Theor Appl Climatol
DOI 10.1007/s00704-016-1901-z

ORIGINAL PAPER

Statistical analysis of short-term water stress conditions at Riggs Creek OzFlux tower site

Mohammad Azmi[1] · **Christoph Rüdiger**[1] · **Jeffrey P. Walker**[1]

Received: 6 December 2015 / Accepted: 12 August 2016

Abstract A large range of indices and proxies are available to describe the water stress conditions of an area subject to different applications, which have varying capabilities and limitations depending on the prevailing local climatic conditions and land cover. The present study uses a range of spatio-temporally high-resolution (daily and within daily) data sources to evaluate a number of drought indices (DIs) for the Riggs Creek OzFlux tower site in southeastern Australia. Therefore, the main aim of this study is to evaluate the statistical characteristics of individual DIs subject to short-term water stress conditions. In order to derive a more general and therefore representative DI, a new criterion is required to specify the statistical similarity between each pair of indices to allow determining the dominant drought types along with their representative DIs. The results show that the monitoring of water stress at this case study area can be achieved by evaluating the individual behaviour of three clusters of (i) vegetation conditions, (ii) water availability and (iii) water consumptions. This indicates that it is not necessary to assess all individual DIs one by one to derive a comprehensive and informative data set about the water stress of an area; instead, this can be achieved by analysing one of the DIs from each cluster or deriving a new combinatory index for each cluster, based on established combination methods.

✉ Mohammad Azmi
mohammad.azmi@monash.edu

Christoph Rüdiger
chris.rudiger@monash.edu

Jeffrey P. Walker
jeff.walker@monash.edu

1 Department of Civil Engineering, Monash University, Clayton, VIC 3800, Australia

1 Introduction

Drought indices (DIs) are often applied in proactive management systems to mitigate the costs of water stress conditions (Meinke and Stone 2005; Bond et al. 2008). DIs are also known as water stress indices when the temporal resolution of evaluations are considered weekly and within weekly (Svoboda et al. 2002; Svoboda et al. 2015). These indices are commonly used indicators and proxies for detecting the water stress conditions in different parts of a water system (e.g. plants, soil, rivers). DIs are generally classified based on the particular application and type of water stress to which they apply. For example, Wilhite and Glantz (1985) and later Wilhite (2005) defined four main drought (water stress) types as (1) hydrological droughts (water stress) affecting surface and ground water levels; (2) meteorological droughts (water stress) resulting in dry weather patterns dominating an area, (3) ecological/agricultural droughts (water stress) which focus on the quantity of plant available water, both within the ground and the biomass itself; and (4) socio-economic droughts (water stress), relating the supply and demand of various commodities to drought. This wide range of definitions led to the development of a large number of environmental indices.

By the late 1990s, indicators such as the Palmer indices (Palmer 1965), Surface Water Supply Index (SWSI) (Shafer and Dezman 1982), Evaporative Fraction (EF) (Shuttleworth et al. 1989), Standardized Precipitation Index (SPI) (McKee et al. 1993, 1995), and Soil Moisture Drought Index (SMDI) (Hollinger et al. 1993) were the most commonly used DIs. While those indices by themselves have the potential to describe specific types of water stress reasonably accurately, the main obstacle in detailed water stress evaluations was the lack of sufficient, and more importantly, accurate data and information across different temporal and spatial scales.

Over the past decade, the advancement in remote sensing, in situ monitoring and modelling technologies has allowed water experts to gain access to unprecedented levels of spatially consistent, high-resolution data products containing high-quality and timely information on surface processes, both in terms of the water and vegetation dynamics, as well as for the land–atmosphere interactions. In recent years, the most commonly used DIs are Standardized Runoff Index (SRI) (Shukla and Wood 2008), Normalized Difference Vegetation Index (NDVI) (Jackson et al. 2004; Maki et al. 2004), Vegetation-Temperature Condition Index (VTCI) (Singh et al. 2003; Patel et al. 2012), Perpendicular Drought Index (PDI) (Ghulam et al. 2007), Soil Moisture Index (SMI) (Hunt et al. 2009) and Water Surplus Variability Index (WSVI) (Gocic and Trajkovic 2015).

Previous studies have shown that no single DI is capable of capturing all weather conditions equally well (Wilhite 2000; Van Loon and Van Lanen 2012). As an example, precipitation-based DIs usually detect water stress conditions as “water stress conditions” in winter in northern Europe; however, there is a significant amount of precipitation in the form of snow and ice, which is often not adequately represented (Van Loon 2013). Therefore, many studies compared different DIs under different land use and or climate conditions to evaluate the meteorological/climatological behaviours of case study areas (Bayarjargal et al. 2006; Mpelasok et al. 2008; Barua et al. 2011; Belal et al. 2012; Liu et al. 2012; Choi et al. 2013; Song et al. 2013; Maccioni et al. 2014). The results of those studies have shown that the most reliable solution to address individual deficiencies of single DIs is to evaluate an appropriate set of them simultaneously. Despite this conclusion, there is no documented study which considers an appropriately large set of spatially and temporally high-resolution DIs with sufficient data diversity. Consequently, this study deals with identifying similarities and conflicts between the nine conventional and contemporary high-resolution DIs to evaluate short-term water stress conditions of the Riggs Creek OzFlux tower site in southeastern Australia. Another objective is to determine a methodology to combine various drought indices to develop a more robust generalized index. This is achieved by (i) making use of spatio-temporally high-resolution data sources to form an appropriate set of DIs, (ii) assessing the advanced statistical characteristics of the members of each individual cluster and (iii) determining the dominant water stress type(s) with its most representative DI.

2 Case study area, variables and data sources

The Riggs Creek OzFlux tower was chosen as the case study for this investigation as it provides both high temporal and spatial resolution data sets, as well as a minimum amount of data gaps across the 1.5 years (from 18 December 2010 to 1 May 2012) of verified within daily data used in this study. The tower is located within the Goulburn-Broken catchment (36° 38.59′ S, 145° 34.21′ E), in northern Victoria, Australia (Fig. 1) (Andrykanus 2011; Beringer 2014). Figure 2 illustrates two photos of Riggs Creek OzFlux tower site along with the tower. The site’s elevation is 152 m above sea level, and the surrounding area is dominated by broadacre farming practices. The main land use in this region is dryland agriculture, with a predominant use as pasture. Based on a nearby weather station operated by the Australian Bureau of Meteorology (Euroa station, BOM ID 82016), Riggs Creek has a mean annual precipitation of 650 mm, a mean maximum air temperatures of 12.3 °C in July and 29.7 °C in February and mean minimum air temperatures between 4.1 °C in July and 15.3 °C in February. Carbon dioxide, water vapour and latent/sensible heat are measured via the open-path eddy flux technique (at height of 2 m). Finally, soil heat flux plates are installed at a depth of 0.08 m to complement the collection of the soil moisture content (at a depth of 0.1 m) using time domain reflectometry.

