MONASH University

Application of High Resolution Data Sources via Data Fusion Processes in Deriving a Comprehensive Drought Index

Mohammad Azmi
B.Sc in Water Engineering
M.Sc in Water Resources Engineering

A thesis submitted for the degree of Doctor of Philosophy at Monash University in 2016
Department of Civil Engineering, Faculty of Engineering, Clayton Campus

Copyright notice

Abstract

The main purpose is to develop a new methodology to address existing shortcomings in spatio-temporally evaluating water stress conditions by employing individual drought indices for a specific location. The proposed methodology is able to monitor water stress conditions of terrestrial ecosystems by objectively linking different aspects of the ecosystem such as water availability and vegetation conditions. The developed index, called Data Fusion-based Drought Index (DFDI), makes use of advanced statistical methods, and also considers eco-meteorological characteristics, such as landuse, land-cover, and climate of an area to determine the water stress conditions at each time step for each specific location.

The capabilities of the DFDI are demonstrated comprehensively making use of data from three OzFlux tower sites in Australia, as well as satellite data to develop a spatially distributed weekly water stress evaluation framework across Victoria, Australia. To achieve this, observations from those OzFlux Tower sites were complemented with nearby synoptic stations, while the spatial data where obtained from the Australian Water Availability Project, as well as the Soil Moisture and Ocean Salinity and the MODIS-Terra satellite missions.

One of the main advantages of the spatial DFDI methodology is to present explicit mathematical equations, ultimately making the process of spatial water stress monitoring at current and future time steps more user-friendly and consequently more accessible for the industry. To achieve this objective, the area is first regionalized according to the two criteria pairs of wetness/dryness and active/non-active vegetation via a K-mean clustering method using mean monthly normalized data of precipitation and the vegetation cover fraction, respectively. Then, for each sub-region of the wet/dry (active/non-active) map, a mathematical equation is developed by employing the Symbolic Regression Method in which the dependent variable is Standardized Aggregated Water Availability Index (or Standardized Aggregated Vegetation Index) and independent variables are selected amongst individual drought indices of the water content (or vegetation conditions) cluster. The final derived mathematical equations have generally shown promising functionalities and accuracies, especially in appropriately detecting the spatial trend of the extreme events.

As the methodology was developed using short *in situ* datasets, a sensitivity analysis has been implemented to investigate the issues arising from this for the performance of DFDI. A particular focus was put on the statistical characteristics of such data, specifically how the short-term observations of the three focus sites compared against long-term model predictions. The

rationale behind this approach is that while ground observations are generally sparse both in space and time, and satellite data are also often not covering sufficiently long periods, models provide the only consistent long-term reference for climate-related studies. As an alternative, short-term observations could potentially provide the same statistical characteristics as long-term observations would. Thus, the 3-year dataset from the three OzFlux towers was compared against a long-term dataset (1911-2016) from the AWAP in terms of both their stationary and dynamic statistical characteristics. The results show that observations and model displayed similar statistical features of long-term water conditions despite their different lengths, suggesting that both may be used for validation purposes.

Publications during enrolment

Azmi, M., Rüdiger, C., Walker, J. (2016) Statistical Analysis of Short Term Water Stress Conditions at Riggs Creek OzFlux Tower Site, Theoretical and Applied Climatology, DOI 10.1007/s00704-016-1901-z.

Azmi, M., Rüdiger, C., Walker, J. (2016) A data fusion-based drought index. *Water Resources Research*, 52: 2222-2239, DOI:10.1002/2015WR017834.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer reviewed journals and three submitted publications. The core theme of the thesis is drought and water stress monitoring. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the department of civil engineering under the supervision of Dr. Christoph Rüdiger and Professor Jeffrey Walker.

In the case of four chapters, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
3	*Statistical analysis of short-term water stress conditions*	*Published*	*75%. Concept, deriving results and writing first draft*	*1) Christoph Rüdiger, Editing and making comments 15%* *2) Jeffrey Walker Editing and making comments 10%*	*No* *No*
4	*A data fusion-based drought index*	*Published*	*75%. Concept, deriving results and writing first draft*	*1) Christoph Rüdiger, Editing and making comments 15%* *2) Jeffrey Walker Editing and making comments 10%*	*No* *No*
5	*Spatial monitoring of weekly water stress in Victoria, Australia*	*Under Review*	*75%. Concept, deriving results and writing first draft*	*1) Christoph Rüdiger, Editing and making comments 15%* *2) Jeffrey Walker Editing and making comments 10%*	*No* *No*
6	*Comparative evaluations of short-term OzFlux tower observations and long-term AWAP data in weekly water stress monitoring*	*Under Review*	*75%. Concept, deriving results and writing first draft*	*1) Christoph Rüdiger, Editing and making comments 10%* *2) Adam B. Smith Inputting data and making comments 10%* *3) Jeffrey Walker Editing and making comments 5%*	*No* *No* *No*

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: Date: 21/06/2016

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: Date: 21/06/2016

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Dr. Christoph Rüdiger and Professor Jeffrey Walker for their support of my Ph.D. project, for their trust, patience, motivation, and immense knowledge. Their guidance continuously helped me in all the time of research and writing of this thesis.

Last but not the least, I would like to thank my family, friends, and especially my lovely wife for supporting me spiritually throughout this long journey.

Doubt is the key to knowledge.
A Persian proverb

Table of Content **Page**

Table of Content **Page**

List of figures **Page**

List of figures **Page**

List of tables **Page**

Chapter 1
Introduction

1.1 Overview

Throughout this research, a set of advanced statistical methods are used in combination to introduce a novel methodology to derive a more robust and generally applicable drought index, enabling a more diversified investigation of the different aspects of water stress conditions across a region or the world. The new index is able to determine the dominant drought type of an area with respect to the current climate conditions and landuse situations (active/non-active seasons) for any specific month. Multivariate methods such as the Independent Component Analysis (ICA), symbolic regressions, and a new clustering methodology are the main techniques applied to achieve this objective. In terms of data sources, *in situ* observations (i.e. OzFlux tower and synoptic stations measurements), satellite data, and modelled data are utilized to evaluate the performance of this new methodology across point (i.e. in locations of some OzFlux tower sites) and spatial scales (across whole state of Victoria in Australia). This study concludes with a sensitivity analysis of the new drought index to assess the effect a dataset's length has on the development of the statistical characteristics required for such an index.

1.2 Importance and motivation of research

Australia is one of the driest countries in the world, consisting mostly of deserts and semi-arid regions. According to the modified Köppen climate classification (Peel et al. 2007), only narrow band along the southern and eastern shoreline areas have humid or temperate weather regimes effectively limiting where intensive agriculture can take place. Figure 1.1 illustrates the modified Köppen climate classifications of Australia provided by the Bureau of Meteorology (BoM). Given these extreme variations in climate regimes across the continent,

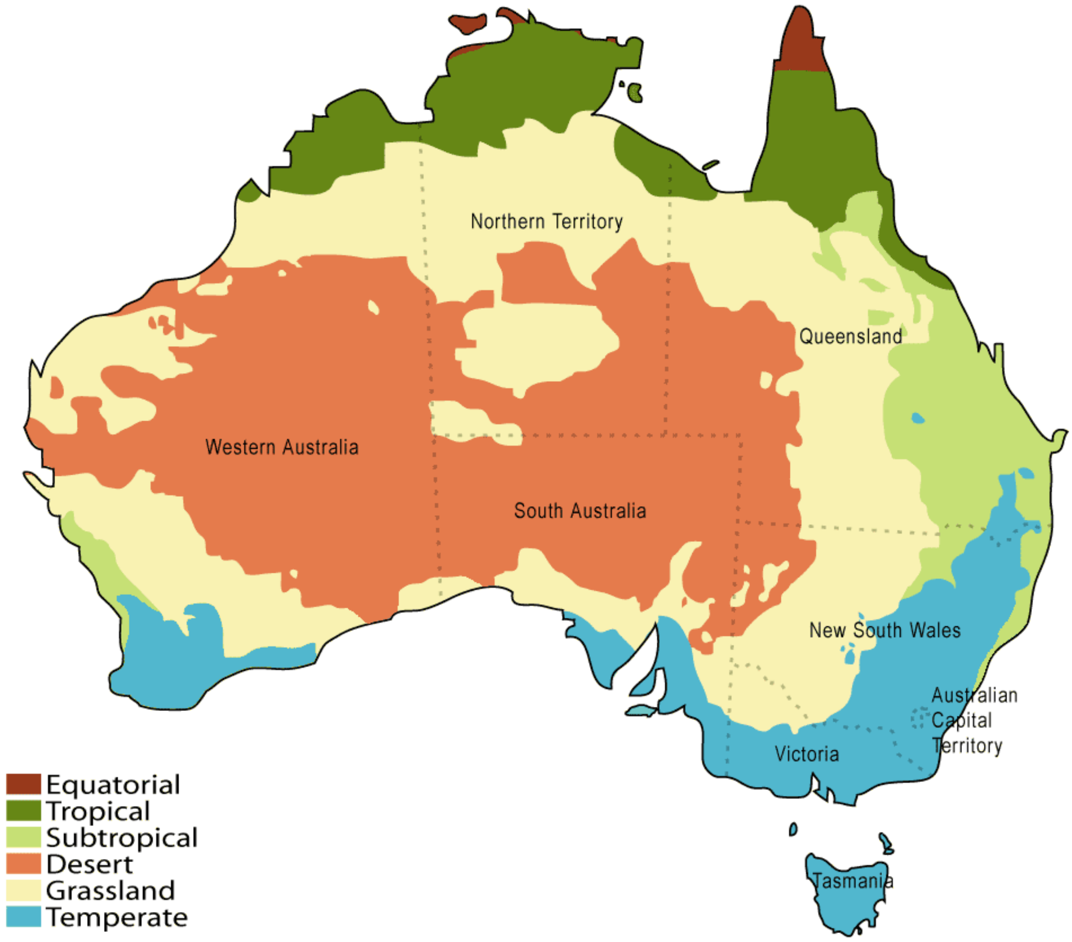


Figure 1.1. Climate map of Australia, based on the Köppen climate classification calculated based on a standard 30-year climatology from 1961 to 1990 (http://www.bom.gov.au/iwk/climate_zones/map_1.shtml).

most agricultural and livestock activities are concentrated within the south-east (Victoria and southern parts of New South Wales) and south-west, due to their fertile soils and relatively moderate temperatures.

Drought in Australia is a recurring and costly natural hazard which is defined by the Australian Bureau of Meteorology (BoM) as rainfall over a three-month period being in the lowest decile of what has been recorded for that region in the past. According to BoM records, a severe drought has occurred in Australia on average once every 18 years (Anderson 2013). In 2000, the most severe drought since settlement – known now as the millennium drought (1997-2010) – was detected and continued until late 2009/early 2010, resulting in total drought relief payments of AU$4.5 billion (Hobday and McDonald 2014). This drought was mainly due to significant precipitation deficits in the south-east and south-west of Australia, where most of the agricultural production is taking place. Figure 1.2 illustrates spatial clustering of

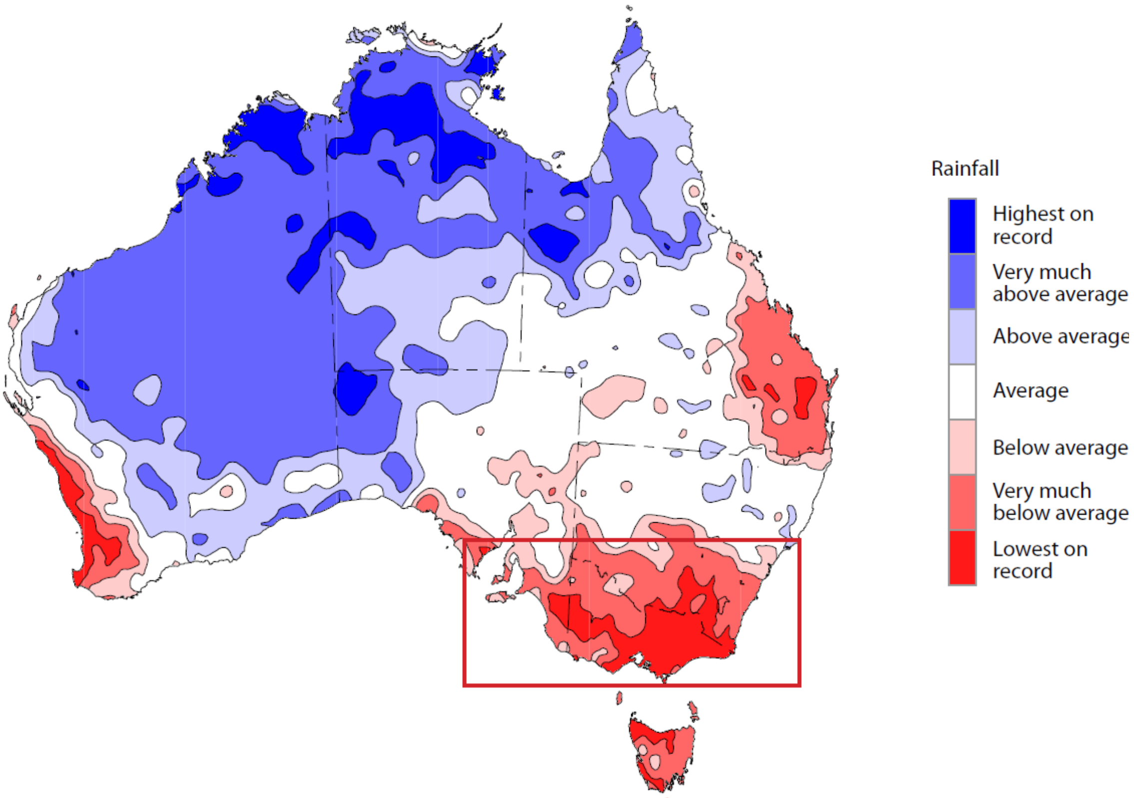


Figure 1.2. Regionalizing Australia and specifically in south east regions (depicted by the red rectangular) in terms of rainfall deciles for a 13-year period (1997-2009) using deciles calculated based on data from 1900 to 2009 (Taken from SEACI 2011).

Australia in terms of precipitation deciles for a 13-year period (1997-2009) using deciles calculated based on data from 1900 to 2009.

Water stress information is a crucial type of information in water resources management, because it is the main input for water resources planning. Given its key role in decision making processes, inaccurate and unreliable information can provide misleading information resulting into management decisions that can intensify the water stress conditions of a region. Over the past century, water stress and drought occurrences and also an increase in water consumption across different sectors have been documented in Australia. This led to a significant concern about the future of this country in terms of plant-available water, implicitly associating this with food security and bushfire hazard prediction skills. In addition, the pressure on arable land due to urbanization and climate change, as well as likely drier conditions and an increase in future extreme events will render this precarious situation even worse.

In order to cope with the natural aspects of this problem, water stress and drought assessments are of primary importance for water resources planning and management (DEPI 2009, FAO 2012). In fact, having sufficient and precise information can help authorities to mitigate the costs of water stress conditions. McColl and Young (2010) provided a list of different policies and tools in managing drought and structural adjustments. They presented that the current biophysical drought indicators, as input of drought (water stress) assessment frameworks, are poor tools for detecting and forecasting purposes. They also suggested that the water authorities have focused in the past more on financial assistance which can often do “more harm than good”, rather than investing into the implementation of drought preparedness to increase resilience.

Overall, this research has been motivated by the above issues to derive a more robust drought or water stress indicator to provide sufficient, timely, and reliable information for water resources planners and decision makers to mitigate and even possibly prevent the costs and damages caused by such natural disasters.

1.3 Statement of research

Many governmental organizations in most parts of the world are required to increasingly pay attention and divert funding to the growing threat of water stress. This has in the past been very apparent in water-limited environments, nevertheless, there are indicators suggesting that this will extend to all countries in the near future (Tsakiris et al. 2013). For instance, Australia has overall sufficient water resources for its human and ecosystem requirements, except that its water resources suffer from a spatio-temporal imbalance (NCEDA 2010) between its tropical and semi-arid regions, and their respective population densities, with the added problem that climate change is expected to decrease annual mean flows by 16 to 48 % by 2100 (Hennessy et al. 2007). While water demand will increase at the same time, this will put an additional strain on the existing water resources (Sahin et al. 2015). To mitigate the impact of the expected water deficits as a consequence of droughts, accurate and reliable spatio-temporal monitoring of water stress conditions is essential for water resources management and planning (Mu et al. 2013).

From a meteor-climatological point of view, the identification and measurement of droughts and water stress conditions is more complex than for other natural hazards, such as floods or storms, especially since drought and water stresses for an area are concepts largely based on deviations from historical conditions. Droughts and water stress conditions, however, can be slowly developing processes and their transition may last longer than other natural, short-term disasters (Logar and van den Bergh 2013). In terms of available data for evaluations, in-situ observations, satellite data, and modelled data are the three main data sources, though they come with their own limitations. For instance, due to the small number of stations and flux towers, useful *in situ* observations that can be employed for analyses are sparse in space and time. In contrast, the volume of modelled data is usually adequate for assessments, but not all that reliable due to model and forcing uncertainties. Satellite data can be available for any place in the world, even remote regions, however the presence of a variety of errors such as atmospheric interferences (i.e. during cloudy days) can be a considerable issue. Moreover, satellites only see a snapshot of the surface at any given time without any ability to forecast the conditions into the future without the use of a complex model framework.

In order to tackle those issues, drought indices have been derived and applied to detect, evaluate, and ultimately mitigate the impact of water stress conditions (AghaKouchak et al. 2015). A variety of drought indices have been introduced and developed subject to different factors and situations, such as data availability (quality-quantity), user expertise, required accuracy, and accessibility to sufficient hardware and software (Kousari et al. 2014). However, these indices often have deficiencies such as arbitrary metrics, lack of historic data, statistical inconsistencies, non-practicality, complexity, and non-comparability (Steinemann et al. 2015). In the recent past, it was repeatedly shown that a single drought index (SDI) is not adequate for consistently monitoring and identifying water stress conditions (Smakhtin and Hughes 2007) and that information from various drought indices needs to be considered simultaneously (Hao and Singh 2015). To achieve this aim, two main approaches have been taken into consideration here: i) framework-based indices (Sheffield et al. 2014, Steinemann et al. 2015, Wood et al. 2015, Mehran et al. 2015), and ii) multi-index drought models (MIDMs) (AghaKouchak et al. 2015).

The main goal of this thesis is to assess and evaluate the water stress and drought conditions via developing a more robust and reliable methodology based on multi-index drought models. The developed methodology will be able to appropriately respond to varying types of droughts in different regions of the world, subject to a wide range of climate conditions and landuse/landcover situations. Moreover, the proposed methodology will be easily transferable into simple mathematical functions, for its application across different types of industries.

1.4 Objectives and scope of research

The main questions addressed throughout this research are as follows:

1- To which hydroclimatic variables, as well as what type of data are the different drought types most sensitive to, and which of those should be chosen for an advanced system?

Hypothesis: The chosen hydroclimatic variables should cover the main elements of the hydrological cycle such as water availability and water stress based vegetation conditions. Therefore, the considered in-situ and satellite products should be able to provide reliable and accurate data for deriving individual drought indices based on the abovementioned hydroclimatic variables. It is assumed that a combination of such individual drought indices can convey better information for the development of a new combination based drought index than each single one alone.

2- Which statistical methods are more robust/reliable when evaluating and comparing drought indices?

Hypothesis 2: Through the employment of cluster analyses, as well as making use of multivariate methods (e.g. factor analysis, and independent component analysis) it will be possible to identify previously unknown statistical and probabilistic characteristics of a set of variables, and determine their similarities with each other.

3- Is there any methodology to combine individual drought indices to derive a more comprehensive one?

Hypothesis 3: By considering the dominant drought types for each single month for an area with a specified landuse situation, it will be possible to merge information from several drought indices to derive a new drought index which can be more reliable and comprehensive than a single common one.

4- Can a single mathematical function describe the water stress conditions adequately, drawing on individual information from other drought indices?

Hypothesis 4: Using optimization algorithms based on genetic programming, it is possible to derive the best explicit mathematical equations between the chosen single drought indices and combination based drought index for a specific area.

Considering the research questions, this research falls into the category of monitoring a terrestrial ecosystem in terms of water stress conditions. The main purpose is to develop a new methodology to address available shortcomings in evaluating drought (water stress) conditions. The proposed methodology will be able to monitor water stress conditions of such ecosystems by objectively linking different aspects of the system, such as water availability and vegetation conditions. The developed methodology will make use of advanced statistical methods (i.e. multivariate methods such as independent components analysis), and also considers the eco-meteorological characteristics (i.e. landuse, land-cover, and climate) of an area to determine the water stress conditions at each time step.

1.5 Outline of research

The project consists of three progressive parts. First, an appropriate set of single drought indices will be derived at a weekly temporal resolution for the Riggs Creek OzFluz tower site, and then comparatively evaluated with statistical methods commonly utilized in drought and water stress assessments. This includes the development of a new approach to comprehensively cluster those single, previously incompatible drought indices. Secondly, the new methodology will be applied to data collected at the three individual sites across different climate regimes (Riggs Creek: Temperate, Howards Springs: Tropical, and Alice Springs: Grassland) as a proof-of-concept study and subsequently compared with common single drought indices derived for those locations. The selection of these three sites ensures that the performance of the new methodology is being tested across different locations with a variety of climate conditions and landuse/landcover scenarios. Thirdly, the new methodology will be extended from single points to spatial scale for the state of Victoria in Australia, and the results assessed for a selection of weeks covering different seasons. Similar to the second part, the results derived from the new methodology will be compared to commonly used drought indices to evaluate the new indicator's accuracy and reliability in a spatial context. Finally, a sensitivity analysis of the proposed drought index will be undertaken to study the impact of time series length on the reliability and stability of the index.

The main research tasks completed are as follows:

a) Compiling and quality controlling an hourly database consisting of the required hydrometeorological data from different data sources such as in-situ observations, satellite data/information, and estimated data from global models;
b) Implementing quality control of hydrometeorological data and filling probable gaps by using linear/nonlinear regression methods;
c) Providing weekly time series for an appropriate set of common drought (water stress) indices;
d) Clustering common drought indices to select the two groups which include water availability and vegetation conditions based indices;
e) Aggregating the indices that fall into each of the two clusters (formed at the previous step) by making the use of multivariate methods;
f) Deriving the final weekly proposed drought index by integrating the two aggregated

indices;

g) Demonstrating the efficiency of the developed methodology on three single sites (three OzFlux tower sites) in Australia as well as spatially across the state of Victoria in Australia.

h) Implementing a sensitivity analysis over the performance of proposed drought index to the length of observations by making use of long-term modelled data.

1.6 Structure of thesis

In addition to the Introduction chapter, six more chapters form the structure of this thesis as follows:

Chapter 2: Literature review

This part consists of common approaches which have been introduced and applied in water stress evaluations. Moreover, it covers information over a wide range of common and state-of-the-art drought indices and also a discussion over their strengths and limitations. Finally, knowledge gaps and methodological development are explained.

Chapter 3: Statistical analysis of short-term water stress conditions at the Riggs Creek OzFlux Tower Site using previously introduced drought indices

Within this chapter, the most common statistics-based methods to compare previously introduced drought indices are explained, followed by a new approach for clustering drought indices. Later, all methods are applied to compare their drought indices at the Riggs Creeks OzFlux tower site at a weekly temporal resolution. The results of this chapter are also presented in form of a journal paper as follows:

- Azmi, M., Rüdiger, C., Walker, J. (2016) Statistical analysis of short-term water stress conditions at Riggs Creek OzFlux tower site. *Theoretical and Applied Climatology*. DOI 10.1007/s00704-016-1901-z.

Chapter 4: Developing a new data fusion based drought index

Within this chapter, new proposed methodology is elaborated and then implemented over three case study areas of Riggs Creek, Howards Springs, and Alice Springs OzFlux tower sites. The weekly results are then compared to commonly used drought indices to examine the performance of the new index. The results of this chapter is also presented in form of a journal paper as follows:

- Azmi, M., Rüdiger, C., Walker, J. (2016) A data fusion-based drought index. *Water Resources Research*, 52: 2222-2239, DOI:10.1002/2015 WR017834.

Chapter 5: Spatial monitoring of weekly water stress conditions in Victoria, Australia, using the DFDI

Here, the methodology proposed in the previous chapters is developed from a single point scale to a coarser spatial scale. The method is then applied to assess the water stress conditions across the state of Victoria in Australia at a weekly temporal resolution. Similar to the previous chapter, the results derived from the proposed methodology are compared with the commonly used drought indices. The results of this chapter are also presented in form of a journal paper as follows:

- Azmi, M., Rüdiger, C., Walker, J. Spatial monitoring of weekly water stress in Victoria, Australia. *Water Resources Management*. Under Review.

Chapter 6: Comparative evaluations of short-term OzFlux tower observations and long-term AWAP data in weekly water stress monitoring

To identify whether the use of a short-term dataset has negatively affected the quality of the new index, the statistical static (i.e. quartiles) and dynamic (i.e. correlations) characteristics of the proposed drought index are derived from short-term observations and long-term modelled data and those are compared with each other. First, the differences and similarities in the statistical characteristics of the short-term observations and the long-term modelled data are discussed, and then the modelled data are employed to derive the time series of the proposed drought index for the locations of the three OzFlux tower sites introduced in Chapter 4. The time series of the proposed drought index derived from the long-term modelled data is evaluated against the time series derived in Chapter 2. The results of this chapter is also presented in form of a journal paper as follows:

- Azmi, M., Rüdiger, C., Smith, A.B., Walker, J. Comparative evaluations of short-term OzFlux tower observations and long-term AWAP data in weekly water stress monitoring. *Geophysical Research Letters.* Under Review.

Chapter 7: Conclusions and future recommendations

This chapter elaborates the findings and conclusions of this research and outlines a range of recommendations for future directions and work.

Chapter 2
Literature Review

2.1 Overview

Throughout this literature review, background information on drought and water stress related issues will be presented, followed by identifying the significance of research in water stress monitoring with a particular focus on Australia. This is further underlined by discussing the key motivations for introducing a new index to more comprehensively evaluate the water stress conditions in a specific region, rather than applying state-of-the-art approaches. At first, the current methods to determine drought and water stress conditions are explained, again within an Australian context. This is followed by an elaboration of the most commonly used drought indices, as well as their capabilities and deficits. Ultimately, the knowledge gaps and required future research are identified and highlighted.

2.2 Approaches to identify and address water stress and drought conditions

Ecosystem response to global and regional climate changes due to the influence of these deriving forces on the water availability and surface air temperature. This is mainly due the relationship between the vegetation activities, vegetation water consumption, and available energy in form of air temperature in an ecosystem. Seddon et al. (2016) stated that at a global scale, terrestrial ecosystems located in different regions respond to changes in water and energy conditions in various ways. Therefore, one of the remaining issues in hydrometeorological research is to identify the sensitivity and resilience of an ecosystem to the changes of water and energy conditions; which indirectly can impact on human well-being (Seddon et al. 2016).

Water stress directly and significantly affects water systems and terrestrial ecosystems. In addition, the processes to identify, measure, and quantify those events is more difficult and complex than for other natural hazards, such as floods or storms, due to their unique features

(Logar and van den Bergh 2013), which can vary significantly between regions. For instance, drought and water stress concepts are relative, depending on deviations from historical conditions in a certain area. They are known as creeping phenomena, the onset and ending thresholds are blurred, and usually last longer than other natural disasters (Logar and van den Bergh 2013). While a variety of definitions and explanations have been presented to date, it can be stated that water scarcity due to insufficient precipitation (as the main input element of a water system) and or high water consumptions (i.e. evapotranspiration and human related consumptions) are likely among the most accessible definitions (Barua 2010).

Water stress conditions in an ecosystem can be defined as a short-term (i.e. daily, weekly), sustained and regionally extensive occurrence of below average natural water availability. Moreover, in case of severe water stress conditions over a long period of time (i.e. monthly to seasonal, or beyond), this natural phenomenon is defined as a drought condition (Svoboda et al. 2002, Tallaksen and van Lanen 2004, Svoboda et al. 2015).

According to Wilhite (2005), there are four main types of drought: i) hydrological droughts occur when the quantity of surface and ground water is less than during average historical conditions; ii) meteorological droughts are related to an insufficient quantity of precipitation; iii) ecological/agricultural droughts focus on the deficit of soil moisture and the available water inside the plants for evapotranspiration and physiological activities; and iv) socio-economic droughts are known as water shortages which affect the health, well-being, and general quality of life of humans.

The above drought types may occur at different times, i.e. there may be time lags between their occurrences. Rasmusson et al. (1993) showed that the presence of insufficient precipitation (meteorological drought) in a specific area for a period of time can lead to the sharp oscillation of the soil water content (agricultural drought), later impact on streamflow (hydrological drought), and eventually on groundwater contents (Figure 2.1) become observable. The result of long-term severe meteorological, hydrological and agricultural droughts in a region leads to socio-economic drought which can cause social disasters such as health issues and mandatory migrations.

Half of the Earth’s terrestrial surfaces are susceptible to water stress and drought conditions (Kogan 1997, Mishra and Singh 2010) and unfortunately almost all of the major agricultural lands are located in those areas (USDA 1994). It has been extensively shown that amongst all natural disasters that happened in the 20th century, water stress and droughts have been the most

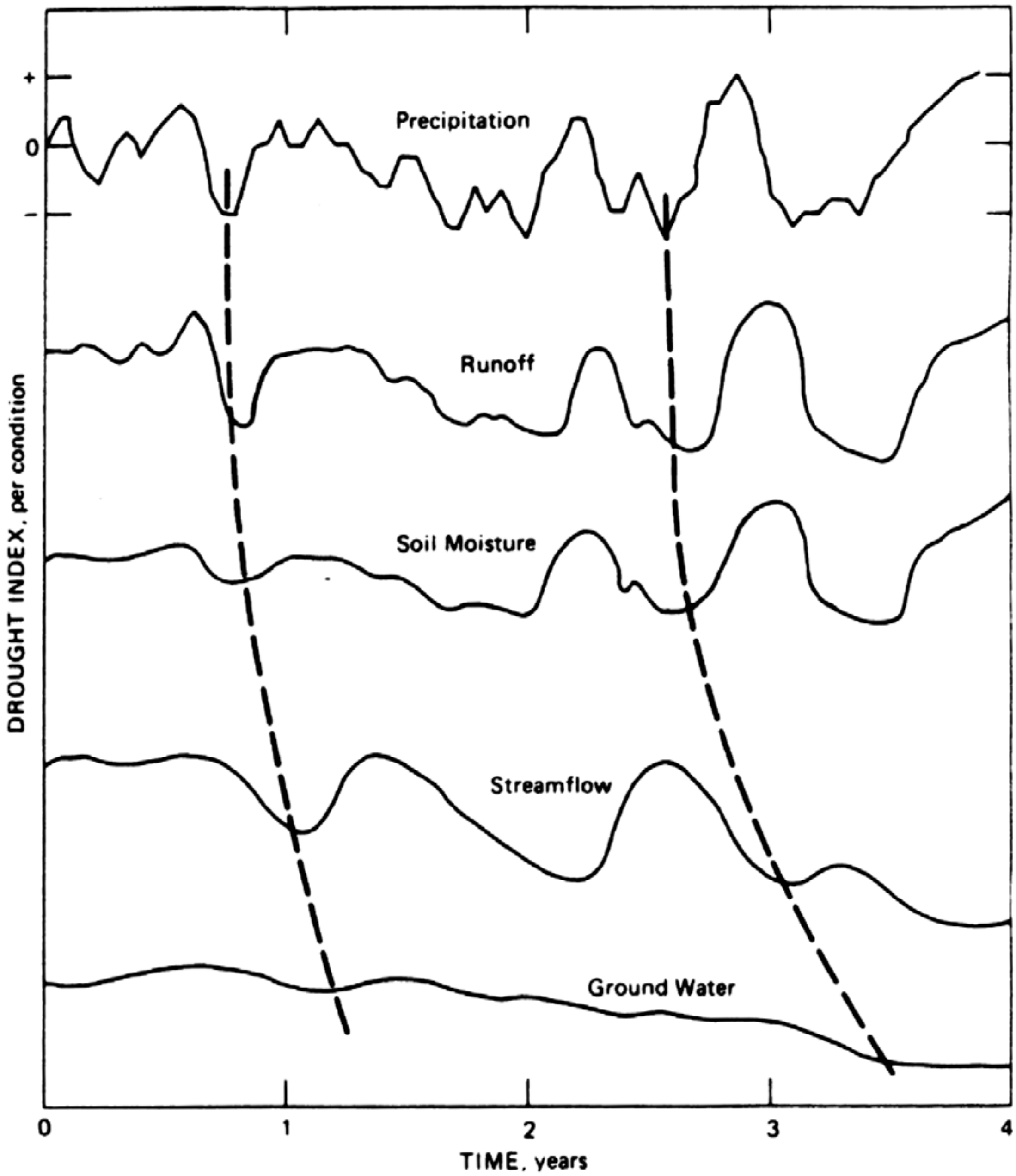


Figure 2.1 Time lags between different elements of a water system affected by a drought condition (Taken from Rasmusson et al. 1993).

costly and damaging (Bruce 1994, Environmental Canada 2004, Cook et al 2007). For instance, the impact of a national scale drought in the USA in 1988 has been assessed at around US$ 40 billion, which is 2 to 3 times greater than the loss caused by the 1989 San Francisco earthquake (Riebsame et al. 1990). In another example, since 1991, the annual economic impact of droughts and water stress in Europe has been €5.3 billion, with the economic damage of the 2003 drought in Europe amounting to at least €8.7 billion (European Commission 2007).

Australia's environment has become increasingly susceptible and vulnerable to water stress conditions and from 1995 to 2009 suffered through some of the most extreme drought conditions since European settlement (Gallant et al. 2013). The profound impacts of the mentioned drought period became evident at various temporal and spatial scales across most parts of southern and eastern Australia (Risbey 2011). For instance, the 2006 drought alone reduced the national winter cereal crop by 36% which cost around AU$ 3.5 billion. Over the past decades, a variety of methodologies, approaches, and policies have been introduced and applied to mitigate and even prevent such costs, however, direct impacts and side effects of this phenomenon shows the necessity of more research in this field of study.

Despite early research showing the need for pro-active measures, the Australian government's response has generally been reactive (Bond et al. 2008), which essentially led to an accumulation of a significant amount of cost for drought relief payments. For instance, such payments for the 1992–1999 period were around AU$100 million per year (DoE&H 2001b). Furthermore, the direct loss of agricultural production and financial assistance for the 2002 and 2006 droughts totalled AU$ 7.36 billion and AU$ 6.2 billion, respectively (ABS 2004, ABARE 2006). Being prepared for such drought conditions is essential to reducing costs, as some of those are preventable, or at least can be mitigated, if there were adequate measures in place. However, to achieve this, a drought prediction framework needs to be in place that consistently and comprehensively describes both water and vegetation conditions.

A number of studies in Australia (Meinke and Hochman 2000, Everingham et al. 2002, Meinke and Stone 2005) have shown that drought prediction systems may give clear indications to agricultural decision makers on a potential drought's severity and impact, even when considering the difficulties and uncertainties of predicting weather and climate conditions. However, more accurate and reliable information is still needed to prepare for droughts and to decrease costs and mitigate negative impacts as a consequence. This shortcoming underlines the need for a proactive management approach including an accurate high-resolution prediction system in Australia. This system should be able to provide tools to decision makers to at least mitigate, if not prevent, the direct consequences of severe droughts on the Australian ecosystem and also their negative side effects on the economy (Bond et al. 2008, Mu et al. 2013).

The core of each drought monitoring/prediction system are indicators such as a drought index (DI), which theoretically should be able to detect, evaluate, and ultimately help to mitigate impacts of waters stress conditions in an area (AghaKouchak et al. 2015). The different variables of a DI often describe the water stress conditions of an area. A DI is generally derived based either on a single or a set of hydroclimatic variables. Depending on the objective of a specific drought study, different DIs may be introduced and applied (Hayes et al. 2012). As a general rule, an ideal DI should be able to represent all different aspects of a drought (or water stress) event in a given location as precisely and accurately as possible. Conversely, an inappropriate DI provides insufficient and sometimes incorrect information, which may lead to misinterpretations of the real water stress conditions (Kiem et al. 2016, Van Loon 2015, Van Loon et al. 2016a , 2016b).

