If $T$ is a straight-set, we call $Z(t)$ a continuous-time stochastic process. All possible values that $Z(t)$ can assume, are called its state space (Chiang, 1968). A Markov chain is a stochastic process having the property that the value of the process at time $t$, $Z_t$ depends only on its previous value $Z_{t-1}$ at time $t-1$, but does not depend on the past values of the process (Haan, 1977).

*2.3.1. Drought conditions as a Markov chain process.* A discrete Markov chain is a random process that describes a sequence of events from a set of finite possible states. In contrast, the current event depends only on the preceding event. It has been commonly used to model uncertain events in various disciplines. Each discrete Markov chain is characterised by a transition probability matrix that represents the probability of transition from one state to another. Shatanawi et al. (2013) reported that the exact prediction of Drought Index (DI) values is impossible with ARIMA models. However, early warning of drought can be detected from monthly Markov transition probabilities. Therefore, in drought modelling context, time-series data SPTI drought classes for a single station can be considered as a sequence of ordinal drought classes.

Consequently, a historical series of drought classification states for a specified station can be embodied as a discrete Markov chain process. Here, we assume that any single class of drought in time series of SPEI depends on its previous class and then proceed to the construction of the transition probability matrix. It is just statistical compliance that allows us to consider each drought class as a first-order Markov chain. However, one can also use a second-order Markov chain, where if each class is assumed to depend on its previous two classes.

*Table 2.* Classification criteria SPEI.

| SPEI values | Class |
|---|---|
| $\geq 2$ | Extremely wet (EV) |
| 1.50 to 1.99 | Severe wet (SW) |
| 1.00 to 1.49 | Moderate wet (MW) |
| .99 to $-.99$ | Near normal (NN) |
| $-1$ to $-1.49$ | Moderate drought (MD) |
| $-1.5$ to 1.99 | Severe drought (SD) |
| $\leq -2$ | Extreme drought (ED) |

*2.3.2. Configuration of weighted Markov chain (WMC) for observed drought classes.* In this scenario, time-series data on drought classes determined by SPEI drought index can be assumed as a series of correlated random variables. Various empirical studies show that self-correlation coefficients in the historical data of drought classes for all the study regions have significant importance for the prediction of future drought classes. This confirms that previous drought classes (on a monthly or yearly basis) can be considered in advance to predict the present drought class. So, in our case, the basic idea behind using WMC is that weighted averages can be made according to the incidence behaviour in the past month. Hence, the probability of present or next drought classes can be inferred and predicted in advance by appropriate configuration of weights to each drought class in WMC framework.

The fundamental steps involved in the proposed method for prediction of drought classes using SPEI drought index under the WMC model are given below.

*2.3.2.1. Classification quantitative values of SPEI for transition probability matrix.* Let $D_1$, $D_2$, …, $D_n$ be the time series of drought classes. Where $D$ can assume the nominal value droughts classes depending on the classification criteria of SPEI drought index (see Table 2). As in each drought class, we can find the transition probability matrix in the following form:

Drought Classes

$$
\begin{array}{cccc}
 & c_1 & c_2 & \ldots & c_n \\
\begin{array}{c} c_1 \\ c_2 \\ \vdots \\ c_n \end{array} &
\left( \begin{array}{cccc}
p_{11} & p_{12} & \cdots & p_{1n} \\
p_{21} & p_{22} & \cdots & p_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
p_{n1} & p_{n2} & \cdots & p_{nn}
\end{array} \right)
\end{array}
$$

where, $c_1$, $c_2$, …, $c_n$ are the drought classes, i.e. too wet, very wet, …, extremely dry.

In this step, we classify SPEI drought index estimated with a one-month time scale according to the classification criteria provided in Table 2. Descriptions on the classification states for drought classes determined by SPEI drought index were shown in Table 2.

Transition probability matrices for each station are computed using the *Markovchain* package to predict the future drought classes (Spedicato et al., 2016) package of R.