According to the report published in 2011 by the South Eastern Australian Climate Initiative (SEACI 2011), the Australian state of Victoria experienced a severe water stress from 1997 to 2009 with an average precipitation rate of 12.4 % below the twentieth century mean. In 2010–2011, because of a strong La Niña phenomenon, significantly higher levels of precipitation were recorded for Victoria with annual

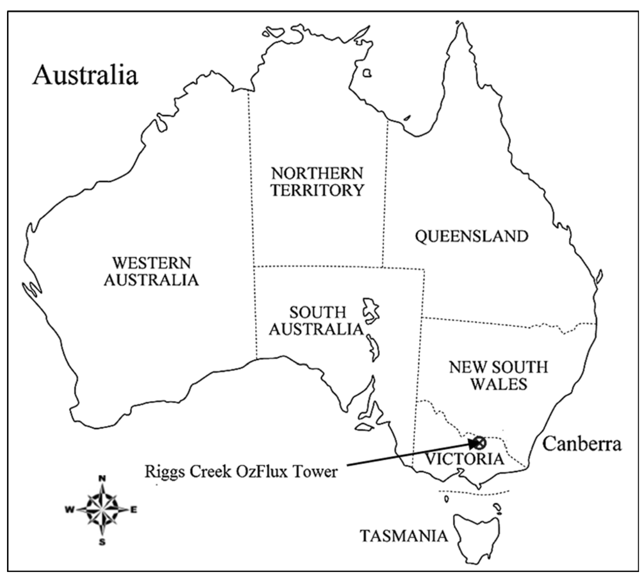


Fig. 1 The geographical location of the Riggs Creek case study site within Australia

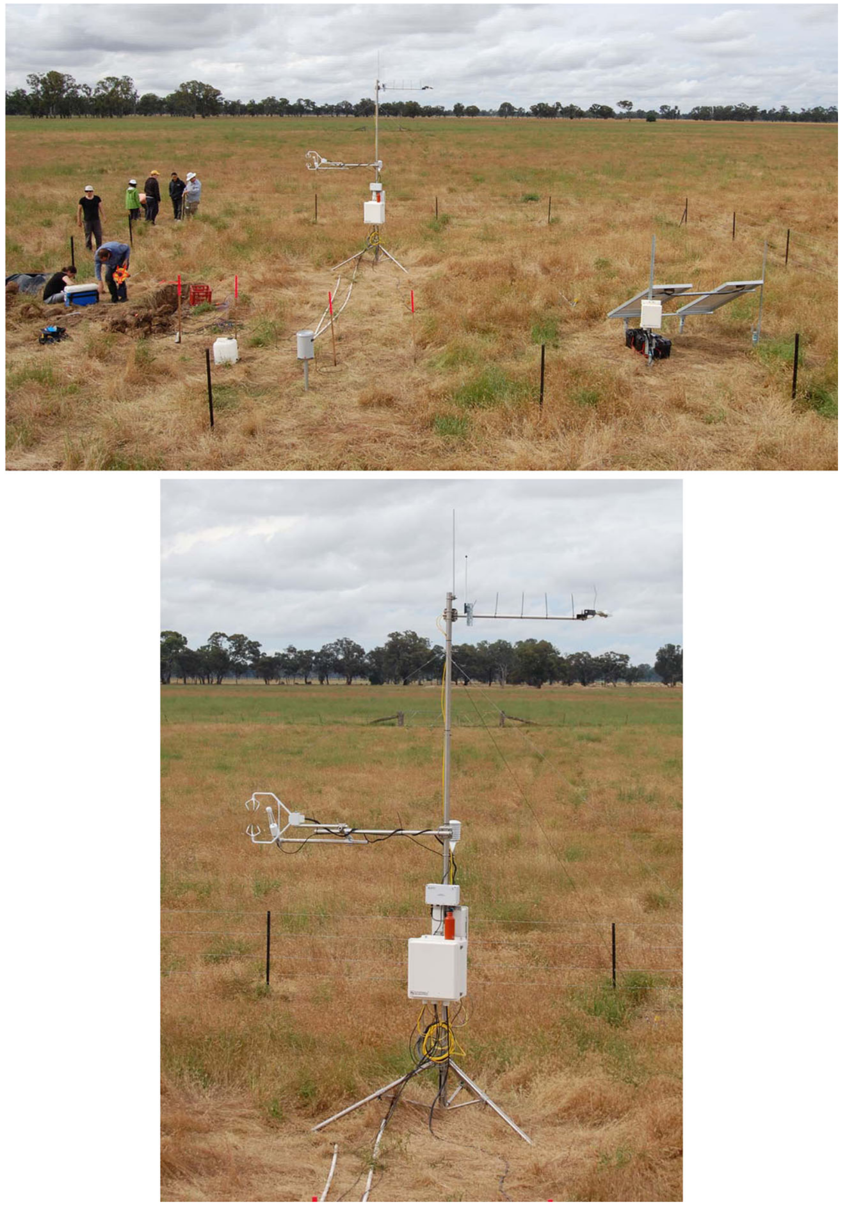


Fig. 2 Two photos of Riggs Creek OzFlux tower site located within a grassland (pasture) along with the tower. Underground-based instruments measure soil moisture at depths of 0.1, 0.2, 0.4, 0.8, 1.2, and 2 m. Ground-based instruments are data logger and rain gauge. Mid-height instruments measure soil temperature (0.08 m) and soil heat flux (0.1 m). Top-height instruments measure atmospheric pressure (2.5 m), net radiation (4 m), open-path CO_2 H_2O, sonic anemometer (2.5 m), air temperature (2.5 m) and relative humidity (2.5 m)

rates exceeding 810 mm (SEACI 2011), while by early 2012, no extreme events (flood and severe drought) had been recorded (Howden 2012).

Considering the main elements of the hydrological cycle as well as water stress types, a diverse set of variables/proxies was taken into consideration to derive DIs for this study (Table 1). For each DI, the category of the utilized data sources (in situ observations, satellite information and/or a combination of them), their spatial and temporal scale and the method of deriving values for the DIs (direct measurement, measurement calculations) are summarized. The three main data sources of this study are as follows:

1. OzFlux is an Australian ecosystem research network of 37 sites set up to provide half hourly high-resolution flux tower measurements of water, energy and carbon. The variables observed through this network are used in hydroclimatic research but may also be employed in validating micrometeorological theories of fluxes and air flows (Finnigan et al. 2003; Finnigan 2004). More information can be accessed at http://www.ozflux.org.au.
2. The Asia-Pacific Water Monitor (APWM) (Van Dijk 2010) coordinated by the Commonwealth Scientific and Industrial Research Organization (CSIRO) provides a database of estimated daily hydrological variables such as precipitation, precipitation, runoff and catchment water storage. The final data derived from APWM is a combination of the output of several sources (in situ observations and satellite data) and models (Australian Water Resources Assessment Landscape models) to achieve

Table 1 Hydroclimatic variables derived from different data sources to calculate drought indices and proxies compared in this study

Category	Data sources	Primary variables	Abbreviation	Units	Temporal scale	Spatial scale	Calculation
In situ observations	OzFlux Tower Network	Soil Moisture content (depth = 10 cm)	SM	mm	0.5 h	N/A	Direct measurement
		Precipitation	P	mm	0.5 h		
		Moisture Flux (latent heat)	MF	W/m^2	0.5 h		
		Evaporative Fraction Index	EFI	No Dim.	0.5 h		$\mathrm{EFI} = \frac{H}{\mathrm{Rn}-G}$
A combination of in situ observations, satellite information and models' output	Asia-Pacific Water Monitor (APWM) Section	Runoff and Surface Soil Moisture	RSSM	mm	1 day	500 m	A combination of the output of several sources and models
satellite information	MODIS-Terra Satellite	Normalized Difference Vegetation Index	NDVI	No Dim.	1 day	250 m	$\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{VIS}}{\mathrm{NIR}+\mathrm{VIS}}$
		Vegetation Condition Index	VCI	No Dim.	1 day	250 m	$\mathrm{VCI} = \frac{\mathrm{NDVI}-\mathrm{NDVI}_{\min}}{\mathrm{NDVI}_{\max}-\mathrm{NDVI}_{\min}}$
		Temperature Condition Index	TCI	No Dim.	1 day	500 m	$\mathrm{TCI} = \frac{\mathrm{Tb}_{\max}-\mathrm{Tb}}{\mathrm{Tb}_{\max}-\mathrm{Tb}_{\min}}$
		Perpendicular Drought Index	PDI	No Dim.	1 day	250 m	$\mathrm{PDI} = \frac{1}{\sqrt{M^2+1}}(\mathrm{VIS} + M \times NIF)$