In addition to *in situ* observations and also satellite information, significant amounts of money are spent annually in Australia on a variety of local data collection efforts such as flux towers. Nonetheless, monitoring and forecasting water stress and drought conditions are still mainly investigated based on meteorological variables, such as precipitation, which does not provide sufficient integrating information to comprehensively evaluate the terrestrial ecosystems (Mpelasok et al. 2008, Risbey 2011), as it does not provide any direct information on the vegetation state or lower soil moisture contents. Thus, the information provided to water resource managers has been of limited use (Wilhite 2000). Consequently, the main motivation of the current research is to introduce a new and more robust methodology to evaluate weekly water stress conditions of the terrestrial ecosystems in Australia, making use of a range of observations covering hydrology and plant-physiology.

2.3 Water stress and drought indices

A variety of DIs have been introduced and applied in drought monitoring and forecasting systems since the 1950s. Those DIs are either based on a single or an incorporation of a range of hydrometeorological variables, however describe mostly very specific conditions. A review of the most common indices is presented in the sections below.

2.3.1 Precipitation based DIs

In the early 1950s, Kohler and Linsley (1951) and also recently Crow and Ryu (2009) introduced and applied the Antecedent Precipitation Index (API) as an indicator for flood and drought evaluations, which is expressed as

$$API = \sum_{t=-1}^{-i} P_t k^{-t} \ , \qquad (2.1)$$

where i is the number of preceding days, P is the amount of precipitation during day t, and k is a decay constant which can range from 0.80 to 0.98, as a function of the soil and general climatic conditions, but which is usually set to 0.9, as suggested by Heggen (2001).

Later, Gibbs and Maher (1967) proposed the Percent of Average (PA) rainfall (runoff or streamflow) as an applicable DI, where the actual precipitation (runoff) for a given period is divided by the historical average of the considered variable. Obviously, "simplicity" and "easy to measure" are the main strengths of this DI. However, this index is independent from other environmental and ecological factors and does not consider seasonal differences in real water consumptions (i.e. actual evapotranspiration), which can lead to misinterpretation in the water stress conditions in different locations.

The Deciles Index is a more complete version of PA (Gibbs and Maher 1967) and is commonly used in Australia to evaluate meteorological drought conditions. To derive this indicator, the distribution of occurrences over a long-term precipitation (runoff) record are divided into tenths of its distribution. The scale of the Decile Index is from 1 to 10, for which extreme droughts are located in the top 10%–20%. Although, the PA and Deciles Index have similar merits, in order to have accurate evaluations a long data record along with local hydrometeorological observations is required, as well as precise frequency distributions. Ignoring those can lead to a reduction in the accuracy and reliability of the water stress evaluations of an area.

The Standardized Precipitation Index (SPI) (McKee et al. 1993, McKee and Edwards 1997) is the most widely applied precipitation-based DI due to the data availability of its input variable (precipitation), compatibility with different timescales, comparability across different regions, and the simplicity of its calculation. To create a time series of SPI for a specific area, the long-term precipitation record is fitted to a probability distribution (generally the Gamma distribution), which is then transformed into a standard Gaussian distribution. Similar to PA, beside the requirement of the long record, the main shortcoming of SPI is neglecting ecological and landuse conditions, and also water consumption of a terrestrial ecosystem during the process of water stress evaluations (Kallis 2008). In other words, evaluating real water requirements of a system only based on the input factors (i.e. precipitation, snow) can lead to misjudgements in water resources planning. This drawback has persuaded researchers to follow models and indices which are on the basis of the water balance in an ecosystem.

2.3.2 Water balance model-based DIs

The most common DIs derived from water balance models are the Palmer Drought Severity Index (PDSI) (Palmer 1965), and Crop Moisture Index (CMI) (Palmer 1968), as discussed below.

The PDSI was initially developed based on a two-layer soil water balance model to estimate the plant available water, subject to the inputs of temperature, precipitation, and water content of the soil for a specific time scale. On a monthly scale, PDSI can be computed via the comparison between actual monthly precipitation and the desirable precipitation which can sustain a normal climatological condition in a region (Z). Finally, the value of PDSI for the i^{th} time step can be determined as follows

$$PDSI_i = 0.897 \times PDSI_{i-1} + 0.333 \times Z_i \ , \tag{2.2}$$

where the weights of 0.897 and 0.333 are “duration factors” which have been empirically derived based on the original study areas in central Iowa and western Kansas, in the USA. This means that Equation 2.2 can only be utilized in regions similar to the above areas in terms of landuse, dominant vegetation, and climate conditions. To apply the PDSI across different regions of the world, the above coefficients should be re-calculated to have a consistency between the model and the conditions of an area (self-calibrating PDSI; Wells et al. 2004). While PDSI is one of the most commonly used DIs on a global scale, the original methodology of this index does not work appropriately in mountainous and snow covered areas (Burke et al. 2006).

Finally, the CMI was introduced based on an experimental framework for short-term drought evaluations (i.e. weekly) across crop-producing regions. The CMI is the summation of the evapotranspiration anomalies and the wetness indicator (which is based on soil moisture). The CMI values are always equal or greater than zero, and are at or near zero at the start of a growing season, while increasing depending on the water stress conditions of an area. The main limitation of this index is its inappropriate performance during long-term drought (water stress) conditions (Hayes 2000). In fact, the CMI tends to rapidly respond to changes in short-term conditions which provides misleading information about long-term situations. For instance, an unexpected, significant rainfall event lasting for a few of days during a 5-year drought period can result in the CMI values to indicate wet conditions for a 2-week period, while the actual long-term drought in that location still persists.

2.3.3 Surface water based DIs

As mentioned above, the PDSI was developed based on the difference between actual monthly precipitation and the desirable precipitation, and was originally meant for regions where local precipitation events were the sole or primary sources of the landscape moisture. However, in mountainous regions, such as Colorado in the western United States, where the main sustainable water resources are in the rivers and reservoirs, the PDSI cannot reflect the water stress conditions appropriately. Therefore, the Surface Water Supply Index (SWSI) (Shafer and Dezman 1982) with a range between -4.2 and 4.2, based on the combination of four hydro-climatological variables of snow water content, streamflow, rainfall, and storage reservoir volume was developed. It is worth noting that during the summer, streamflow replaces the snow water content.

The SWSI equation is expressed as:

$$SWSI = \frac{[(a \times P(rs) + b \times P(sf \;\; or \;\; sn) + c \times P(pr)) - 50]}{12} \,, \tag{2.3}$$

where, P represents the non-exceedance probability (%) of historical records of reservoir storage (*rs*), streamflow (*sf*), snowpack (*sn*) and precipitation (*pr*). The parameters of a, b and c are coefficients that determine the contribution of that component to surface water resources, but which are specified subjectively. The limitation of the SWSI is that its results and processes are quite subjective and depend on the regional conditions and the experience of the user. Moreover, in case of adding new stations and or changing water resources policies, the entire algorithm must be re-formulated.

One of the main elements of the SWSI is the snow water content, which plays a significant role in regions such as northern Europe where snow is a significant contributor to the water resources in the form of snowmelt. A deficit in the winter snowpack potentially leads to hydrological, and consequently, agricultural droughts in spring and summer.

2.3.4 Soil moisture-based DIs

The first introduced soil moisture-based drought index was the Keetch-Byram Drought Index (KBDI) (Keetch and Byram 1968). However, originally developed for fire potential assessments, the KBDI essentially measures the amount of precipitation necessary to return the soil to its field capacity. The original KBDI ranges from zero (completely saturated) to 800 units (wilting point for vegetation) which is indicative of the moisture regime ranging from zero to ~20cm (eight inches in the original document) of the top soil layer. KBDI is in fact an empirical approximation of soil moisture using maximum daily temperature, total daily precipitation and average annual precipitation, and was developed based on the climate and landuse conditions of various regions in the USA. However, it has been used for other parts of the world without any re-calibrations through which associated errors are inevitable (Ganatsas et al. 2011).

Two more soil moisture-based DIs are the Standardized Soil Moisture Index (SSMI) (Hao and AghaKouchak 2013) and the Soil Moisture Percentile (SMP) (Wang et al. 2009). Time series of these two DIs are calculated similarly to the algorithm used for the SPI. The soil moisture observations as inputs to these DIs can be obtained from *in situ* observations, land surface model simulations, or satellite estimations (Wang and Qu 2009). The main limitation of soil moisture-based indices is the lack of sufficient and reliable datasets (either spatially or temporally sparse), as satellites are generally providing low-resolution data, and *in situ* stations are mere point measurements and may not be representative of the surrounding areas (Yee et al. 2016).

Another issue with this type of indices is that they only consider the water content in top soil layer (usually to the depth of 1 m) which necessarily cannot be as representative of water stress in an area. For instance, if a region with landuse of forest has relatively light soil texture, the water easily percolates into deeper layers. In this case, soil moisture based indices detect water stress conditions at top layers, however, the terrestrial ecosystem has sufficient water content to be taken up by the roots of trees.

Considering the abovementioned limitations, researchers have found a tendency towards employing soil moisture indices as auxiliary indicators to evaluate and even validate the behaviour of newly developed DIs. For instance, Sur et al. (2015) recently considered a satellite-based soil moisture index as a reference to validate the monthly performance of two new DIs across the Korean peninsula: an energy-based water deficit index, and the standalone Moderate Resolution Imaging Spectroradiometer-based evaporative stress index. The results showed a moderate linear correlation between the soil moisture behaviour and the two other developed DIs. They finally stated that due to the uncertainties and errors of satellite data and also the complexities of the process of deriving the proposed satellite based DIs, the two proposed indices may only be valuable for monthly, regional water stress monitoring when the climate data are not available or are sparsely distributed.

2.3.5 Evaporation-precipitation-based DIs

In addition to soil moisture as an input variable to calculate DIs, evapotranspiration is an effective element of the water and energy cycles, reflecting mass and energy exchange between ecosystems and the atmosphere (Wang and Dickinson 2012). Amongst a number of evapotranspiration-based DIs, the Crop Water Stress Index (CWSI) (Jackson et al. 1981) and Evaporative Stress Index (ESI) (Anderson et al. 2013) are two of the most commonly used water stress indicators. The CWSI is calculated based on the ratio of the actual ET (AET) to the potential ET (PET) expressed as

$$CWSI = 1 - \frac{AET}{PET}, \tag{2.4}$$

The *PET* depends on climate conditions of a region, however, *AET* reflects the actual water consumptions of a region’s dominant vegetation. Therefore, two locations with the same climate condition (same PET), yet with a different landcover (different AET) would have different thresholds for CWSI. To address the mentioned issue, CWSI is transformed into a normalized index called ESI expressed as

$$ESI = \frac{\left(\frac{AET}{PET}\right)_i - \left(\frac{AET}{PET}\right)_{mmi}}{\sigma}, \tag{2.5}$$

where $\left(\frac{AET}{PET}\right)_i$ is the ratio of actual evapotranspiration to potential evapotranspiration at month i, the subscript *mmi* refers to the monthly mean of this ratio for month i; and σ is the standard deviation of all its values for month i.

Furthermore, Tsakiris et al. (2007) introduced the Reconnaissance Drought Index (RDI) for a given month (k) of a year expressed as:

$$r_k = \frac{\sum_{i=1}^{k} P_i}{\sum_{i=1}^{k} PET_i}, \tag{2.6}$$

where P_i and PET_i are precipitation and potential evapotranspiration for month i, respectively, and k can be determined at 3, 6, 9 or 12 (months) intervals. In case of choosing k equal to 12, the RDI would be equivalent to the Aridity Index defined by the Food and Agriculture Organization (FAO) (Tsakiris et al. 2007).

The Standardized Precipitation Evapotranspiration Index (SPEI) was developed as a multi-scalar drought index, which is based on the standardized difference (D) between monthly precipitation and potential evapotranspiration (Vicente-Serrano et al. 2010). The process of standardizing D is similar to what is required for the SPI (see section 2.3.1). Considering the fact that RDI and SPEI do not use actual ET, their input data can appropriately be provided either from satellite-retrieved air temperature data (Dalezios et al. 2012) or in-situ observations.

The DIs of this category are derived based on fixed mathematical equations linking precipitation and evapotranspiration variables. However, they have been used primarily for water stress detection, and due to the lack of a variety of hydrometeorological variables from different drought types (soil moisture from agricultural drought type or surface water from hydrological drought type), they cannot be regarded as comprehensive indices.

2.3.6 Groundwater and terrestrial water storage-based DIs

Because of existing complexities in evaluating water storage across a terrestrial ecosystem, including aquifers, only a few drought indices such as the Terrestrial Water Storage (TWS) have been proposed and applied to assess an integrated measure of water availability. TWS is based on the combination of land-surface models with the Gravity Recovery and Climate Experiment (GRACE) data (Rodell and Famiglietti 2002). TWS provides anomalies in the total water column, consisting of groundwater, moisture stored in vegetation, soil moisture, surface water, and snow/ice (Rodell 2012). Having TWS, groundwater changes can then be approximated via a balance model of

$$\Delta G = \Delta TWS - \Delta SM - \Delta SWE, \tag{2.7}$$

where Δ represents changes to a variable; and SM and SWE represent soil moisture and snow water equivalent, respectively (Rodell et al. 2007). The GRACE-based data and information have been utilized for drought monitoring over different regions, including the Canadian Prairie (Yirdaw et al. 2008), Australia (Leblanc et al. 2009, vanDijk et al. 2013), the Amazon River basin (Chen et al. 2009), and western and central Europe (Li et al. 2012). Also, for the 2011

Texas drought, Long et al. (2013) found that TWS was a valuable index in water stress and drought monitoring.

Nevertheless, the main shortcomings of this index are the limited period of data (since 2002), as well as their coarse temporal (~30 days) and spatial resolution (>150000 km^2; Houborg et al. 2012). Using data with low spatio-temporal resolutions may lead to a loss of information over an area, as it integrates the positive and negative anomalies over time, essentially ignoring the natural complexities of the underlying geographical and landcover features within a terrestrial ecosystem. In turn, this may result in errors in the final interpretation due to an imprecise picture of the real conditions. In order to circumvent those problems, data assimilation methods have been shown to result in acceptable scales (Zaitchik et al. 2008, Thomas et al. 2014). The downscaled data can then be used as input to local water stress monitoring and forecasting systems for water resources and agricultural water managements.

2.3.7 Vegetation-based DIs

Vegetation-based DIs can usually evaluate ecological and health situations of vegetation in a region affected by water stress conditions (Nemani et al. 2009). However, the main drawbacks of these indices are during cold seasons when vegetation is dormant, or because of extensive cloud cover during rainy seasons, when the surface is obscured. Also, external factors such as crop maturity or senescence, which may be unrelated to hydrological conditions, may significantly influence vegetation indices (Kallis 2008).

The Normalized Difference Water Index (NDWI) (Gao 1996) is a satellite-derived index from the Near-Infrared (NIR-MODIS Band 2) and Short Wave Infrared (SWIR- MODIS Band 6) channels and can be derived from instruments on satellites such as Terra/Aqua and Landsat. Sentinel-2 and others. This index can be a good indicator in retrieving the vegetation water content (Ceccato et al. 2001) and has a negative correlation with leaf water content (Tucker 1980). NDWI can be expressed as

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR}, \tag{2.8}$$

The main limitations of NDWI are the effects of soil background for partially vegetated covers and also snow cover on the accuracy of the NDWI's results (Gao 1996, Ceccato et al. 2001). AS mentioned above, other factors can cause a decrease in NDWI, either due to changes in land cover or the impact of pests and diseases, which all can be responsible for large variations of NDWI. The latter issue would also be true for other vegetation-based drought indices.

The Normalized Difference Vegetation Index (NDVI) introduced by Rouse et al. (1974) is one of the most applied vegetation based DI. The NDVI utilizes spectral measurements to detect greenness and health of the plants, indicating a sufficient level of plant-available water for evapotranspiration (Karnieli et al. 2010). Therefore, this index can implicitly represent the water stress conditions of vegetation in a terrestrial ecosystem and is expressed as

$$NDVI = \frac{NIR - VIS}{NIR + VIS}, \tag{2.9}$$

where *VIS* is the spectral reflectance measurements acquired in the visible (Red) regions (MODIS band 1). However, for high NDVI values, it tends to saturate and therefore detailed information on the vegetation is not accessible anymore.

Some other vegetation-based DIs were subsequently derived from the original NDVI, with the most common including the Vegetation Condition Index (Kogan and Sullivan 1993), Temperature Condition Index (TCI) (Kogan 1995), Vegetation Health Index (VHI) (Kogan 1995), Standardized Vegetation Index (SVI) (Park et al. 2008), and Perpendicular Vegetation Index (PDI) (Ghulam et al. 2007).

For instance, the value of SVI for time *i* can be calculated based on transferring the NDVI's values into Z scores as follows:

$$SVI_i = \frac{NDVI_i - NDVI_{Ave}}{\sigma}, \tag{2.10}$$

where $NDVI_{Ave}$ and σ are the mean and standard deviation of a *NDVI* time series, respectively.

The VCI is a standardized form of the NDVI in which

$$VCI_i = \frac{NDVI_i - NDVI_{\min}}{NDVI_{\max} + NDVI_{\min}}, \tag{2.11}$$

where $NDVI_i$ is the value of *NDVI* at the i^{th} time step (and or grid point); while $NDVI_{max}$ and $NDVI_{min}$ are the maximum and minimum NDVI for a given time series (and or grid points), respectively. However, this is not easily applicable for northern hemisphere vegetation types, where leaves are often shed from trees.

The TCI applies an equation similar to that used for the VCI using brightness temperature (Tb) reflected from the surface as follows:

$$TCI = 100 \times \frac{Tb_{\max} - Tb}{Tb_{\max} - Tb_{\min}}, \tag{2.12}$$

where Tb_{max} and Tb_{min} are the maximum and minimum brightness temperature (MODIS Band 4) for a given time series (and or grid point), respectively.

Further, a linear regression of VCI and TCI leads to deriving

$$VHI = \alpha \times VCI + (1-\alpha) \times TCI\ , \tag{2.13}$$

where α refers to the relative contribution of the *VCI* and *TCI*, and which is determined subject to an expert's understanding and knowledge of the ecosystem. Similar to the NDVI, a weekly time series of VHI can be utilized for short-term water stress monitoring in different regions (Seiler et al. 1998, Kogan 2001).

The TCI and VHI were initially derived based on a hypothesis that increasing temperatures negatively affect vegetation activity and increase the water stress inside the vegetation. However, Karnieli et al. (2006) showed that for high latitude regions of the Northern hemisphere, in that case Mongolia, rising temperature positively influences the vegetation physiological activities, because of the predominantly cold climate. Therefore, the VHI values should be interpreted with caution in such regions, as the optimal temperature range for gross primary production is unlikely to be exceeded and plant water is generally available. This shows the lack of adaptability of VHI and TCI as individual indices.

As a further development, the PDI uses a regression model between spectral measurements to determine water stress conditions in an area expressed as:

$$PDI = \frac{1}{\sqrt{M^2+1}}(VIS + M \times NIF)\ , \tag{2.14}$$

where *VIS* and *NIF* are visible (MODIS band 1) and near infrared (MODIS band 2) reflectance measurements, respectively; and *M* is the slope of the soil line (line connecting B and C in Figure 2.2) in a NIR-VIS spectral feature space. The values of PDI range from 0 to 1, with larger values representing higher water stress. Nevertheless, the comparative results of PDI

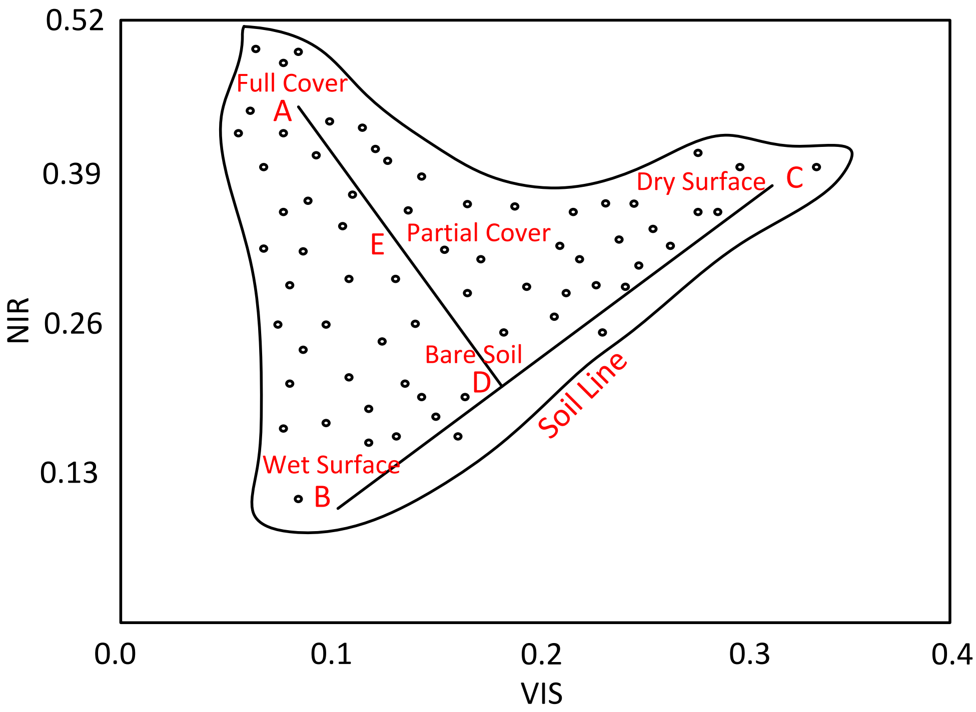


Figure 2.2. A sketch map of PDI

with ground evidences show that this index is found to lack accuracy over densely vegetated fields such as pastures and agricultural fields (Ghulam et al. 2008). Moreover, the index assumes land cover and soil types are homogeneous across the satellite image in the area of study (Ghulam et al. 2008), which is not always certain.

2.3.8 Combination- and aggregation-based DIs

There are a variety of studies (Wilhite 2000, Hao and Aghakouchak 2013, AghaKouchak et al. 2015, Hao and Singh 2015) which have indicated that no single DI can perform appropriately in all circumstances, and that most individual DIs cannot comprehensively evaluate the water stress conditions of a single terrestrial ecosystem.

To overcome these drawbacks, a number of studies suggested to make use of combination and aggregation approaches to derive new DIs, ultimately leading to more accurate and reliable indices than individual DIs alone (Keyantash and Dracup 2004, Balint and Mutua 2011, Barua et al. 2012, Zhang and Jia 2013, Li et al. 2014). The main aim of these approaches is to develop an inclusive DI that is more accurate and reliable than the individual DIs. Some of the most common combination- and aggregation-based DIs are presented below.

1- Blending objective and subjective indices

This method blends objective indicators with subjective information from local experts/institutes to develop more comprehensive indices. The two most applied models based on this method are the U.S. Drought Monitor (USDM) (Svoboda et al. 2002) and the North American Drought Monitor (NADM) (Lawrimore et al. 2002). The NADM is a continent-scale program to cover Canada, Mexico, and the continental USA, whereas the USDM focusses almost entirely on the USA.

The USDM project commenced in 1999 with the goal of spatial, weekly monitoring and tracking of water stress conditions across the USA by benefitting from various types of information and data. It consists of water stress indicators/proxies, regional and local climatic observations, water expert opinions across the country, and the output of numerical models.

After integrating all information and data, the USDM bins the outcome into one of five drought categories, including abnormally dry (D0), moderate drought (D1), severe drought (D2), extreme drought (D3), and exceptional drought (D4). The methodology of categorization is on the basis of a percentiles concepts for which D0, D1, D2, D3, and D4 correspond to the 30^{th}, 20^{th}, 10^{th}, 5^{th}, and 2^{nd} percentiles, respectively. The main considered single DIs and hydrometeorological variables within the USDM are PDSI, SPI, KBDI, SSMI, 7-day average streamflow, and precipitation anomalies which are combined based on input from local and national experts across the US. To blend the individual DIs, two linear regressions are derived for short- and long-term monitoring in which the coefficients are fixed as follows:

1- short-term: 35% Palmer Z-Index; 25% 3-Month Precipitation; 20% 1-Month Precipitation; 13% Climate Prediction Center Soil Moisture Model; and 7% Palmer (Modified) Drought Index.

2- long-term: 25% Palmer Hydrologic Drought Index; 20% 12-Month Precipitation; 20% 24-Month Precipitation; 15% 6-Month Precipitation; 10% 60-Month Precipitation; 10% Climate Prediction Center Soil Moisture Model.

The output of the above linear regressions for each grid point results in the local USDM value which finally led to providing maps such as Figure 2.3. Although the USDM has drawbacks in detecting drought conditions of different time scales, as well as in capturing local drought conditions, it can generally provide a “bigger picture” of water stress conditions across coarse spatial scales, which is useful for the public, media, and policy makers (Steinemann et al. 2005, Wilhite 2005). More information about USDM and the description of categories are available at http://droughtmonitor.unl.edu/aboutus/ classificationscheme.aspx.

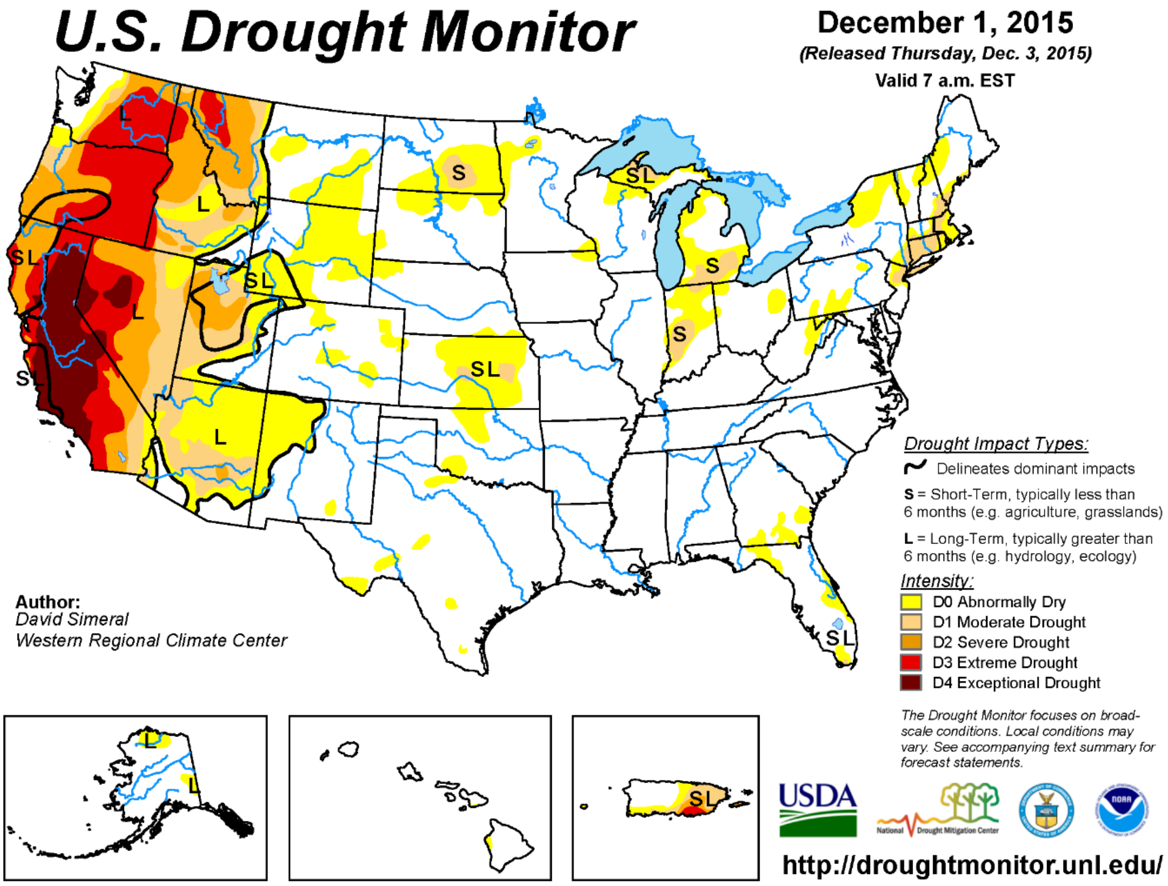


Figure 2.3. The USDM map of USA in December 1, 2015, available at http://droughtmonitor.unl.edu/

Recently, the Vegetation Drought Response Index (VegDRI) (Tadesse et al. 2015) has been introduced to assess the water stress of terrestrial ecosystems (with 1-km spatial resolution) every fortnight. This index employs a regression tree analysis between a set of DIs consisting of NDVI, SPI, self-calibrating PDSI, and five biophysical variables of the environment including landuse/landcover, soil available water-holding capacity, ecoregion type, elevation, and percent of irrigated agriculture. VegDRI can be used independently for water stress monitoring or as an input to the USDM.

In addition to needing a wide range of spatial and temporal information and data, it is also suffering from low spatial precision in areas without dense station networks. VegDRI also relies on spatially interpolated climate data input (Tadesse et al. 2015), which can introduce further uncertainties. Most importantly, the regression trees combine all input variables, regardless whether there is any physical and meaningful relationship between them. For instance, integrating the elevation information with PDSI is not necessarily meaningful from a hydrometeorological point of view.

2- Linear combination-based DIs

Indices of this group are based on considering a weighted average between a number of single DIs to derive a new combined DI. The considered weights may subjectively be chosen or make use of optimization methods. The most widely used linear combination-based DIs, are the Optimal Blended NLDAS Drought Index (OBNDI) (Xia et al. 2014a,b), Grand Mean Index (GMI) (Mo and Lettenmaier 2013), Microwave Integrated Drought Index (MIDI) (Zhang and Jia 2013), and Combined Drought Index (CDI) (Balint and Mutua 2011), which are explained in some more detail here.

The OBNDI was developed to form a bridge between the North American Land Data Assimilation System (NLDAS) and USDM products. The OBNDI allows the derivation of weights to be used in the USDM for the linear combination of the DIs used for operational drought characterization. The error function used for obtaining the optimum weights is

$$E = \frac{1}{MT}\sum_{t=1}^{MT}\sqrt{\frac{1}{C}\sum_{c=1}^{C}(A_{t,c} - O_{t,c})}\ , \tag{2.15}$$

where MT is the total number of months, C is the number of drought categories, and $A_{t,c}$ and $O_{t,c}$ are the drought area percentages derived from NLDAS and USDM, respectively. In comparison to the original USDM, the OBNDI can detect onset, duration, and magnitude of drought events more accurately. However, depending on different climate conditions and landuse situations the results of USDM outweigh the OBNDI's outputs (Xia et al. 2014 a,b).

The GMI is an arithmetic average between three drought indices of the 6-month Standardized Precipitation Index, total Soil Moisture Percentiles, and 3-month Standardized Runoff Index. The thresholds used for the classification of drought events are exactly the same as what is considered for the USDM. The main limitations of GMI are pre-defined input variables and coefficients of linear regression, and also the lack of any indicator analysing the actual vegetation conditions at different growing stages of plants from a water stress point of view.

Satellite remote sensing products can complement the data/information provided from in-situ observations and land surface model outputs. Therefore, over the last few years there has been an increasing tendency towards making use of those products in a variety of applications, such as flood and drought monitoring and forecasting. For example, Zhang and Jia (2013) introduced MIDI on the basis of multi-sensor microwave remote sensing for monitoring short-term water stress conditions expressed as

$$\mathrm{MIDI} = \alpha_1 \times \mathrm{PCI} + \alpha_2 \times \mathrm{SMCI} + (1 - \alpha_1 - \alpha_2) \times \mathrm{TCI} \ , \tag{2.16}$$

where, α_i is a coefficient determined subject to the user's knowledge of the system, PCI is the Precipitation Condition Index, derived from actual precipitation observations provided by the Tropical Rainfall Measuring Mission, SMCI is the soil moisture condition index derived from space-borne products, and TCI is the land surface Temperature Condition Index. The soil moisture product is generally obtained from operational, gridded remotely sensed soil moisture products. The land surface temperature is obtained from 37 GHz vertically polarized brightness temperature measurements. The main drawback of MIDI is the coarse spatial resolution (0.25^0) of the gridded single DIs (PCI, SMCI, and TCI). Consequently, while the MIDI may be applicable for regional water stress monitoring, it cannot be helpful at the farm scale.

The CDI was initially developed for drought monitoring based on spatio-temporal 10-day or monthly data and was first applied to conditions from 1992 to 2007 in Kenya and Somalia (Balint and Mutua 2011). It is a combined drought index based on the weighted average between the precipitation drought index (with a weight of 0.5), the temperature drought index (with a weight of 0.25), and the vegetation drought index (with a weight of 0.25). The thresholds for this index range from 1, which shows no drought condition, to less than 0.4, which indicates extreme drought conditions. Modified versions of CDI have also been applied in other regions such as Europe (Sepulcre-Cantó et al. 2012) and Australia (GrainGrower 2016). The main limitations of this index is the use of fixed weights for independent variables, the impact of growing and non-growing seasons on the results of CDI, and finally the influence of atmospheric errors (i.e. on cloudy days) on the satellite based data of vegetation drought index which can affect CDI considerably.

Overall, the linear combination-based methods are relatively simple to implement, but in case that the physical meaning and texture of the primary DIs are different, combining and interpreting them can sometimes be unreasonable and difficult (AghaKouchak et al. 2015, Hao and Singh 2015, Hao et al. 2016). Moreover, weights (coefficients) of linear regressions are hard to determine objectively (because they are dependent on various hydro-climatological and natural geographical conditions of a region), which is why they are mostly specified empirically for a given region, however, would not be accurate and reliable enough to employ to other regions with different climate and vegetation conditions (AghaKouchak et al. 2015, Hao and Singh 2015).

3- Joint distribution-based DIs

Some of the common joint multivariate distribution based DIs are the Joint Drought Index (JDI) (Beersma and Buishand 2004), Multivariate Standardized Drought Index (MSDI) (Hao and AghaKouchak 2014), and Standardized Palmer Drought Index (SPDI)-based Joint Drought Index (SPDI-JDI) (Ma et al. 2014).

The SPDI-JDI determines first the joint probability (or percentile) between marginal distributions of precipitation and streamflow variables via a fitted joint distribution function. To construct the joint distribution function, one of the Copula family functions (Nelsen 2006), such as empirical Copula, is utilized to model dependence structures (Kao and Govindaraju 2010). Then, using the Kendall distribution function (F_k), the cumulative probability for different events with the joint deficit of less or equal to a certain probability threshold (P_0) are found. Finally, the latter probabilities are put into an inverse standard normal distribution to derive the final values of SPDI-JDI. A similar methodology has been implemented for precipitation and soil moisture which led to introducing the MSDI.

Ma et al. (2014) showed that it is possible to integrate more than two single DIs (with independent single DIs) using the joint probability approach by making use of multivariate Gaussian and t-copula functions, and then combining the Standardized Palmer Drought Index across different time windows to derive the SPDI-JDI.

A valuable outcome of this approach is in drought risk analysis by deriving time periods of extreme occurrences (Beersma and Buishand 2004, Kao and Govindaraju 2010). In terms of shortcomings, the main limitation is to deal with the high number of input variables (single DIs). Essentially, by increasing the number of independent variables the modelling of joint distribution functions becomes much more complicated, and sometimes impossible due to an increase in their dimensions (Hao and Singh 2013). Moreover, this approach only takes into account statistical properties between primary DIs, while it cannot consider the physical process (i.e. time lags between different types of droughts) between individual DIs (Hao and Singh 2015).

4- Aggregation-based DIs using multivariate methods

There are also some multivariate methods such as the linear Principal Component Analysis (PCA) (Hidalgo et al. 2000), which have the ability of dimension (variable/information) reduction without losing significant variability (variance) of the original data/information. Keyantash and Dracup (2004) utilized PCA to aggregate six hydroclimatological variables of precipitation, soil moisture, streamflow, snow water content, reservoir storage, and potential evapotranspiration. The normalized first principal component (PC1) of the aggregated variables were then defined as the Aggregated Drought Index (ADI).

It is worth noting that the linear PCA follows assumption of linearity of the data transformation between primary and aggregated variables. To address the shortcomings related to the linear assumption, Barua et al. (2012) replaced the linear PCA with a nonlinear PCA method (Linting et al. 2007) to combine the six abovementioned hydroclimatological variables, called the nonlinear aggregation drought index (NADI). Recently, Rajsekhar et al. (2015) used the kernel entropy component analysis (KECA) to aggregate four hydroclimatological variables (precipitation, soil moisture, runoff, and potential evapotranspiration). The KECA has two main advantages in comparison to previous aggregation-based DIs, consisting of : 1) using a nonlinear combination process during the aggregation, and 2) considering all PCs for deriving a final aggregation variable, meaning that the maximum amount of information of input variables are kept in deriving the KECA.