*2.3.2.2. Construction of the transition probability matrix.* Let $Y_{ij}^{(t)}$ the number of transitions from the state $S_i$ to state $S_j$ through $t$ steps in time series length of drought classes $X_k$ calculated from a one-month time scale. Here, the transition probabilities for various time steps and various drought classes are computed using the following equation.

$$
P_{ij}^{(t)} = \frac{Y_{ij}^{(t)}}{Y_i}, \ i, \ j = 1, 2, \ldots\ldots\ m \tag{6}
$$

where $Y_i$ is the total number of individual drought classes, $t$ represents the order of Markov chain. Further, the transition probability matrix for various drought classes can be obtained as;

$$
P^{(t)} = \begin{bmatrix}
p_{11}^{(t)} & p_{12}^{(t)} & . & . & p_{1m}^{(t)} \\
p_{21}^{(t)} & p_{22}^{(t)} & . & . & p_{2m}^{(t)} \\
. & . & . & . & . \\
. & . & . & . & . \\
p_{m1}^{(t)} & p_{m2}^{(t)} & . & . & p_{mm}^{(t)}
\end{bmatrix} \tag{7}
$$

Transition probability matrices for Astore and Balakot stations from the time series data on drought classes determined by SPEI drought index are shown in Table 3.

*2.3.2.3. Computation weights using autocorrelations.* The weights $(w_i)$ for the weighted Markov chain model can be computed by standardising the self-correlation coefficient $(r_i)$. The formula for weights $(w_i)$ and self-correlation coefficient $(r_i)$ are provided in Eqs. (8) and (9), respectively.

$$
w = \frac{|r_p|}{\sum_{p=1}^{m} |r_p|} \tag{8}
$$

$$
r = \frac{\sum_{p=1}^{n-p} (Z_{(p)} - \overline{Z})(Z_{(p+1)} - \overline{Z})}{\sum_{p=1}^{n} (Z_{(p)} - \overline{Z})^2} \tag{9}
$$

*2.3.2.4. Prediction of the probability of the occurrence of a drought class using weighted Markov chain.* In this step, we assume the occurrence of drought classification states in the very last month as an initial drought class $\Psi_i$ and combine it with the row vectors of their corresponding transition probability matrix. Here, we assumed that the inaccurate drought class is probable due to the estimation of SPEI time series data. Hence, weighting the drought classes according to Eq. (10), we arrive at the prediction probabilities $P_i$ for the next drought classes.

*Table 3.* One step transition probability matrix.

| | | Astore | | | | | | | Balakot | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MD | MW | NN | SD | SW | | | MD | MW | NN | SD | SW |
| | MD | 0.136 | 0.045 | 0.636 | 0.045 | 0.136 | | MD | 0.125 | 0.063 | 0.813 | 0 | 0 |
| | MW | 0.068 | 0.169 | 0.661 | 0.085 | 0.017 | | MW | 0.025 | 0.1 | 0.8 | 0.025 | 0.05 |
| $P_{ij}^1$ | NN | 0.038 | 0.118 | 0.698 | 0.121 | 0.025 | $P_{ij}^1$ | NN | 0.028 | 0.072 | 0.747 | 0.113 | 0.041 |
| | SD | 0 | 0 | 0.453 | 0.538 | 0.009 | | SD | 0.032 | 0.113 | 0.5 | 0.29 | 0.065 |
| | SW | 0 | 0.067 | 0.933 | 0 | 0 | | SW | 0.045 | 0.091 | 0.727 | 0.091 | 0.045 |

*Table 4.* Percentage frequencies of drought class.

| | Percentages of drought classes | | | | |
|---|---|---|---|---|---|
| Stations | MD | MW | NN | SD | SW |
| Astore | 0.09 | 0.03 | 0.71 | 0.15 | 0.02 |
| Balakot | 0.09 | 0.04 | 0.75 | 0.09 | 0.03 |
| Kotli | 0.12 | 0.03 | 0.64 | 0.17 | 0.03 |
| Chirat | 0.07 | 0.03 | 0.69 | 0.15 | 0.05 |
| Chilas | 0.08 | 0.02 | 0.67 | 0.17 | 0.06 |
| Islamabad | 0.11 | 0.02 | 0.66 | 0.15 | 0.07 |
| Gupis | 0.11 | 0.02 | 0.67 | 0.16 | 0.04 |
| Peshawar | 0.1 | 0.02 | 0.67 | 0.18 | 0.04 |
| Saidu-Shareef | 0.09 | 0.04 | 0.69 | 0.15 | 0.03 |
| Muzafarabad | 0.08 | 0.04 | 0.71 | 0.14 | 0.02 |
| Bunji | 0.1 | 0.03 | 0.68 | 0.17 | 0.04 |
| DIKhan | 0.09 | 0.03 | 0.66 | 0.16 | 0.05 |
| Drosh | 0.08 | 0.03 | 0.7 | 0.14 | 0.04 |
| Gari-Dupata | 0.07 | 0.04 | 0.7 | 0.16 | 0.03 |
| Dir | 0.07 | 0.03 | 0.71 | 0.16 | 0.03 |
| Gilgit | 0.08 | 0.03 | 0.67 | 0.16 | 0.06 |
| Kakul | 0.12 | 0.02 | 0.63 | 0.17 | 0.05 |