NIR spectral reflectance measurements acquired in the near-infrared regions (700–1100 nm), *VIS* spectral reflectance measurements acquired in the visible (Red) regions (400–700 nm), $NDVI_{max}$ *and* $NDVI_{min}$ maximum and minimum *NDVI* for a given time series, *Tb* brightness temperature of the spectral reflectance measurements acquired with band 4 of MODIS, Tb_{max} and Tb_{min} maximum and minimum brightness temperature, respectively, *M* the slope of soil line in the NIR-Red spectral feature space, *H* latent heat flux (Wm^{-2}), *Rn* net radiation (Wm^{-2}), *G* ground/soil heat flux (Wm^{-2})

the best possible estimations (Van Dijk 2010). More information is available at http://eos.csiro.au/apwm/apwm.html.

3. The MODerate resolution Imaging Spectroradiometer (MODIS) measures 36 bands of spectral reflectance. The obtained data are used in deriving ecological, hydrological and oceanic products on daily to monthly scales. The MODIS instrument is installed on board NASA's Terra satellite and has been in operation since 18 December 1999. More information about this source can be found at http://terra.nasa.gov.

Due to the fact that plants are the main elements of the terrestrial ecosystems and reflect the water stress (wilting point) with a lag time of around 1 week (especially cultivated pasture and rainfed agricultural areas), daily data products show decorrelated results between the indices and the surface conditions. Moreover, information, with longer time spans (monthly and seasonally), of water stress monitoring cannot be practical for ecological water management due to the wilting point time span (which is around 1 week) of many species of vegetation (Svoboda et al. 2002; Heim Jr 2002). Therefore, weekly temporal resolutions were considered for evaluating water stress conditions in this study.

To derive weekly DIs, first, the few data gaps that occurred during the considered period were in-filled using a univariate linear regression, for gaps lasting several hours (in situ observations), and multivariate linear/nonlinear regressions for day-long gaps (mainly for the satellite observations). Next, the time scale of all data was converted to a daily basis using arithmetic averaging. Then, daily DIs were calculated based on the latter data set and converted into weekly time scale by averaging, and finally, weekly DIs were standardized.

3 Standardizing hydroclimatic variables and proxies

As different hydroclimatic variables and proxies represent different quantities, they are not directly comparable (e.g. because of different dimensions/units). Moreover, even though they are measured at the same time, they may not be related to the same event due to the existence of lag times. Therefore, in order to compare the different statistical aspects of the hydroclimatic variables, it is essential to employ a spatio-temporal standardization approach.

The most practical and acceptable method to standardize hydroclimatic variables and their proxies, for statistical comparisons, is based on an equiprobability transformation, as presented by Panofsky and Brier (1958) (McKee et al. 1993, 1995; Shukla and Wood 2008). The transformation maintains the probability of a given value within its primary sample group to be the same probability as in the transformed normally distributed variate (Edwards and McKee 1997). It is worth noting that this method of standardizing retains the main statistical characteristics (e.g. skewness) of the primary time series.

To follow this method, first, the most appropriate cumulative distribution function (CDF) of each variable is chosen by fitting different CDFs over the variable's data and using two goodness-of-fit tests, the Anderson-Darling and p value (Stephens 1974), to evaluate the statistical relationship of each data pair.

After choosing the most appropriate CDF, the occurrence probability of each observation is extracted, and finally, the corresponding value of the extracted probability derived from a standard Gaussian CDF ($\mu = 0$, $\sigma = 1$) (McKee et al. 1993, 1995; Shukla and Wood 2008). After this process, the hydroclimatic variables and proxies become standardized drought indices (SDIs). Thus, this standardizing method adequately retains the statistical characteristics of the primary time series. To analyse normal and extreme conditions, thresholds based on a variety of percentiles of the SDI's time series (2.5th, 5th, 25th, 45th, 50th, 55th, 75th, 95th and 97.5th percentiles) were considered.

4 Comparison between SDIs

The above statistical/probabilistic comparisons allow determining the likelihoods of events of a given magnitude to occur under the same climatic conditions. As a result, two more aspects of the considered variables may be investigated: (i) the characteristics of the extreme events subject to each SDI and (ii) the clustering of the SDIs, i.e. the grouping of similar indices into categories of different water stress types. This information can eventually be used as an indicator of the statistical dis-/similarities between SDIs under normal and extreme conditions.

4.1 Extreme events

Extreme events, as identified by SDIs, can be studied for the behaviour of their three main parameters: *duration*, *severity* and *magnitude*. Keyantash and Dracup (2002) defined that if the absolute values of a SDI from time t to $t + k$ (where k is the continuous duration of an event) are equal or more than the absolute value of the predefined extreme threshold (ET), then the *severity* of such an event (*Sev*(t,k)) can be derived as:

$$\mathrm{Sev}(t,k) = (\mathrm{Median}\{\mathrm{Mag}(i)\}) \times k, \tag{1}$$

where

$$Mag(i) = SDI_i - NCT, i = t, \ldots, t + k, \tag{2}$$

where Mag(i) is the *magnitude* of an extreme event at time i; SDI_i the value of the SDI at time i; NCT the normal condition threshold (50th percentile of the time series of the SDI); and Median(Mag(i)) the median of all Mag(i) from time t to $t+k$. The values of dry and wet extreme thresholds for an SDI were defined as 2.5th, and 97.5th percentiles of the entire time series of the SDI, respectively.

4.2 Clustering SDIs

Hydrological and meteorological drought types describe the availability of water resources in a water system, while agricultural/ecological drought types describe the water stress conditions inside plants. Therefore, analysing the performance of these drought types is necessary for comprehensive water resource management of a terrestrial ecosystem. It is worth noting that economy/social water stress type is not considered here due to a lack of relevant data at the case study area.

Clustering methods are mostly used to categorize the SDIs into the different water stress clusters. Amongst a variety of clustering methods, the agglomerative hierarchical clustering method (AHC) is recommended in temporal and spatial hydroclimatic classification issues (Santos et al. 2010). The AHC consists of two main parts: (i) the linkage criterion and (ii) the distance/similarity function. As such, the linkage criteria determine how the algorithm and process of the distance between two clusters are defined and the distance/similarity functions measure the respective distances. The single criterion (distance between two clusters is the minimum distance between an item in one cluster and an item in the other cluster), average criterion (distance between two clusters is the mean distance between an item in one cluster and an item in the other cluster) and the Ward criterion are linkage criteria alternatives (criterion for choosing the pair of clusters to merge at each step based on the optimal value of an objective function of error sum of squares); similarly, the Euclidean (Euclidean distance), Pearson correlation (linear correlation) and Spearman correlation (ranked correlation) are the most common choice for the distance/similarity functions (Soltani and Modarres 2006; Santos et al. 2010; Sarmadi and Shokoohi 2015; Sarmadi and Azmi 2016).