Similar to the joint distribution-based approach, it is worth noting that while statistically any selected set of variables can be aggregated via this approach, combining the variables without physical and logical similarities may not be reasonable (AghaKouchak et al. 2015, Hao and Singh 2015). Thus, to avoid this issue, the independent variables (individual DIs) should be considered consistent, in terms of physics and nature, before starting the aggregation and combination steps.

2.4 Key points of the presented DIs

A variety of drought indices and proxies have been presented within this chapter which have distinct properties and characteristics as well as specific abilities and shortcomings. In order to present them in a comprehensive and critical summary, this section discusses those indices more closely with references to each other.

Drought indices such as SPI and SSMI which are constructed based on single hydrometeorological variables are commonly applied in accordance with their data availability and computational simplicity. However, such DIs can only concentrate on one aspect of the water stress conditions (i.e. only from meteorological point of view) within a terrestrial ecosystem, and therefore cannot be considered as a robust tool for decision making and operational systems (Wilhite 2005).

The USDM and VegDRI, from the category of the blending objective and subjective indicators, assess water stress conditions across country and continental scales by integrating information from various physical indicators and inputs from local experts. These indices have shown their unique functionalities for operational and research purposes in the USA. However, besides the need for the large amount of information and data as input, the main limitations would be in their applications across different time scales and also their limited use for decision making processes at higher spatial resolutions such as the farm scale (Wilhite 2005).

Drought indices such as PDSI and CMI incorporate hydrometeorological variables via water balance models. They were initially developed for some regions of the continental USA and showed adequate performances in drought assessments. The main drawbacks are the considered assumptions and simplifications in the water balance models, and the applicability in other parts of the world with different climate and landuse/landcover conditions (Hao and Singh 2015, AghaKouchak et al. 2015).

According to the literature, it is possible to make use of mathematical transformations to combine hydrometeorological variables, such as precipitation and evapotranspiration, to derive univariate drought indices such as CWSI, ESI, and RDI. However, during the process of combination, physical and meaningful relationships between hydrometeorological variables should be considered. This limitation can lead to the decrease in the degree of flexibilities to integrate the indicators of interests (Hao and Singh 2015, AghaKouchak et al. 2015).

To be able to combine DIs regardless of their physical relations, using linear regression methods seems to be the simplest and easiest choice to derive new DIs such as GMI, MIDI and CDI. The main problem lies in the fact that in most of the cases relations and dependencies between hydrometeorological variables (consequently DIs) are essentially nonlinear, and moreover, weights are determined subjectively (Hao and Singh 2015, AghaKouchak et al. 2015).

To resolve the issues introduced by the assumption of linearity and introduction of subjective weights through the linear regression-based combination methods, joint distribution-based models are employed to introduce DIs such as JDI, MSDI, and SPDI-JDI. Throughout this approach, complex dependency structures such as tail or extreme dependencies between single DIs are modelled to analyze statistical characteristics. The selection of the most suitable multivariate distribution in case of the presence of a large number of single DIs would be the most challenging part (Hao and Singh 2015).

In order to address the limitations of joint distribution-based models, common multivariate methods, such as PCA, are utilized to reduce the information in high dimensions. Two of the introduced DIs based on this approach are ADI and NADI. The potential issues are associated with the appropriate selection of single DIs with physical and meaningful relations, and also the assumption that the first component with the maximum variance can convey sufficient water stress information (Hao and Singh 2015, AghaKouchak et al. 2015).

2.5 Knowledge gaps

The main shortcomings of the historical DIs presented in this chapter are:

1- Different individual DIs provide different, if not conflicting, information under various climate conditions, landcover, and the perspective of the application. Therefore, no single DI works appropriately in all different circumstances.

2- The main drawbacks of the introduced combination- and aggregation-based DIs consist of i) reliance on subjective selection of inputs, ii) limited statistical and mathematical frameworks to objectively link or combine a variety of DIs from different drought types (i.e. hydro-meteorological droughts, and ecological/agricultural droughts) into one new DI, iii) to derive the values of the combination based DI for the current and future time steps, it is required to re-implement the whole process of combination from the beginning steps.

3- The process of validating DIs is generally undertaken by simply comparing the behaviour of new indices to the traditional DIs such as SPDI and SPI as the ultimate benchmark. However, if new indices are compared against older ones for validation, what is the purpose of new indices? However, Hao and Singh (2015) stated that no reliable “ground truth” (for a drought index) exists that may be used as an ultimate reference for the exact validation of a new index, so the question truly is how the accuracy of an index can be reliably assessed. One pathway to address this issue is to benefit from reports gathered from observed evidences of fields or drought assessments' claims received from stakeholders such as farmers. Converting the mentioned information to quantitative forms (i.e. exceptional circumstances maps in Queensland, Australia) gives the opportunity of validating newly developed drought indices and also comparing their performances with commonly applied DIs properly. In case of the lack of such abovementioned "ground truth", another solution to alleviate this problem is to objectively include a larger range of observed variables. Thus, the main concern in this field of study is to develop a new and more robust methodology to address the above shortcomings for accurately evaluating water stress conditions in a specific area. The proposed methodology should be able to monitor water stress conditions of terrestrial ecosystems by objectively linking different aspects of the system such as water availability and vegetation conditions. The methodology to be developed may make use of advanced statistical methods (i.e. multivariate methods such as independent components analysis), and also consider the eco-meteorological characteristics (i.e. landuse, land-cover, and climate) of an area to state the ultimate water stress conditions at each time step. In order to address any issues in the current validation processes of DIs, the behaviour of the new DI should be logically compared with the physics, nature (landuse/landcover information) and climate conditions (i.e. wet and dry seasons based on total downfalls) of the considered terrestrial ecosystem. Finally, for verification and generalization, the new DI needs to be examined and tested in different regions with a diverse range of surface and atmospheric conditions.

2.6 Chapter summary

In Australia, the drought indices most widely applied in industry and government organizations are precipitation based indices, such as the Decile Index (Rahmat et al. 2015, Jain et al. 2015), which indicate anomalies from a long-term trend. As previously mentioned within this chapter, in order to have accurate evaluations using precipitation-based drought indices, a long data record along with local hydrometeorological observations is required, as well as precise frequency distributions of precipitation. Ignoring those can lead to a reduction in the accuracy and reliability of the water stress evaluations of an area. Moreover, similar to all other individual drought indices, they can only reflect one aspect of drought phenomena (meteorological drought) which is not sufficiently inclusive for robust evaluations. As another commonly applied dataset in Australia, GRACE data are used in conjunction with models, has a limited data set length (since 2002), as well as coarse temporal (~30 days) and spatial resolutions (>150000 km2). Using data with low spatio-temporal resolutions may lead to a loss of information over an area, as it integrates the positive and negative anomalies over time, essentially ignoring the natural complexities of the underlying geographical and landcover features within a terrestrial ecosystem. In turn, this may result in errors in the final interpretation due to an imprecise picture of the real conditions. This shows a necessity of developing new indices which are able to address the currently drawbacks and provide a holistic analysis over the short- and long-term water stress monitoring and forecasting.

This chapter presented a review of the two common approaches of “reactive” and “proactive” procedures to address water stress and drought conditions in Australia, and their economic benefits. It then discussed the most common and state-of-the-art water stress and drought indices (DIs). Subject to the utilized hydrometeorological variables to derive DIs, the considered indices were categorized into eight main groups of: i) precipitation-based DIs, ii) water balance model-based DIs, iii) surface water-based DIs, iv) soil moisture-based DIs, v) evaporation-precipitation-based DIs, vi) terrestrial water storage-based DIs, vii) vegetation-based DIs, and viii) combination- and aggregation-based DIs. The latter group also was further split into four sub-categories of: i) blending objective and subjective indices, ii) linear combination-based DIs, iii) joint distribution-based DIs, and iv) aggregation-based DIs using multivariate methods. Then, strengths and limitations of the presented DIs were elaborated; and the main knowledge gaps available in this field of research were stated and finally solutions were briefly proposed, essentially highlighting the need for a new, more robust, and comprehensive drought index.

Chapter 3

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

3.1 Overview

This chapter will use a range of spatio-temporally high-resolution (daily and sub- daily) data sources to evaluate a number of drought indices (DIs) for the Riggs Creek OzFlux tower site in south-eastern Australia. Therefore, the main aim is to evaluate the statistical characteristics of commonly used 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 of this chapter has been accepted as a journal paper as follow:

Azmi, M., Rüdiger, C., Walker, J. (2016) Statistical Analysis of Short Term Water Stress Conditions at Riggs Creek OzFlux Tower Site, *Theoretical and Applied Climatology*, DOI 10.1007/s00704-016-1901-z.

3.2 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 18-12-2010 to 1-5-2012) of verified sub-daily data used in this chapter. The tower is located within the Goulburn-Broken catchment (36° 38.59' S, 145° 34.21' E), in northern Victoria, Australia (Figure 3.1) (Andrykanus 2011, Beringer 2014). Figure 3.2 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 landuse 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 temperature of 12.3 ^{0}C in July and 29.7 ^{0}C in February, and mean minimum air temperatures between 4.1 ^{0}C in July and 15.3 ^{0}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 state of Victoria experienced severe water stress from 1997 to 2009 with an average precipitation rate of 12.4% below the 20th century mean. In 2010-11, because of a strong La Nina phenomena, significantly higher levels of precipitation were recorded for Victoria with annual rates exceeding 810 mm (SEACI 2011); while by early 2012 no extreme events (flood and severe drought) had been recorded (Howden 2012).

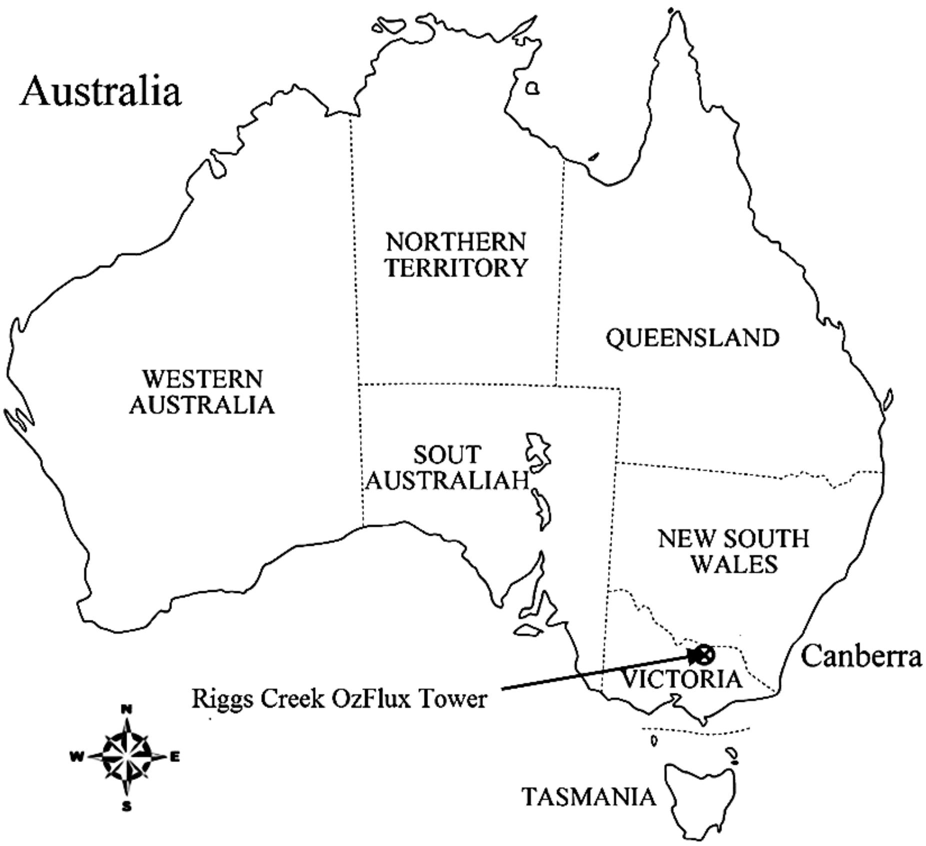


Figure 3.1.The geographical location of the Riggs Creek case study site within Australia.

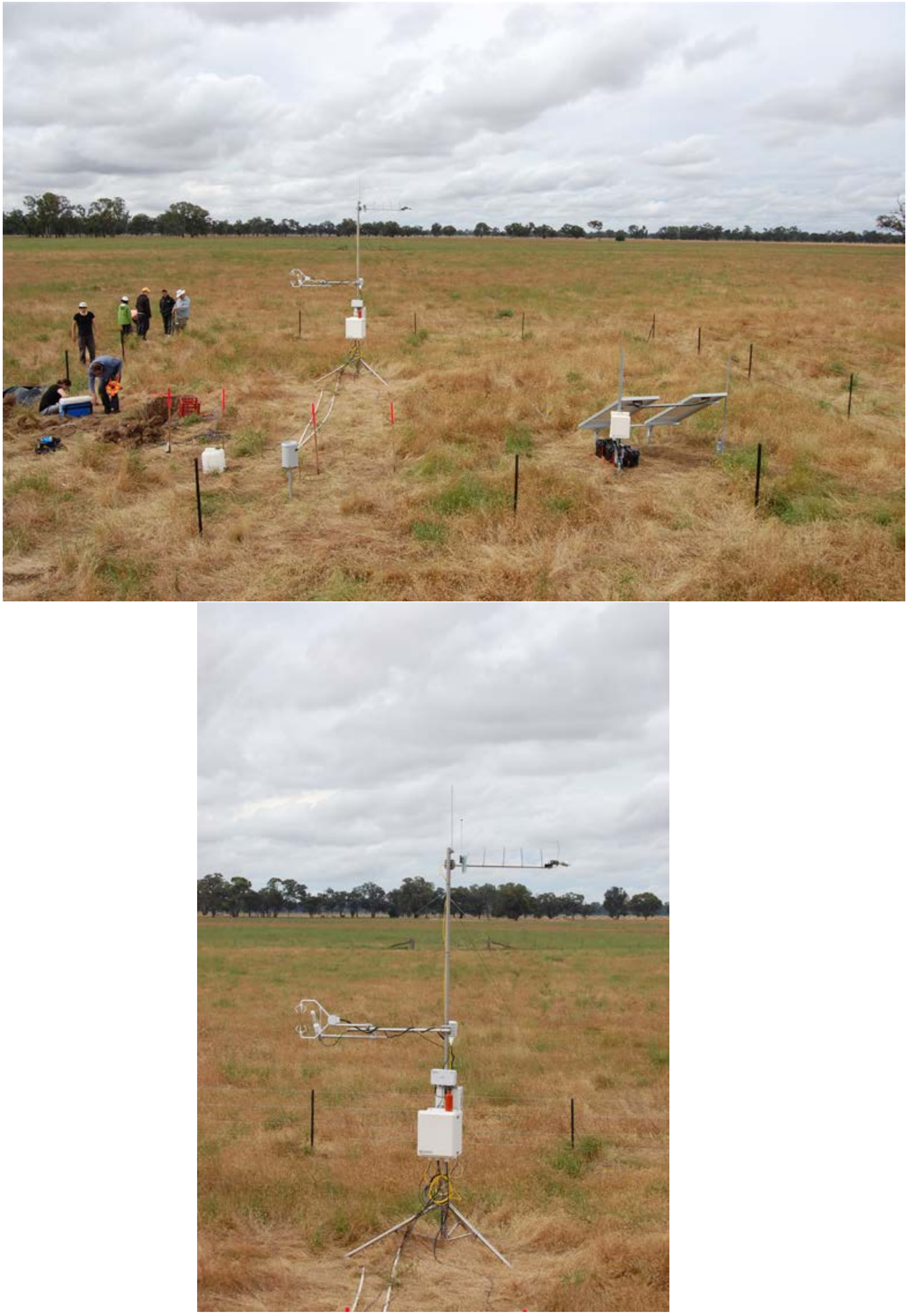


Figure 3.2. Two photos of Riggs Creek OzFlux tower site located within a grassland (pasture) along with the tower. Underground based instruments measure soil moisture at depths of 0.1m, 0.2m, 0.4m, 0.8m, 1.2m, and 2 m. Ground based instruments are data logger and rain gauge. Mid-height instruments measure soil temperature (0.08m) and soil heat flux (0.1m). Top-height instruments measure atmospheric pressure (2.5m), net radiation (4m), open path CO_2 H_2O, sonic anemometer (2.5m), air temperature (2.5m), and relative humidity (2.5m).

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 3.1). 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 chapter 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 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 from which data were obtained for this study 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 one 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 one week) of many species of vegetation (Svoboda et al. 2002, Heim Jr 2002). Therefore, weekly temporal resolution were considered for evaluating water stress conditions in this chapter.

Table 3.1. Hydroclimatic variables derived from different data sources to calculate drought indices and proxies compared in this chapter.

Category	Data Sources	Primary Variables	Abbreviation	Units	Temporal Scale	Spatial Scale	Calculation
In-situ Observations	OzFlux Towers Network	Soil Moisture Content (Depth=10cm)	SM	mm	0.5 hr	N/A	Direct Measurement
		Precipitation	P	mm	0.5 hr		
		Moisture Flux (Latent Heat)	MF	W/m^2	0.5 hr		
		Evaporative Fraction Index	EFI	No Dim.	0.5 hr		$EFI = \frac{H}{Rn - G}$
A Combination of In-situ Observations, Satellite Information and Models'sOutput	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	$NDVI = \frac{NIR - VIS}{NIR + VIS}$
		Vegetation Condition Index	VCI	No Dim.	1 day	250 m	$VCI = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}$
		Temperature Condition Index	TCI	No Dim.	1 day	500 m	$TCI = \frac{Tb_{max} - Tb}{Tb_{max} - Tb_{min}}$
		Perpendicular Drought Index	PDI	No Dim.	1 day	250 m	$PDI = \frac{1}{\sqrt{M^2 + 1}}(VIS + M \times NIF)$

NIR: the spectral reflectance measurements acquired in the near-infrared regions (700-1100 nm), *VIS*: the 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}).

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. Finally, daily DIs were calculated based on the latter data set, and converted into weekly time scale by averaging; and finally weekly DIs were standardized as described below.

3.3 Standardizing hydroclimatic variables and proxies

As different hydroclimatic variables and proxies represent different physical 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 analyze 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.

3.4 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.

3.4.1 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:

$$Sev(t,k) = \left(Median\ \{Mag(i)\}\right) \times k\ , \tag{3.1}$$

where

$$Mag(i) = SDI_i - NCT \qquad i = t,...,t+k \quad , \tag{3.2}$$

Where, $Mag(i)$: 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.

3.4.2 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 resources management of a terrestrial ecosystem. It is worth noting that the economic/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 for 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 is 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 expert 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 characteristics 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 reaches 24 (here primary SDIs are equal to 8 variables and lagged SDIs would be equal to16 variables);

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 that 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 the different cluster sizes, linkage methods, and distance functions, the 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) ($PS_{i,j}$) as follows:

$$PS_{i,j} = \frac{n}{N} \times 100, \tag{3.3}$$

where n is the number of times that two variables (i,j) are located in the same cluster; 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, that 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 remainder then represents 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% ($-0.67<Z<+0.67$) reflect moderate conditions or in other words a "moderate similarity", and consequently, PS values greater than 60% ($Z>+0.67$) and less than 40% ($Z<-0.67$) show "strong and weak similarities", respectively (extreme conditions). All SDIs with a "strong similarity" can definitely be located in one cluster; while single SDIs which have a "weak similarity" with all 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).

3.5 Results and discussion

For the perpendicular drought index (PDI), the slope of the soil line (M) is determined using all pairwise values of the NIR-VIS (Near InfraRed-VISible) relationship for the location of the Riggs Creek OzFlux tower. Those data were obtained from daily MODIS observations on board NASA's Terra satellite. In Figure 3.3, 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 triangle 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, as 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 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 pair-wise distribution towards points B and C shows the conditions of wet and dry surfaces, respectively. Figure 3.3 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 Figure 3.4. The operational functionality of the standardization of the indices is apparent in the values of SNDVI and SVCI (the standardized values of NDVI and VCI, respectively). As VCI is itself a standardized form of NDVI (Table 3.1), 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 SSMI, SRSSMI and 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. Surprisingly, no particular pattern was discernible for the last index (SPDI), which is generally used as a benchmark for new indices, showing the necessity for advanced statistical analyses to highlight inter-index differences.

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

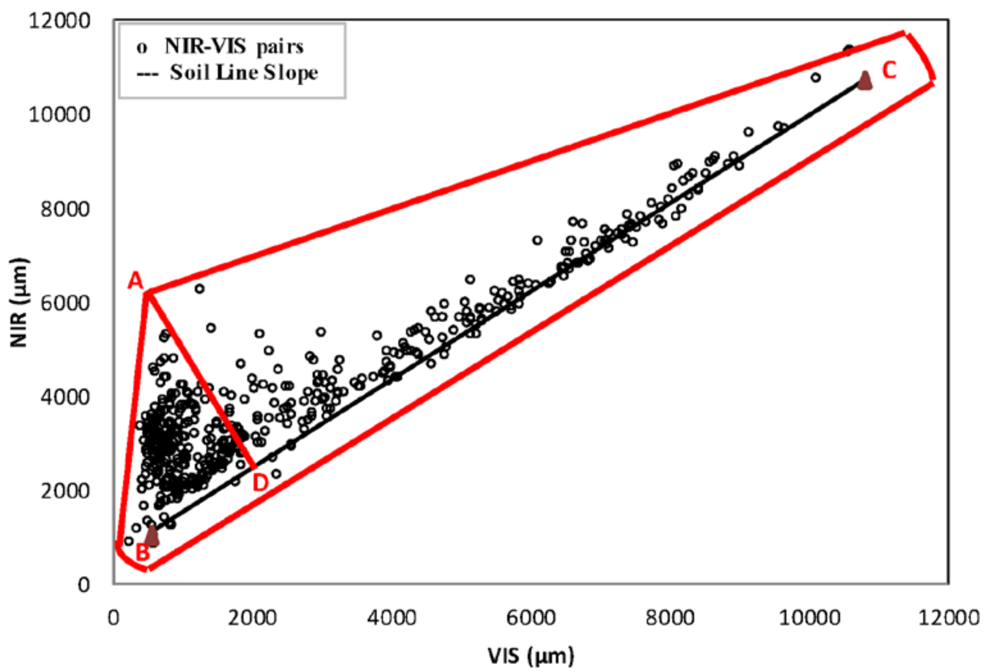


Figure 3.3. Pairwise values of NIR-VIS for the location of Riggs Creek OzFlux Tower for deriving the slope of soil line (M).

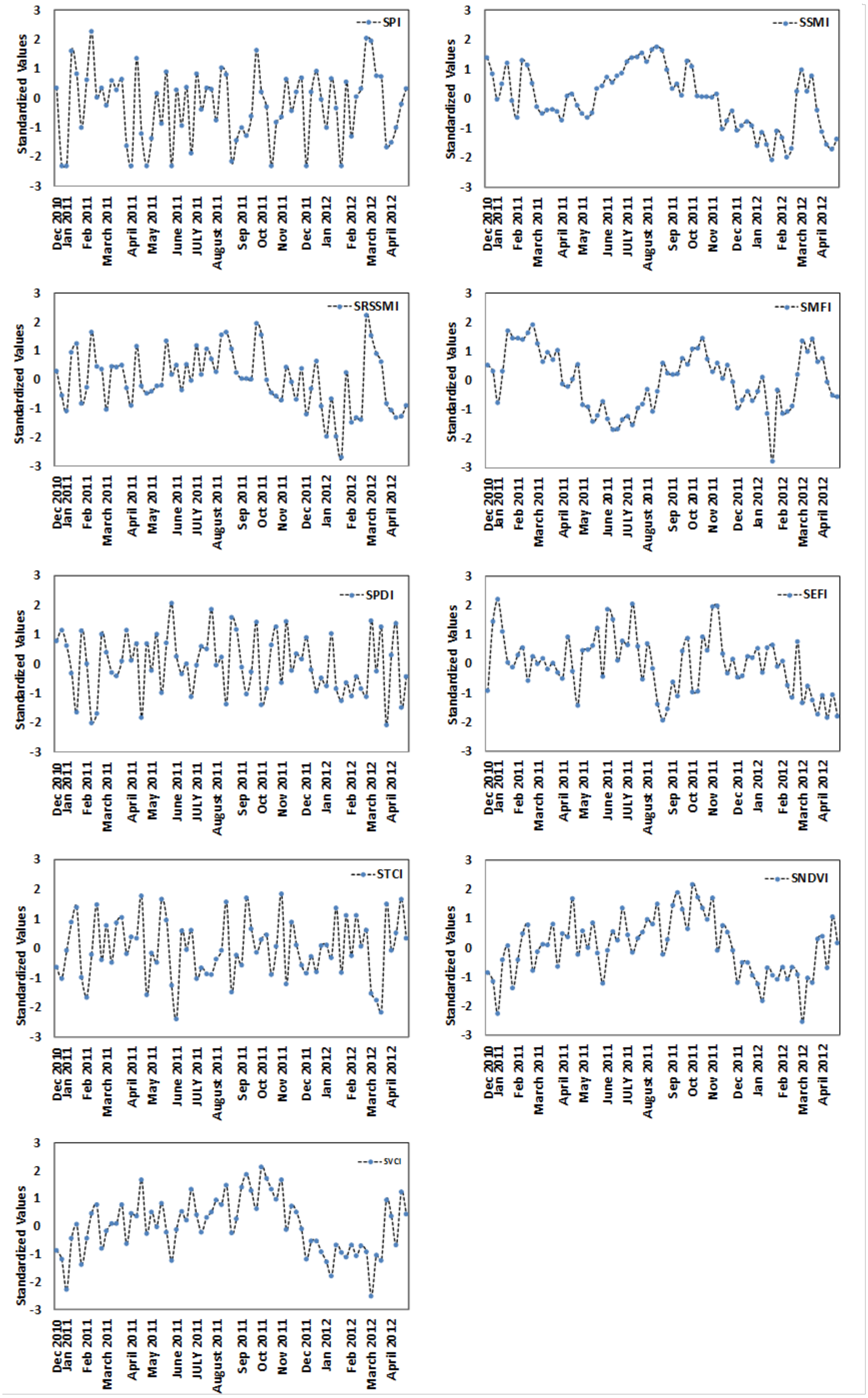


Figure 3.4. Weekly time series of SDIs at Riggs Creek OzFlux Tower from 18-Dec-2010 to 1-May-2012.

Considering the goodness-of-fit analysis (Anderson-Darling and p-value tests), the log-normal 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 3.2). 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. Figure 3.4. Weekly time series of SDIs at Riggs Creek OzFlux Tower from 18-Dec-2010 to 1-May-2012.

Omitting zero values can markably 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 3.4.1 were used to determine the respective durations of the extreme events. In addition, the severities (Sevs) and magnitudes (Mags) of the extreme conditions were calculated using equations (3.1) and (3.2). The durations, Sevs and Mags of extreme dry as well as wet events for all SDIs are presented in Table 3.3 and 3.4, respectively. Table 3.3 shows that for most of the cases durations of extreme dry events were 1, meaning that extreme dry conditions lasted generally one week or less. From a climatological point of view, an extreme event with the duration of one week cannot be considered as a drought or flood condition. However, one 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 one 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 3.4 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, Figure 3.5 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.

Table 3.2. The main statistical parameters, the best cumulative distribution function, and different percentiles for each considered SDI.

Standardized Variables	SDIs	The best chosen *cdf* for standardizing	Skewness	Kurtosis	Min.	2.5^{th}Percentile	5^{th}Percentile	25^{th}Percentile	45^{th}Percentile	50^{th}Percentile	55^{th}Percentile	75^{th}Percentile	95^{th}Percentile	97.5^{th}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

Table 3.3. Magnitude and Severity of SDIs under extreme dry conditions (2.5th percentile of the entire time series of the SDI) calculated based on equations 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 3.4. Magnitude and Severity of SDIs under extreme wet conditions (97.5th percentile of the entire time series of the SDI) calculated based on equations 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				

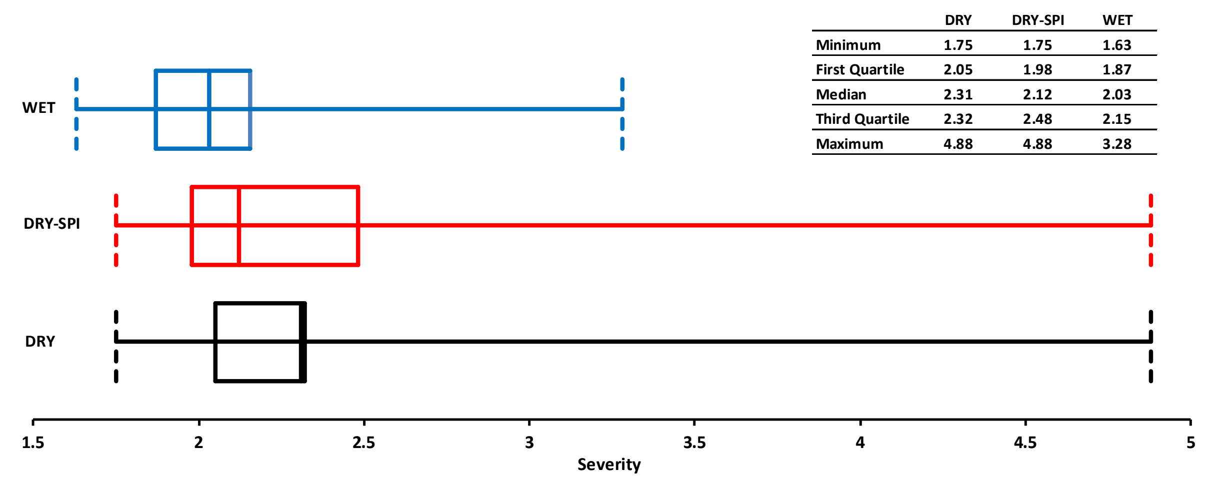


Figure 3.5. Box plot of comparison between extreme dry and wet events severities at Riggs Creek Flux Tower site from 18-Dec-2010 to 1-May-2012.

In other words, 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 to be expected, and the limited data set reflected this through its 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 3.5, and the final PS between each SDI pair is indicated in Table 3.6. According to Table 3.6, 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 (SPDI, SPI, SSMI, SRSSMI and SMFI) with an average PS of 62% can be categorized as a cluster (representing water availability). Three SDIs (STCI, SNDVI and SVCI) with an average PS of 75.3% can be categorized as a second cluster (representing vegetation conditions). The index of SEFI with an average PS of 17.8% (average of all 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 the probabilistic-based clustering. The values of α for the first and second groups were 0.63 and 0.83, respectively, which validates the derived clusters. To confirm that SEFI is not well

Table 3.5.The probabilistic similarities between SDIs without time lags, as well as with time lags 1 and 2 calculated based on equation 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.

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 conditions (STCI, SNDVI, SVCI), 2- water availability (SPI, SSMI, SRSSMI, SPDI and SMFI), and 3- water consumption (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 established

Table 3.6. Final probabilistic similarities derived from the average between all probabilistic similarities of a pair of SDIs mentioned in Table 3.5 (values are in percentage).

	STCI	SNDVI	SVCI	SPDI	SPI	SSMI	SROI	SSMFI	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
SROI								70.37	31.48
SSMFI									11.11
SEFI									

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 for a given area requires the consideration of all derived information (probabilistic behaviours, extreme dry/wet conditions and etc.) of the representative SDIs.

3.6 Chapter summary

This chapter presented a set of statistical methods to evaluate short term water stress conditions for a study site in south-eastern Australia. The required data were obtained from the Riggs Creek flux tower (OzFlux Network), the Asia-Pacific Water Monitor (APWM), and the MODIS instrument on board NASA’s Terra satellite. A 72-week period spanning across the years 2010-12 was chosen based on the low level of missing data. The temporal scale of the data was from 0.5 hr 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 landuse and surrounding topography. The results showed that by using a probabilistic standardizing method, the main statistical parameters of the SDIs could be derived and analyzed. 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 analyzed. Finally, dominant water stress types along with possible representative SDIs were presented. In this chapter, SDIs were grouped into three main clusters, the first group represented availability of water, the second group reflected plant response to water supply across different seasons, and the last one represented the water consumption across the whole water system. 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. However, as stated before, single DIs may not be able to fully describe the variations in water stress, and therefore, it is recommended to develop a more detailed clustering method.

In the next chapter, a drought index based on the clustering method will be introduced and then applied to the Riggs Creek data set to combine drought indices of each derived cluster and finally allow conclusions on the water stress conditions. The same approach will also be employed to using data from a further two OzFlux towers located in different climate and landcover conditions. Finally, the results of the proposed index will be comparatively evaluated with common drought indices such as SPI and SSMI.

Chapter 4

Proposing a new data fusion-based drought index for short-term water stress monitoring

4.1 Overview

In the course of the previous chapter, the most common drought indices were applied and comparatively evaluated for short-term water stress monitoring at the Riggs Creek OzFlux tower site. A new clustering methodology was also introduced which will be used throughout the current chapter as one of the components of the proposed methodology. Within this chapter, a data fusion-based drought index (DFDI) will be developed and its performance analyzed for three different locations in Australia. The proposed index comprehensively considers all types of drought through a selection of indices and proxies associated with each drought type. By considering a set of individual DIs, this chapter develops a methodology to monitor water stress conditions of terrestrial ecosystems by objectively linking water availability and vegetation conditions. The combination methodology makes use of advanced statistical methods (i.e. multivariate methods such as independent components analysis), and also considers the eco-meteorological characteristics (i.e. landuse, land-cover, and climate) of an area to determine the ultimate water stress conditions at each time step. In order to test the ability of the new approach to generalize a range of DIs, three case study areas, each with different combinations of landuse and climate regimes, and subsequently a diverse range of surface and atmospheric conditions, are presented. The results of this chapter have been published as a journal paper as follow:

Azmi, M., Rüdiger, C., Walker, J. (2016), A data fusion-based drought index, *Water Resources Research*, 52:2222-2239, DOI: 10.1002/2015WR017834.

4.2 Data fusion-based drought indices (DFDIs)

4.2.1 Commonly used DFDIs and their main drawbacks

Before discussing the commonly used DFDIs, it is necessary to separate the monthly conditions of terrestrial ecosystems into two distinct aspects:

- Wet/Dry months: a month can be identified as a wet month if the mean of the historical precipitation for that month is higher than the total mean of all months of the year; otherwise it would be considered to be a dry month.

- Active/Non-Active months: during active months, the mean air temperature is usually between 6^{o} to 40^{o} C with most plant types being physiologically active and growing (Singh and Dhillon 2006); however the degree of the plant activity depends on the plant available water (Singh and Dhillon 2006), as well as the vegetation type. In contrast, during a non-active month, mean air temperatures are generally lower than 6^{o} or higher than 40^{o} C, with plants having a minimum physiological activity or even shutting down, unless the plant has adapted to the prevailing local conditions (Brut et al. 2010).

The relationship between the status of a month in terms of dryness/wetness and active/non-active vegetation, as well as the selection of the appropriate drought indices can be described based on one of the following categories:

1. If time *i* falls within a wet month, no matter that the month is plant-physiologically active or non-active, the water stress monitoring can be appropriately evaluated by drought indicators reflecting the water contents, such as the standardized precipitation index (SPI). This is because the water balance of the system is then much more sensitive to inputs such as precipitation and streamflow than outputs such as real evapotranspiration.

2. If time *i* falls within a dry and active month, the water stress monitoring can be assessed by drought indicators such as the Normalized Difference Vegetation Index (NDVI), which can be seen as a proxy reflecting the water consumption and vegetation growth. This is because the water balance of the system is much more sensitive to outputs such as actual evapotranspiration rather than inputs such as the amount of precipitation, which are usually reduced during this period.

3. If time i falls within a dry and non-active month, the water stress monitoring can be appropriately evaluated by drought indicators which reflect the stored water such as the soil moisture drought index. This is because plants are usually in shut-down conditions and their water consumption would be insignificant, in addition to there being no remarkable input to the system in these cases.

The various drought indices presented in the literature can be divided into commonly used state-of-the-art DFDIs as summarised below (see also Chapter 2 for more detailed descriptions):

- Linear Aggregated Drought Index (LADI) (Keyantash and Dracup 2004):
  LADI is based on the combination of six hydro-meteorological variables of precipitation, potential evapotranspiration, streamflow, reservoir storage, soil moisture content, and snow water content using the multivariate method of linear principal component analysis (LPCA) (Hidalgo et al. 2000). LADI utilizes only the first principal component as it explains the largest fraction of the variance described by the full members.

- Nonlinear Aggregated Drought Index (NLADI) (Barua et al. 2012):
  NLADI is an extended form of LADI. Essentially, NLADI employs the nonlinear principal component analysis (NLPCA) (Linting et al. 2007) to combine the six hydro-meteorological variables used in LADI. The combination process is similar to LADI.