*Table 5.* Standardised weights of transition from one state to another stat.

| | Lags | | | | |
|---|---|---|---|---|---|
| Stations | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ |
| Astore | 0.249 | 0.226 | 0.201 | 0.175 | 0.149 |
| Balakot | 0.230 | 0.215 | 0.200 | 0.186 | 0.170 |
| Kotli | 0.247 | 0.222 | 0.199 | 0.177 | 0.154 |
| Chirat | 0.236 | 0.217 | 0.200 | 0.183 | 0.164 |
| Chilas | 0.237 | 0.221 | 0.202 | 0.181 | 0.159 |
| Islamabad | 0.236 | 0.216 | 0.198 | 0.182 | 0.167 |
| Gupis | 0.240 | 0.222 | 0.201 | 0.179 | 0.159 |
| Peshawar | 0.229 | 0.213 | 0.199 | 0.186 | 0.173 |
| Saidu-Shareef | 0.240 | 0.221 | 0.201 | 0.179 | 0.159 |
| Muzafarabad | 0.248 | 0.225 | 0.199 | 0.175 | 0.153 |
| Bunji | 0.229 | 0.216 | 0.201 | 0.185 | 0.170 |
| DIKhan | 0.215 | 0.207 | 0.200 | 0.193 | 0.185 |
| Drosh | 0.228 | 0.213 | 0.200 | 0.187 | 0.172 |
| Gari-Dupata | 0.233 | 0.215 | 0.198 | 0.184 | 0.170 |
| Dir | 0.226 | 0.213 | 0.200 | 0.187 | 0.173 |
| Gilgit | 0.257 | 0.230 | 0.200 | 0.171 | 0.143 |
| Kakul | 0.239 | 0.216 | 0.196 | 0.183 | 0.167 |

$$P_i = \sum_{i=1}^{m} w_i P_{ij}^{(t)} \qquad (10)$$

In the above equation $P_i$ are the weighted probabilities for incoming drought events. However, the predicted drought class is the drought class having max $\{P_i, i \ \varepsilon \ S\}$ under the weighted Markov Chain method.

## 3. Results

For exploratory analysis, the percentage of occurrences of various drought classes in the historical time series data on SPEI-1 with a one-month time scale of the selected study regions are presented in Table 4. A high proportion of near-normal drought classes were found in most of the stations. Moreover, the test of equality of proportions, as suggested in Marden (1996), is performed using easyanova Arnhold (2014) package in R.

Additionally, multiple comparison tests show that the near-normal drought class has a significantly high proportion as compared to other drought classes. However,

the severely dry drought class has a significant difference from the severely wet class. Besides, the proportion of the moderate dry class is significantly greater than for moderate wet.

Here, transition probability matrices are used for further prediction of future drought classes using equation 10. Weights from step one to step five associated with each station are shown in Table 5. These weights are calculated by normalising autocorrelation coefficients (see Eq. (7)). Table 6 shows the predicted probabilities for each drought class in December 2017 for Astore station. For the near-normal drought class, the predicted probability of this class using the weighted Markov chain is 0.712, which is very high concerning other drought classes. The actual drought condition is also near to normal, which indicates that the prediction is correct, and the chance of errors is very low. On the same line, the interpretation can be made for the rest of the months and stations as well. One month ahead prediction probabilities of the drought classes for the rest of the stations are

*Table 6.* Predicted probabilities for the month of December 2017 at Astore.