Previous hydroclimatic classification studies have only used a pair of linkage and distance functions, with their selection generally based on preconceived expertise ideas and/or some other case studies (Soltani and Modarres 2006; Santos et al. 2010; Sarmadi and Shokoohi 2015; Sarmadi and Azmi 2016). However, due to the fact that each combination of these pairs can change the final clustering (due to linear and nonlinear relationships between SDIs), a variety of combinations should be considered in the clustering in order to gain a clearer and more accurate understanding of the general similarities between different SDIs. In addition, due to the complex relationships between SDIs, it seems that categorizing SDIs with deterministic methods like dendrograms is not sufficient when considering all aspects of the statistical behaviours of SDIs. Therefore, a probabilistic-based algorithm is proposed here to derive pairwise similarities between SDIs as follows:

1. Produce two sets of SDIs based on imposed time lags of 1 and 2 on the primary set of SDIs (it is worth noting again that the temporal scale of the current study is weekly). These two data sets along with the primary set of SDIs are considered as input data sets, meaning that the total number of variables (here primary SDIs are equal to 8 variables and lagged SDIs would be equal to 16 variables) reaches 24.
2. Use two sizes of clusters equal to 3 (three main water stress types of hydrological, meteorological and agricultural/ecological) and 4 (to represent an interstitial group may be located between the main groups).
3. Employ three linkage methods (Single, Average, Ward) and three distance functions (Euclidean, Pearson correlation, Spearman correlation).
4. Consider all different combinations of cluster sizes, linkage methods and distance functions (2 cluster sizes × 3 linkage methods × 3 distance functions = 18 different combinations). Obviously, subject to considering different cluster sizes, linkage methods and distance functions, total number of combinations would be changed.
5. Undertake the clustering of the 24 variables based on the 18 combinations of step 4.
6. Determine the probabilistic similarities between a pair of variables (with and without time lags) ($\mathrm{PS}_{i,j}$) as follows.

$$\mathrm{PS}_{i,j} = \frac{n}{N} \times 100, \tag{3}$$

where n is the number of times that two variables (i,j) are located in the same cluster and N is the total number of all different combinations of clustering (here, it would be equal to 18 according to the step 4).

7. Calculate the average between all probabilistic similarities (PSs) of a pair of SDIs (with and without time lags) to derive a final probabilistic similarity for each pair of SDIs.
8. Develop a deterministic clustering for SDIs by defining thresholds for PSs: "strong similarity": PSs greater than 60 %; "moderate similarity": PSs between 40 to 60 %; and "weak similarity": PSs less than 40 %.

To define the thresholds, it is first assumed that the degree of consistency between members of a cluster has a linear correlation with the average PSs of a set of variables. Further, the PSs between a set of variables statistically follow a standard normal probabilistic distribution function. Therefore, Z scores between −0.67 and +0.67 cover 50 % of all events which are moderate conditions and the rest can represent extreme conditions. Transferring this range of Z scores to a 0–100 scale, the values would fall within a band of 40–60 out of 100. Therefore, PSs between 40 and 60 % ($-0.67 < Z < +0.67$) can state moderate condition or in other words "moderate similarity", and consequently, PSs greater than 60 % ($Z > +0.67$) and less than 40 % ($Z < -0.67$) can show "strong and weak similarities", respectively (extreme conditions). All SDIs with a "strong similarity" can definitely be located in one cluster, while SDIs which have "weak similarity" with others may be considered as a single-member cluster.

9. Validate the results of step 8 by performing Cronbach's alpha (α) test (Cronbach 1951). This test shows the consistency between the members of a cluster: $\alpha \geq 0.9$ shows very good consistency; $0.6 \leq \alpha < 0.9$, good consistency; $0.5 \leq \alpha < 0.6$, poor consistency; and finally $\alpha < 0.5$, an unacceptable (inconsistent) cluster (Kline 2000).

5 Results and discussions

For the perpendicular drought index (PDI), the slope of the soil line (M) is determined using all pairwise values of near-infrared-visible (NIR-VIS) for the location of the Riggs Creek OzFlux tower from daily MODIS observations on board NASA's Terra satellite. In Fig. 3, according to Ghulam et al. (2007), after manually defining the triangular region bounded by the observed NIR-VIS pairs, the slope of the base of the triangular is drawn and then considered as the slope of the soil line (M), which is found to be equal to 1.024 for this data set. Care must be taken, the height of the soil line (line $\overline{AD}$) describes the vegetation condition (from full cover at A to bare soil at D), meaning that A describes the area as fully covered by plants, which will result

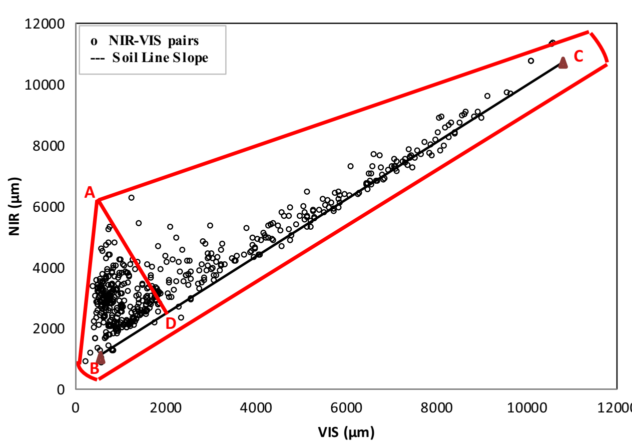


Fig. 3 Pairwise values of NIR-VIS for the location of Riggs Creek OzFlux Tower for deriving the slope of soil line (M)

in the vegetation to be active for a longer time, while any point between A and D indicates reduced vegetation levels. In addition, the skewness of a pairwise distribution towards points B and C shows the conditions of wet and dry surfaces, respectively. Figure 3 shows for Riggs Creek that the area has received significant levels of water during the study period (with its skewness towards point B). Nonetheless, the area is only partially covered (relatively short $\overline{AD}$ line), which can be explained with continuous grazing as well as seasonal harvesting of the pasture areas.

The weekly time series of SDIs are presented in Fig. 4. The operational functionality of the standardization of the indices is apparent in the values of SNDVI and SVCI. As VCI is itself a standardized form of NDVI (Table 1), those two normalized indices will generally match. Both SNDVI and SVCI are indicative of plant dynamics (growing and senescence seasons), and therefore, their maximum and minimum values are found at the beginning and end of growing seasons, respectively (Singh et al. 2003; Jackson et al. 2004). Moreover, the three indices Standardized Soil Moisture Index (SSMI), Standardized Runoff and Surface Soil Moisture Index (SRSSMI) and Standardized Moisture Flux Index (SMFI) follow similar patterns, as they are strongly dependent on precipitation events as well as different seasonal conditions (air temperature and consequently actual evapotranspiration). STCI shows roughly a seasonal correlation with air temperature, which was predictable because they reflect temperature conditions. No particular pattern was describable for the last index (SPDI) which shows the necessity for advance statistical analyses.

Due to difficulties in accurately measuring and calculating the real evapotranspiration, it is recommended to benefit from EFI (Kustas et al. 1993; Nishida et al. 2003). Here, SEFI is the only index which states the temporal variabilities of water consumptions, and this is the reason why it has an opposing behaviour in comparison with indices such as SSMI which shows the availability of water.