- Weighted Average based Drought Index (WADI) (Balint and Mutua 2011):
  WADI is a combined drought index based on the weighted average method between precipitation drought index (with the weight of 0.5), temperature drought index (with the weight of 0.25), and vegetation drought index (with the weight of 0.25).

- Arithmetic Average based Drought Index (AADI) (Zhang and Jia 2013):
  AADI is an index based on arithmetically averaging standardized variables of precipitation, soil moisture and air temperature.

The different drought indices have their individual advantages and disadvantages. In particular, LADI and NLADI only consider hydroclimatological variables which reflect the water content of a system, which is insufficient for analyzing the water stress of a terrestrial ecosystem as it does not consider the level of plant-physiological activity, especially during the growing months. Consequently, it is necessary to broaden the considered data set and choose indices that also recognize the physiological stress of the vegetation, which may not be directly related to the water content based indices. In addition, the aggregation of the variables is performed by only using the PCA method. One of the main assumptions of the PCA is the Gaussianity of its input variables and the underlying linear regressions to derive any aggregated variables. To rely solely on this multivariate method for combining data is a limiting approach, as most hydro-meteorological variables do not necessarily follow Gaussian probabilistic distribution functions and therefore, complex relationships between the variables can be better assessed assuming nonlinear relationships. Moreover, these two DFDIs are already the first principal component (PC1) of six fixed hydroclimatological variables aggregated by the PCA, irrespective of the variances of the entire data set that can be covered by PC1. Finally, they cannot be applied easily for estimating and predicting waster stress in an area, because assessing the water stress for following time steps requires re-performing all combination stages from the beginning of the observations. Conversely, WADI is based on a weighted average combination in which the weights are determined subjectively (or through calibration) for a specific case study area. The advantage is that the index values are inter-comparable between sites. However, by choosing standard weights they may lack physical meaning, as the relative importance of the included variables may change across seasons. Thus, there is no guarantee to get reliable results for any area with different climate and landuse. Further, WADI and AADI are both derived based on averaging single DIs, an approach which may neither by naturally nor physically consistent.

4.2.2 Proposed DFDI

To address the abovementioned deficits of commonly used DFDIs, a new index is proposed at this point. The workflow schematic of the proposed algorithm is presented in Figure 4.1. First, an appropriate set of indices and proxies is determined based on the available data. Those consist of a pool of variables covering water contents, water consumption, and vegetation conditions. Then, the indices and proxies derived from the previous step are standardized based on an equiprobability transformation (Shukla and Wood 2008) and are consequently labelled Standardized Drought Indices (SDIs), and are finally clustered via the Probabilistic Similarities (PSs) method which was previously elaborated in Chapter 3. In particular two of the formed clusters, which consist of i) water content-based SDIs, such SPI, and ii) vegetation condition-based SDIs, such as NDVI, can provide sufficiently relevant information to precisely assess the water stress of an area, all subsequent steps are only applied to the two abovementioned clusters. It is well understood that the water stress conditions inside plants are closely related to the level of plant-available water of an area. Consequently, considering indices from both groups can provide a more comprehensive evaluation of the ultimate water stress situations of an ecosystem, and potentially help exclude vegetation states that could be seen as drought-related from remotely sensed data sources, but are in fact false positives, caused by vegetation cover affected by pests and diseases.

In order to investigate the efficiency of performing multivariate methods to aggregate the SDIs located within the two main clusters, the Kaiser-Meyer-Olkin (KMO) test (Kaiser and Rice 1974) is employed. Using this method, a value greater than 0.5 indicates the efficiency of multivariate methods in combining SDIs. Seeing that the clusters derived from the PSs method consist of similar indices (either water or plant-physiology driven), it is expected that the multivariate methods are appropriate to aggregate the SDIs located within the same cluster. This aggregation is then achieved, using three common multivariate methods, namely the Principal Component Analysis (PCA) (Hotelling 1933), Factor Analysis (FA) (Kim and Mueller 1978) and Independent Component Analysis (ICA) (Hyvarinen and Oja 2000). FA and ICA have the advantage of assuming the aggregated variables to not only be uncorrelated but also statistically independent, non-Gaussian, and also consider nonlinear regressions between variables (Hyvarinen et al. 2001).

The indices obtained through the aggregation within the clusters are then standardized by dividing them by their individual standard deviations. To specify the best number of aggregated variables derived for the two mentioned clusters when using the different multivariate methods, the Kaiser1 method (Kaiser 1960) is used. For the Kaiser1 method, the aggregated variables in which their Eigenvalues are greater than 1 are selected. Considering that the PSs method forms consistent clusters, the first principal component (PC1) of each cluster will dominantly cover the highest amount of variance for all variables located in a group with Eigenvalues greater than 1; or the difference between the Eigenvalues of the first and second principal components is quite large enough to only consider PC1 as the aggregated variable of each cluster.

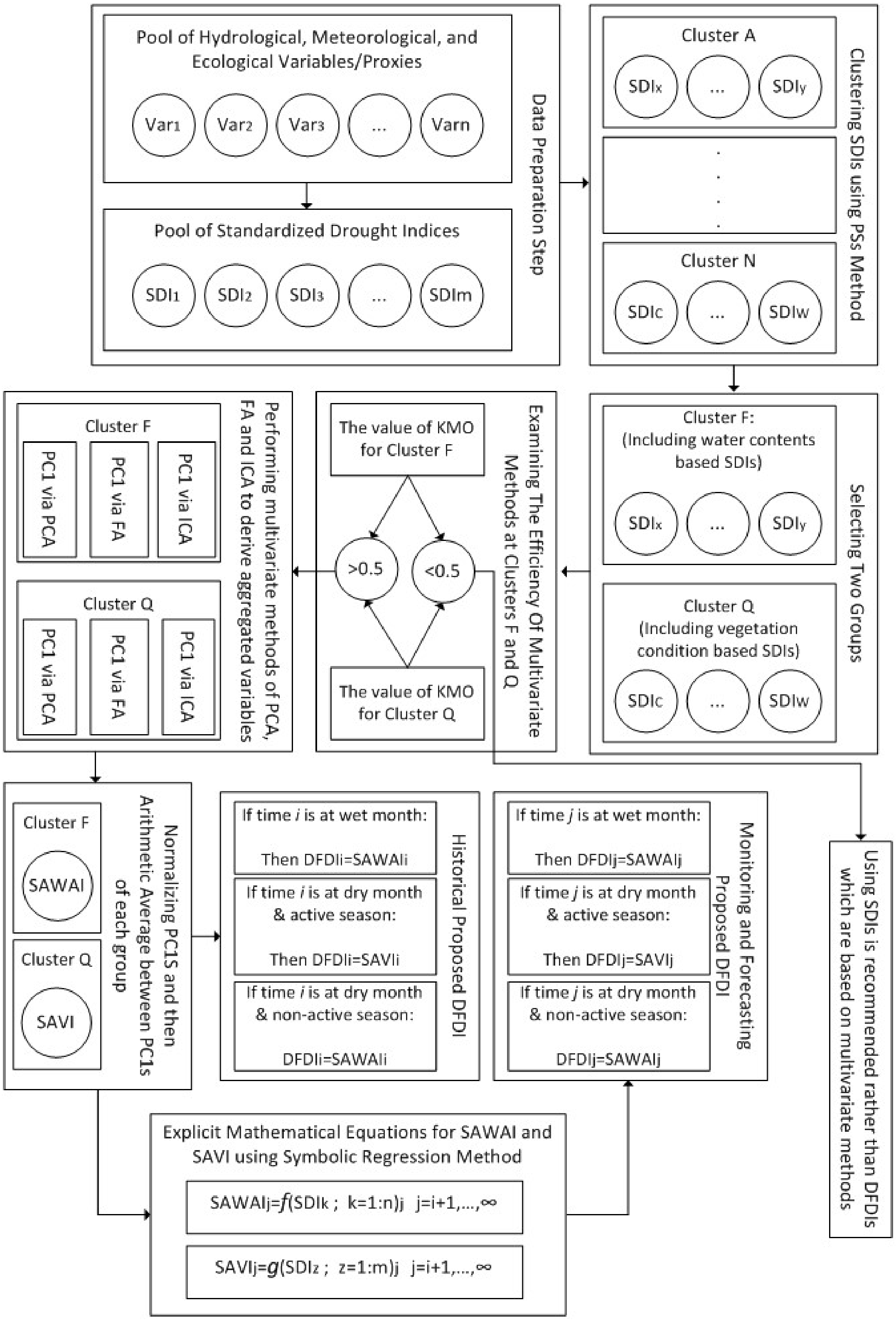


Figure 4.1. Schematic outline of the proposed drought index.

After standardizing the PC1s by dividing them by their standard deviations; an arithmetic average between the PC1s of the two mentioned clusters derived from different multivariate methods (PCA, FA, and ICA) is calculated, giving the final aggregated variable of that specific cluster. Here, as those two clusters represent either water or vegetation, they are named Standardized Aggregated Water Availability Index (SAWAI), and Standardized Aggregated Vegetation Index (SAVI), respectively.

In order to make the current methodology more applicable in a real-time scenario, explicit mathematical equations are derived for SAWAI and SAVI by employing the Symbolic Regression Method (Koza et al. 2003). The independent variables for the mathematical equations are drawn from a set of SDIs from within the SAWAI and SAVI clusters. Emphasis is given to select more readily available SDIs as independent variables, such as the standardized precipitation index (SPI), rather than more complex ones, in order to make the final drought index as practical and directly applicable as possible. The final proposed data fusion based drought index (DFDI) may be expressed as follows:

If time *i* is at a wet month, irrespective of whether that month is physiologically active or non-active, the water stress monitoring can be appropriately evaluated by SAWAI:

$$DFDI_i = SAWAI_i; \;\; \& \;\; DFDI_j = SAWAI_j = f(SDI_k; k = 1:n)_j \quad j = i+1,...,\infty \tag{4.1}$$

If time *i* is during a dry month, as well as an active season, the water stress monitoring should be assessed by SAVI, which reflects the water consumption and vegetation growth, therefore

$$DFDI_i = SAVI_i; \;\; \& \;\; DFDI_j = SAVI_j = g(SDI_z; z = 1:n)_j \quad j = i+1,...,\infty \tag{4.2}$$

If time *i* is during a dry month, as well as a non-active season, the water stress monitoring should be evaluated by SAWAI which reflects the amount of stored and plant-available water, again following Equation 4.1.

In the above equations, *i* includes all previous time steps until the current time; *j* is all time steps after *i*; *f* and *g* are the optimum mathematical functions between the selected SDIs to calculate $SAWAI_j$ and $SAVI_j$, respectively; SDI_k and SDI_z are *k* and *z* selected SDIs as independent variables to derive SAWAI and SAVI, respectively.

Due to the fact that the output of the considered aggregating methods (PCA, FA, and ICA) is based on Gaussian distribution functions, thresholds of dry, normal, and wet events can also be defined based on the Gaussian variates of the standard deviations (Barua et al. 2012, Keyantash and Dracup 2004). Following the justifications for the thresholds used for the Probabilistic Similarities mentioned, the thresholds proposed are as shown in Table 4.1.

In terms of temporal resolution, weekly data sets, derived by averaging the daily information of that week, are considered for the evaluation of water stress in the present chapter. This aggregated temporal window is necessary as the physiological conditions of plants are the main elements of the terrestrial ecosystems and reflect, or integrate, the water stress (wilting point) with a lag time of around one week. In fact, daily data products may show decorrelated results between the indices and the surface conditions for this same reason. Moreover, information of water stress monitoring with longer time spans (monthly, seasonally) cannot be practical for ecological water management due to the wilting point onset for many vegetation species (Svoboda et al. 2002, Heim Jr 2002).

4.3 Case study sites and data sources

Due to the fact that characteristics of an ecosystem’s water stress mainly depend on its climate and landuse, three case study sites have been chosen (Figure 4.2) to evaluate the capacity of the proposed methodology, with each having different climate and landuse conditions.

Table 4.1. Proposed thresholds for the DFDI.

<table>
<tr><th>Threshold Range</th><th colspan="4">Gaussian Function Curve Areas</th><th>Classifications</th></tr>
<tr><td>DFDI < -1.65</td><td>5%</td><td rowspan="2"></td><td></td><td rowspan="7">100%</td><td>Extreme Dry</td></tr>
<tr><td>-1.65 ≤ DFDI < -1.15</td><td>12.5%</td><td rowspan="5">90%</td><td>Severe Dry</td></tr>
<tr><td>-1.15 ≤ DFDI < -0.67</td><td>12.5%</td><td rowspan="3">75%</td><td>Moderate Dry</td></tr>
<tr><td>-0.67 ≤ DFDI ≤ 0.67</td><td>50%</td><td>Normal</td></tr>
<tr><td>0.67 < DFDI ≤ 1.15</td><td>12.5%</td><td>Moderate Wet</td></tr>
<tr><td>1.15 < DFDI ≤ 1.65</td><td>12.5%</td><td rowspan="2"></td><td>Severe Wet</td></tr>
<tr><td>1.65 < DFDI</td><td>5%</td><td></td><td>Extreme Wet</td></tr>
</table>

- **Riggs Creek OzFlux tower site (RG)**

The Riggs Creek OzFlux tower is located within the Goulburn-Broken catchment (Lat: -36.650°, Lon: 145.576°), in northern Victoria, Australia (Beringer 2014). The predominant landuse in this temperate region consists of dryland agriculture and pasture. Carbon dioxide, water vapour and latent/sensible heat are measured via the open-path eddy flux technique (at height of 2 m). The soil moisture contents and soil temperature are collected using installed sensors every 0.1 m across the profile. The utilized time series of verified data from this OzFlux tower is 1.5 years (having a minimum amount of data gaps) starting in December 2010.

- **Alice Springs OzFlux tower site (AS)**

The Alice Springs OzFlux tower is located on Pine Hill cattle station (Lat: -22.287°, Lon: 133.640°), near Alice Springs in the Northern Territory, Australia (Cleverly 2011). The landuse of this area is woodland characterized by a Mulga canopy in a generally arid to semiarid climate. The soil is overlying a 49m deep water table. The tower is 13.7m tall; and carbon dioxide, water vapour and heat measurements are collected via the open-path eddy covariance technique at 11.6m. Soil moisture and temperature measurements are collected in bare soil, Mulga, and understory habitats. Here, considering the availability of verified data from this OzFlux tower with a minimum of gaps, data from 2010-09-03 to 2013-06-30 are considered for this site.

- **Howard Springs OzFlux Tower Site (HS)**

The Howard Springs flux station is located in the Black Jungle Conservation Reserve in the Northern Territory, Australia (Lat: -12.495°, Lon: 131.150°) (Beringer 2013). The flux tower site is categorized as an open woodland Savanna (average tree height is 14–16m), and is found within a tropical climate. The tower is 23 m tall and instruments are installed at approximately 10 m above the ground. Carbon dioxide, water vapour and heat measurements are collected via the open-path eddy covariance technique. Soil moisture content is also measured at the site. Again considering the data gaps, for this OzFlux tower, data from 2011-01-01 to 2013-12-31 are applied.

Except for the observed data from the OzFlux network (Finnigan et al. 2003, 2004), the remaining two data sources used in this chapter are obtained from model output (APWM) (van Dijk 2010) and satellite data from MODIS on NASA’s Terra (Table 4.2). In the current chapter, standard, processed, freely available satellite data products are being used. Data quality flags, e.g.

for cloudiness or significant retrieval errors, were considered and the respective data points were filtered out. As this chapter uses point observations, it is assumed that the MODIS products used here (with a native resolutions of 250-500m) are representative of the location where the towers are found, and that no further processing is required in that regard. Considering the main elements of the hydrological cycle as well as drought types, a diverse set of variables/proxies was taken into consideration to derive the single DIs for this chapter (Table 4.2). 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 in Table 4.2.

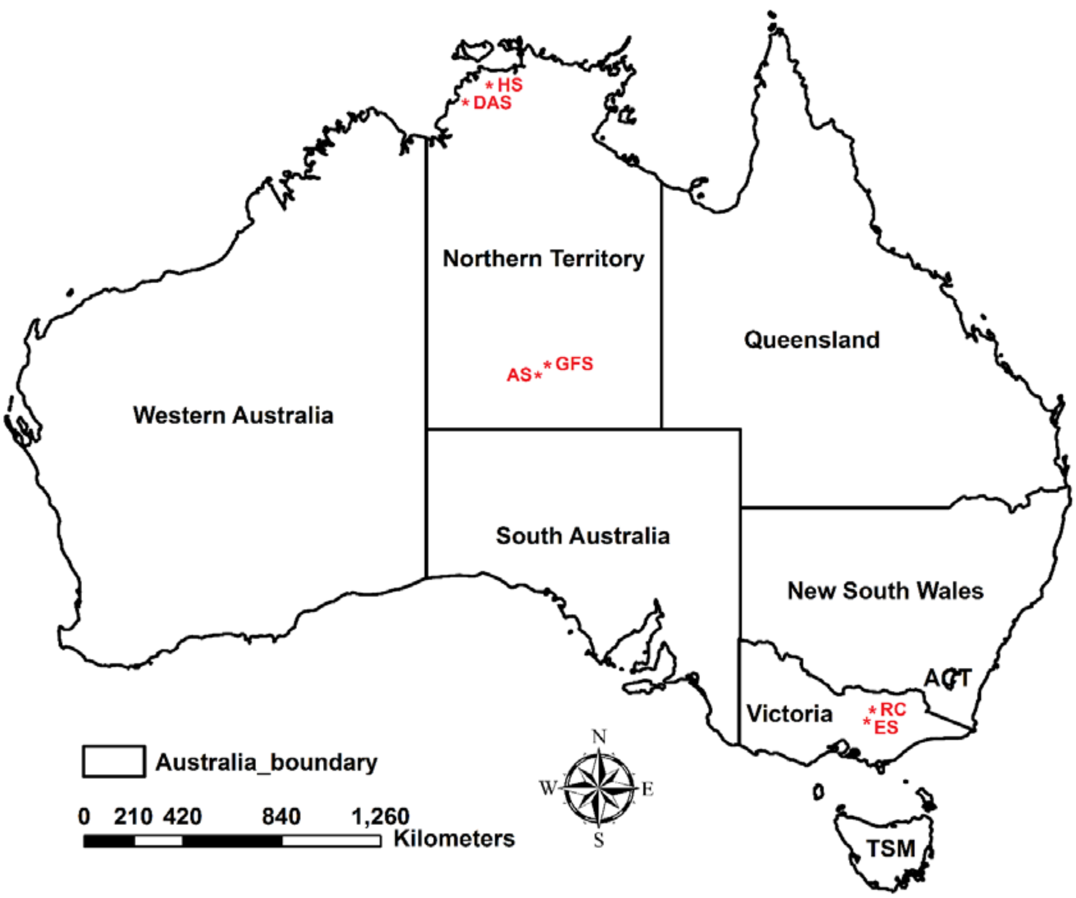


Figure 4.2. Locations of the three OzFlux tower sites and adjacent synoptic stations used in this study. HS: Howard Springs OzFlux Tower Site, DAS: Darwin Airport Station, AS: Alice Springs OzFlux Tower Site, GFS: Grape Farm Station, RC: Riggs Creek, ES: Euroa Station.

Table 4.2. Hydroclimatic variables/proxies used in this chapter.

Category	Data Sources	Primary Variables	Abbreviation	Units	Temporal Scale	Spatial Scale	Calculation
In situ observations	OzFlux Towers Network	Soil Moisture Content (Depth at 10cm)	SM	mm	0.5 hr	5* m	Direct Measurement
		Precipitation	P	mm	0.5 hr		
		Moisture Flux (Latent Heat)	MF	W/m^2	0.5 hr		
		Evaporative Fraction Index (Shuttleworth et al. 1989)	EFI	-	0.5 hr		$EFI = \frac{H}{Rn - G}$
A combination of *in situ* observations, satellite information and model output	Asia-Pacific Water Monitor (APWM) Section	Runoff & Surface Soil Moisture	ROI	mm	1 day	500 m	A combination of the output of several sources and models
Satellite information	MODIS-Terra Satellite	Normalized Difference Vegetation Index (Maki et al. 2004)	NDVI	-	1 day	250 m	$NDVI = \frac{NIR - VIS}{NIR + VIS}$
		Vegetation Condition Index (Patel et al. 2012)	VCI	-	1 day	250 m	$VCI = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}$
		Temperature Condition Index (Patel et al. 2012)	TCI	-	1 day	500 m	$TCI = \frac{Tb_{max} - Tb}{Tb_{max} - Tb_{min}}$
		Perpendicular Drought Index (Ghulam et al. 2007)	PDI	-	1 day	250 m	$PDI = \frac{1}{\sqrt{M^2 + 1}}(VIS + M \times NIF)$

NIR: the spectral reflectance measurements acquired in the near-infrared regions (700-1100 nm), *VIS*: the 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-VIS spectral feature space. H: latent heat flux (Wm^{-2}), Rn: Net radiation (Wm^{-2}), G: ground/soil heat flux (Wm^{-2}). *This is an approximation.

- **Data limitations**

The generally accepted approach to calculate long-term drought and water stress conditions requires time series data that extend beyond 30 years, so as to cover the full dynamic range of the local hydrology, while retaining sufficient data for validation purposes. The development of a general drought or water index, such as presented here, would therefore require a complete set of variables, including soil moisture, land surface water and energy fluxes, amongst others. However, such a data set does not exist *in situ* for such an extended range of variable and length of time.

The only observational data available for such a study come from comprehensive tower sites, such as the FluxNet network (ORNL DAAC 2016), which provides data for selected sites of up to 15 years, but with most sites in the range of 4-7 years of data availability. Hence, it is necessary to develop new algorithms with the limited data that are currently available. For all sites, the period chosen has covered a wide dynamical range of the local hydrology including the 2010 period when Australia transitioned from an extreme drought condition to a significantly wetter climate with significant precipitation events in the following years. It was therefore possible to capture almost the full range of conditions within a short time window. While not optimal, the simplifying assumption is made that this wide range of conditions has the same statistical properties as a longer time series, therefore allowing for the development of the drought and water stress index presented here. This aspect is taken up again in the Chapter 6 to address this limitation for future studies.

4.4 Results and discussion

For the calculation of the time series of the Perpendicular Drought Index (PDI), the values of the soil line slope in the NIR-Red spectral feature space were found to be 1.4, 0.94, and 1.02 for the Howard Springs, Alice Springs, and Riggs Creek tower sites, respectively. Depending on the environmental conditions and the available data (and therefore the derived DIs), the number of DI clusters may be more than two. This is actually expected, in particular for environmental conditions that are significantly different to the arid- and semi-arid examples in this chapter, as the underlying physical drivers may be different from region to region. In addition, inaccurate primary selection of individual DIs (i.e. selecting irrelevant indices with water stress issue which leads to forming more clusters) and/or inconsistencies stemming from data errors may influence the clustering. For the present chapter, considering only two clusters consisting each of water content and vegetation conditions, is considered sufficient for the appropriate evaluation of an ecosystem's water stress conditions. For example, Table 4.3 shows the SDIs of RG categorized into three clusters, of which Cluster 3 had only a single member (SEFI). This suggests that the information contained within SEFI (at least at RG), can be considered redundant or even irrelevant for the derivation of the main index, maybe due to data errors rooted within the measurements. As SEFI basically shows the output of a terrestrial ecosystem in form of evapotranspiration, it should consequently be located within the cluster representing water content-based indices. As for AS and HS, the SDIs were distributed into two clusters, in which Cluster 1 includes indices generally representing water availability, while Cluster 2 consists of proxies which describe vegetation conditions. The average values of PSs between members of each cluster are greater than 60%, underlining the accuracy of categorizing SDIs. Further, having values of the KMO test greater than 0.5 shows that performing multivariate methods on each cluster to derive aggregated variables was effective. In addition, with the Eigenvalues of PC1 (all exceeding 1) as well as variable variances covered by PC1 (with the minimum value of 70%), it can be stated that the PC1 derived from each cluster appropriately represents that cluster, as it can properly cover the majority of the statistical characteristics of the members of that particular cluster.

Table 4.3. Primary statistical information of individual SDIs at different regions.

Regions	Clusters	$\overline{PSs}$	KMO Test	Eigenvalues of PC1	Variance of variables covered by PC1
RG	Cluster 1: SPI, SSMI, SMFI, SROI, SPDI	62	0.59	2.3	77%
	Cluster 2: SNDVI, SVCI, STCI	75	0.61	2.2	70%
	Cluster 3: SEFI	N/A			
AS	Cluster 1: SPI, SSMI, SMFI, SEFI, SPDI	80	0.78	3.5	72%
	Cluster 2: SNDVI, SVCI, STCI	62	0.56	2.1	70%
HS	Cluster 1: SPI, SSMI, SMFI, SEFI, SPDI	87	0.84	3.7	75%
	Cluster 2: SNDVI, SVCI, STCI	62	0.59	2.4	81%

* $\overline{PSs}$: The average of PSs between members of each cluster; PC1: first principal component of variables located at a same cluster; RG: Riggs Creek; AS: Alice Springs; HS: Howard Springs; N/A: non-applicable for a single member cluster.

In this chapter, the lack of runoff data at the Alice Springs (AS) and Howard Springs (HS) OzFlux tower sites meant that the standardized runoff and surface soil moisture index (SROI) was only calculated and subsequently used to derive the SAWAI for the Riggs Creek (RG) OzFlux tower site. To determine the sensitivity of SAWAI to SROI at Riggs Creek, the process of deriving SAWAI was repeated but this time without considering the SROI (named SAWAI*). The four goodness-of-fit criteria of Spearman rank correlation coefficients (Scc), Pearson correlation coefficients (Pcc), Root Mean Square Error (RMSE), and Volume Error (VE) were then derived to compare SAWAI and SAWAI*.

The values of Scc=0.96, Pcc=0.97, RMSE=0.028, and VE=0.44 suggest that the impact of missing the runoff variable in deriving SAWAI can be mitigated by including other variables such as precipitation (eg. Standardized Precipitation Index (SPI)) and soil moisture (or Standardized Soil Moisture Index (SSMI)). Based on the presented sensitivity analysis and given that the main drivers of runoff are precipitation and antecedent soil moisture conditions, the assumption has been made that SROI may not be needed for the two other case studies, as SPI and SSMI implicitly contain the key information provided by SROI. A further analysis of the validity of this assumption and its transferability to other climatic zones, in particular temperate and humid regions is required, but it is suggested to be valid here, given the closeness of the surface hydrological characteristics of the three sites.

In order to implement the proposed methodology, the wet and dry, as well as the plant-physiologically active and non-active months were determined for each of the case study areas (Table 4.4) according to Section 4.2.1. For that purpose, 30 years of monthly data were used to classify active/non-active months. At Riggs Creek, the monthly historical data of precipitation, mean minimum and mean maximum air temperature of the synoptic station of Euroa (1980-2010), as well as historical information of monthly pasture growth in Victoria (DEPI 2015) were used for this purpose. Also, historical data of the synoptic stations at Grape Farm and Darwin Airport (1980-2010) were used to define wet/dry as well as active/non-active months for Alice Springs and Howard Springs, respectively (Table 4.4).

Time series of SAVI, SAWAI and the proposed DFDI for the study sites are shown in Figure 4.3. Figures 4.3A, 4.3D and 4.3G present the time series of SAWAI along with SSMI at the different locations. For all sites, SSMI and SAWAI appear to have similar trends; however,

Table 4.4. Specification of wet/dry months, active/non-active months, and proposed DFDI based on SAWAI/SAVI for each month at the different locations.

Month	Riggs Creek			Alice Springs			Howard Springs		
	Wet/Dry	Active/ Non-Active	Proposed DFDI	Wet/Dry	Active/ Non-Active	Proposed DFDI	Wet/Dry	Active/ Non-Active	Proposed DFDI
January	Dry	Non-Active	SAWAI	Wet	Active	SAWAI	Wet	Active	SAWAI
February	Dry	Non-Active	SAWAI	Wet	Active	SAWAI	Wet	Active	SAWAI
March	Dry	Non-Active	SAWAI	Dry	Active	SAVI	Wet	Active	SAWAI
April	Dry	Active	SAVI	Dry	Active	SAVI	Dry	Active	SAVI
May	Wet	Active	SAWAI	Dry	Non-Active	SAWAI	Dry	Active	SAVI
June	Wet	Non-Active	SAWAI	Dry	Non-Active	SAWAI	Dry	Active	SAVI
July	Wet	Non-Active	SAWAI	Dry	Non-Active	SAWAI	Dry	Active	SAVI
August	Wet	Non-Active	SAWAI	Dry	Non-Active	SAWAI	Dry	Active	SAVI
September	Wet	Active	SAWAI	Dry	Active	SAVI	Dry	Active	SAVI
October	Wet	Active	SAWAI	Dry	Active	SAVI	Dry	Active	SAVI
November	Dry	Active	SAVI	Wet	Active	SAWAI	Wet	Active	SAWAI
December	Dry	Non-Active	SAWAI	Wet	Active	SAWAI	Wet	Active	SAWAI

SSMI usually reflects water stress conditions more gradually and smoothly throughout a specified time span, and consequently cannot show significant sensitivities to short-term water stress fluctuations. This pattern is due to the long-term hydrologic memory of soil moisture (equivalent to low-frequency variability), in particular in comparison with other hydrological variables such as precipitation, which is spectrally white (Blender and Fraedrich 2006). This is one of the main reasons why drought indices such as SNDVI and DFDIs are preferred at weekly temporal resolution (Keyantash and Dracup 2004, Barua at al. 2011, 2012). Further, Figures 4.3B, 4.3E and 4.3H show that the time series of SAVI are close to SNDVI at all three case study sites. The reason of this high similarity lies in this fact that all SDIs which are located in Cluster 2 are derived from spectral measurements directly related to the vegetation conditions. In general, the highest values for SAVI are associated with growing months and the lowest ones are related to non-growing months or growing months with high water stress. In terms of memory, both SNDVI and SAVI have a relatively short-term memory when compared to SSMI.

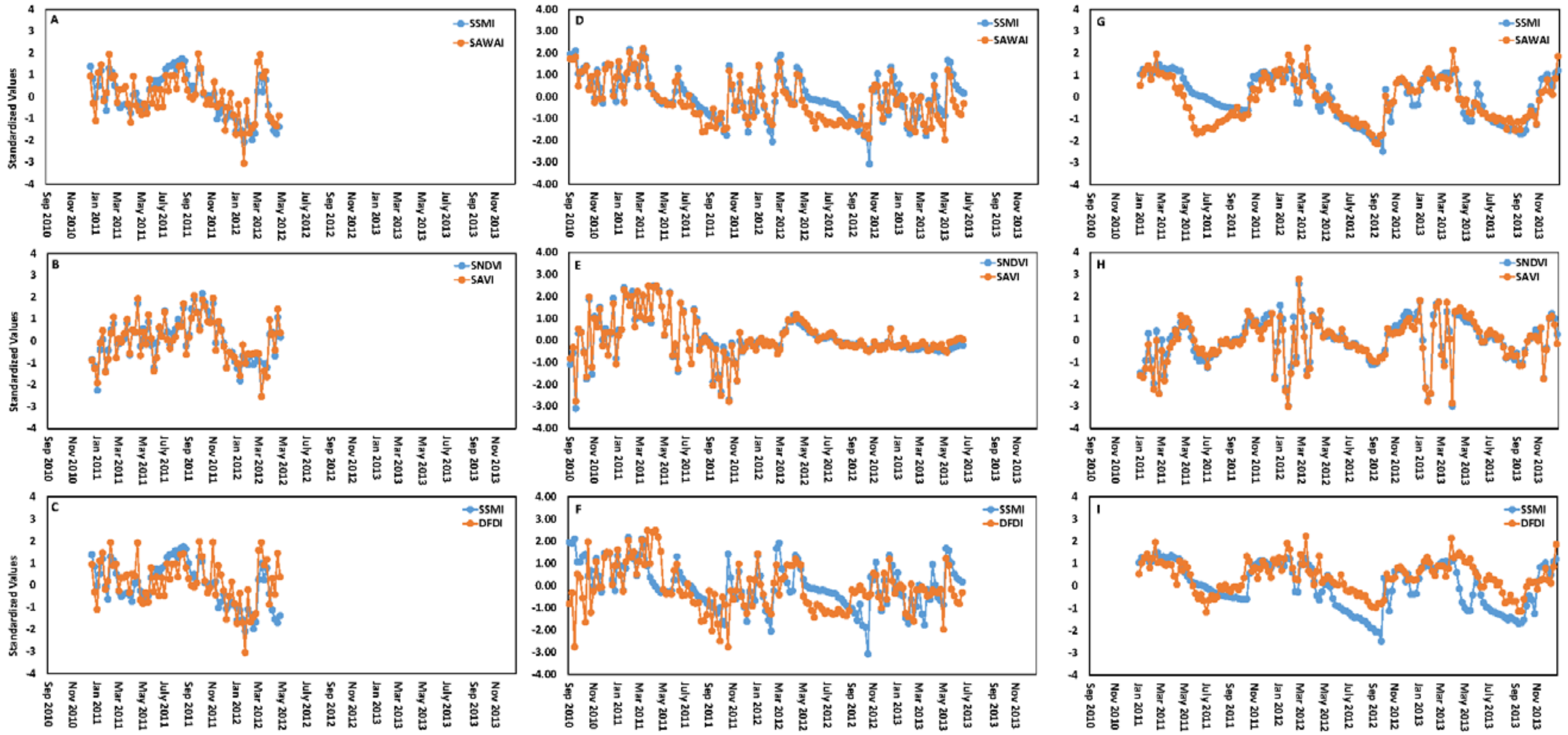


Figure 4.3. Time series of A) SAVI compared with SSMI, B) SAWAI compared with SNDVI and C) the proposed DFDI compared with SSMI at Riggs Creek OzFlux tower site. Descriptions of Panels D, E and F are same as A, B and C but for the Alice Springs OzFlux tower site. Descriptions of Panels G, H and I are same as A, B and C but for the Howard Springs OzFlux tower site.

After deriving the time series of the aggregated variables SAWAI and SAVI, the final proposed DFDI can be derived based on the results of Table 4.4 and Equations 4.1 and 4.2. Therefore, according to Table 4.4, at Riggs Creek the proposed DFDI was equal to SAVI for the months of April and November, and equal to SAWAI for the remainder of the year. While the soil moisture conditions in April and November may be described as “normal” (Figure 4.3A), the vegetation water stress are between moderate wet and severe wet (Figure 4.3B). The reason for this difference lies in the water consumption of the natural system. In fact, for both of these months, mean maximum air temperature is usually mild (at Riggs Creek it is between 10 ^{0}C and 25 ^{0}C) and therefore evapotranspiration is moderate and energy limited, rather than water limited.

For Riggs Creek, the drought conditions in April and November cannot comprehensively be described by water content-based indices such as SSMI and SAWAI, because at Riggs Creek these two months are categorized as dry and active months (Table 4.4), and consequently, it is necessary to employ a vegetation condition based index such as NDVI or SAVI to obtain more accurate information on the prevailing conditions (Figure 4.3C).

At Alice Springs, the DFDI values were equal for the months March, April, September and October based on SAVI. Performing cross-correlation analyses between SAWAI and SSMI showed a 2-month time lag which can also be noticed in Figure 4.3D. This is explained with the vegetation conditions of this area in September and October 2011 being associated with the soil moisture of July and August 2011, or rather the deeper root-zone conditions that are a consequence of the antecedent surface soil moisture. This is also reflected in an upward trend for DFDI between September and October 2011, while SSMI drops remarkably (Figure 4.3F).

The overall climate conditions and vegetation cover at Howard Springs may be considered as being between that of Riggs Creek and Alice Springs. Thus, the water stress conditions for this area should be analyzed based on a combination of reasons already considered at Riggs Creek and Alice Springs. According to Table 4.4, the months of April to October are active and dry. Therefore, according to Equations 4.1 and 4.2, the proposed DFDI was determined based on SAVI. In this region, the average mean maximum and mean minimum air temperatures from April to October are approximately 22 ^{0}C and 31 ^{0}C, respectively.