| Initial month | Drought State of December 2017 | Step | Weights | MD | MW | NN | SD | SW |
|---|---|---|---|---|---|---|---|---|
| July 2017 | NN | $P^{(5)}$ | 0.149 | 0.023 | 0.084 | 0.783 | 0.094 | 0.016 |
| August 2017 | NN | $P^{(4)}$ | 0.175 | 0.020 | 0.080 | 0.804 | 0.083 | 0.014 |
| September 2017 | NN | $P^{(3)}$ | 0.201 | 0.015 | 0.073 | 0.832 | 0.068 | 0.012 |
| October 2017 | SD | $P^{(2)}$ | 0.226 | 0.001 | 0.006 | 0.254 | 0.739 | 0.001 |
| November 2017 | NN | $P^{(1)}$ | 0.249 | 0.005 | 0.040 | 0.923 | 0.028 | 0.005 |
| December 2017 | $P_i$ | | | 0.011 | 0.052 | 0.712 | 0.216 | 0.009 |

*Table 7.* One month ahead forecast probabilities of various drought classes.

| Stations | Initial states | Drought classes probabilities | | | | |
|---|---|---|---|---|---|---|
| | | MD | MW | NN | SD | SW |
| Astore | NN | 0.011 | 0.052 | 0.712 | 0.216 | 0.009 |
| Balakot | NN | 0.009 | 0.033 | 0.895 | 0.055 | 0.008 |
| Kotli | NN | 0.018 | 0.075 | 0.754 | 0.139 | 0.013 |
| Chirat | NN | 0.021 | 0.115 | 0.755 | 0.064 | 0.045 |
| Chilas | MD | 0.083 | 0.111 | 0.684 | 0.074 | 0.048 |
| Islamabad | NN | 0.011 | 0.018 | 0.633 | 0.335 | 0.003 |
| Gupis | NN | 0.014 | 0.042 | 0.870 | 0.063 | 0.010 |
| Peshawar | MW | 0.213 | 0.292 | 0.225 | 0.011 | 0.259 |
| Saidu-Shareef | NN | 0.007 | 0.035 | 0.573 | 0.376 | 0.009 |
| Muzafarabad | MD | 0.013 | 0.060 | 0.804 | 0.102 | 0.021 |
| Bunji | NN | 0.005 | 0.026 | 0.427 | 0.533 | 0.009 |
| DIKhan | NN | 0.025 | 0.056 | 0.617 | 0.217 | 0.084 |
| Drosh | MD | 0.150 | 0.321 | 0.395 | 0.022 | 0.113 |
| Gari-Dupata | NN | 0.010 | 0.067 | 0.821 | 0.088 | 0.013 |
| Dir | NN | 0.009 | 0.062 | 0.841 | 0.073 | 0.016 |
| Gilgit | MW | 0.010 | 0.043 | 0.798 | 0.130 | 0.019 |
| Kakul | NN | 0.013 | 0.056 | 0.662 | 0.259 | 0.010 |

shown in Table 7. Except for Drosh and Peshawar, the probabilities of near-normal drought categories are very high in all the stations. It is due to the high frequency of near-normal drought classes in the qualitative vectors of drought classes in all stations.

## 4. Discussion

Results associated with this test show that there is a significant difference in the proportion of each drought class. These results are consistent from the ecological and climatological point of view since, in these areas, most of the stations bear longer Moonsoon and precipitation periods. In most stations, near-normal drought class found a high probability of occurrences. However, the moderate wet drought class has a low probability as compared to a severely dry drought class. However, the near-normal drought class has a high probability in most of the study region. However, predicted probabilities for each drought class in Peshawar, Bunji, and Drosh have somehow different behaviour. In Peshawar, drought classes have almost equal probabilities except for the Severely dry class. Bunji shows a 0.533 probability of occurrence for the severely dry class. In Drosh station, almost equivalent probabilities are found for near normal and moderate wet.