Considering the goodness-of-fit analysis (Anderson-Darling and p value tests), the lognormal probability function was determined as the best CDF for most of the SDIs (6 out of 9), the 3-p Gamma probability function was chosen for SPI and SMFI and the 3-p Weibull probability function was chosen for SPDI (Table 2). It is worth noting that the minimum, 2.5th and 5th percentile values of SPI are identical, which is due to having a significant amount of zero precipitation values on the weekly scale. Omitting zero values can remarkably reduce the skewness of precipitation data; nonetheless, they need to be kept for representing extreme dry conditions of the area, in terms of precipitation, throughout the process of water stress evaluations. With the exception of the SPI, the values of percentiles of the remaining SDIs were close to each other, suggesting similar behaviour under extreme conditions. This finding is discussed in more detail below.

The 2.5th, 50th and 97.5th percentiles defined in Section 4.1 were used to determine the respective durations

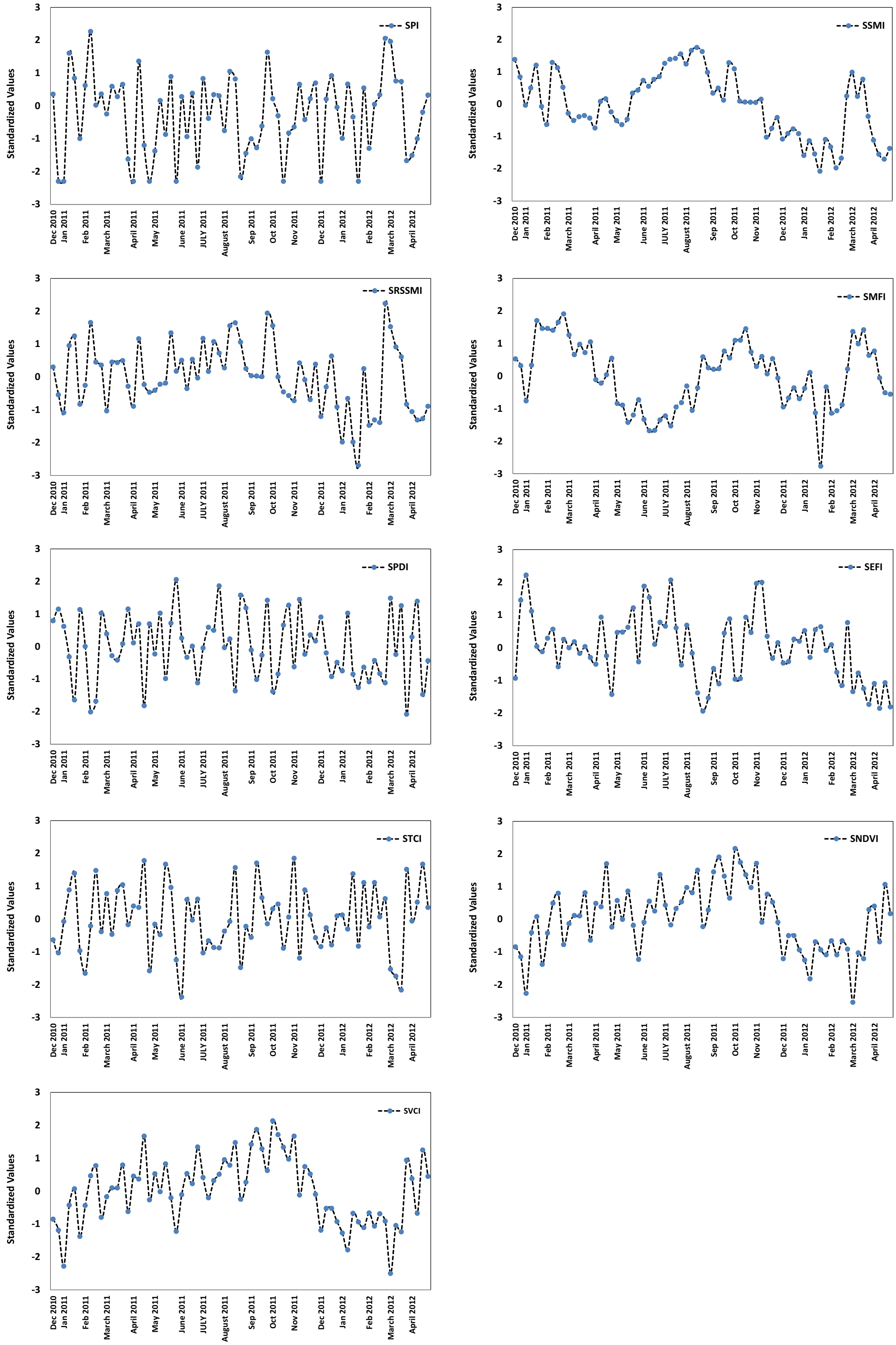


Fig. 4 Weekly time series of SDIs at Riggs Creek OzFlux Tower from 18-Dec-2010 to 1-May-2012

Table 2 The main statistical parameters, the best chosen cumulative distribution function and different percentiles for each considered SDI

Standardized variables	SDIs	The best chosen *cdf* for standardizing	Skewness	Kurtosis	Min.	2.5th percentile	5th percentile	25th percentile	45th percentile	50th percentile	55th percentile	75th percentile	95th percentile	97.5th percentile	Max
Standardized SM	SSMI	3-P lognormal	−0.13	−0.90	−2.08	−1.77	−1.63	−0.74	−0.05	0.07	0.14	0.81	1.49	1.64	1.75
Standardized P	SPI	3-P gamma	−0.17	−0.72	−2.30	−2.30	−2.30	−1.10	−0.27	0.02	0.21	0.64	1.62	1.98	2.27
Standardized MF	SMFI	3-P gamma	−0.017	−0.56	−2.77	−1.67	−1.48	−0.083	−0.08	0.07	0.22	0.73	1.46	1.67	1.91
Standardized RSSM	SRSSMI	3-P lognormal	−0.08	−0.17	−2.69	−1.98	−1.43	−0.70	−0.14	0.01	0.17	0.57	1.60	1.73	2.24
Standardized NDVI	SNDVI	3-P lognormal	−0.04	−0.29	−2.54	−1.93	−1.31	−0.73	−0.11	0.00	0.15	0.61	1.71	1.75	2.16
Standardized VCI	SVCI	3-P lognormal	−0.06	−0.38	−2.50	−1.91	−1.33	−0.74	−0.14	−0.02	0.16	0.68	1.67	1.75	2.13
Standardized TCI	STCI	3-P lognormal	−0.06	−0.54	−2.38	−1.85	−1.62	−0.72	−0.16	−0.07	0.02	0.71	1.67	1.73	1.85
Standardized PDI	SPDI	3-P Weibull	−0.07	−0.78	−2.08	−1.87	−1.66	−0.79	−0.21	−0.03	0.10	0.76	1.46	1.65	2.06
Standardized EF	SEFI	2-P lognormal	0.13	−0.36	−1.94	−1.82	−1.64	−0.69	−0.1	0.04	0.17	0.61	1.92	2.01	2.22

of the extreme events. In addition, the severities (Sevs) and magnitudes (Mags) of the extreme conditions were calculated using Eqs. (1) and (2). The durations, Sevs and Mags of extreme dry as well as wet events for all SDIs are presented in Tables 3 and 4, respectively. Table 3 shows that for most of the cases durations of extreme dry events were 1, meaning that extreme dry conditions lasted generally 1 week or less. From a climatological point of view, an extreme event with the duration of 1 week cannot be considered as a drought or flood condition. However, 1 week is long enough for pasture and rainfed agriculture areas to get remarkable damages due to water stress especially in the growing season. As a result, all extreme events need to be considered even with duration of 1 week.