Those local conditions result in a moderate level of evapotranspiration, which can also be defined as the water consumption of the full terrestrial system, which is a combination of the lack of precipitation and lower average daily temperatures in winter. Based on the available data, both SAWAI and SSMI indicate significant levels of water stress, while the ecosystem is now defined as slightly under stress by the DFDI (within the range of 0 to -1). Consequently, throughout the period April to October, SSMI (or SAWAI) determined the water stress of this region to be higher than that proposed by DFDI (Figure 4.3I). Similar to Alice Springs, a 2-month time lag was also found between SAVI and SAWAI. For example, in March 2011 SAWAI started decreasing until May 2011, before increasing again until July 2011. This behaviour was imitated by SAVAI from May 2011 with a minimum in June 2011. However, during this period, SSMI was only decreasing (Figure 4.3I), showing that individual SDIs such as soil moisture (SSMI) cannot be used to evaluate and interpret all different aspects of water stress of a terrestrial ecosystem (Wilhite 2000, Van Loon and Van Lanen 2012).

- **Deriving mathematical equations for SAVI and SAWAI**

To complete the proposed methodology, explicit mathematical formulations for SAVI and SAWAI were derived using the symbolic regression method (Table 4.5), such that the proposed index could be applied in a practical way. As mentioned in the methodology section, the SDIs should be selected as independent variables of SAVI and SAWAI, as these indices are usually easy to quantify and have high similarities with dependent variables. At the study sites of Alice Springs and Howard Springs, considering the time series of available SDIs (as independent variables) and dependent variables (SAWAI and SAVI), three indices of SPI, SSMI and SEFI were applied to estimate SAWAI values. The SNDVI was used for deriving SAVAI values. At Riggs Creek, because of omitting SEFI after the clustering step, only two indices of SPI and SSMI were considered as independent variables to estimate the values of SAWAI.

In order to derive the final equations for the DFDI, the “Eureqa Formulize” software developed by Schmidt and Lipson (2009) was chosen to derive explicit mathematical equations between independent and dependent variables. In recent years, this software has been utilized in a variety of environmental issues as a reliable and accurate tool to evaluate symbolic regression based problems (Abrahart and Beriro 2012, Drobot et al. 2014).

Table 4.5. Derived mathematical equations along with goodness-of-fit criteria calculated on validation data for SAVI and SAWAI at the different locations.

Regions	Dependent Variable	Independent Variables	Mathematical Equations	Goodness-Of-Fit values calculated on validation data			Formula Evaluations	Confidence Stability	Confidence Maturity
				r^2	Max. Error	Mean Abs. Error			
RG*	SAVI	SNDVI	SAVI = SNDVI - 0.19*sin[delay(SNDVI,3) + sin(110.79*SNDVI)]	96%	0.34	0.14	$3.4e^{10}$	91.2%	99.5%
	SAWAI	SPI, SSMI	SAWAI = 0.58*SPI + 0.54*SSMI + 0.15*delay(SSMI,2) + 0.15*SPI^2 - 0.09*SSMI^2	93%	0.66	0.19	$8\ e^{10}$	93.3%	98.2
AS*	SAVI	SNDVI	SAVI = SNDVI + 0.07*SNDVI*sin(2.13*SNDVI + 1.15*SNDVI^2)	94%	0.36	0.11	$5.5\ e^{10}$	98.4%	99.5%
	SAWAI	SPI, SSMI, SEFI	SAWAI = 0.13 + 0.52*SEFI + 0.27*SSMI + 0.23* SPI	96%	0.37	0.08	$5.3\ e^{10}$	99.6%	99.8%
HS*	SAVI	SNDVI	SAVI = SNDVI -0.003/(1+ sin(1.17 + 130.17*SNDVI - 11.48*SNDVI^2))	95%	0.70	0.17	$3e^{11}$	94.5%	96.5%
	SAWAI	SPI, SSMI, SEFI	SAWAI = 0.06 + 0.55*SEFI + 0.34*SSMI + 0.07*SPI*SEFI + 0.09*SSMI^2 - 0.12*SPI^2	96%	0.30	0.10	$6.5\ e^{10}$	99.8%	99.8%

* RG: Riggs Creek; AS: Alice Springs; HS: Howard Springs; Formula Evaluation is the total number of equations which are examined; "Confidence Stability" refers to how long the solution list has not changed; "Confidence Maturity" relates to how much computational effort has been put into the current listed solutions. (Adapted from Eureqa User-Manual/website, http://formulize.nutonian.com/documentation/eureqa/).

Five main mathematical operators, exponential, trigonometry, and delay functions to derive the optimum mathematical equations subject to a variety of the regression complexities have been considered. The level of equation complexities grows when the number of contributing independent variables and mathematical functions increases. The optimum solution is an equation in which by increasing its complexity, the degree of its accuracy would not be changed meaningfully. This optimum explicit equation can be easily used to derive the values for the dependent variable. According to suggestions of previous research in modelling data by data-driven methods (Azmi et al. 2010, Araghinejad et al. 2011, Gharun et al. 2015), 50% of the entire data should be considered for the calibration stage, 25% for the validation stage and 25% for the verification stage; and for all these three stages both normal and extreme events should be present amongst the considered data set. As discussed above, only limited temporal data coverage is available for this study, which affects the approach that can be taken to derive and validate these equations.

In order to include a diverse range of hydrologic and vegetation conditions, a partial sampling approach is therefore considered here for the different stages of calibration, validation and verification (Araghinejad et al. 2011). This approach also overcomes the problem of overfitting the regression to dynamic regimes that may not be representative of the entire data set, as would be the case for a short and very dynamic data range when applying the traditional split sampling technique, and generally results in a better performance when compared to other approaches. The advantages and limitations of different sampling strategies for hydrological models have been presented in other papers, such as Juston et al. (2009), and therefore are not discussed in detail here.

Table 4.5 presents selected mathematical equations for modelling SAVI and SAWAI in the different regions. The goodness-of-fit results, confidence stability, and confidence maturity criteria show that the selected solutions were accurate and reliable. Amongst the selected equations, the regression which belongs to SAWAI at Howard Springs had the highest accuracy (Deterministic coefficient (r^2)=96%, maximum error=0.3, and mean absolute error=0.1) and reliability (confidence stability=99.8%, and confidence maturity=99.8%). Confidence stability and maturity values close to 100% show that the final list of solutions is strongly reliable (Eureqa User-Manual/website, http://formulize.nutonian.com/documentation/eureqa/). By using these explicit mathematical equations, users can easily derive the values of the proposed DFDI for future water stress monitoring over the considered case study areas; which presents a user-friendly and practical characteristic of the proposed methodology.

For the Riggs Creek OzFlux tower site, the accuracy of the mathematical solution versus its complexity is presented in Figure 4.4A and 4.4B. Figure 4.4A indicates that the optimum equation for SAWAI at Riggs Creek has a complexity size of 26 with mean absolute error of 0.19; while Figure 4.4B presents the optimum equation for SAVI at Riggs Creek as having a complexity size of 38 with mean absolute error of 0.14. Further, the plot of observations versus predicted values of SAWAI and SAVI are shown in Figure 4.4C and 4.4D respectively. These scatter plots show that the selected equations have the ability to model both non-extreme as well as extreme events appropriately, which is a desirable outcome for water stress monitoring issues.

It is expected to have different mathematical equations for areas with different climate regimes and/or landuse-landcover conditions. To apply this methodology spatially (i.e. national scale), the following steps need to be undertaken: 1- following the proposed methodology of this chapter, SAWAI/SAVI values would be derived for each grid point, 2- the area would be regionalized once, subject to climate regimes and also according to the landcover/landscape (the outcome of this step would be two regionalization maps), 3- for each regionalized map, a mathematical equation would be derived between SAWAI (SAVI) and the corresponding independently observed variables.

- **Proposed DFDI versus commonly used DFDIs**

So far, the proposed DFDI was derived on the basis of addressing the deficits of commonly used SDIs and DFDIs, and their advantages presented in comparison to single SDIs. Here, a quantitative comparison between the proposed DFDI and previously used DFDIs is presented. Time series of four commonly used DFDIs and the proposed DFDI are presented in Figure 4.5 for the three case studies. It is worth noting that to be able to compare the time series, each of the commonly used DFDIs were standardized by their standard deviation. To comparatively evaluate between commonly used DFDIs and the proposed DFDI, Spearman Rank Correlation (constrained between ±1) were used for determining the best aggregate correlation (Wilks 1995, Keyantash and Dracup 2004), as it removes the problem of non-Gaussian distributions. According to Table 4.6, the highest rank correlations were between LADI and NLADI (97%), as well as between WADI and AADI (89%), due to their similar base and foundations. The average rank correlations between the proposed DFDI and other DFDIs ranged from 60% (proposed DFDI vs NLADI) to 65% (proposed DFDI vs AADI), which shows a moderate correlation.

According to Figure 4.5, all DFDIs present similar trends from a general point of view. Interestingly, under circumstances in which the proposed DFDI is considered equal to SAWAI, the values of the proposed DFDI are mostly located between the maximum and minimum values derived from other DFDIs. Nevertheless, in other cases in which the DFDI is equal to SAVI, the proposed DFDI mostly presents maximum or minimum values with reference to others.

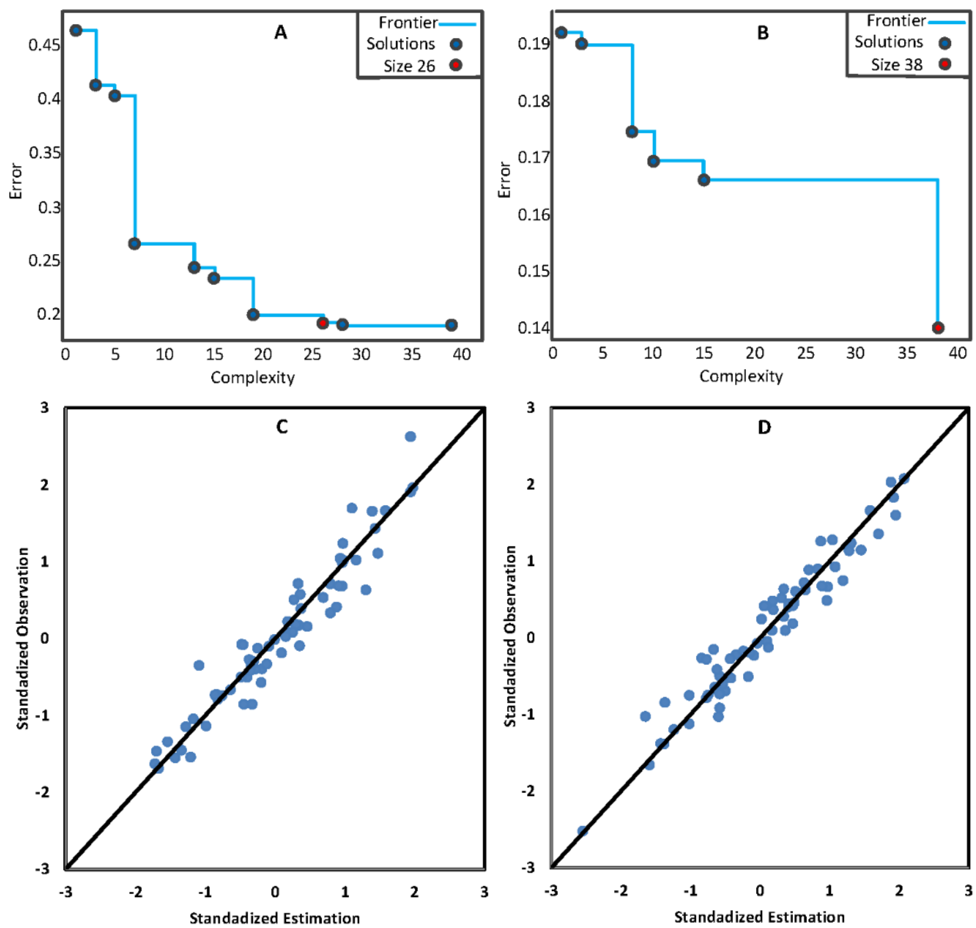


Figure 4.4. A & B) Accuracy of the mathematical solution versus complexity within the Symbolic Regression method for SAWAI and SAVI at Riggs Creek respectively; C & D) observations versus predicted values derived from the selected mathematical equation for SAWAI and SAVI at Riggs Creek respectively.

Table 4.6. Comparison between commonly used data fusion based drought indices and the proposed drought index using the Spearman rank correlation coefficient at the different locations.

DFDIs	LADI			NLADI			WADI			AADI		
	RG	AS	HS	RG	AS	HS	RG	AS	HS	RG	AS	HS
PDFDI	0.72	0.47	0.72	0.72	0.42	0.67	0.55	0.67	0.67	0.58	0.62	0.76
LADI				0.97	0.97	0.97	0.55	0.54	0.73	0.68	0.66	0.87
NLADI							0.44	0.46	0.71	0.58	0.59	0.84
WADI										0.85	0.91	0.91

*RG: Riggs Creek; AS: Alice Springs; HS: Howard Springs

For all three cases, AADI and WADI tend to reflect minimum values from April to September in comparison with others. In fact, the above-mentioned period covers the end of autumn, winter and early spring in the southern hemisphere, therefore the values of air temperature are the lowest, and consequently the values of AADI and WADI which are based on a weighted average of air temperature and two other hydrological variables drop regardless of the actual water availability situation and vegetation conditions.

Considering that LADI and NLADI are in fact aggregated drought indices of a set of SDIs derived from water contents variables, it was expected to have high similarities between these two indices and the proposed DFDI in months in which it is calculated based on SAWAI. However, during other months (proposed DFD is equal to SAVAI) remarkable discrepancies are apparent at times. This occurs because the reference of the proposed DFDI to detect the water stress of an area during those months is based on the vegetation conditions which in turn depends on the water availability of the area during the immediately preceding time steps (here previous weeks) due to the presence of a lag time between water availability and its influence on vegetation. One limitation of any statistical model is the requirement for data that describe the diverse hydroclimatological and plant-physiological variables, both in space and time. This study makes use of data collected at a variety of stations that provide such local datasets. In areas without advanced instruments, the current algorithm can be applied using input data from nearby synoptic stations (suitable where there is a comparatively low spatial variability in the surface conditions), along with global distributed data sources such as model outputs (e.g. reanalysis datasets such as NCEP/NCAR) and remotely sensed data products (e.g. MODIS and Landsat).

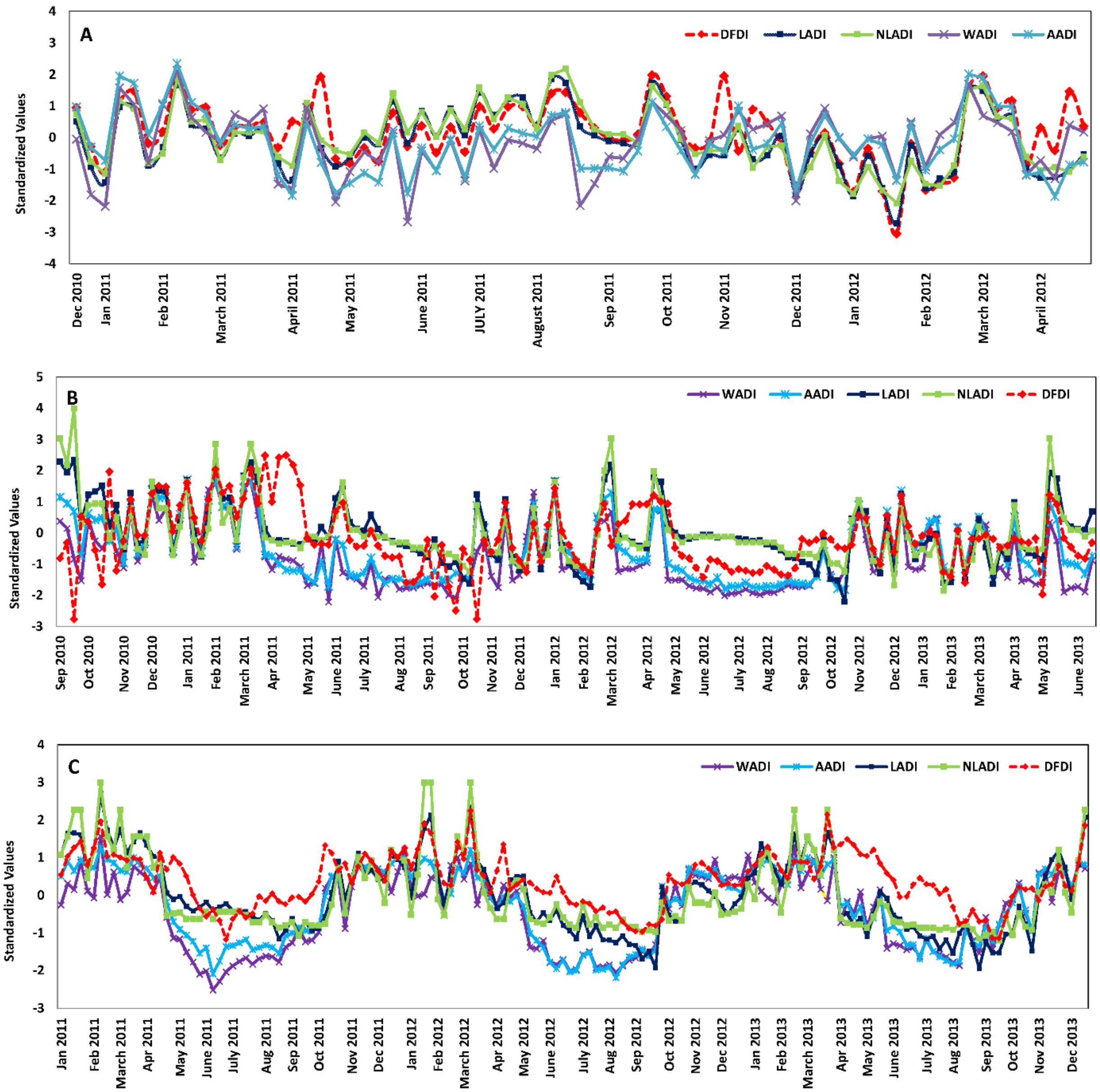


Figure 4.5. Comparison between commonly used and proposed DFDIs at A) Riggs Creek, B) Alice Springs, and C) Howard Springs.

4.5 Chapter summary

A comprehensive, robust and user-friendly water stress (drought) index is required by decision support system models to obtain information for an improved water resources management and planning. Nonetheless, evaluation and monitoring of water stress over a terrestrial ecosystem is much more complicated than using single drought indices such as SSMI or PDSI. In the current chapter, a data fusion-based drought index (DFDI) has been introduced to describe water stress based on coincident information from different single drought indices. In this way, the proposed DFDI uniquely describes water stress conditions beyond the traditional individual meteorological, hydrological, and agricultural subcategories. The proposed DFDI considers simultaneously the water availability of the system and the water stress conditions of vegetation to analyze different situations. Three case studies, with different climate and landuse conditions were selected to describe the ability to generalize the proposed methodology. The high resolution data and information required were provided from three sources; the OzFlux Network, the Asia-Pacific Water Monitor (APWM), and the MODIS-Terra satellite.

In the next chapter, the proposed drought index will be used for the spatial monitoring of short-term water stress conditions across Victoria, Australia. To validate the results of the spatial DFDI product, weekly DFDI time series of the Riggs Creek OzFlux tower site (from December 2010 to June 2012) derived in the current chapter will be used as a benchmark. Further, the derived DFDI's spatial maps will be comparatively evaluated with the maps derived from common drought indices such as SPI and SSMI.

Chapter 5

Spatial monitoring of weekly water stress across Victoria using DFDI

5.1 Overview

Throughout the previous chapter, DFDI was applied to evaluate weekly water stress conditions for three different point locations with a spatially limited data set. In the current chapter the previously developed approach will be applied to more extensive spatio-temporal data set with a longer time series to investigate its applicability in a spatial context. The study in this chapter demonstrates a spatial monitoring approach for weekly water stress across Victoria, Australia, using three data sources: i) the Australian Water Availability Project (AWAP), ii) the Soil Moisture and Ocean Salinity (SMOS), and iii) MODIS-Terra mission. Under diverse climate conditions and landuse/landcover situations, the DFDI maps are capable of showing meaningful results across different seasons throughout the state. Further, a set of mathematical equations have been derived for the state of Victoria to facilitate water stress forecasting. The contents of this chapter has been submitted as a journal paper:

Azmi, M., Rüdiger, C., Walker, J. Spatial Monitoring of Weekly Water Stress in Victoria, Australia, *Water Resources Management*. Under review.

5.2 Case study area and data sources

The state of Victoria is located in south-eastern Australia (from 33.97° S, 140.96° E to 39.2° S, 149.96° E), and is considered as the study area for this research. It is the second most populated state in Australia despite having the second-smallest land area (237,629 km^2) (Nedovic-Budic et al. 2004). The capital city of Victoria, Melbourne, is the second largest city of Australia with four distinct seasons and highly variable weather. In Victoria, spring generally lasts from September to November, and summer from December to February, while autumn conditions occur during March to May, and finally winter from June to August. From a climatological point of view, this state contains wet and temperate climate in the southeast, in winter snow-covered Victorian alpine areas, and extensive semi-arid regions in the west and northwest. This shows that precipitation events vary significantly across the state (Abdul Rauf and Zeephongsekul 2014). According to the Australian Valuation Property Classification Code (AVPCC), there are nine primary categories of land use in Victoria (VRO 2015) with a high ratio of essential industrial sectors suggesting that Victoria will be susceptible to adverse climatological conditions, and requires a comprehensive water stress monitoring framework.

This study utilizes three main data sources: i) spectral reflectance data from the MODerate resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite (Salomonson et al. 2001), ii) gridded precipitation and vegetation cover fraction datasets from the Australian Water Availability Project (AWAP) (Raupach et al. 2007), and iii) the gridded soil moisture dataset from Centre Aval de Traitement des Donnees SMOS (CATDS) (Jacquette et al. 2010). More information about each data source is presented as follows.

The MODIS instrument observes 36 spectral reflectance bands for hydrological, oceanic, and ecological applications with temporal resolutions ranging from daily to monthly. The MODIS/Terra instrument has been in operation since 2000 (Salomonson et al. 2001). Here, Version 5 of the daily spectral reflectance bands 1, 2, and 4 were used to derive the related drought indices with a nominal spatial resolution of 1 km (Table 5.1).

The original daily precipitation data for AWAP were supplied by the Australian Bureau of Meteorology (BoM), and are based on a 34-year observation period (1980-2014) collected from synoptic stations throughout the country. The data are already gap filled and verified, and gridded onto a 0.05^o continental grid by the AWAP team (Raupach et al. 2007) (Table 5.1). In terms of vegetation cover fraction, this product is globally available at $\sim 0.04^o$ spatial and monthly resolutions (from 1998-2005) (Gorbon et al. 2002), however it has been resampled to 0.05^o to suit AWAP (Table 5.1).

CATDS has been providing gridded SMOS L3/L4 products since 2010. Here, the daily gridded ($\sim 0.25^o$) surface soil moisture (L3) product of the ascending orbit was selected (Table 5.1). This product provides an estimate of the soil moisture for grid points with a temporal coverage of at least every 2-3 days (Jacquette et al. 2010).

Table 5.1 The main data sources and their primary products that are used in this chapter.

Data Sources	Category	Primary Variables	Abbreviation	Units	Temporal Scale	Spatial Scale
Australian Water Availability Project (AWAP) under the supervision of CSIRO Marine and Atmospheric Research (CMAR) in Australia	Combination of *in situ* observations, satellite information and model output	Precipitation	P	mm	1 day	0.05º
		Vegetation Cover Fraction	VCF	%	Mean Monthly	
Soil Moisture and Ocean Salinity (SMOS)	Satellite information	Soil Moisture Content (depth= 5 cm)	SMC	$\frac{m^3}{m^3}$	1 day	0.25º
MODIS-Terra Satellite		Band 1: Near-Infrared Region	NIR	nm		1 km^2
		Band 2: Visible (Red) Region	VIS			
		Band 4: Thermal Infrared	TIR			

5.3 Methodology of spatial water stress monitoring using DFDI

This chapter provides a spatial application of the DFDI methodology introduced by Azmi et al. (2016), and also discussed in the previous chapter, to demonstrate its value for future global applications. Figure 5.1 shows the detailed algorithm flow-path applied to the current study. After collecting the required hydroclimatological variables/proxies (Table 5.1), the spatial and temporal scales of all variables were aggregated to match the SMOS 0.25^{o} grid (by using a nearest neighbour interpolation approach to resample between cells with finer spatial scales to the sample size of 0.25^{o}) and a weekly basis, respectively. Next, six drought indices (Table 5.2) covering water content, water consumption, and vegetation conditions were derived for each grid cell and then standardized based on an equiprobability transformation (Shukla and Wood 2008). These standardized drought indices (SDIs) were then clustered via the Probabilistic Similarities (PSs) method introduced by Azmi et al. (2016). In particular, the clusters consisting of i) water content and ii) vegetation condition based SDIs have been found to provide relevant comprehensive information to reliably evaluate the water stress of an area. Thus, all subsequent steps only used the two above-mentioned clusters. Next, the multivariate method of the principal component analysis (PCA; Hotelling 1933), and Factor Analysis (FA; Kim and Mueller 1978) were performed on each of the clusters, and the first principal component (PC1) was derived. After standardizing the PC1s by dividing by their standard deviations, an arithmetic average between the PC1s of the two mentioned clusters derived from PCA and FA multivariate methods was calculated giving the final aggregated variable of that specific cluster. Hereafter, the final aggregated variables for clusters including water content and vegetation conditions based SDIs are labelled as standardized aggregated water availability index (SAWAI) and standardized aggregated vegetation index (SAVI) respectively, as defined in the previous chapters.

Following the defined steps, historical time series of the DFDI for each grid point q were determined subject to the conditions of time step i in terms of “wetness/dryness” and “activeness/non-activeness” (Figure 5.1). According to Azmi et al. (2016), for grid point q, a month is wet if the historical mean of the precipitation at that particular location and for that month is higher than the historical annual mean, otherwise it would be considered a dry month. For grid point q, a month is defined as photosynthetically active if the available plants are physiologically active and growing, otherwise it would be non-active. To allow for a qualitative interpretation of the DFDI, the thresholds introduced by Azmi et al. (2016) were utilized as shown in Table 5.3.

According to Figure 5.1, the final proposed data fusion based drought index (DFDI) was calculated as follows:

If time t is during a dry month and an active season, the water stress monitoring is assessed by SAVI, which reflects the water consumption and vegetation growth, according to

$$DFDI_t = SAVI_t;\ \&\ DFDI_j = SAVI_j = g(SDI_z; z = 1:n)_j \quad j = t+1,...,\infty\,, \tag{5.1}$$

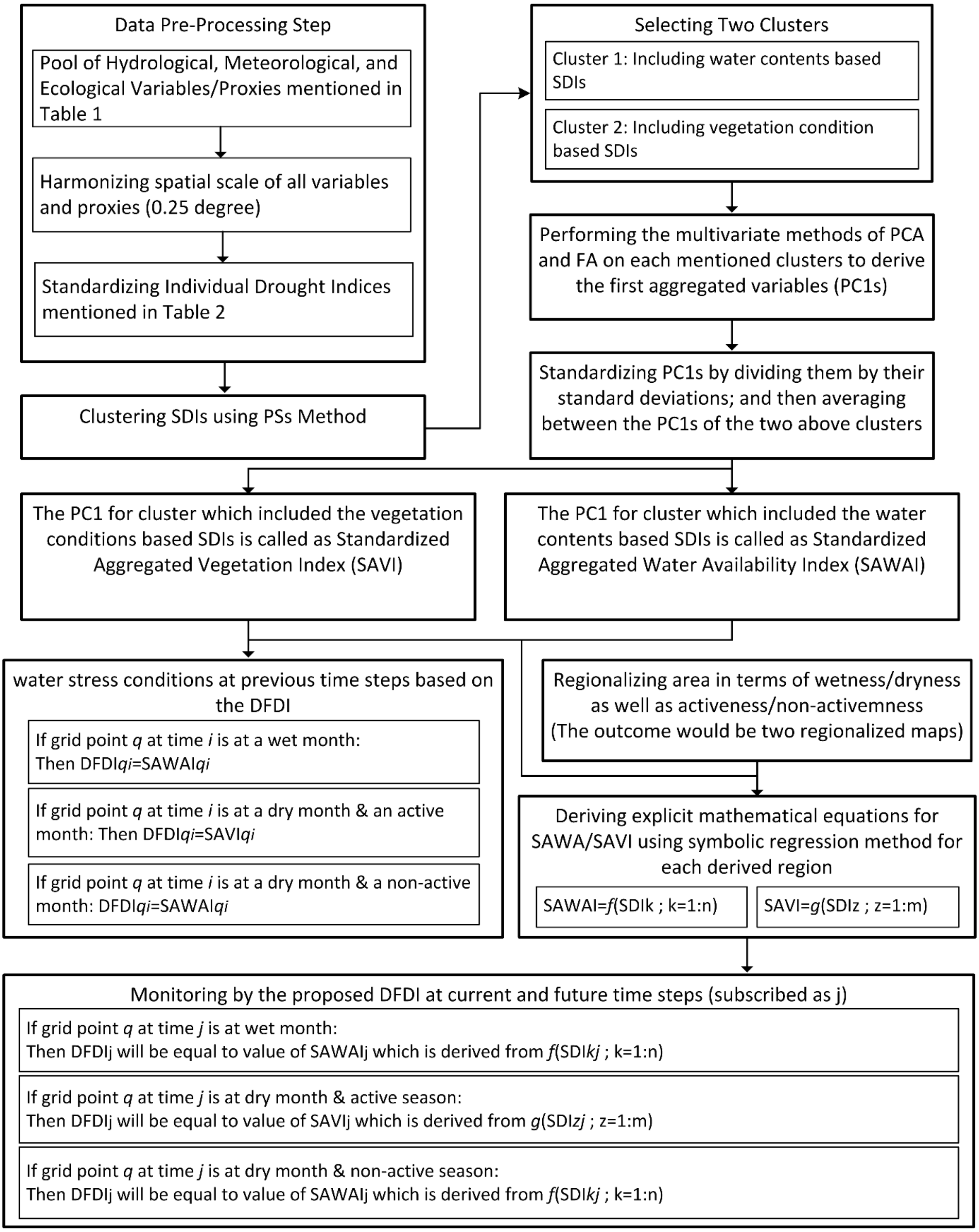
If time t is during a wet month, irrespective of whether that month is physiologically active or non-active, the water stress monitoring is assessed by SAWAI:

$$DFDI_t = SAWAI_t;\ \&\ DFDI_j = SAWAI_j = f(SDI_k; k = 1:n)_j \quad j = t+1,...,\infty\,, \tag{5.2}$$

If time t is during a dry month and a non-active season, the water stress monitoring is evaluated by SAWAI which reflects the amount of stored and plant-available water following Equation 5.2. One of the advantages of the DFDI methodology is to present explicit mathematical equations which make the process of water stress monitoring at current and future time steps more user-friendly.

To achieve this objective, the area is first regionalized according to the two criteria of wetness/dryness and active/non-active vegetation via a K-mean clustering method using mean monthly normalized data of precipitation and the vegetation cover fraction (or other vegetation-based indices such as NDVI which can detect the growth and activity of plants), respectively. The outcome of this step is to spatially cluster the state into regions where all points falling into each cluster follow similar behaviour in terms of wetness/dryness and active/non-active vegetation criteria in different months. Then, for each sub-region of the wet/dry (active/non-active) map, a mathematical equation is developed by employing the Symbolic Regression Method (Koza et al. 2003) in which the dependent variable is SAWAI (or SAVI) and independent variables are selected amongst members (SDIs) of the water content (or vegetation conditions) cluster.

Emphasis is given to select more readily available SDIs as independent variables, rather than more complex ones, in order to make the water stress monitoring as practical and directly applicable as possible. Finally, following the last step of the algorithm shown in Figure 5.1, DFDI values at the current and future time steps are calculated.



**f* and *g* are mathematical functions between dependent variables and their related independent variables. Subscribe *k* and *z* show the selected independent variables with total number of *n* and *m* respectively.

Figure 5.1. Schematic outline of water stress monitoring in spatial scales using DFDI.

Table 5.2. Overview of the individual drought indices used in this chapter.

Primary Variables	**Abbreviation**	**Temporal Scale**	**Spatial Scale**	**Calculation**	**Time period used to derive SDIs**
Antecedent Precipitation Index (Kohler and Linsley 1951)	API	weekly	0.25°	$API = \sum_{t=-1}^{-i} P_t k^{-t}$	1980-2014
Standardized Precipitation Index (McKee et al. 1993)	SPI			Standardizing Measurements	
Soil Moisture Index	SMI			Direct Measurements	There is no need to use a historical time series for calculating these indices. Values of SMC, NIR, VIS, and TIR are used to calculate daily SMI, NDVI, TCI, and PDI. These SDIs are then rescaled to weekly scales by calculating the arithmetic average over seven days.
Normalized Difference Vegetation Index (Maki et al. 2004)	NDVI			$NDVI = \frac{NIR - VIS}{NIR + VIS}$	
Temperature Condition Index (Patel et al. 2012)	TCI			$TIR = \frac{TIR_{max} - TIR}{TIR_{max} - TIR_{min}}$	
Perpendicular Drought Index (Ghulam et al. 2007)	PDI			$PDI = \frac{1}{\sqrt{M^2 + 1}}(VIS + M \times NIF)$	

Note: *i:* the number of antecedent days (here set to 7 days), *P*: the mount of precipitation during day *t*, *k*: a decay constant which can range from 0.80 to 0.98 (at current study *k* is set to 0.9 as suggested by Heggen (2001). *NIR*: the spectral reflectance measurements acquired in the near-infrared regions (700-1100 nm), *VIS*: the spectral reflectance measurements acquired in the visible (Red) regions (400-700 nm). *TIR*: Thermal Infrared of the spectral reflectance measurements acquired with band 4 of MODIS, TIR_{max} and TIR_{min}: maximum and minimum Thermal Infrared values within the considered area, respectively. *M*: the slope of soil line in the NIR-VIS spectral feature space (Ghulam et al. 2007) calculated as 1.024 for Victoria.

Table 5.3 Proposed thresholds for the DFDI (after Azmi et al. 2016).

Threshold Range	**Classifications**
DFDI < -1.65	Extreme Dry
-1.65 ≤ DFDI < -1.15	Severe Dry
-1.15 ≤ DFDI < -0.67	Moderate Dry
-0.67 ≤ DFDI ≤ 0.67	Normal
0.67 < DFDI ≤ 1.15	Moderate Wet
1.15 < DFDI ≤ 1.65	Severe Wet
1.65 < DFDI	Extreme Wet

- **Validation of DFDI maps**

To validate the results of the spatial DFDI product, weekly DFDI time series derived by Azmi et al. (2016) from data collected at the Riggs Creek OzFlux tower site (from December 2010 to June 2012) were used as a benchmark reference (also see previous chapter). The location of this tower is within the Goulburn-Broken catchment (Lat:-36.650^o, Lon:145.576^o) in northern Victoria, Australia (Beringer 2014). The closest map value to the OzFlux site is subsequently compared with those reference values. In addition to the qualitative comparisons which will be presented based on the thresholds shown in Table 5.3, a quantitative comparison via the index of agreement (d) (Willmott 1984) was also performed as follows:

$$d = 1 - \frac{\sum_{i=1}^{N} (ref_i - est_i)^2}{\sum_{i=1}^{N} (|est_i - ref_{mean}| + |ref_i - ref_{mean}|)^2}, \qquad (5.3)$$

where ref_i and est_i are the i^{th} reference (observed at Riggs Creek site) and estimated (derived from spatial maps) values of the considered variable (SAWAI/SAVI); and N is the total number of samples of the considered variable. The index of agreement (d) varies between zero and one; in which d=1 indicates a perfect agreement between reference and estimated values. Due to the fact that the aggregated hydroclimatological variables derived from the PCA method follow the standard Gaussian distribution function after standardization, by dividing by their standard deviations (Keyantash and Dracup 2004), the mean of the SAWAI/SAVI time series for Riggs Creek (ref_{mean}) can be set to zero for the current study.

5.4 Results and discussion

Considering the relatively short length of the available DFDI time series at Riggs Creek for validation, the availability of all used data sources, errors caused by atmospheric factors such as cloud cover, and different seasons in Victoria (leading to different combinations of wetness/dryness and activeness/non-activeness situations), eight time steps including the 1st and 2nd weeks of April 2011, the 1st and 2nd weeks of July 2011, the 1st week of October 2011, the 2nd week of January 2012, and the 1st and 2nd weeks of March 2012 were selected to derive DFDI maps for Victoria.