## 5. Conclusion

Prediction and forecasting play a vital role, especially in early warning situations. Consequently, accurate and precise techniques of drought forecasting may reduce their severe effect by making effective drought mitigation policies. In this article, the SPEI-1 drought index being a more comprehensive drought monitoring procedure is used to classify historical monthly drought profile for seventeen meteorological stations of Pakistan. To predict the future drought classes, the standardised self-correlation coefficient in the time series data of SPEI-1 index is used as weights in WMC method (see Table 5). Outcomes associated with the WMC method for prediction of drought classes show that this forecasting method is flexible enough to incorporate change of drought condition, just by changing the transition probability matrix and the autocorrelation structure. It is observed that the probability of the near-normal drought class is higher in most stations of the study area (see Table 7). Further, in a global warming context, the situation of a significant increase in trend from wet drought classes to dry drought classes should be alarming for water resources and management authorities.

The limitation of the study includes:

1. The current study is based on SPEI index. Other indices such as SPI and SPEI can be incorporated. Further, their comparative assessment should be made for better understanding.
2. In this manuscript, we have used quantitative time-series data. Consequently, our weights are based on autocorrelation. In the future, the qualitative time

series of drought can be used with our novel weighting scheme for WMC (see Ali et al. 2018).

## References

Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Hussain, T. and co-authors. 2017. Forecasting drought using multilayer perceptron artificial neural network model. *Adv. Meteorol.* **2017**, 1–9. doi:10.1155/2017/5681308

Ahmad, Z., Hafeez, M. and Ahmad, I. 2012. Hydrology of mountainous areas in the Upper Indus Basin, Northern Pakistan with the perspective of climate change. *Environ. Monit. Assess.* **184**, 5255–5274. doi:10.1007/s10661-011-2337-7

Ali, Z., Hussain, I., Faisal, M., Almanjahie, I. M., Ahmad, I. and co-authors. 2019. A probabilistic weighted joint aggregative drought index (PWJADI) criterion for drought monitoring systems. *Tellus. Dyn. Meteorol. Oceanogr.* 1–21.

Ali, Z., Hussain, I., Faisal, M., Almanjahie, I. M., Ismail, M. and co-authors. 2018. A new weighting scheme in weighted Markov model for predicting the probability of drought episodes. *Adv. Meteorol.* **2018**, 1–10. doi:10.1155/2018/8954656

Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Abd-el Moemen, M. and co-authors. 2017. A novel multiscalar drought index for monitoring drought: the standardized precipitation temperature index. *Water Resour. Manage.* **31**, 4957–4969. doi:10.1007/s11269-017-1788-1

Archer, D. R. and Fowler, H. J. 2004. Spatial and temporal variations in precipitation in the Upper Indus Basin, global teleconnections and hydrological implications. *Hydrol. Earth Syst. Sci.* **8**, 47–61. doi:10.5194/hess-8-47-2004

Arnhold, E. 2014. Easyanova: analysis of variance and other important complementary analyzes. *R package version*, 4.

Awan, S. A. 2002. The climate and flood risk potential of northern areas of Pakistan. *Science Vision* **7**, 100–109.

Bacanli, U. G., Firat, M. and Dikbas, F. 2009. Adaptive neuro-fuzzy inference system for drought forecasting. *Stoch. Environ. Res. Risk Assess.* **23**, 1143–1154. doi:10.1007/s00477-008-0288-5

Benoit, G. 2005. Weighted Markov chains and graphic state nodes for information retrieval. *Proc. Am. Soc. Info. Sci. Tech.* **39**, 115–123. doi:10.1002/meet.1450390113

Chen, J. and Yang, Y. 2012. SPI-based regional drought prediction using weighted Markov chain model. *Res. J. Appl. Sci. Eng. Technol.* **4**, 4293–4298.

Chiang, C. L. 1968. *Introduction to Stochastic Processes in Biostatistics*, Vol. 95. Wiley, New York.

Crommelin, D. and Vanden-Eijnden, E. 2006. Fitting timeseries by continuous-time Markov chains: a quadratic programming approach. *Comput. Phys.* **217**, 782–805. doi:10.1016/j.jcp.2006.01.045

De-di, L. and Chen X.-H. 2006. Annual precipitation forecasting based on the weighted Markov chain in Beijing River Basin. *Journal of China Hydrology* **6**, 006.

Edwards, B., Gray, M. and Hunter, B. 2009. A sunburnt country: the economic and financial impact of drought on rural and regional families in Australia in an era of climate change. *Aust J Labour Econ.* **12**, 109.