Under dry conditions, the variable Mags had a narrow range from −2.84 to −1.75, however Sevs may range from −4.88 to −1.75. Table 4 indicates that Mags of extreme wet events had a range from 1.64 to 2.25, with Sevs ranging from 1.63 to 3.28. To sum up the results of extreme dry/wet conditions, Fig. 5 provides a comparison of all extreme dry and wet events at Riggs Creek, showing that the area experienced much more severe extreme dry events in comparison with extreme wet events. Moreover, the middle box plot (Dry-SPI) highlights that even when extreme dry events of SPI were removed from the set of extreme dry events of all SDIs, the severities of extreme dry events of other SDIs were still larger than extreme wet events. Put differently, this means that, considering different drought types, the area generally suffered more severe extreme dry events in comparison with wet events during the time period of this study. Actually, because two dry seasons and one wet season were considered, this was expectable, and our limited data set reflected appropriate analysis. While such a figure cannot necessarily be considered as a definitive decision-making tool, it can help water managers to gain a more comprehensive understanding of the historical water conditions of an area, due to its simplicity and potential to summarize and present data and information descriptively.

The probabilistic similarities between all SDIs considered in this study are presented in Table 5, and final PS between each SDI pair is indicated in Table 6. According to Table 6, the minimum PS (11.11 %) is found between the pairs SEFI-SNDVI, SEFI-SVCI and SEFI-SMFI, while the maximum value (87.04 %) is found for the pair SSMI-SRSSMI, as it would be expected due to their both including soil moisture observations. Five SDIs of SPDI, SPI, SSMI, SRSSMI and SMFI with an average PSs of 62 % can be categorized as a cluster (representing water availability). Three SDIs of STCI, SNDVI and SVCI with an average PS of 75.3 % can be categorize as a cluster (representing vegetation conditions). The index of SEFI with an average PSs of 17.8 % (average of PSs between SEFI and other SDIs) can be considered as a single-member cluster (representing water consumption). The Cronbach alpha (α) test was applied to validate the results of

Table 3 Magnitude and severity of SDIs under extreme dry conditions (2.5th percentile of the entire time series of the SDI) calculated based on Eqs. 1 and 2

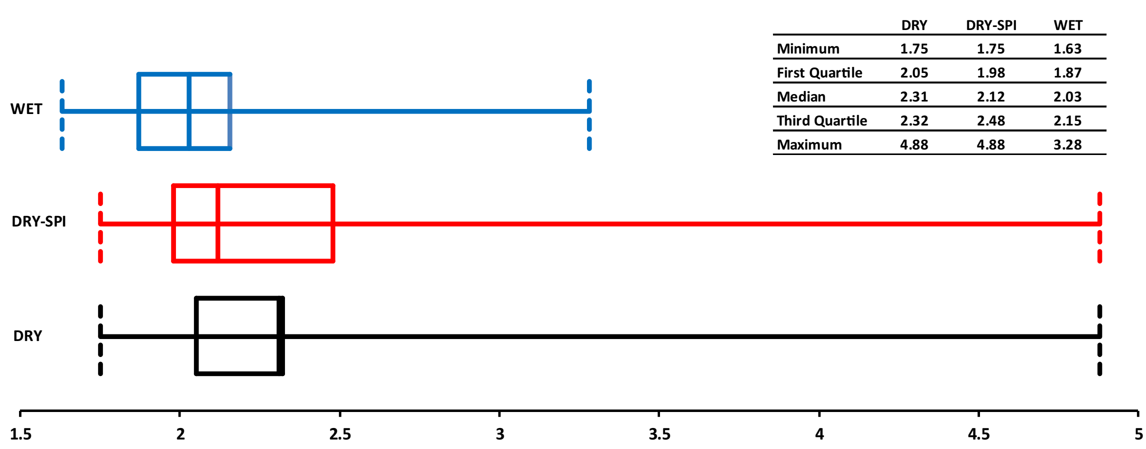
SDI	Duration (weeks)	Magnitude	Severity	SDI	Duration (weeks)	Magnitude	Severity
SPI	2	−2.32	−4.64	SRSSMI	1	−2.08	−2.08
	1	−2.32	−2.32		2	−2.44	−4.88
	1	−2.32	−2.32	SNDVI	1	−2.26	−2.26
	1	−2.32	−2.32		1	−2.54	−2.54
	1	−2.32	−2.32	SVCI	1	−2.26	−2.26
	1	−2.32	−2.32		1	−2.48	−2.48
	1	−2.32	−2.32	STCI	1	−2.31	−2.31
SSMI	1	−2.15	−2.15		1	−2.09	−2.09
	1	−2.04	−2.04	SPDI	1	−1.98	−1.98
SMFI	1	−1.75	−1.75		1	−2.05	−2.05
	1	−2.84	−2.84	SEFI	1	−1.98	−1.98
					1	−1.90	−1.90

Table 4 Magnitude and severity of SDIs under extreme wet conditions (97.5th percentile of the entire time series of the SDI) calculated based on Eqs. 1 and 2

SDI	Duration (weeks)	Magnitude	Severity	SDI	Duration (weeks)	Magnitude	Severity
SPI	1	2.25	2.25	SMFI	1	1.63	1.63
	1	2.04	2.04		1	1.84	1.84
SSMI	2	1.64	3.28	SRSSMI	1	1.84	1.84
SPDI	1	2.09	2.09		1	2.14	2.14
	2	1.89	1.89	SNDVI	1	1.91	1.91
STCI	1	1.85	1.85		2	2.16	2.16
	2	1.92	1.92	SVCI	1	1.89	1.89
SEFI	1	2.18	2.18		2	2.15	2.15
	1	2.03	2.03				

probabilistic-based clustering. The values of α for first and second groups were 0.63 and 0.83, respectively, which validates the derived clusters. To confirm that SEFI is not well correlated with the other indices, it was located once in cluster 1 and then in cluster 2. For the two cases, the derived values of α were 0.55 and 0.52, respectively, showing the poor consistency of this variable in relation to others.

Overall, the results show that the monitoring of water stress at this case study area can be done based on evaluating the individual behaviour of three clusters: (1) vegetation



	DRY	DRY-SPI	WET
Minimum	1.75	1.75	1.63
First Quartile	2.05	1.98	1.87
Median	2.31	2.12	2.03
Third Quartile	2.32	2.48	2.15
Maximum	4.88	4.88	3.28

Fig. 5 Box plot of comparison between extreme dry and wet events severities at Riggs Creek Flux Tower site from 18-Dec-2010 to 1-May-2012

Table 5 The probabilistic similarities between SDIs without time lags, as well as with time lags 1 and 2 calculated based on Eq. 3 (values are in percentage)