For instance, Figure 5.2 shows the maps of the traditional SDIs for the 2nd week of January (the middle of summer in Victoria) 2012 underlining the significant variations between the different indices. Here, larger positive values of SDIs such as the standardized SMI (SSMI) and NDVI (SNDVI), as well as SPI reflect a historically wet ecosystem as well as a higher level of activity for this vegetation and point in time, while more negative values show more severe water stress conditions, with the exception of the Standardized PDI (SPDI) in which the situation is reversed (more positive values indicate more severe water stress conditions). An initial evaluation of this figure suggests that during this period the entire state underwent a water content deficit (most clearly according to SSMI). However, at the same time the south-east and north-west appear to experience wetter conditions (according to SPI and SAPI). From a vegetation condition point of view (values of SNDVI, STCI, and SPDI), the vegetation in the southern parts of Victoria is in a relative sense more subjected to water stress, while the vegetation water stress conditions in the north-east region has eased. Available discrepancies in the evaluation and interpretation of water stress conditions subject to considering different SDIs is the main motivation to use new indices such as DFDI, which is based on the combination and integration of a set of SDIs to derive more comprehensive and consistent information.

As mentioned in the methodology section, the PSs method is employed for the SDIs to categorize them into different clusters. The aim of this step is to determine SDIs which will be the members of two clusters, representing water content and vegetation conditions, respectively. The results of this process are presented in Table 5.4, which shows that three groups have been formed. The members included in the first two groups have very high similarities (averaged PSs=100%, and α>0.9) with each other, in which the first group includes vegetation condition-based indices of STCI and SNDVI. The second group consists of the water content-based indices SPI, SAPI, and SSMI; and the third group consists of a single member (SPDI).

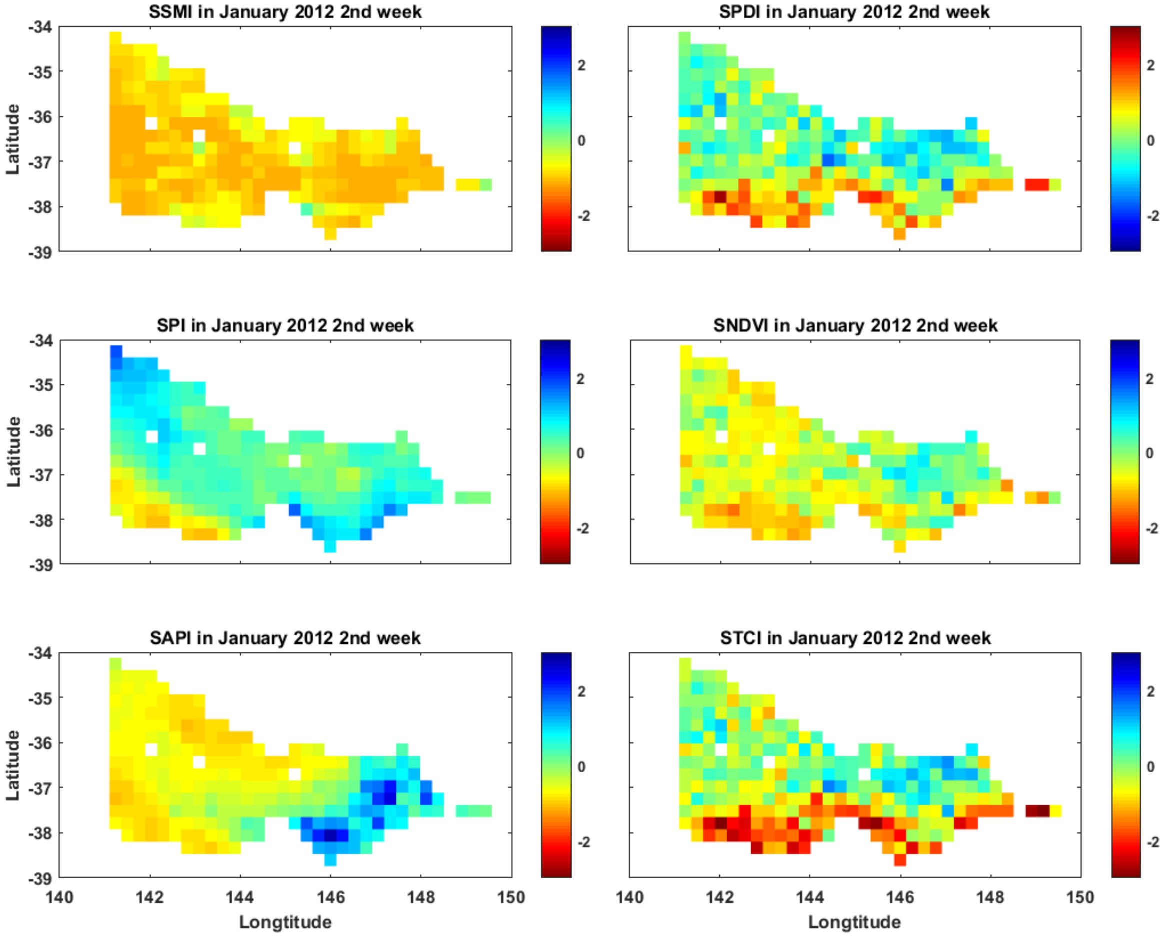


Figure 5.2. SDI maps for the 2nd week of January 2012 in Victoria. Dark blue colour shows the wetness of the ecosystem and or activeness of vegetation, while dark red colour indicates higher levels of the ecosystem dryness and or non-activeness of vegetation. The white grid points within the spatial data are due to extensive water bodies (i.e. wetlands).

Table 5.4. Probabilistic similarities, and Cronbach α test amongst members of derived clusters, and also statistical charactristics of PC1s derived from each cluster.

Clusters	$\overline{PS}s$	**α test**	**Eigenvalues of PC1**	**Variance of members by PC1**
STCI, SNDVI	100	0.92	1.76	77%
SPI, SAPI, SSMI	100	0.91	1.66	75%
SPDI	N/A			

*N/A: Non-Aplicable.

According to the equation of SPDI (Table 5.2), this index holds some information related to both water content and vegetation conditions, and that is why it was detected as a new cluster with redundant, overlapping, and more importantly partly inaccurate information. In terms of accuracy, SPDI was found to lack accuracy on densely vegetated fields such as pastures and agricultural fields (Ghulam et al. 2007, Ghulam et al. 2008), which are common in northern and western Victoria. In order to assess the impact of SPDI on the cohesiveness of each cluster, it was added to two other variable groups, and an α test was performed, which led in both cases to values of less than 0.4. This shows that merging SPDI with other clusters of the current study can decrease the accuracy of those clusters remarkably, and consequently SPDI was ignored for the remaining steps. After aggregating the SDIs of each cluster representing vegetation water content conditions, the aggregated variables which are representative of all SDIs were selected. To achieve this goal, the Eigenvalues (>>1) as well as variances of indices covered by the PC1s (>0.75) were considered according to the Kaiser1 method (Kaiser 1960). Subsequently, it was concluded that the PC1s of the first two groups can appropriately contain the charactristics and information of the entire cluster, and consequently can be regarded as representatives of water content and vegetation conditions-based SDIs, respectively.

Following the clustering of the SDIs, the combination process was performed to derive two aggregated variables of SAWAI and SAVI. For example, Figure 5.3 presents the results of this process for the 2nd week of July 2011. Considering the left panels of this figure, water content based SDIs (SSMI, SPI, and SAPI) indicate drier conditions in the north, south-east and particularly north-west in comparison to other parts of Victoria, which is well reflected in the SAWAI (lower left panel). In addition, the right panels represent the water stress of the vegetation particularly across the eastern half of the state, with the exception of the very east and south-east regions due to precipitation events during the 1st and 2nd weeks of July 2011 across those areas which provided enough water vegetation activities. As for the SAWAI, SAVI is able to represent the spatial patterns reasonably well.

It is worth noting that these vegetation-based indices show some inconsistencies in the maps, which can be due to a spatially diverse landuse-landcover in Victoria, and also spectral measurement errors in deriving the vegetation based SDIs such as NDVI and STCI at an aggregate level.

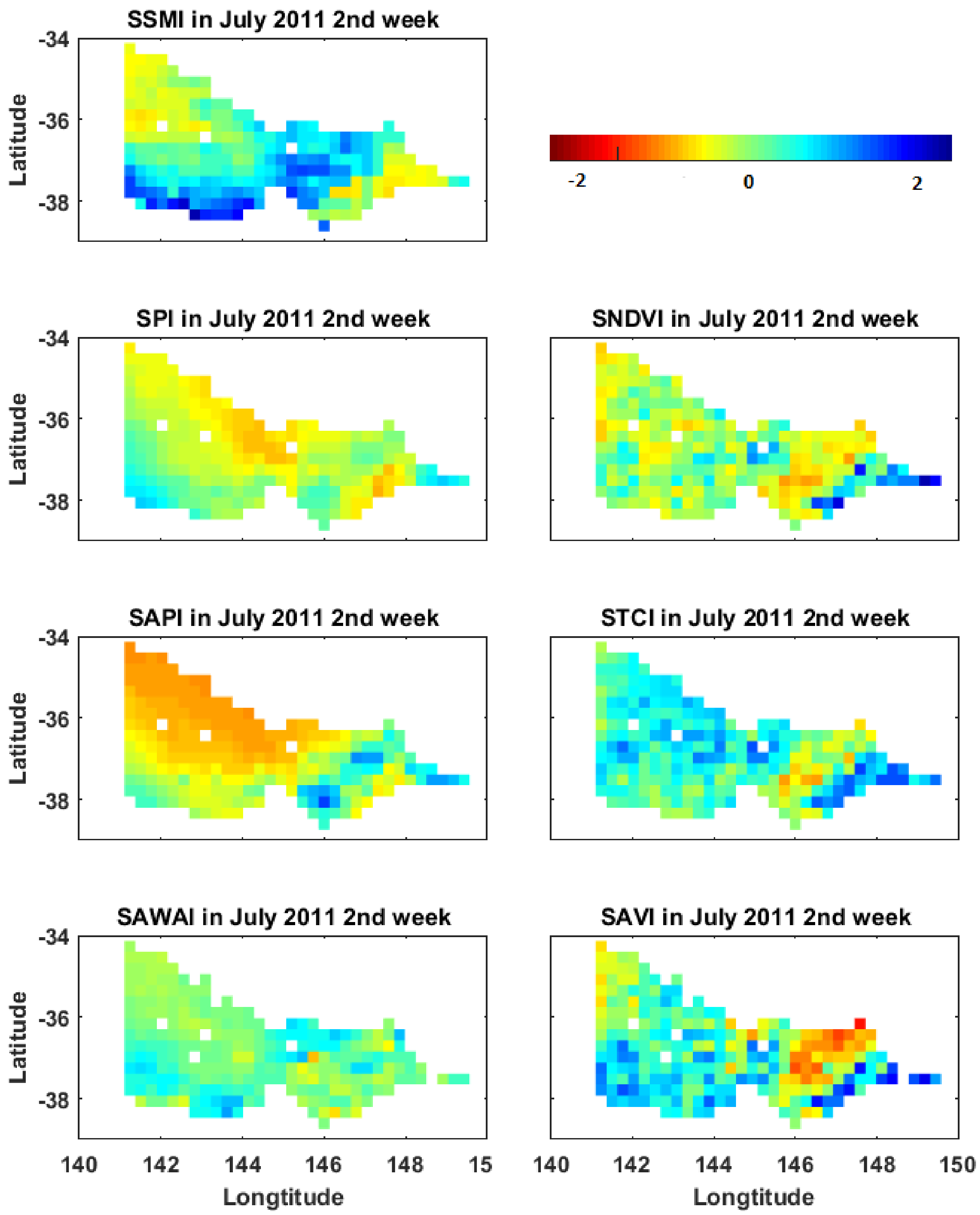


Figure 5.3. SAWAI map (bottom-left panel) derived from the aggregation of SSMI, SPI and SPAI maps, as well as SAVI map (bottom-right panel) derived from the aggregation of SNDVI and STCI maps for the 2nd week of July 2011. The white grid points within the spatial data are due to extensive water bodies (i.e. wetlands, and natural lakes).

Figure 5.4 presents the results of regionalization with respect to active/non-active vegetation (three regions of north-west, center and some parts of west, and finally east and some parts of south-west) and wet/dry conditions (seven regions from north-west to very end east) criteria. As mentioned before, the regionalization is implemented via the K-mean clustering method using mean monthly normalized data of precipitation and the vegetation cover fraction for wetness/dryness and active/non-active criteria respectively. Considering the modified Köppen climate classification (Peel et al. 2007) and employing 30 years of data from 1961-1990 (BoM 2016), in Figure 5.4 (the top panel) Region 1 can be regarded as grassland and Region 2 and 3 as temperate. According to Table 5.5, wet months mostly occur from the end of autumn (May) to the end of spring season (November), while in terms of vegetation activity, September and October (spring season) are considered as active months for the whole of Victoria.

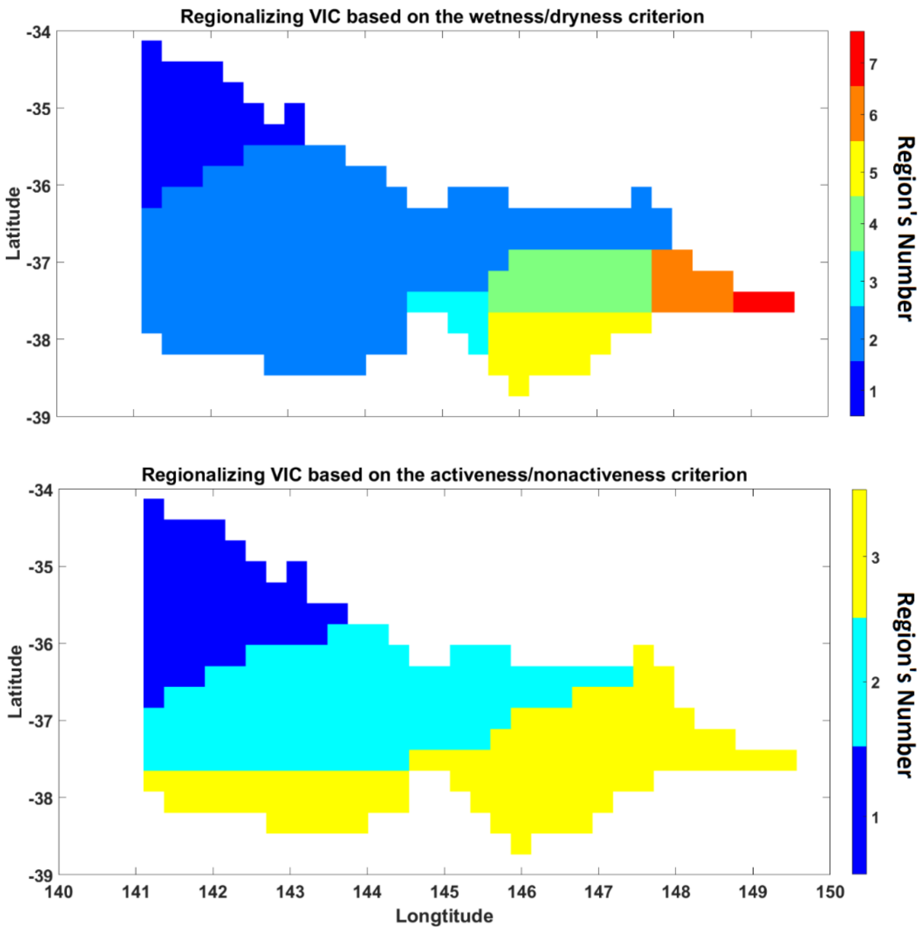


Figure 5.4. Regionalization of Victoria in terms of wet/dry (top panel), and active/non-active (bottom panel) criteria.

Nonetheless, in southern and north-eastern Victoria (Region 3 based on active/non-active criterion), active months start from July and end in November. By comparing the lower panel of Figure 5.4 with the landuse/land-cover information of Victoria, Region 1 is covered by sparse woody plants and wheat and cereal fields, Region 2 covered almost entirely by cereals and legumes fields as well as pastures, and finally Region 3 mostly included softwood and native woody plants.

Finally, based on the results of the two aggregated variables SAVI and SAWAI, as well as the information provided in Figure 5.4 and Table 5.5, the final DFDI maps for the considered time steps are derived and presented in Figure 5.5. A qualitative assessment of those maps shows consistent spatial patterns throughout. In Victoria, April is a non-active month throughout, therefore DFDI is derived based on the values of SAWAI. The calculations for the 1st week of April 2011 indicate severe and extreme dry conditions for the north and south-west while the remainder of the state is in normal conditions.

Table 5.5. Monthly wetness/dryness and activeness/non-activeness situations of derived regions in Victoria.

Month	**Regions regionalized based on the Wetness/Dryness Criterion**							**Regions regionalized based on the Act./Non-Act. Criterion**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**1**	**2**	**3**
Jan.	W	D	D	D	D	D	D	NA	NA	NA
Feb.	D	D	D	D	D	W	D	NA	NA	NA
March	D	D	D	D	W	D	D	NA	NA	NA
Apr.	D	D	W	D	D	D	W	NA	NA	NA
May	W	W	W	D	D	D	W	NA	NA	NA
June	W	W	W	D	W	W	W	NA	NA	NA
July	W	W	W	W	W	D	W	NA	NA	A
Aug.	W	W	W	W	W	D	D	NA	A	A
Sep.	W	W	W	W	W	D	W	A	A	A
Oct.	W	W	W	W	W	W	W	A	A	A
Nov.	W	D	W	W	W	W	W	NA	NA	A
Dec.	W	D	W	W	W	W	D	NA	NA	NA

*W: Wetness, D: Dryness, A: Activeness, NA: Non-Activeness.

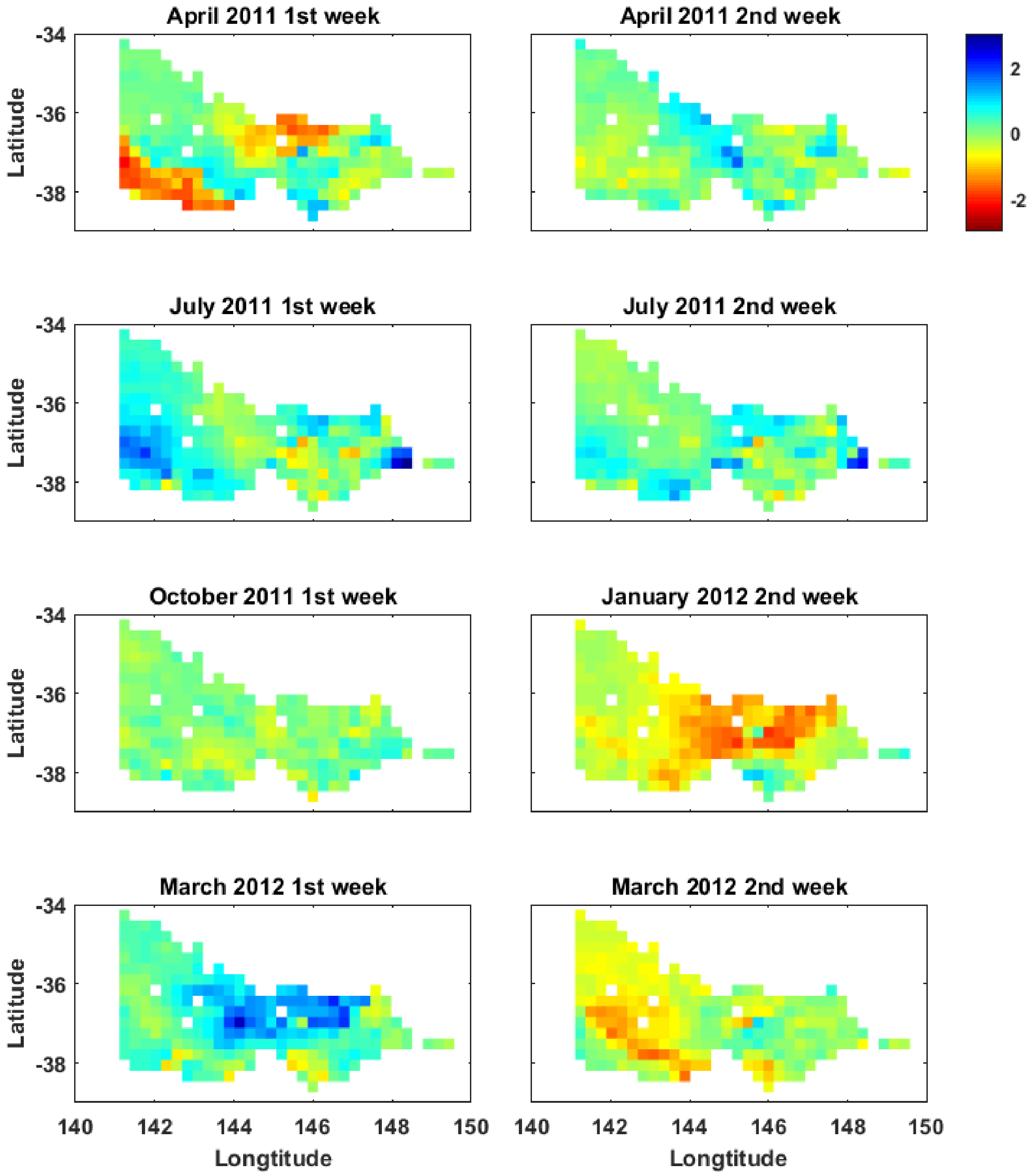


Figure 5.5 Final water stress maps of Victoria based on the DFDI for the eight different time steps. The white grid points within the spatial data are due to extensive water bodies (i.e. wetlands, and natural lakes).

With respect to precipitation events during the 2nd week of April, severe and extremely dry conditions in northern and south-western Victoria have turned into moderately dry and normal conditions in the 1st week of April. In contrast, for the 1st week of March 2012, Victoria generally experienced less water stress conditions compared to the following week, due to the lack of precipitation events in the 2nd week of March and also the presence of water consumption in form of evaporation and possibly percolation. Considering Figure 5.5, water stress conditions throughout Victoria during the first two weeks of July 2011 generally indicated a range of normal (north and north-west) to extreme wet conditions (small areas in east and south-west), which were caused by precipitation events in winter and a generally low level vegetation water consumption.

In terms of validating these spatial maps, comparisons were made to the Riggs Creek OzFlux tower site. Table 5.6 shows the comparative evaluations between the values of DFDI derived from the different panels of Figure 5.5 with the DFDI time series of Riggs Creek OzFlux tower site derived by Azmi et al. (2016). The index of agreement (*d*) of 0.81 suggests a strong fit between the point and spatial data, supporting the assumption that spatial maps can be dervied from satellite data.

Table 5.6 Comparison between the DFDI values derived from spatial maps at the same geographical location as the Riggs Creek OzFlux tower site with the time series of Riggs Creek OzFlux tower data derived by Azmi et al. (2016).

Events	**Areas**	**DFDI's value**	**Qualitative Conditions**
April 2011, week 1	Riggs Creek	-0.36	Normal Conditions
	Spatial Maps	0.50	Normal Conditions
Apr. 2011, week 2	Riggs Creek	0.08	Normal Conditions
	Spatial Maps	0.41	Normal Conditions
July 2011, week 1	Riggs Creek	0.03	Normal Conditions
	Spatial Maps	0.97	Moderate Wet Conditions
July 2011, week 2	Riggs Creek	0.09	Normal Conditions
	Spatial Maps	0.27	Normal Conditions
Oct. 2011, week 1	Riggs Creek	0.74	Moderate Wet Conditions
	Spatial Maps	1.30	Severe Wet Conditions
Jan. 2012, week 2	Riggs Creek	-0.79	Moderate Dry Conditions
	Spatial Maps	-1.6	Moderate Dry Conditions
March 2012, week 1	Riggs Creek	0.78	Moderate Wet Conditions
	Spatial Maps	1.90	Extreme Wet Conditions
March 2012, week 2	Riggs Creek	0.60	Normal Conditions
	Spatial Maps	0.91	Moderate Wet Conditions

The qualitative interpretation of the DFDIs following Table 5.3 shows that in 50% of the cases, the descriptors from both spatial maps and Riggs Creek tower are identical; while in other cases spatial maps tend to reflect the water stress conditions more extreme than what has been derived by Azmi et al. (2016) at the location of Riggs Creek OzFlux tower site. As the Riggs Creek OzFlux tower is the only available point scale reference in Victoria located in relatively open conditions, it is assumed here that the accuracy of the spatial DFDI maps with respect to this reference can be appropriately accepted.

In order to derive the final DFDI equations for future time steps, the "Eureqa Formulize" software developed by Schmidt and Lipson (2009) was used. Five main mathematical operators together with exponential, logical, and trigonometric functions were considered to derive the optimum mathematical equations, subject to a variety of regression complexities. The level of equation complexities grows when the number of contributing independent variables and mathematical functions increases. The optimum solution is an equation in which by increasing its complexity, the degree of its accuracy would not be changed meaningfully. Table 5.7 presents the optimum derived mathematical equations for each region along with their goodness-of-fit values calculated with the validation data set. According to this table, the best values of goodness-of-fit criteria belong to equations of SAVI (r^2=0.85, maximum error=0.7, and mean absolute error=0.25) for Region 1, and SAWAI (r^2=0.80, maximum error=0.7, and mean absolute error=0.23) for Region 4. While the worst results are r^2=0.81, maximum error=1.3, and mean absolute error=0.46 for SAVI for Region 3 as well as r^2=0.5, maximum error=1.8, and mean absolute error=0.41 for SAWAI for Region 2.

Using the derived mathematical equations, DFDI maps can be provided for different time steps as shown in Figure 5.6. Comparing Figures 5.5 and 5.6 then indicates that the central north and south-west regions of Victoria have the highest differences. This point is highlighted through the maps of the 1st week of April 2011, 2nd week of January 2012, and 2nd week of March 2012. However, the estimated maps have been able to appropriately track the spatial trends of the extreme events. These differences are mostly due to highly diverse spatial patterns of landuse/landcover and climate regimes in these parts of Victoria and the significantly lower spatial scales of the input data and their associated uncertainties, when compared to local point measurements.

Table 5.7 Derived mathematical equations along with goodness-of-fit criteria for SAVI and SAWAI calculated on the validation stage of modelling mathematical equations at different regions of Victoria.

Categories	Regions Based on Categories	Dependent Variable	Independent Variables	Total No. of Obs.*	Mathematical Equation	Goodness-Of-Fit Values		
						r^2	ME	MAE
Activeness/ Non-Activeness	1	SAVI	SNDVI, STCI	448	SAVI = (-25.417 - 1.845×SNDVI)/(STCI - 5.854) - 4.630	0.85	0.7	0.25
	2			832	SAVI = 0.595×SNDVI + 0.573×STCI - 0.097	0.79	1.1	0.33
	3			1168	SAVI = 0.218 + 0.565×SNDVI + 0.427×STCI + 0.026×STCI×$SNDVI^2$ - 0.071×$STCI^2$	0.81	1.3	0.46
Wetness/ Dryness	1	SAWAI	SPI, SSMI, SAPI	312	SAWAI = 0.368 + SSMI/(2.327 + SAPI)	0.72	0.5	0.14
	2			1400	SAWAI = SSMI/(1.483 + SPI^2 + SSMI×if(SSMI>0, SSMI, SPI) - 1.941×SSMI×SPI)	0.52	1.8	0.41
	3			104	SAWAI = 0.823 + 0.479×SPI + 0.096×SPI×SAPI + 3.042/(-2.491 - SSMI)	0.62	0.7	0.21
	4			184	SAWAI = SPI^2 + 0.317×if(SPI>0, if(SAPI>0, $SSMI^2$, SSMI), SAPI) - 0.125 - 0.145×0.152^{SSMI} - 0.302×SPI^3×if(SSMI>0, SPI^2, SPI)	0.80	0.7	0.23
	5			285	SAWAI = 0.280×SSMI×SPI + 0.280×SPI×SAPI - 0.120	0.49	0.9	0.28
	6			83	SAWAI = 0.649 + 0.740×SSMI	0.66	0.6	0.11
	7			80	SAWAI = 0.053 + 1.026×SAPI + 0.621×SSMI×SAPI - 1.262×SPI - 0.311×SSMI×SPI^3	0.64	0.4	0.18

* Total No. of Obs: Total number of observations (samples) for each dependent/independent variable used to derive the regressions (including calibration, verification and validation stages). MAE: Mean Absolute Error. ME: Maximum Error.

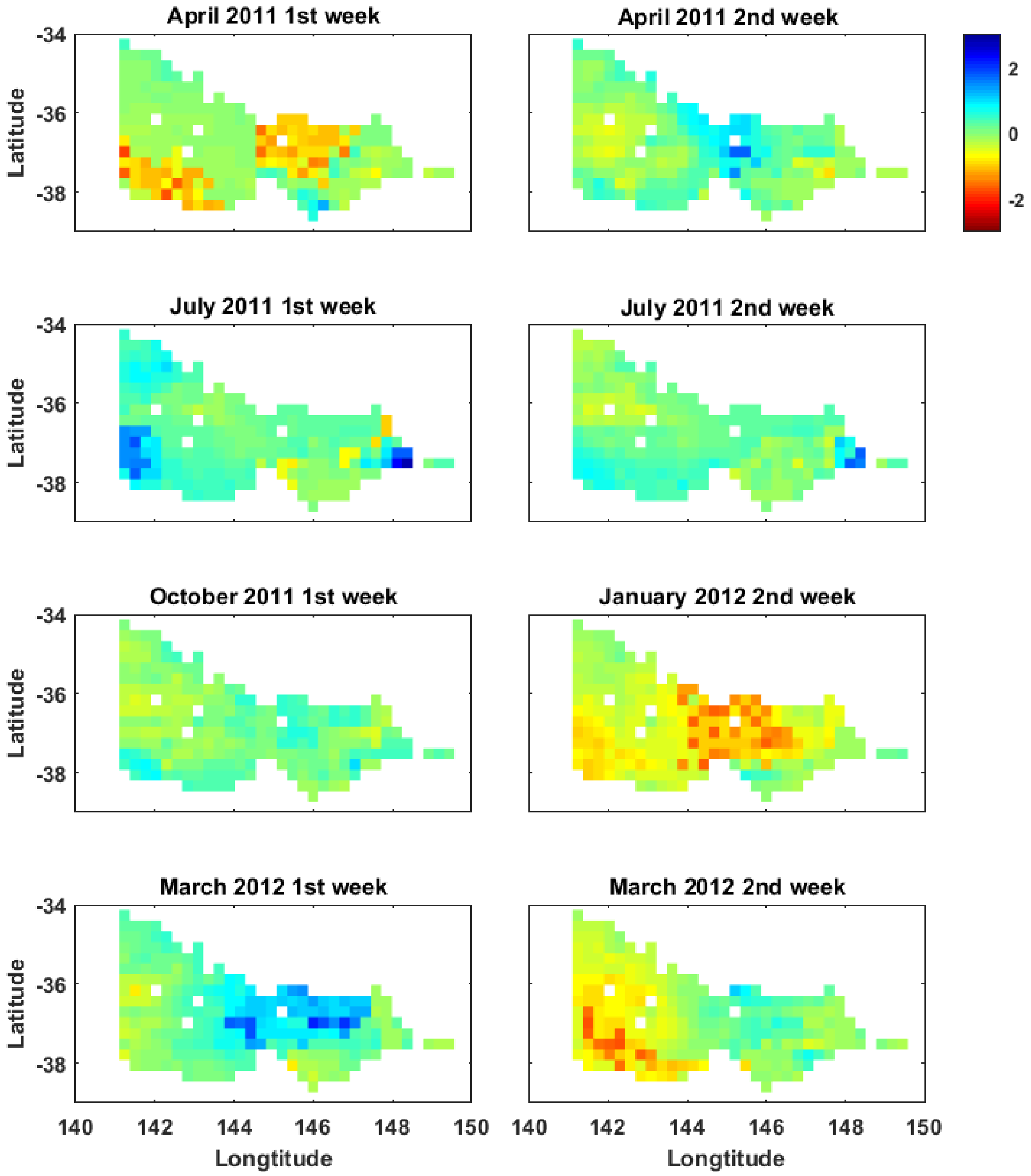


Figure 5.6. Final water stress maps of Victoria based on the estimated DFDI for the eight different time steps. The white grid points within the spatial data are due to extensive water bodies (i.e. wetlands, and natural lakes).

This issue needs to be investigated more thoroughly in the future by splitting the central north and south-west parts of Victoria into more homogeneous sub-regions by considering other regionalizing techniques and methodologies. In fact, a variety of more sophisticated geostatistical methods, as well as different sets of data (i.e. using NDVI instead of vegetation cover fraction) can be considered for regionalizing a specific area which may lead to different rationalizing maps with different uncertainties. However, this is well beyond the scope of this thesis, as other items that would also have to be considered then are the observational uncertainties and heterogeneities of each pixel, the annual changes, in particular for agricultural zones, and how this can be incorporated into the DFDI.

In order to use this methodology for practical operations such as spatial water stress predictions, forecast values from global/local distributed land surface or climatological model outputs may be converted into SDIs. Those are then substituting the historical/current observations in the derived equations to determine future levels of SAVI and SAWAI. Eventually, using Table 5.5 and also Equations 5.1 and 5.2, the final DFDI values are determined much in the same way as with the satellite data.

Since there is a lack of real ground-truth across Victoria in the form of accessible maps and information, the validation process is limited to comparing the DFDI maps with commonly used drought indices, such as soil moisture and precipitation. Nonetheless, Exceptional Circumstances data available through the Queensland government can be regarded as an acceptable ground truth for validating any new indices such as the DFDI. First results have shown promise in terms of the accuracy of the DFDI in spatio-temporally detecting drought events in comparison with the commonly used drought indices (e.g. SPI and soil moisture). Nonetheless, further analysis would be one of the main future research directions. For instance, Figure 5.7 presents spatial DFDI map along with Rainfall Percentile Ranking (RPR) and an EC map for March 2013. Comparing the maps obviously illustrates that DFDI has been able to detect the water stress conditions much more properly in comparison with RPR which is currently utilized by the Australian Bureau of Meteorology, in particular across the northern parts of the continent. An in-depth quantitative comparison between the accuracy and reliability of different drought indices in Australia would be one of the future directions.

5.5 Chapter summary

In order to develop a more inclusive and efficient water resources management and planning system, a comprehensive and reliable drought index is required to detect and even predict the water stress conditions of an area. Azmi et al. (2016) introduced a data fusion based drought index (DFDI) that combines different single drought indices by considering simultaneously the water availability of the system and the water stress conditions of the

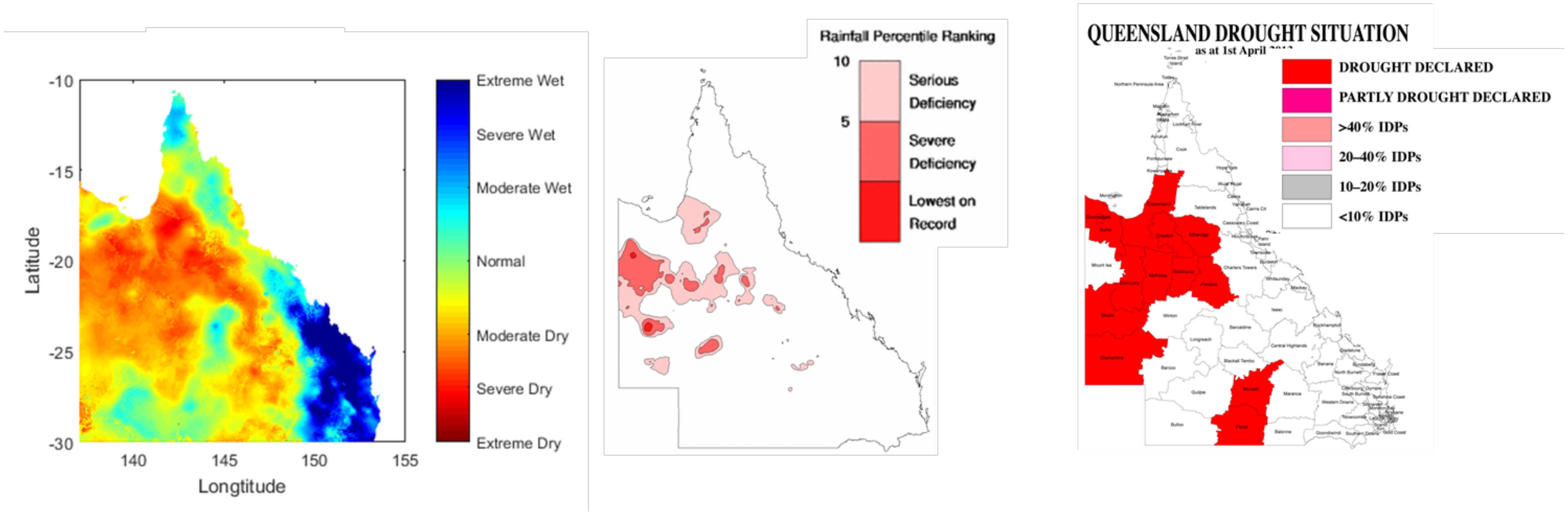


Figure 5.7. Comparison between DFDI (left panel), Rainfall Percentile Ranking (middle panel) taken from http://www.bom.gov.au/, and Queensland drought situation map (right panel) taken from https://www.longpaddock.qld.gov.au, for March 2013.

vegetation. They showed that at a point scale (OzFlux Tower sites) DFDI can be more accurate and reliable than commonly used drought indices such as SPI. This research investigated the suitability of the DFDI across larger spatial scales, such as the state of Victoria in Australia, using spatio-temporally sparse datasets.