Gong, Z., Chen, C. and Ge, X. 2014. Risk prediction of low temperature in Nanjing city based on grey weighted Markov model. *Nat. Hazards* **71**, 1159–1180. doi:10.1007/s11069-013-0690-2

Gui, Y. and Shao, J. 2017. Prediction of precipitation based on weighted Markov chain in Dangshan. In *Proceedings of the International Conference on High Performance Compilation, Computing and Communications*, ACM, 81–85.

Haan, C. T. 1977. *Statistical Methods in Hydrology*. The Iowa State University Press, Iowa.

Hargreaves, G. H. 1994. Defining and using reference evapotranspiration. *J. Irrig. Drain. Eng.* **120**, 1132–1139. doi:10.1061/(ASCE)0733-9437(1994)120:6(1132)

Kaliakatsos-Papakostas, M. A., Epitropakis, M. G. and Vrahatis, M. N. 2011. Weighted Markov chain model for musical composer identification. In *European Conference on the Applications of Evolutionary Computation*, Springer, 334–343.

Lange, K. 2010. *Numerical Analysis for Statisticians*. Springer, Berlin.

Lee, H.-Y. and Chen, S.-L. 2006. Why use Markov-switching models in exchange rate prediction? *Econ. Model.* **23**, 662–668. doi:10.1016/j.econmod.2006.03.007

Le-Tian, X. 2005. Prediction of plum rain intensity based on index weighted Markov chain. *J. Hydraul. Eng.* **8**, 988–993.

Lohani, V. K. and Loganathan, G. 1997. An early warning system for drought management using the palmer drought index1. *JAWRA Journal of the American Water Resources Association*, **33**, 1375–1386.

Marden, J. I. 1996. *Analyzing and Modeling Rank Data*. CRC Press, Boca Raton.

McKee, T. B., Doesken, N. J. and Kleist, J. 1993. The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, Vol. 17, American Meteorological Society Boston, MA, pp. 179–183.

Paulo, A. A. and Pereira, L. S. 2007. Prediction of SPI drought class transitions using Markov chains. *Water Resour. Manage.* **21**, 1813–1827. doi:10.1007/s11269-006-9129-9

Peng, Z., Bao, C., Zhao, Y., Yi, H., Xia, L. and co-authors. 2010. Weighted Markov chains for forecasting and analysis in incidence of infectious diseases in Jiangsu Province. *J. Biomed. Res.* **24**, 207–214. doi:10.1016/S1674-8301(10)60030-9

Sen, Z. 1990. Critical drought analysis by second-order Markov chain. *J. Hydrol. (Amsterdam)* **120**, 183–202. doi:10.1016/0022-1694(90)90149-R

Shatanawi, K., Rahbeh, M. and Shatanawi, M. 2013. Characterizing, monitoring and forecasting of drought in Jordan River Basin. *JWARP.* **05**, 1192–1202. doi:10.4236/jwarp.2013.512127

Spedicato, G. A., Kang, T. S., Yalamanchi, S. B. and Yadav, D. 2016. The Markovchain package: package for easily handling discrete Markov chains in r. Accessed Dec.

Svoboda, M. and Fuchs, B. 2016. *Handbook of Drought Indicators and Indices. Integrated Drought Management Tools and Guidelines Series*. World Meteorological Organization, National Drought Mitigation Center, University of Nebraska-Lincoln, Nebraska-Lincoln.

Takahashi, K., Morikawa, K., Takeda, D., Mizuno, A. 2007. Inventory control for a Markovian remanufacturing system with stochastic decomposition process. *Int. J. Prod. Econ.* **108**, 416–425. doi:10.1016/j.ijpe.2006.12.023

Thornthwaite, C. W. 1948. An approach toward a rational classification of climate. *Geogr. Rev.* **38**, 55–94. doi:10.2307/210739

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

Vicente-Serrano, S. M., Lo'pez-Moreno, J. I. and Van der Hurk, B. 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.* **9**, 523–533.

White, G. F. 1974. *Natural Hazards, Local, National, Global*. Oxford University Press, Oxford.

Zhou, P., Zhou, Y., Jin, J., Liu, L., Wang, Z. and co-authors. 2011. An improved Markov chain model based on autocorrelation and entropy techniques and its application to state prediction of water resources. *Chin. Geogr. Sci.* **21**, 176–184. doi:10.1007/s11769-011-0447-3