	STCI	SNDVI	SVCI	SPDI	SPI	SSMI	SRSSMI	SMFI	SEFI
STCI		88.9	88.9	33.3	61.1	72.2	72.2	66.7	11.1
SNDVI			100.0	33.3	72.2	83.3	83.3	77.8	11.1
SVCI				33.3	72.2	83.3	83.3	77.8	11.1
SPDI					22.2	33.3	33.3	27.8	33.3
SPI						88.9	88.9	83.3	0.0
SSMI							100.0	94.4	11.1
SRSSMI								94.4	11.1
SMFI									11.1
SEFI									
	STCI-1	SNDVI-1	SVCI-1	SPDI-1	SPI-1	SSMI-1	SRSSMI-1	SMFI-1	SEFI-1
STCI	44.4	55.6	55.6	33.3	50.0	66.7	66.7	66.7	11.1
SNDVI	55.6	66.7	66.7	33.3	61.1	77.8	77.8	77.8	11.1
SVCI	55.6	66.7	66.7	55.6	61.1	77.8	77.8	77.8	11.1
SPDI	33.3	44.4	11.1	11.1	44.4	33.3	33.3	55.6	33.3
SPI	44.4	55.6	55.6	44.4	66.7	77.8	77.8	83.3	0.0
SSMI	55.6	66.7	66.7	38.9	66.7	88.9	88.9	94.4	11.1
SRSSMI	55.6	66.7	66.7	38.9	66.7	88.9	88.9	94.4	11.1
SMFI	50.0	61.1	61.1	38.9	61.1	83.3	83.3	100.0	11.1
SEFI	16.7	11.1	11.1	11.1	22.2	22.2	22.2	11.1	100.0
	STCI-2	SNDVI-2	SVCI-2	SPDI-2	SPI-2	SSMI-2	SRSSMI-2	SMFI-2	SEFI-2
STCI	50.0	55.6	55.6	44.4	38.9	66.7	38.9	66.7	11.1
SNDVI	44.4	66.7	66.7	33.3	38.9	77.8	38.9	77.8	11.1
SVCI	44.4	66.7	66.7	33.3	38.9	77.8	38.9	77.8	11.1
SPDI	33.3	44.4	11.1	33.3	50.0	33.3	50.0	50.0	33.3
SPI	33.3	55.6	55.6	88.9	33.3	77.8	33.3	83.3	0.0
SSMI	44.4	66.7	66.7	100.0	38.9	88.9	38.9	94.4	11.1
SRSSMI	44.4	66.7	66.7	100.0	38.9	88.9	38.9	94.4	11.1
SMFI	38.9	61.1	61.1	94.4	33.3	83.3	33.3	100.0	11.1
SEFI	11.1	11.1	11.1	11.1	61.1	22.2	61.1	11.1	100.0

SDIs-1 SDIs with time lag equal to 1, *SDIs-2* SDIs with time lag equal to 2

conditions (STCI, SNDVI, SVCI), (2) water availability (SPI, SSMI, SRSSMI, SPDI and SMFI) and (3) water consumptions (SEFI). This indicates that it is not necessary to assess all SDIs one by one to derive a comprehensive and informative information set about water stress of an area but rather by analysing one of the SDIs from each cluster (e.g. SVCI from group 1, SSMI from group 2 and SEFI from group 3) or deriving a new index for each group, based on

Table 6 Final probabilistic similarities derived from the average between all probabilistic similarities of a pair of SDIs mentioned in Table 5 (values are in percentage)

	STCI	SNDVI	SVCI	SPDI	SPI	SSMI	SRSSMI	SMFI	SEFI
STCI		70.37	70.37	48.77	43.21	56.17	51.23	51.85	12.96
SNDVI			85.19	53.70	52.47	70.99	60.49	66.67	11.11
SVCI				53.70	52.47	70.99	60.49	66.67	11.11
SPDI					37.65	53.70	46.91	53.70	18.52
SPI						67.28	65.43	59.26	27.78
SSMI							78.40	87.04	18.52
SRSSMI								70.37	31.48
SMFI									11.11
SEFI									

established combination methods (Svoboda et al. 2002; Keyantash and Dracup 2004; Balint and Mutua 2011; Barua et al. 2011) can be quite enough. In other words, monitoring, predicting and finally planning/management of water stress situations at an area requires the consideration of all derived information (probabilistic behaviours, extreme dry/wet conditions, etc.) of the representative SDIs.

6 Summary and conclusion

This paper presented a set of statistical methods to evaluate short term water stress conditions for a study site in southeastern Australia. The required data were obtained from the Riggs Creek flux tower (OzFlux Network), the APWM and the MODIS instrument on board NASA's Terra satellite. A 72-week period in 2010–2012 was chosen based on the low level of missing data. The temporal scale of the data was from 0.5 h to daily, depending on the observed variable, allowing highly representative data to be derived at the weekly scale. In addition, the spatial scale of the different data sources was acceptable and consistent with this case study area, which also had a high level of homogeneity in its land use and surrounding topography. This study showed that by using a probabilistic standardizing method, the main statistical parameters of SDIs could be derived and analysed. The considered SDIs were compared in terms of their normal and extreme events. Then, using a proposed methodology, probabilistic similarities between standardized drought indices were presented and analysed. Finally, dominant water stress types along with possible representative SDIs were presented. At this study, SDIs were grouped into three main clusters, first group represented availability of water, second group reflected reaction of plants based on water statuses as well as different seasons and the last one showed the water consumptions in the water system.

The current paper is the proof-of-concept of a more comprehensive water stress study using advanced statistical methods and sources in which limited point data are available at the flux stations throughout Australia and definitely will be applied using longer term satellite data sets in future studies. Despite the limited length of data, the results are nevertheless promising and show the value of applied methods in the case of data shortages. As any approach based on statistics, larger data sets will make this type of studies more robust in its predictive skills. In fact, despite the limited amount of available data for this study, it was shown that by using appropriate statistical analysis methods, it is possible to derive initial useful information which can be applied for evaluating water stress issues. Moreover, the results showed that only relying on few common indices such as standardized precipitation index cannot be sufficient to evaluate all different aspects of water stress issues, which this can be quite persuasive why it is need to consider an appropriate set of indices and proxies from different drought types. Further, extreme conditions information described by descriptive statistical methods can provide clear, quick and uncomplicated visions to responsible authorities to have the better understanding of extreme water stress situations over an area.

Acknowledgments Mohammad Azmi acknowledges Monash University to fund this research in the form of a PhD stipend and tuition fee scholarship (MGS and FEIPRS). APWM and the Australian OzFlux network, especially Professor Jason Beringer from the University of Western Australia, are recognised for providing access to their data and also the photos of tower.

References

Andrykanus R (2011) Riggs Creek OzFlux tower site OzFlux: Australian and New Zealand flux research and monitoring. hdl: 102.100.100/6970

Balint Z, Mutua FM (2011) Drought monitoring with the combined drought index, FAO-SWALIM, Nairobi, Kenya

Barua S, Ng AWM, Perera BJC (2011) Comparative evaluation of drought indexes: case study on the Yarra River catchment in Australia. *J Water Resour Plan Manag* 137:215–226

Bayarjargal Y, Karnieli A, Bayasgalan M, Khudulmur S, Gandush C, Tucker CJ (2006) A comparative study of NOAA–AVHRR derived drought indices using change vector analysis. Remote Sens Environ 105:9–22

Belal AA, El-Ramady HR, Mohamed ES, Saleh AM (2012) Drought risk assessment using remote sensing and GIS techniques. Arab J Geosci. doi:10.1007/s12517-012-0707-2

Beringer J (2014) Riggs Creek OzFlux tower site OzFlux: Australian and New Zealand flux research and monitoring. hdl: 102.100.100/14246

Bond NR, Lake PS, Arthington AH (2008) The impacts of drought on freshwater ecosystems: an Australian perspective. Hydrobiologia 600:3–16. doi:10.1007/s10750-008-9326-z

Choi M, Jacobs JM, Anderson MC, Bosch DD (2013) Evaluation of drought indices via remotely sensed data with hydrological variables. J Hydrol 476:265–273

Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. Psychometrika 16(3):297–334. doi:10.1007/bf02310555

Edwards DC, McKee TB (1997) Characteristics of 20th century drought in the United States at multiple time scales. Climatology Rep. 97–2, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, 155 pp

Finnigan JJ (2004) A re-evaluation of long-term flux measurement techniques. Part II: coordinate systems. Bound-Layer Meteorol 113:1–41