In order to evaluate this ability, modelled data such as AWAP, and satellite data such as SMOS, and MODIS-Terra were used to provide daily data for evaluating weekly water stress conditions for eight different weeks during 2011-12, and compared against *in situ* benchmark values. A set of mathematical equations were also derived to facilitate the process of determining DFDI values for future time steps (water stress forecasting).

Throughout the last three chapters, limited observations (less than four years of weekly data) were successfully employed to evaluate water stress conditions. However, the question is whether this set of data is sufficient for this type of applications or whether longer datasets are needed to be applied. The next chapter provides an analysis of long- and short-term data sets available in Australia (and in part used in this study) to address this issue. The statistical characteristics of those data sets will be compared to answer the above question.

Chapter 6

Comparative evaluations of short-term OzFlux tower observations and long-term AWAP model data

6.1 Overview

Studies in a variety of fields have dealt with determining sufficient sample size (or the length of time series) in order to accurately describe the statistical properties of a dataset population (i.e. McDonald and Green 1960, Farajzadeh et al. 2010). For hydro-meteorological applications, a 30-year period is generally considered sufficient to detect all extreme and normal conditions across a region (McDonald and Green 1960, Shahabfar and Eitzinger 2013, Jain et al. 2014, Bayissa et al. 2015). However, for observed variables such as soil moisture and evapotranspiration such long-term data are not available, making it necessary to use short-term observations, which is possible if the temporal stationarity of the statistical characteristics is ensured (Seneviratne et al. 2010, 2012). In contrast to the commonly stipulated data length requirement of 30 years for water stress (drought) monitoring, a number of studies have suggested that the sufficient length of a variable depends on its nature, temporal dynamics and stationarity of their statistical characteristics. For instance, Porth et al. (2001) indicated that under certain conditions, considering five years of streamflow is sufficient to estimate 50% of the return intervals correctly. Similarly, Wu et al. (2005) discussed that if the short-term data sets (in their case, precipitation) can convey the statistical probabilistic characteristics (i.e. probabilistic distribution functions and quartiles) of long-term climatic conditions of a region, the short-term data may be deemed acceptable for hydro-meteorological evaluation and validation scenarios. This is supported by a recent study developing a coastal drought index which also examined the effect of different time series lengths (Conrads et al. 2016). There, it was shown that only 5 to 10 years of data to derive the index can be sufficient to provide useful information on coastal drought and freshwater conditions.

Here, 3-year datasets from three distinct OzFlux towers are compared against a long-term dataset (1911-2016) from the Australian Water Availability Project in terms of stationary and dynamic statistical characteristics. To compare the data sources, box plots and empirical cumulative distributions functions are applied to analyze the stationary statistical characteristics, and also four further criteria (Pearson coefficient, Spearman coefficient, Nash-Sutcliffe coefficient, and Index of agreement) are employed to assess the dynamic similarities between two different datasets. The results of this chapter have been submitted as a journal paper as follows:

Azmi, M., Rüdiger, C., Smith, A.B., Walker, J. Comparative evaluations of short-term OzFlux tower observations and long-term AWAP data in weekly water stress monitoring. *Geophysical Research Letters.* Under review.

6.2 Data and methodology

6.2.1 Data

The *in situ* observations were prepared by aggregating high temporal resolution (0.5h) data collected at the three OzFlux towers of Riggs Creek (Beringer 2014; data coverage Dec-2010 to May-2012), Alice Springs (Cleverly 2011; Sep-2010 to June-2013), and Howard Springs (Beringer 2013; Jan-2011 to Dec-2013). These three study sites were chosen because of their different climate and landuse conditions, representing temperate to arid climate regimes. Overall, five variables were considered here: soil moisture at 10 cm, accumulated precipitation, moisture flux (latent heat), evaporative fraction, and runoff, which would be aggregated to derive the hydrological cluster's representative of the DFDI (Standardized Aggregated Water Availability Index – SAWAI; Azmi et al. 2016).

The long-term (1911-2016) daily modelled data were provided by the Australian Water Availability Project (AWAP; Raupach et al. 2007). The data are continuous and gridded onto a regular 0.05^{o} grid. Here, considering the available data and information, five hydro-meteorological variables of accumulated precipitation, real evapotranspiration, runoff, and soil moisture from depth zero to 10 cm and from 10cm to 100cm were aggregated to derive SAWAI based on the AWAP dataset. To collocate the data, the values of the closest grid points with respect to the locations of the OzFlux tower sites were chosen.

It is worth noting that according to the Azmi et al. (2016), the hydrological variables which are employed for aggregation are not fixed and only need to cover the main hydrological variables of a terrestrial ecosystem.

6.2.2 Methodology

For each hydro-meteorological variable, three different populations were obtained, consisting of 1) short-term, standardized observations from the OzFlux towers, 2) long-term, standardized model data, and 3) short-term, standardized model data with the same length of data as observations. The mean and standard deviation of each variable's population is standardized as follows:

$$Z_i = \frac{x_i - \mu}{\sigma}, \tag{6.1}$$

where x_i is the value of variable x at time step i; and μ and σ are the respective mean and standard deviation of the associated population.

Using the long-term standardized data of the AWAP dataset, SAWAI time series were derived and compared with time series which were already available from the tower observations presented in Azmi et al. (2016). To calculate the SAWAI, three multivariate methods of the Principal Component Analysis (PCA) (Hotelling 1933), Factor Analysis (FA) (Kim and Mueller 1978) and Independent Component Analysis (ICA) (Hyvarinen and Oja 2000) were utilized for aggregating the considered standardized hydro-meteorological variables. An arithmetic average of the first Principal Components derived from the above multivariate methods was then calculated to determine the SAWAI time series (for more detailed information about the process of deriving the SAWAI, refer to Azmi et al. (2016)).

To derive the SAWAI time series, all abovementioned hydro-meteorological variables were utilized. However, for the assessment of the statistical characteristics between variables with different populations, only precipitation and soil moisture were considered as representatives of whole system.

In order to evaluate the probabilistic behaviour of standardized precipitation and soil moisture calculated from different populations, box plots (quartiles, mean, median and ranges of data) and empirical cumulative distributions functions (CDFs) were used. Moreover, four criteria were applied to analyze the consistency, similarities and dynamic relationships between the tower observations and long-term modelled time series with the same time windows as observations. The abovementioned criteria are the Pearson correlation (*Pe*), the Spearman rank correlation (*Ra*), the Nash-Sutcliffe efficiency (*NSE*), and the Index of Agreement (*d*).

The Pearson correlation coefficient (Pearson 1895) is defined as

$$Pe = \frac{Cov(Obs, Mod)}{\sigma_{Obs}\sigma_{Mod}}, \tag{6.2}$$

where $Cov(Obs,Mod)$ is the covariance between observations and modelled data; σ_{Obs} and σ_{Mod} are the standard deviations of observations and modelled data, respectively.

As the Pearson correlation only considers linear relationships between two variables, the Spearman rank correlation coefficient (Spearman, 1904) was introduced to consider non-linear relationships between variables. It is expressed as

$$Ra = \frac{Cov(rg_{Obs}, rg_{Mod})}{\sigma_{rg_{Obs}}\sigma_{rg_{Mod}}}, \tag{6.3}$$

where $Cov(rg_{Obs},rg_{Mod})$ is the covariance between the rank observations and rank modelled data; $\sigma_{rg_{Obs}}$ and $\sigma_{rg_{Mod}}$ are the standard deviations of the ranked observations and ranked modelled data, respectively.

In order to include the impact of bias between two datasets, the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) was introduced to evaluate the ability of models to represent the dynamics of a system with the correct magnitude. The *NSE* is defined as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{N}(Obs_i - Mod_i)^2}{\sum_{i=1}^{N}(Obs_i - Obs_{mean})^2}, \tag{6.4}$$

where Obs_i and Mod_i are the i^{th} observation and modelled data of the considered variable; N is the total number of samples of the considered variable; and Obs_{mean} is the mean of the observed variable over the entire time series.

Finally, Willmott (1984) proposed the Index of Agreement (*d*) to overcome the insensitivity of the Nash-Sutcliffe efficiency and Pearson/Ranking correlations to differences in means and variances of observed and model data as follows:

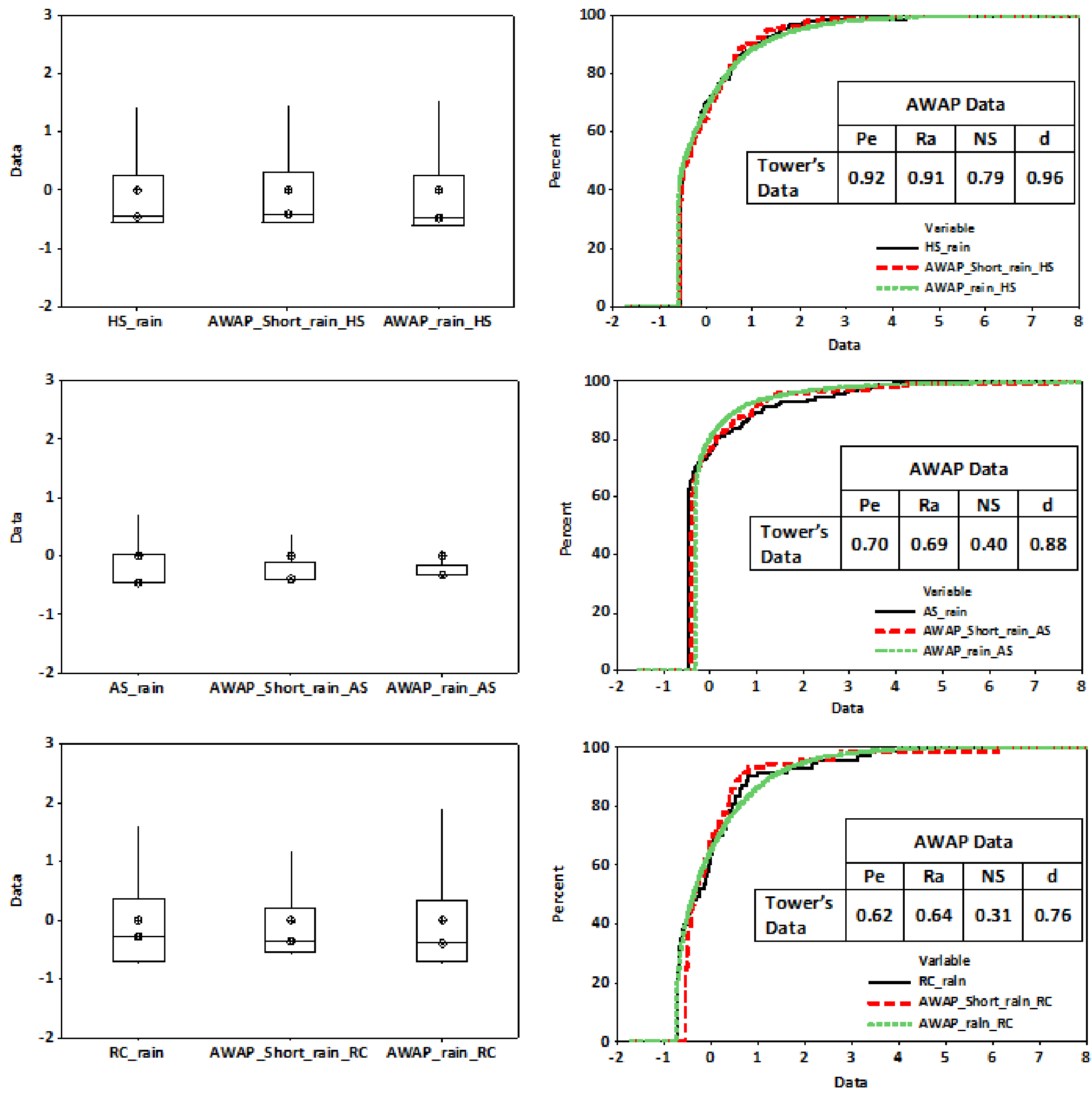
$$d = 1 - \frac{\sum_{i=1}^{N} (Obs_i - Mod_i)^2}{\sum_{i=1}^{N} (|Mod_i - Obs_{mean}| + |Obs_i - Obs_{mean}|)^2} , \quad (6.5)$$

where Obs_i and Mod_i are the i^{th} observation and modelled data of the considered variable; and *N* is the total number of samples of the considered variable. The Index of Agreement (*d*) varies between zero and one; in which d=1 indicates a perfect agreement between reference and estimated values, as it is the case for the other coefficients.

6.3. Results and discussion

In the current study, two hydro-meteorological variables, namely precipitation and soil moisture (top 10 cm), are evaluated as representatives of all available hydro-meteorological variables. Figure 6.1 and 6.2 show the box and CDF plots of the two variables (observed and modelled). Moreover, the statistical charactristics of different datasets including the results of K-S test over two datasets are presented in Table 6.1. The statistics of the observed data include all available observations, whereas those of the modelled data are either based on the same length as the observations (AWAP_short) or the entire modelled data range. As discussed above, the standardized long-term AWAP data were calculated using the entire modelled data set, and the dynamic statistics (four criteria) presented here are calculated between observations and long-term data with a time window length identical to the one of the observations.

In both figures and also Table 6.1, the difference between the mean, median and also skewness values across all sites implies a biasness in the distribution of both precipitation and soil moisture, highlighting that the data do not follow a Gaussian distribution. As the mode of all data is less than zero, there is a persistency and domination of dry events across all sites. However, the distribution tails have very high positive values which shows that wet conditions are more likely to be in the form of extreme conditions. Most important is the fact that, according to the results of K-S test (Table 6.1), the CDFs graphs for precipitation and soil moisture across all sites appear to be relatively consistent, irrespective of the time length, suggesting that both the long-term modelled data and short-term observations similarly represented the probabilistic characteristics of a location.



* RC: Riggs Creek, AS: Alice Springs, HS: Howards Springs, AWAP: Australian Water Availability Project, Pe: Pearson Correlation, Ra: Ranking Correlation, NS: Nash–Sutcliffe Correlation, d: Index of agreement.

Figure 6.1. Left panels: Box plots (Outliers are not displayed) of standardized weekly precipitation based on the observations of OzFlux towers (left box), short-term data from AWAP for the same time period as towers (middle box), and long-term data from AWAP (right box); Right panels: CDFs of standardized weekly precipitation based on the observations of OzFlux towers (black line), short-term data from AWAP for the same time period as towers (red dash line), and long-term data from AWAP

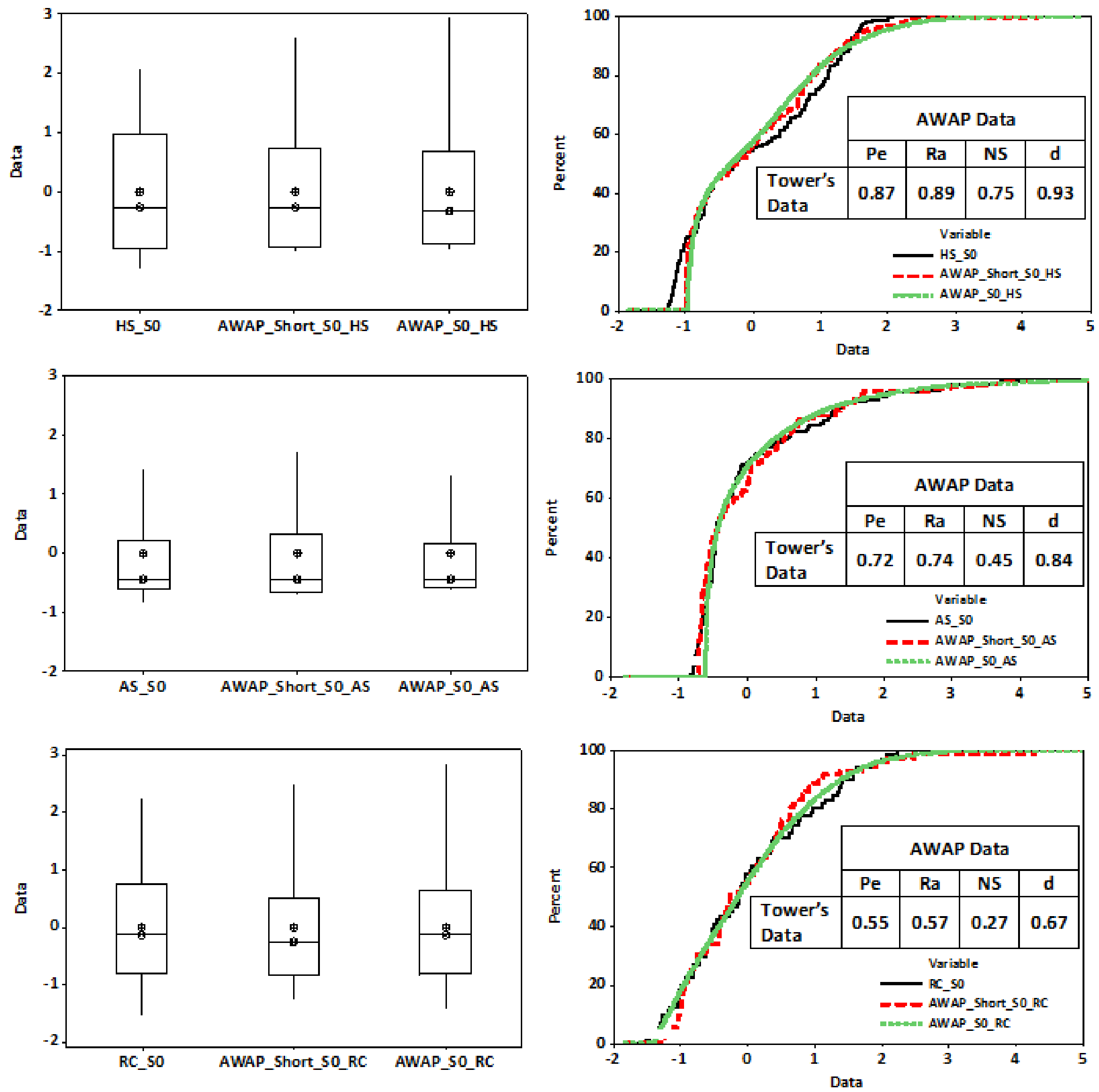


Figure 6.2. As for Fig. 1, but for Soil Moisture.

(green dash line). Tables inside the right panels represent the goodness-of-fit criteria between the observations of OzFlux towers and long-term data from AWAP with the same time window as towers.

According to Figure 6.1, the precipitation data for Howard Springs are very close, both in terms of dynamic and stationary characteristics for any of the three datasets, with the box plots and CDFs of this location being almost identical. The goodness-of-fit criteria of the Pearson and Rank correlations, as well as the d index are greater than 0.9, while the Nash-Sutcliffe is just under 0.8, underlining the strong dynamic similarities between the different datasets in this region.

Table 6.1. Statistical charactristics between different datasets at three OzFlux tower sites.

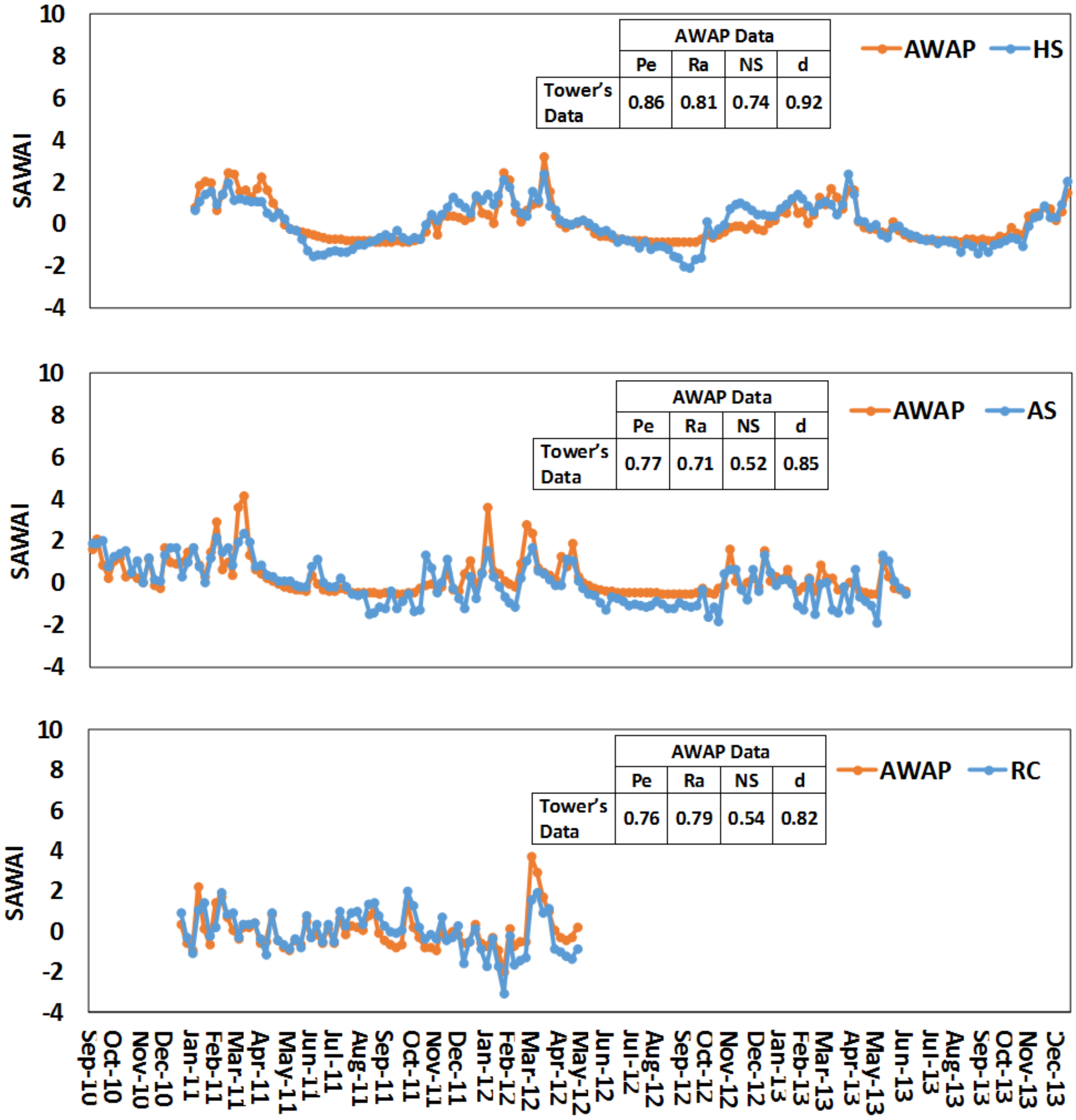
Tower site	Climate regime	Landuse	Observations (dataset 1) vs short term modelled data (dataset 2)			$\begin{bmatrix}P1 & S1\\ P2 & S2\\ P3 & S3\end{bmatrix}$ where P_i and S_i are the statistical charactristics of precipitation and soil moisture for different datasets					
			Goodness-of-fit	Rainfall	Soil moisture	1st quartile	Median	3rd quartile	IQ Range	Skewness	Two-sample K-S test between datasets 1 and 3
HS	Tropical	Woodland savanna	Pe	0.92	0.87	-0.96,-0.56 -0.93,-0.56 -0.89,-0.60	-0.45,-0.26 -0.43,-0.26 -0.47,-0.32	0.86,0.25 0.74,0.30 0.69,0.25	1.92,0.81 1.67,0.86 1.55,0.85	4.10,2.33 4.80,0.89 3.73,1.02	Similar
			Ra	0.91	0.89						
			NS	0.79	0.75						
			d	0.96	0.93						
AS	Arid to Semiarid	Mulga canopies	Pe	0.70	0.72	-0.47,-0.60 -0.40,-0.66 -0.36,0.58	-0.41,-0.45 -0.39,-0.43 -0.37,-0.44	-0.09,0.21 -0.10,0.25 -0.16,0.19	0.48,0.82 0.31,0.96 0.27,0.76	3.14,2.05 4.43,2.31 4.91,2.84	Similar
			Ra	0.69	0.74						
			NS	0.40	0.45						
			d	0.88	0.84						
RC	Temperate	Dryland agriculture and pasture	Pe	0.62	0.55	-0.71,-0.81 -0.61,-0.83 -0.71,-0.82	-0.28,-0.15 -0.36,-0.21 -0.38,-0.17	0.33,0.75 0.28,0.59 0.33,0.64	1.07,1.56 0.85,1.34 1.04,1.46	2.23,0.59 3.99,1.10 2.45,0.77	Similar
			Ra	0.64	0.57						
			NS	0.31	0.27						
			d	0.76	0.67						

HS: Howards Springs, AS: Alice Springs, RC: Riggs Creek, Pe: Pearson Correlation, Ra: Ranking Correlation, NS: Nash–Sutcliffe Correlation, d: Index of agreement.

The Riggs Creek tower site observations appear to have similar stationary characteristics when compared to the long-term AWAP data, rather than over the short-term AWAP data. However, this is only the case for the extreme conditions, and the values of the median for all Box plots are close. Despite this difference, the CDFs of the observations and the short-term AWAP data follow quite similar patterns, with a distinct deviation from the long-term curve around the 80th percentile data. While the statistical coefficients show lower values for the Riggs Creek site compared to Howard Springs, and therefore indicating a lesser dynamic relationship between the datasets, the correlations are still acceptable.

In contrast to the similarities at the other sites, the Alice Springs tower site indicates significant differences between precipitation observations and modelled data. However, the soil moisture patterns on Figure 6.2 show strong similarities. As there has been a significant level of variation in the locations of the rain gauges in this region this is assumed to be the main factor for this difference. However, while the data range indicates a narrower distribution of the precipitation events, the overall dynamics appear to be well represented (i.e. Pe=0.72, Ra=0.74, d=0.88).

Figure 6.2 presents the same approach used for precipitation applied to the soil moisture data (top 10 cm). The box plots show a high consistency and similarity between the different datasets for all sites, in particular for the first and third quartiles, and the median. The highest level of discrepancies between two CDFs are found at Howard Springs, where observations indicate drier conditions than AWAP, while in the wet conditions the modelled data has specified more extreme events, pointing towards a wet bias in the model. This may be a result of surface soil moisture decoupling in the observed data, which may not be reflected as such in the modelled data. The CDFs for the observations and short-term modelled data of the Howard Springs site reveal that during 2011-13 the area experienced several extreme conditions resulting in an apparent bi-modal distribution, which is shown in the curvatures in the CDFs around its median. The analysis of the Riggs Creek site shows a considerable difference between the CDFs derived from observations and short-term modelled data, in particular for data between the 70th and 90th percentiles, where the short-term modelled data shows water conditions drier than the observations, which are even indicating generally wetter conditions. In contrast, for data below the 20th percentile, short-term modelled data display wetter conditions than their long-term counterpart and observations indicate, which may be due to soil parameterizations and an artificially high wilting point.



* SAWAI: Standardized Aggregated Water Availability Index, RC: Riggs Creek, AS: Alice Springs, HS: Howards Springs, AWAP: Australian Water Availability Project, Pe: Pearson Correlation, Ra: Ranking Correlation, NS: Nash–Sutcliffe Correlation, d: Index of agreement.

Figure 6.3. Comparison between SAWAI time series calculated from OzFlux towers observations and long-term AWAP dataset. Tables inside the panels represent the goodness-of-fit criteria calculated between two time series with the same time windows.

In general, the results shown in Figures 6.1 and 6.2 indicate that for the considered time period, the aggregated weekly observations of precipitation and soil moisture at the towers have appropriate consistencies and similarities with the long-term AWAP modelled data. Since these two hydro-meteorological variables are the two key variables in the hydrological cycle of the terrestrial eco-systems, it can be argued that the periods contain sufficiently similar statistical characteristics of the overall long-term conditions to reliably inform the DFDI. The available differences in the results are likely due to uncertainties in modelling gridded datasets, measurements and observation errors propagated through atmospheric model forcing and *in situ* observation equipment, but most importantly due to the discrepancies of data representation at different scales (point observations against 5km model resolution).

To confirm that the differences in the time series is not significant, SAWAI values were calculated from long-term AWAP modelled data and compared with the SAWAI time series already derived from tower observations by Azmi et al. (2016). Figure 6.3 shows the SAWAI time series for all three OzFlux tower sites based on observations as well as long-term modelled data. As it could be expected from the previous analysis, Howard Springs achieved the highest values across the four statistical criteria. Despite lower scores, Alice Spring and Riggs Creek also have close time series similarities, with most extreme events being identified. One caveat is that the CDFs of Figures 6.1 and 6.2, as well as the time series of Figure 6.3, identify a slight positive bias of the modelled data against the observations, which is akin to a wetter overall state of the system when comparing the period 2010-2013.

6.4. Chapter Summary

This study evaluated the difference in the statistical properties of weekly hydrological variables obtained by both *in situ* monitoring sites and model predictions, across three different climate conditions and landuse situations. While the OzFlux towers provided the short-term observational dataset (less than 4 years between 2010 and 2013), the AWAP modelled data were used to derive a long-term dataset (106-year data). Results showed that observations and model displayed similar statistical features of long-term water conditions despite their different lengths, suggesting that both may be used for validation purposes. Furthermore, the evaluation of a water stress index using the two datasets indicates high consistencies and similarities between the indices across all towers. Consequently, it can be assumed that the data used in the previous chapters to develop the DFDI are providing the necessary statistical properties to derive indices that resemble those developed using long-term data.

Chapter 7

Conclusions and future directions

7.1 Overview

Throughout this dissertation, commonly used drought indices and statistical methods to assess water stress and drought conditions were explained and applied. Moreover, this research introduced a new methodology which is able to evaluate different aspects of water stress conditions for different months throughout the year, which are subject to a variety climate and landuse/landcover situations. The contributions of the presented research to the general knowledge have been presented across two main categories: i) water stress evaluation at the point-scale, and ii) spatial analysis of water stress conditions, with a focus on the state of Victoria, Australia. In addition to those points, future directions of this work have been suggested to continue the current study in the previous chapters, and are being discussed at the end of this chapter, as well.

7.2 Conclusions

7.2.1 Water stress evaluation at the point scale

The proof-of-concept of a more comprehensive water stress study using advanced statistical methods and sources was presented in Chapter 3. There, limited point data were obtained from three flux stations across different climate regimes in Australia, to compare, develop, and test existing and new ways to classify droughts or water stress. Despite the limited length of data, the results were promising and showed the value of the applied methods in the case of data shortages. As any approach based on statistics, larger data sets would make this type of studies more robust in its predictive skills. In fact, despite the limited amount of available data, 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 the Standardized Precipitation Index, cannot be sufficient to evaluate all different aspects of water stress issues. It was therefore argued that there is a need to consider an appropriate set of indices and proxies from different drought types, rather than single indices alone. Furthermore, the information obtained from such an approach on extreme events by deploying descriptive statistical methods will provide clear, quick, and uncomplicated predictions of extreme water stress situations over an area to the responsible authorities, allowing them to develop a better understanding of the current and future conditions.

Chapter 4 confirmed the hypothesis developed from the previous studies that individual SDIs are not sufficient to be applied in all circumstances, particularly in the presence of diverse land-use and climate conditions. Moreover, some DFDIs, such as WADI and AADI, which are formed on the basis of empirical equations along with subjective assumptions, cannot be generalized as a comprehensive index. Furthermore, two DFDIs (LADI and NLADI) always require fixed hydro-meteorological variables/proxies during their aggregating process. However, human or measurement errors may render some of those variables inconsistent with others, due to their fixed parameters, leading to a degradation of the prediction and monitoring capabilities of those indices. The approach taken in this chapter was then to filtering the SDIs using a Probabilistic Similarity Method to allow for the removal of inaccurate or unreliable data.

The presence of more SDIs during the combination process has been shown to lead to a more robust DFDI, and automated selective ignorance of some SDIs does not appear to affect the results significantly, as long as their characteristics are implicitly mapped through other indices. For instance, the inclusion of the standardized runoff and surface soil moisture index (SROI) at Riggs Creek showed that the proposed DFDI does not have a high sensitivity to this indicator due to the presence of individual precipitation and soil moisture indices. Therefore, the unavailability of the SROI at Alice Springs and Howards Springs was assumed not to be an issue for the results of the proposed DFDI, as long as the other indices are available.

One of the main advantages of the proposed methodology is the explicit inclusion of vegetation condition-based indices alongside water content-based indices to derive the developed DFDI. Regarding the physical concepts of terrestrial ecosystems during dry and active months, employing only water content based SDIs during the combination process is quite insufficient. Therefore, to have a comprehensive DFDI, it is necessary to consider both aggregated indices, that were presented as the Standardized Aggregated Water Availability Index (SAWAI) and Standardized Aggregated Vegetation Index (SAVI), which are subject to different wet/dry and active/non-active conditions throughout a year.

It is worth noting that determining wet/dry and active/non-active months is based on the historical climatologies which, technically, could mean that a shift in the current climate would not be reflected properly. However, this type of process is currently implicitly taken into consideration. If the SAVI identifies a low level of vegetation activity during a climatologically active month, the first step must be to check the values of the SAWAI to investigate whether the low rate of activity is subject to water stress in the system or not. In case that the SAWAI does not show water stress conditions, air and even soil temperature would be the next alternative to identify the reason. Therefore, the DFDI can also handle the mentioned situation simply by checking the air temperature as an auxiliary variable.

In order to make the proposed DFDI as user-friendly and applicable as possible, a symbolic regression approach was used to derive explicit mathematical equations for SAVI, SAWAI and consequently the proposed DFDI. Here, the three considered case studies resulted in average values of the goodness-of-fit criteria for SAVI and SAWAI as follows: r^2=95%, Maximum Error=0.45, and Mean Absolute Error=0.13. In this case, users such as water managers and other decision makers can rely on the equations to derive values of the proposed DFDI at any specific time for the considered areas.

In the past, new proposed indices have been validated based on their degrees of similarities (correlation) when compared to SPI or PDSI (Keyantash and Dracup 2004, Balint and Mutua 2011, Barua et al. 2012, Zhang and Jia 2013, Li et al. 2014). I.e., in case the new indices followed the behaviour of SPI and or PDSI, they were considered as validated. However, this negates the development of new indices, as they then simply mimic the behaviour of the existing ones. According to a recent review (Hao and Singh 2015), no reliable “ground truth” exists (for a drought index) that may be used as an ultimate reference for the exact validation of a new index. In particular, the traditional indices (eg. SPI and PDSI) have limitations and follow assumptions which have led to the development of new indices (Keyantash and Dracup 2002, Van Loon and Van Lanen 2012). The results of Chapter 4 validated the behaviour of the proposed DFDI by comparing the physics and nature of the considered terrestrial ecosystems in addition to the climate conditions (classifying each month to wet and dry) and landuse of the area (clustering each month into active and non-active, subject to the type of dominant plants of the area). For this purpose, the two main elements of the terrestrial ecosystem, consisting of the water balance (input water, storage water, output water) and the vegetation growth, were considered for each month to identify the most appropriate driver responsible for the actual system's water stress. According to the proposed methodology, SAWAI can specify the status of the water balance in terms of input (i.e. SPI), the storage (i.e. SSMI) and the output (i.e. SEFI and SROI) elements together, and SAVI can reflect the period of active vegetation (i.e. SNDVI), which together may indicate the water stress of the ecosystem.

The methodology section of Chapter 4 has presented as the first step the acquisition of an appropriate set of hydro-climatological variables to cover different aspects of water contents, plant water consumption, and vegetation conditions. It does not mean that the current set of the considered individual DIs is a fixed and strict selection. The selected indices will ultimately be driven by data availability, but will need to represent those three main domains. Moreover, regional differences may be found, as the underlying hydrological and plant-physiological drivers are unlikely to be the same between water- and energy-limited environments.