Finnigan JJ, Clements R, Malhi Y, Leuning R, Cleugh HA (2003) A re-evaluation of long-term flux measurement techniques. Part I: averaging and coordinate rotation. Bound-Layer Meteorol 107: 1–48

Ghulam A, Qin Q, Zhan Z (2007) Designing of the perpendicular drought index. Environ Geol 52:1045–1052

Gocic M, Trajkovic S (2015) Water surplus variability index as an indicator of drought. J Hydrol Eng 20(2):04014038. doi:10.1061/(ASCE)HE.1943-5584.0001008

Heim RR Jr (2002) A review of twentieth-century drought indices used in the United States. Am Meteor Soc 1149–1165

Hollinger SE, Isard SA, Welford MR (1993) A new soil moisture drought index for predicting crop yields. In: Preprints, Eighth Conf. on Applied Climatology, Anaheim, CA, Am Meteor Soc, pp. 187–190
Howden S (2012) It's official: Australia no longer in drought. 27 April 2012, http://www.brisbanetimes.com.au/environment/weather/its-official-australia-no-longer-in-drought-20120427-1xpsp.html
Hunt ED, Hubbard KG, Wilhite DA, Arkebauer TJ, Dutcher AL (2009) The development and evaluation of a soil moisture index. *Int J Climatol* 29(5):747–759
Jackson JT, Chen D, Cosh M, Li F, Anderson M, Walthall C, Doriaswamy P, Hunt ER (2004) Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sens Environ* 92:475–482
Keyantash JA, Dracup JA (2002) The quantification of drought: an evaluation of drought indices. Bull Am Meteorol Soc 83(8):1167–1180
Keyantash JA, Dracup JA (2004) An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resour Res* 40:W09304. doi:10.1029/2003WR002610
Kline P (2000) The handbook of psychological testing, 2nd edn. Routledge, London, p. 13
Kustas WP, Schmugge TJ, Humes KS, Jackson TJ, Parry R, Weltz MA, Moran MS (1993) Relationships between evaporative fraction and remotely sensed vegetation index and microwave brightness temperature for semiarid rangelands. J Appl Meteorol 32:1781–1790
Liu L, Hong Y, Bednarczyk CN, Yong B, Shafer MA, Riley R, Hocker JE (2012) Hydro-climatological drought analyses and projections using meteorological and hydrological drought indices: a case study in Blue River Basin, Oklahoma. Water Resour Manag 26:2761–2779
Maccioni P, Kossida M, Brocca L, Moramarco T (2014) Assessment of the drought hazard in the Tiber River Basin in Central Italy and a comparison of new and commonly used meteorological indicators. J Hydrol Eng 1943–5584. doi:10.1061/(ASCE)HE
Maki M, Ishiara M, Tamura M (2004) Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data. *Remote Sens Environ* 90:441–450
McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA
McKee TB, Doesken NJ, Kleist J (1995) Drought monitoring with multiple time scales. Paper Presented at 9th Conference on Applied Climatology. American Meteorological Society, Dallas, Texas
Meinke H, Stone RC (2005) Seasonal and inter-annual climate forecasting: the new tool for increasing preparedness to climate variability and change in agricultural planning and operations. Clim Chang 70: 221–253
Mpelasok F, Hennessy K, Jones R, Bates B (2008) Comparison of suitable drought indices for climate change impacts assessment over Australia towards resource management. *Int J Climatol* 28:1283–1292
Nishida K, Nemani RR, Running SW, Glassy JM (2003) An operational remote sensing algorithm of land surface evaporation. J Geophys Res 108(D9):4270. doi:10.1029/2002JD002062
Palmer WC (1965) Meteorologic drought. US Department of Commerce, *Weather Bureau*, Research Paper No. 45, p. 58
Panofsky HA, Brier GW (1958) Some applications of statistics to meteorology. Mineral Industries Extension Services, College of Mineral Industries, Pennsylvania State University. PUB ID: 101–188-880
Patel NR, Parida BR, Venus V, Saha SK, Dadhwal VK (2012) Analysis of agricultural drought using vegetation temperature condition index (VTCI) from Terra/MODIS satellite data. *Environ Monit Assess* 184: 7153–7163
Santos JF, Pulido-Calvo I, Portela MM (2010) Spatial and temporal variability of droughts in Portugal. *Water Resour Res* 46:W03503. doi:10.1029/2009WR008071
Sarmadi F, Azmi M (2016) Regionalizing mean air temperature in Iran by multivariate analysis and L-moment methods. J Hydrol Eng 10: 05015018. doi:10.1061/(ASCE)HE.1943-5584.0001280
Sarmadi F, Shokoohi A (2015) Regionalizing precipitation in Iran using GPCC gridded data via multivariate analysis and L-moment methods. Theor Appl Climatol. doi:10.1007/s00704-014-1292-y
SEACI (South Eastern Australian Climate Initiative) (2011) The millennium drought and 2010/11 floods. July 2011, http://www.seaci.org/publications/documents/SEACI-2Reports/SEACI2_Factsheet2of4_WEB_110714.pdf
Shafer BA, Dezman LE (1982) Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. In: Preprints, Western Snow Conf., Reno, NV, Colorado State University, pp. 164–175
Shukla S, Wood AW (2008) Use of a standardized runoff index for characterizing hydrologic drought. *Geophys Res Lett* 35:L02405. doi:10.1029/ 2007GL032487
Shuttleworth WJ, Gurney RJ, Hsu AY, Ormsby JP (1989) FIFE: the variation in energy partition at surface flux sites. *IAHS Publ* 186: 67–74
Singh RP, Roy S, Kogan F (2003) Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. Int J Remote Sensing 24(22):4393–4402
Soltani S, Modarres R (2006) Classification of spatio-temporal pattern of rainfall in Iran using a hierarchical and divisive cluster analysis. J Spatial Hydrol 6(2)
Song X, Li L, Fu G, Li J, Zhang A, Liu W, Zhang K (2013) Spatial–temporal variations of spring drought based on spring-composite index values for the Songnen Plain, Northeast China. Theor Appl Climatol. doi:10.1007/s00704-013-0957-2
Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. *J Am Stat Assoc* 69(347):730–737
Svoboda MD, Lecomte D, Hayes M, Heim R, Gleason K, Angel J, Rippey B, Tinker R, Palecki M, Stooksbury D, Miskus D, Stephen S (2002) The drought monitor. Am Meteorol Soc. 1181–1190.
Svoboda MD, Fuchs BA, Poulsen CC, Nothwehr JR (2015) The drought risk atlas: enhancing decision support for drought risk management in the United States. J Hydrol 526:274–286
Van Dijk AIJM (2010) AWRA technical report 3: landscape model (version 0.5) technical description. WIRADA / CSIRO report
Van Loon A (2013) On the propagation of drought: how climate and catchment characteristics influence hydrological drought development and recovery. PhD thesis, Wageningen University, Wageningen, NL
Van Loon A, Van Lanen H (2012) A process-based typology of hydrological drought. *Hydrol Earth Syst Sci* 16:1915–1946
Wilhite DA (2000) Drought as a natural hazard: concepts and definitions. Drought: a global assessment. In: D. Wilhite (Ed.), Vol. 1, 3–18
Wilhite DA (2005) Drought and water crises: science, technology, and management issues. Taylor and Francis Group (CRC Press), Book. pp. 432, vol. 86
Wilhite DA, Glantz MH (1985) Understanding the drought phenomenon: the role of definitions. Water Int 10(3):111–120