Overall, Chapter 4 is to be considered as a proof-of-concept study, developing a robust index for a comprehensive water stress monitoring system by using advanced statistical methods and sources. The temporal coverage of the full range of variables is very limited throughout the world, making the development and validation of such an approach challenging. This study focused on data collected from flux stations throughout Australia, providing a broad range of climatological and hydrological conditions across their short observation period. Despite the limited length of data, the results of the validation process are nonetheless promising and show the value of the methodology when confronted with significant data deficits. As with any statistics-based approach, longer data sets will make this type of index more robust in its predictive skills to estimate water stress conditions at any one time. However, no *in situ* monitoring site provides observations across the full spectrum of variables for a sufficiently long period to allow the development of index parameters for all potential conditions. Consequently, future studies should assess the use of modelled data from climatological reanalyses to provide a consistent, long-term data set. This approach would then allow quantifying the uncertainty introduced by the initial assumption that the long-term statistics of the hydrological dynamics throughout the monitoring stations were indeed covered by the short period of available data. Despite the limitations of the data set used in Chapter 4, it was shown that it is possible to derive useful information and create meaningful variable clusters by using an appropriate drought index which can be further developed for evaluating water stress based issues, provided such long-term data sets are accessible.

To address the data limitations faced in Chapters 3 and 4, the AWAP modelled data were used to obtain a long-term dataset (106-year data), as discussed above (It is worth noting that Chapter 5 deals with spatial monitoring of water stress conditions which its highlights will be elaborated in section 7.2.2). Hence, Chapter 6 evaluated the difference in the statistical properties of weekly hydrological variables obtained by both *in situ* monitoring sites and model predictions, across the three different climate conditions and landuse situations of the chosen OzFlux tower sites.

In spite of available discrepancies between modelled data and observations due to errors caused by models uncertainties, equipment errors, and different spatial resolutions of AWAP data and towers (Tozer et al. 2012, Parana Manage et al. 2016, King et al. 2016); the results showed that during the considered time period, data of the hydro-meteorological variables based on the short-term observations could appropriately convey the long-term statistical stationary and dynamic features of the selected regions. Consequently, there were strong similarities and consistencies between the weekly time series of SAWAI calculated from short-term observations and long-term modelled data. Therefore, it can be concluded that given any length of time series may be used for deriving water stress or drought indices, if a dataset appropriately reflects the dynamic and stationary statistical specifications of the long-term conditions. In the present case, this was applied to the SAWAI of the DFDI, which is indicative of the water stress conditions of the regions. It was shown that the differences of this index derived from observations and long-term datasets, respectively, are negligible.

The results of this assessments can have important implications for the validation procedures of future indices. For instance, the statistical properties of short-term datasets can either be validated by using high-quality, long-term simulations, or those simulations may be used to derive robust indices, after validating their performance against short-term point observations. Further, such research can be implemented to determine and then adjust the bias of modelled data to prepare more reliable and accurate long-term datasets to be employed for analyzing historical water stress conditions.

7.2.2 Spatial analysis of water stress conditions across Victoria, Australia

After successfully applying the DFDI to evaluate water stress conditions at the three OzFlux tower sites, the same methodology was employed across the Victoria, Australia, for eight different events across the period 2010-12. Geographically, Victoria consists of a range of diverse climate conditions and landuse/landcover situations. Nonetheless, the DFDI maps have shown very consistent and realistic results in comparison to individual drought indices, natural geography, and landuse/landcover situations. The final derived mathematical equations have generally shown promising functionalities and accuracies, especially in appropriately detecting the spatial trend of the extreme events. Nevertheless, some areas showed less accurate results than others. It was found that these differences are mostly due to highly diverse and inconsistent spatial patterns within the landuse/landcover and climate regimes in these parts of Victoria. This can be addressed by either splitting the central northern and south-western parts of Victoria into smaller more homogeneous sub-regions by considering other regionalizing techniques, or using other methods to derive the mathematical equations. However, given the accurate representation of the spatio-temporal patterns in most regions, the latter may not be an effective solution. Overall, Chapter 5 successfully applied the original approach developed for single points across larger spatial scales. While some additional work remains to be addressed (i.e. spatial heterogeneity and influence of time lags), the results highlighted the value of DFDI to provide a spatio-temporally consistent product to assess water stress across large scales.

7.3 Future directions

The future directions of this research can be categorized as:

1- In addition to updating the DFDI time series of the Riggs Creek, Howard Spring and Alice Springs towers observations, attempts can be made to derive DFDI time series for other available OzFlux towers and also equipped synoptic stations in Australia which are located in other types of climate and landuse/landcover situations. This can help the community to accept the reliability of this new index across the continent to gradually replace the commonly used DIs such as SPI, by comparing their respective merits and performance against locally experienced drought and water stress conditions.

2- As mentioned in the conclusion section, one of the most important issues in developing and introducing new drought indices is the process of validation. By utilizing long-term data and information provided by local authorities, it would be likely feasible to compare the results of DFDI against “ground truths” such as drought relief payments and agricultural productivity yield maps to validate its performance.

3- Determining the sources and magnitudes of errors in OzFlux towers and AWAP data in different parts of Australia can be invaluable for other researchers to address the degree of uncertainties in water stress evaluations in spatial scales. Using validated satellite data coupled with successful downscaling methods besides *in situ* observations can be one of the solutions to encounter this issue.

4- Output of land surface water-energy flux models may be a source for validating drought indices such as DFDI. However, these kinds of models usually have considerable uncertainties (inevitable input errors, poorly defined parameters, inadequate model structures) and need a variety of parameters and information (i.e. moisture flux, heat flux) which are not commonly available. Nonetheless, at least for areas including flux tower sites they can be implemented and results be utilized for examining the performance of DFDI.

5- In case of deriving monthly time series of DFDI in different parts of Australia, deriving joint probabilities between DFDI time series and global climate indices such as SOI can provide worthwhile information about the interactions between local water stress conditions and global climate cycles. To derive the joint probabilities, Copula functions with different degrees of complexities can be one of the best choices.

6- In case of a forecast, dealing with risk levels would be inevitable because of available uncertainties. So our forthcoming aim is to minimize the negative impacts of such errors sources by i) properly communicating with other researchers and also ii) using the best data sets available and maybe work with ensemble forecasts, to look into the variation of this uncertainty.

7- Available uncertainties due to inconsistencies between datasets in terms of spatio-temporal resolutions and sources (satellite, in-situ, and modelled data) should comprehensively be considered. Further, performance metrics (Koch et al. 2016) of spatial indices, in terms of accuracy and spatial patterns, along with comparing results with ground truths (such as drought assessment reports and information from end users (i.e. farmers)) would more directly help validating the efficiency and precision of this new drought index. Furthermore, there would be a need of a synthetic study aiming at quantifying the effects of the error assumptions, potentially assessing the implicit uncertainties in the DFDI over time, based on a dynamic progression through time.

References

Abdul Rauf, U.F., Zeephongsekul, P. (2014) Analysis of rainfall severity and duration in Victoria, Australia using non-parametric copulas and marginal distributions. *Water Resources Management*, 28:4835-4856, DOI:10.1007/s11269-014-0779-8.

Abrahart, R.J., Beriro, D.J. (2012) How much complexity is warranted in a rainfall-runoff model? Findings obtained from symbolic regression, using Eureqa. *Geophysical Research Abstracts*, 14:8012.

AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow, B.D., Hain, C.R. (2015) Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53:452-480, DOI:10.1002/2014RG000456.

Anderson, D. (2014) Endurance: Australian stories of drought. CSIRO publication. ISBN: 9781486301201.

Anderson, M.C., Cammalleri, C., Hain, C.R., Otkin, J., Zhan, X., Kustas, W. (2013) Using a diagnostic soil-plant-atmosphere model for monitoring drought at field to continental scales, *Procedia Environmental Sciences*, 19:47–56, DOI:10.1016/j.proenv.2013.06.006.

Andrykanus, R. (2011) Riggs Creek OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring. hdl:102.100.100/6970.

Araghinejad, S., Azmi, M., Kholghi M. (2011) Application of artificial neural network ensembles in probabilistic hydrological forecasting. *Journal of Hydrol*ogy, 407:94-104, DOI:10.1016/j.jhydrol.2011.07.011.

Azmi, M., Araghinejad, S., Kholghi, M. (2010) Multi model data fusion for hydrological forecasting using K-nearest neighbour method. *Iranian Journal of Science & Technology, Transaction B, Engineering*, 34:81-92.

Azmi, M., Rüdiger, C., Walker, J. (2016) A data fusion-based drought index. *Water Resources Research*, 52:2222-2239, DOI:10.1002/2015WR017834.

Australian Bureau of Agricultural and Resource Economics (ABARE) (2006) Drought update: Australian crop and livestock report. Technical report, ABARE, Canberra, Australia.

Australian Bureau of Statistics (ABS) (2004) Australian economic indication: impact of 2003-4 farm season on Australia production. Technical report, ABS, Canberra, Australia.

Balint, Z., Mutua, F.M. (2011) Drought monitoring with the combined drought index, FAO-SWALIM, Nairobi, Kenya. http://www.faoswalim.org/resources/site_files/Combined_ drought_ index.pdf.

Barua, S. (2010) Drought assessment and forecasting using a nonlinear aggregated drought index. PhD thesis, School of Engineering and Science Faculty of Health, Engineering and Science, Victoria University, Australia.

Barua, S., Ng, A.W.M., Perera, B.J.C. (2011) Comparative Evaluation of Drought Indexes: Case Study on the Yarra River Catchment in Australia. *Journal of Water Resources Planning and Management*.137:215-226.

Barua, S., Ng, A.W.M., Perera, B.J.C. (2012) Artificial neural network–based drought forecasting using a nonlinear aggregated drought index. *Journal of Hydrologic Engineering*, 17:1408-1413.

Bayissa, Y.A., Moges, S.A., Xuan, Y., Andel, S.J.V., Maskey, S., Solomatine, D.P., Griensven, A.V., Tadesse, T. (2015) Spatiotemporal assessment of meteorological drought under the influence of varying record length: the case of Upper Blue Nile Basin, Ethiopia, *Hydrological Sciences Journal*, 60:1927-1942, DOI: 10.1080/02626667.2015.1032291.

Beersma, J.J., Buishand, T.A. (2004) Joint probability of precipitation and discharge deficits in the Netherlands. Water Resources Research 40:W12508, DOI:10.1029/2004WR003265.

Beringer, J. (2013) Howard Springs Understory OzFlux tower site OzFlux: Australian and New Zealand Flux research and monitoring hdl:102.100.100/14240.

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

Blender, R., Fraedrich, K. (2006) Long-term memory of the hydrological cycle and river runoffs in China in a high-resolution climate model. *International Journal of Climatology* 26:1547-1565, DOI: 10.1002/joc.1325.

Bond, N.R., Lake, P.S., Arthington, A.H. (2008) The impacts of drought on freshwater ecosystems: an Australian perspective. *Hydrobiologia*, 600:3–16, DOI:10.1007/s10750-008-9326-z.

Brut, A., Rüdiger, C., Lafont, S., Roujean, J-L., Calvet, J.-C., Jarlan, L., Gibelin, A-L., Albergel, C., Le Moigne, P., Soussana, J-F., Klumpp, K. (2009) Modelling LAI at a regional scale with ISBA-A-gs: comparison with satellite-derived LAI over southwestern France. *Biogeosciences*, 6:1389-1404, DOI: 10.5194/bg-6-1389-2009.

Bruce, J.P. (1994) Natural disaster reduction and global change. *Bulletin of the American Meteorological Society*. 75:1831–1835.

Bureau of Meteorology (BoM) (2016) Climate classification map of Australia. Retrieved from www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=kpngrp #info. Last Access 21 January 2016.

Burke, E.J., Brown, S.J., Christidis, N. (2006) Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model. *Journal of Hydrometeorology,* 7:1113–1125.

Ceccato, P., Flasse, S., Tarantola, S., Jacquemond, S., Gregoire, J.M. (2001) Detecting vegetation water content using reflectance in the optical domain. *Remote Sensing of Environment,* 77: 22–33.

Chen, J., Wilson, C., Tapley, B., Yang, Z., Niu, G. (2009) 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *Journal of Geophysical Research*, 114:B05404, DOI:10.1029/2008JB006056.

Cleverly, J. (2011) Alice Springs Mulga OzFlux site OzFlux: Australian and New Zealand Flux research and monitoring hdl:102.100.100/14217.

Conrads, P., Tufford, D.L., Darby, L.S. (2016) Critical Aspects of the Coastal Drought Index: Length of Salinity Data Record and Ecological Response Data. 30th Conference on Hydrology and the 28th Conference on Climate Variability and Change. January 2016. New Orleans, Louisiana, USA.

Cook, E.R., Seager, R., Cane, M.A., Stahle, D.W. (2007) North American drought: reconstructions, causes, and consequences. *Earth-Science Reviews*, 81:93–134.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika,* 16:297–334, DOI:10.1007/bf02310555.

Dalezios, N., Blanta, A., Spyropoulos N. (2012) Assessment of remotely sensed drought features in vulnerable agriculture, *Natural Hazards and Earth System Science,* 12: 3139-3150, DOI:10.5194/nhess-12-3139-2012.

Department of Environment and Primary Industries (DEPI) (2009) Managing water stress in grape vines in greater Victoria. AG1074. ISSN 1329-8062.

Department of Environment and Primary Industries (DEPI), State Government Victoria (2014) depi.vic.gov.au/agriculture-and-food/livestock/beef/breeding/management-beef-breeding-cows. Last access: February, 2016.

DoE&H. (Department of the Environment and Heritage) (2001b) Australia State of the Environment, Independent Report to the Commonwealth Minister for the Environment and Heritage. http:/www.ea.gov.au/soe/2001.

Drobot, R., Dinu, C., Draghia, A., Adler, M.J., Corbus, C., Matreata, M. (2014) Simplified approach for flood estimation and propagation. Automation, Quality and Testing, Robotics, IEEE International Conference; DOI: 10.1109/AQTR.2014.6857921.

Edwards, D.C., McKee, T.B. (1997) Characteristics of 20th century drought in the United States at multiple time scales. Climatology Report No: 97–2, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado.

Environment Canada (2004) Threats to water availability in Canada. National Water Research Institute, Burlington, Ontario. NWRI Scientific Assessment Report Series No. 3 and ACSD Science Assessment Series No. 1:128.

European Commission (2007) Communication from the Commission to the Council and the European Parliament, Addressing the challenge of water scarcity and droughts in the European Union. Brussels, 18.07.07, COM (2007) 414 final. http://eur-lex.europa.eu/ LexUriServ/LexUriServ.do?uri=COM:2012:0673:FIN:EN:PDF.

Everingham, Y.I, Muchow, R.C., Stone, R.C, Inman-Bamber, N.G, Singels, A., Bezuidenhout, C.N. (2002) Enhanced risk management and decision-making capability across the sugarcane industry value chain based on seasonal climate forecasts. *Agricultural Systems*. 74:459-477.

Farajzadeh, M., Rahimi, M., Kamali, G.A. (2010) Modelling apple tree bud burst time and frost risk in Iran. *Meteorological Applications*, 17: 45–52. DOI:10.1002/met.159.

Finnigan, J.J. (2004) A re-evaluation of long-term flux measurement techniques. Part II: Coordinate Systems. *Boundary-Layer Meteorology*, 113:1-41.

Finnigan, J.J., Clements, R., Malhi, Y., Leuning, R., Cleugh, H.A. (2003) A re-evaluation of long-term flux measurement techniques. Part I: Averaging and coordinate rotation. *Boundary-Layer Meteorology*, 107:1-48.

Food and Agriculture Organization (FAO) (2012) Coping with water scarcity; an action framework for agriculture and food security. FAO water report 38. ISBN 978-92-5-107304-9.

Gallant, A.J.E., Reeder, M.J., Risbey, J.S., Hennessy, K.J. (2013) The characteristics of seasonal-scale droughts in Australia, 1911–2009. *International Journal of Climatology,* 33: 1658–1672.

Ganatsas, P., Mantzavelas, A., Tsakaldimi, M. (2011) Development of an adapted empirical drought index to the Mediterranean conditions for use in forestry. *Agricultural and Forest Meteorology*, 151:241-250.

Gao, B.C. (1996) NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58: 257-266.

Gharun, M., Azmi, M., Adams, M.A. (2015) Short-term forecasting of water yield from forested catchments after Bushfire: A case study from southeast Australia. *Journal of Water*, 7: 599-614, DOI:10.3390/w7020599.

Ghulam, A., Qin, Q., Zhan, Z. (2007) Designing of the perpendicular drought index, *Environmental Geology*, 52:1045–1052, DOI:10.1007/s00254-006-0544-2.

Ghulam, A., Qin, Q., Kusky, T., Li, Z-L. (2008) A re-examination of perpendicular drought indices, *International Journal of Remote Sensing*, 29: 6037-6044, DOI: 10.1080/ 01431160802235811.

Gibbs, W.J., Maher, J.V. (1967) Rainfall deciles as drought indicators. *Commonwealth Bureau of Meteorology Bulletin*. 48, Melbourne.

Gorbon, N., Pinty, B., Melin, F., Taberner, M., Verstraete, M.M. (2002) Sea wide field-of-view sensor (SeaWIFS) - an optimized FAPAR algorithm - theoretical basis document. JRC Publication No. EUR 20148 EN, Institute for Environment and Sustainability, Ispra, Italy, 20 pp.

GreenGrower (2016) Drought monitoring in Australia using combined drought index. http://www.graingrowers.com.au/products-services/drought-monitor. Last Access: 19 May 2016.

Hao, Z., AghaKouchak, A. (2013) Multivariate standardized drought index: A parametric multi-index model, *Advances in Water Resources*, 57:12–18, DOI:10.1016/j.advwatres .2013.03.009.

Hao, Z., AghaKouchak, A. (2014) A multivariate multi-index drought monitoring framework. *Journal of Hydrometeorology*, 15:89–101.

Hao, Z., Hao, F., Singh, V.P., Xia, Y., Ouyang, W., Shen, X. (2016) A theoretical drought classification method for the multivariate drought index based on distribution properties of standardized drought indices. *Advances in Water Resources*, 92:240-247, DOI: 10.1016/j.advwatres. 2016.04.010.

Hao, Z., Singh, V.P. (2013) Modeling multi-site streamflow dependence with maximum entropy copula. *Water Resources Research*. 49:7139-7143, DOI: 10.1002/wrcr.20523.

Hao, Z., Singh, V.P. (2015) Drought characterization from a multivariate perspective: A review. *Journal of Hydrology*, 527:668-678, DOI:10.1016/j.jhydrol.2015.05.031.

Hayes, M.J. (2016) Comparison of Major Drought Indices, http://drought.unl.edu/Planning/Monitoring/ComparisonofIndicesIntro/CMI.aspx. Last access: April, 2016.

Hayes, M.J., Svoboda, M.D., Wardlow, B.D., Anderson, M.C., Kogan, F. (2012) Drought monitoring historical and current perspectives. In: Wardlow, B., Anderson, M., Verdin, J. (Eds.), Remote Sensing of Drought: Innovative Monitoring Approaches. CRC Press, Boca Raton, pp. 1–19.

Heim Jr, R.R. (2002) A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83:1149-1165.

Heggen, R.J. (2001) Normalized antecedent precipitation index, *Journal of Hydrological Engineering*, 6:377-381.

Hennessy, K., Fitzharris, B., Bates, B.C., Harvey, N., Howden, S.M., Hughes, L., Salinger, J., Warrick, R. (2007) Australia and New Zealand. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II , in Fourth Assessment Report of the Intergovernmental Panel on Climate Change, eds. M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Cambridge University Press, Cambridge, UK.

Hidalgo, H.G., Piechota, T.C., Dracup J.A. (2000) Alternative principal components regression procedures for dendrohydrologic reconstructions. *Water Resources Research*, 36: 3241-3249.

Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24: 498– 520.

Houborg, R., Rodell, M., Li, B. Reichle, R., Zaitchik, B.F. (2012) Drought indicators based on model-assimilated Gravity Recovery And Climate Experiment (GRACE) terrestrial water storage observations, *Water Resources Research*, 48:W07525, DOI: 10.1029 /2011WR011291.

Hyvarinen, A., Karhunen, J. Oja, E. (2001) Independent Component Analysis. John Wiley & Sons, INC. ISBN 0-471-22131-7.

Hyvarinen, A., Oja, E. (2000) Independent Component Analysis: Algorithms and Applications. *Neural Networks*, 13:411-430.

Jackson, T.J., Chen, D., Cosh, M., Li, F., Anderson, M., Walthall, C., Doriaswamy, P., Hunt, E.R. (2004) Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sensing of Environment*, 92:475–482.

Jackson, R.D., Idso, S., Reginato, R., Pinter Jr, P. (1981) Canopy temperature as a crop water stress indicator, *Water Resources Research*, 17: 1133–1138, DOI:10.1029/WR017 i004p01133.

Jacquette, E., Al Bitar, A., Mialon, A., Kerr, Y., Quesney, A., Cabot, F., Richaume, P. (2010) SMOS CATDS level 3 global products over land. In C.M.U. Neale, & A. Maltese (Eds.), Proceeding of SPIE, Remote Sensing for Agriculture, Ecosystems, and Hydrology XII. DOI: 10.1117/12.865093.

Jain, V.K., Jain, M.K., Pandey, R.P. (2014) Effect of the Length of the Streamflow Record on Truncation Level for Assessment of Streamflow Drought Characteristics. *Journal of Hydrological Engineering*, 19(7): 1361-1373.

Jain, V.K., Pandey, R.P., Jain, M.K., Byun, H-R. (2015) Comparison of drought indices for appraisal of drought charactristics in the Ken River Basin. *Weather and Climate Extremes*, 8, 1-11.

Juston, J., Seibert, J., Johansson, P.O. (2009) Temporal sampling and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrological Processes*, 23:3093-3109, DOI: 10.1002/hyp.7421.

Kaiser, H.F. (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20:141-151.

Kaiser, H.F., Rice J. (1974) Little Jiffy Mark IV, *Educational and Psychological measurement*, 34:111-117.

Kallis, G. (2008) Droughts. Annu. Rev. Environ. Resour. 33:85–118.

Kao, S.C., Govindaraju, R.S. (2010) A copula-based joint deficit index for droughts. *Journal of Hydrology*, 380:121–134.

Karnieli, A., Agam, N., Pinker, R.T., Anderson, M., Imhoff, M.L., Gutman, G.G., Panov, N., Goldberg, A. (2010) Use of NDVI and land surface temperature for drought assessment: Merits and limitations, *Journal of Climatology*, 23: 618–633, DOI:10.1175/2009 JCLI2900.1.

Karnieli, A., Bayasgalan, M., Bayarjargal, Y., Agam, N., Khuldulmur, S., Tucker, C.J. (2006) Comments on the use of the Vegetation Health Index over Mongolia. *International Journal of Remote Sensing*, 27:2017-2024, DOI: 10.1080/01431160500121727.

Keetch, J.J, Byram, G. (1968) A drought index for forest fire control. Res. Paper SE-38. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 32 pp.

Keyantash, J.A., Dracup, J.A. (2002) The quantification of drought: An evaluation of drought indices, *Bulletin of the American Meteorological Society*, 83:1167-1180.

Keyantash, J.A., Dracup, J.A. (2004) An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resources Research*, 40:W09304, DOI:10.1029/2003WR002610.

Kiem, A.S., Johnson, F., Westra, S., van Dijk, A., Evans, J.P., O'Donnell, A., Rouillard, A., Barr, C., Tyler,J., Thyer, M., Jakob, D., Woldemeskel, F., Sivakumar, B. and Mehrotra, R. (2016) Natural hazards in Australia: droughts. Climatic Change, doi:10.1007/s10584-016-1798-7.

Kim, J.O., Mueller, C.W. (1978) Factor analysis: statistical methods and practical issues. Beverly Hills, Calif.: Sage Publications. HA29.Q35/VOL 14.

King, A. D., Alexander, L.V. Donat, M.G. (2013) The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia. *Int. J. Climatol.*, 33, 2376-2387, doi: 10.1002/joc.3588.

Kline, P. (2000) The handbook of psychological testing (2nd ed.). London: Routledge, page 13.

Koch, J., Siemann, A., Stisen, S., Sheffield, J. (2016) Spatial validation of large-scale land surface models against monthly land surface temperature patterns using innovative performance metrics, *J. Geophys. Res. Atmos.*, 121, doi:10.1002/2015JD024482.

Kogan, F. (1995) Application of vegetation index and brightness temperature for drought detection, *Advances in Space Research*, 15: 91–100, DOI:10.1016/0273-1177(95)00079-T.

Kogan, F.N. (1997) Global drought watch from space. *Bulletin of the American Meteorological Society*. 78:621–636.

Kogan, F., Sullivan, J. (1993) Development of global drought-watch system using NOAA/ AVHRR data, *Advances in Space Research*, 13: 219–222, DOI:10.1016/0273-1177(93) 90548-P.

Kohler, M.A., Linsley, R.K.Jr (1951) Predicting runoff from storm rainfall. Res. Paper 34, U.S.Weather Bureau,Washington, D.C.

Kousari, M.R., Dastorani, M.T., Niazi, Y., Soheini, E., Hayatzadeh, M., Chezgi, J. (2014) Trend detection of drought in arid and semi-arid aegions of Iran based on implementation of Reconnaissance Drought Index (RDI) and application of non-parametrical statistical method. *Water Resources Management*, 28:1857-1872, DOI:10.1007/s11269-014-0558-6.

Koza, J.R., Keane, M.A., Streeter, M.J. (2003) Evolving Inventions, Scientific American, p. 40-47, ISSN 0036-8733.

Kustas, W.P., Schmugge, T.J., Humes, K.S., Jackson, T.J., Parry, R., Weltz, M.A., Moran, M.S. (1993) Relationships between evaporative fraction and remotely sensed vegetation index and microwave brightness temperature for semiarid rangelands. *Journal of Applied Meteorology*, 32:1781-1790.

Lawrimore, J., Heim Jr., R.R., Svoboda, M., Swail, V., Englehart, P.J. (2002) Beginning a new era of drought monitoring across North America. *Bulletin of the American Meteorological Society*, 83:1191–1192.

Leblanc, M.J., Tregoning, P., Ramillien, G., Tweed, S.O., Fakes, A. (2009) Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, *Water Resources Research*, 45:W04408, DOI:10.1029/2008WR007333.

Li, B., Rodell, M., Zaitchik, B.F., Reichle, R.H., Koster, R.D., van Dam, T.M. (2012) Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe, *Journal of Hydrology*, 446:103-115, DOI:10.1016/j.jhydrol.2012.04.035.

Li, Q., Li, P., Li, H., Yu, M. (2014) Drought assessment using a multivariate drought index in the Luanhe River basin of Northern China. *Stochastic Environmental Research and Risk Assessment*, 29:1509-1520, DOI:10.1007/s00477-014-0982-4.

Linting, M., Meulman, J.J. Groenen, P.J.F., van der Kooij, A.J. (2007) Non-linear principal components analysis: Introduction and application. *Psychological Methods Journal*, 12:336–358.

Logar, I., van den Bergh, J.C.J.M. (2013) Methods to Assess Costs of Drought Damages and Policies for Drought Mitigation and Adaptation: Review and Recommendations. *Water Resources Management*, DOI:10.1007/s11269-012-0119-9.

Ma, M., Ren, L., Singh, V.P., Yang, X., Yuan, F., Jiang, S. (2014) New variants of the Palmer drought scheme capable of integrated utility. *Journal of Hydrology*, 519:1108-1119.

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 Sensing of Environment*, 90:441–450.

McColl, J.C., Young, M.D. (2010) Briefing: drought and structural adjustment in Australia. *Journal of Water Management*, 163:191-195, DOI: 10.1680/ensu.2010.163.4.191.

McDonald, J.E., Green, C.R. (1960) Effects of inhomogeneity and record length on estimates of correlation and variability of precipitation data. *Journal of Geophysical Research*. 56: 2375-2381.

McKee, T.B., Doesken, N.J. Kleist, J. (1993) The relationship of drought frequency and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology, Anaheim, California, USA, 179-184.

McKee, T.B., Doesken, N.J., Kleist, J. (1995) Drought Monitoring with Multiple Time Scales, Paper Presented at 9th Conference on Applied Climatology. American Meteorological Society, Dallas, Texas.

McKee, T.B., Edwards, D.C. (1997) Characteristics of 20th century drought in the United States at multiple time scales, *Atmospheric Science*, 634:97–92.

Mehran, A., Mazdiyasni, O., AghaKouchak, A. (2015) A hybrid framework for assessing socioeconomic drought: Linking climate variability, local resilience, and demand, *Journal of Geophysical Research: Atmospheres*, 120:7520-7533, DOI:10.1002/2015JD023147.

Meinke, H., Hochman, Z. (2000) Using seasonal climate forecasts to manage dryland crops in northern Australia', in Hammer, G. L., Nicholls, N. and Mitchell, C. (eds.), Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems: The Australian Experience, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 149–165.

Meinke, H., Stone, R.C. (2005) Seasonal and inter-annual climate forecasting: the new tool for increasing preparedness to climate variability and change in agricultural planning and operations. *Climatic Change*, 70: 221–253.

Mishra, A.K., Singh, V.P. (2010) A review of drought concepts, *Journal of Hydrology*, 391:202-216, DOI:10.1016/j.jhydrol.2010.07.012.

Mo, K.C., Lettenmaier, D.P. (2013) Objective drought classification using multiple land surface models. *Journal of Hydrometeorology*, 15:990–1010.

Mpelasok, F., Hennessy, K., Jones, R., Bates, B. (2008) Comparison of suitable drought indices for climate change impacts assessment over Australia towards resource management. *International Journal of Climatology*, 28: 1283–1292.

Mu, Q., Zhao, M., Kimball, J.S., McDowell, N.G., Running, S.W. (2013) A remotely sensed global terrestrial drought severity index. *Bulletin of the American meteorological society*, 94:83-98, DOI:10.1175/BAMS-D-11-00213.1.

Nash, J.E., Sutcliffe, J.V. (1970) River flow forecasting through conceptual models part I: A discussion of principles. *Journal of Hydrology*, 10:282–290.

NCEDA (National Centre of Excellence in Desalination) (2010) Australian desalination research roadmap. Murdoch University, Rockingham, Western Australia.

Nedovic-Budic, Z., Feeney, M., Rajabifard, A., Williamson, I. (2004) Are SDIs serving the needs of local planning? Case study of Victoria, Australia, and Illinois, USA. *Computers, Environment and Urban Systems*, 28:329–351.

Nelsen, R.B. (2006) An Introduction to Copulas. Springer, New York.

Nemani, R., Hashimoto, H., Votava, P., Melton, F., Wang, W., Michaelis, A., Mutch, L., Milesi, C., Hiatt, S., White M. (2009) Monitoring and forecasting ecosystem dynamics using the Terrestrial Observation and Prediction System (TOPS), *Remote Sensing of Environment*, 113:1497–1509, DOI:10.1016/j.rse.2008.06.017.

Nishida, K., Nemani, R.R., Running, S.W., Glassy, J.M. (2003) An operational remote sensing algorithm of land surface evaporation. *Journal of Geophysical Research: Atmospheres*, 108:4270, DOI:10.1029/2002JD002062.

Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (2015) FLUXNET Web Page. Available online [http://fluxnet.ornl.gov] from ORNL DAAC, Oak Ridge, Tennessee, U.S.A. Last Access: January, 2016.

Palmer, W.C. (1965) Meteorological Drought. Research Paper No. 45, U.S. Department of Commerce Weather Bureau, Washington, D.C.

Palmer, W.C. (1968) Keeping track of crop moisture conditions, nationwide: The new Crop Moisture Index. Weatherwise 21:156–161.

Panofsky, H.A., Brier, G.W. (1958) Some applications of statistics to meteorology, Mineral Industries Extension Services, College of Mineral Industries, Pennsylvania State University. PUB ID: 101-188-880.

Park, J.S., Kim, K.T., Choi, Y.S. (2008) Application of vegetation condition index and standardized vegetation index for assessment of spring drought in South Korea, in Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, 3:774–777, IEEE, Boston, Mass.

Parana Manage, N., Lockart, N., Willgoose, G.R., Kuczera, G., Kiem, A.S., Chowdhury, A.F.M.K., Zhang, L., Twomey, C. (2016) Statistical testing of dynamically downscaled rainfall data for the Upper Hunter region, New South Wales, Australia. Journal of Southern Hemisphere Earth System Science, 66(2), 203-227.

Patel, N.R., Parida, B.R., Venus, V., Saha, S.K., Dadhwal, V.K. (2012) Analysis of agricultural drought using vegetation temperature condition index (VTCI) from Terra/MODIS satellite data. *Environmental Monitoring and Assessment*, 184:7153–7163.

Pearson, K. (1895) Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58: 240–242.

Peel, M.C., Finlayson, B.L., McMahon, T.A. (2007) Updated world map of the Koppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11:1633–1644.

Porth, L.S., Boes, D.C., Davis, R.A., Troendle, C.A., King, R.M. (2001) Development of a technique to determine adequate sample size using sub sampling and return interval estimation. *Journal of Hydrology*, 251: 110–116.

Rahmat, S.N., Jayasuriya, N., Bhuiyan, M. (2015) Assessing droughts using meteorological drought indices in Victoria, Australia. *Hydrology Research*, 46(3):463.

Rajsekhar, D., Singh, V.P., Mishra, A.K. (2015) Multivariate drought index: an information theory based approach for integrated drought assessment. *Journal of Hydrology*, 526:164–182.

Rassmussen, E.M., Dickinson, R.E., Kutzbach, J.E. Cleaveland, M.K. (1993) Climatology. In: Maidment, D.R. ed. Handbook of Hydrology. McGraw-Hill, Inc., New York, NY, 2.1-2.44.

Raupach, M., Briggs, P., King, E., Paget, M., Trudinger, C. (2007) Australian Water Availability Project (AWAP), CSIRO Marine and Atmospheric Research Component: Final Report for Phase 3, CSIRO Marine and Atmospheric Research, Canberra, Australia.

Riebsame, W.E., Changnon, S.A., Karl, T.R. (1990) Drought and natural resource management in the United States: impacts and implications of the 1987–1989 drought, Westview Press, p. 174.

Risbey, J.S. (2011) Dangerous climate change and water resources in Australia. *Regional Environmental Change*, 11:197–203.

Rodell, M. (2012) Satellite gravimetry applied to drought monitoring, in remote sensing of drought: innovative monitoring approaches, pp. 261–277, CRC Press.

Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J., Wilson, C.R. (2007) Estimating groundwater storage changes in the Mississippi River Basin (USA) using GRACE, *Hydrogeology Journal*, 15: 159–166.

Rodell, M., Famiglietti, J. (2002) The potential for satellite-based monitoring of groundwater storage changes using GRACE: The high plains aquifer, central US, *Journal of Hydrology*, 263: 245–256.

Rouse, J., Haas, R., Schell, J., Deering, D., Harlan, J. (1974) Monitoring the Vernal Advancement and Retrogradation (Greenwave Effect) of Natural Vegetation, 362 pp., Texas A & M Univ., Remote Sens. Cent., College Station, Tex.

Sahin, O., Stewart, R.A., Helfer, F. (2015) Bridging the water supply–demand gap in Australia: coupling water demand efficiency with rain-independent desalination supply. *Water Resources Management*, 29:1-20, DOI:10.1007/s11269-014-0794-9.

Salomonson, V., Guenther, B., Masuoka, E. (2001) A summary of the status of the EOS Terra Mission Moderate Resolution Imaging Spectroradiometer (MODIS) and attendant data product development after one year of on-orbit performance. In: Proceedings of the International Geoscience and Remote Sensing Symposium/IGARSS'2001, Sydney, Australia, 9–13 July.

Santos, J.F., Pulido-Calvo, I., Portela, M.M. (2010) Spatial and temporal variability of droughts in Portugal, *Water Resources Research*, 46:W03503, DOI:10.1029/2009WR008071.

Sarmadi, F., Shokoohi, A. (2015) Regionalizing precipitation in Iran using GPCC gridded data via multivariate analysis and L-moment methods. *Theoretical and Applied Climatology*, 122:121-128, DOI: 10.1007/s00704-014-1292-y.

Sarmadi, F., Azmi, M. (2016) Regionalizing mean air temperature in Iran by multivariate analysis and L-moment methods. *Journal of Hydrologic Engineering*, 21: 05015018, DOI: 10.1061/(ASCE)HE.1943-5584.0001280.

Schmidt, M., Lipson, H. (2009) Distilling free-form natural laws from experimental data. *Science*, 324:81-85.

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.

Seddon, A.W.R., Macias-Fauria, M., Long, P.R., Benz, D., Willis, K. (2016) Sensitivity of global terrestrial ecosystems to climate variability. *Nature*, 531:229-232, DOI:10.1038/nature16986.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J. (2010) Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99: 125–161, DOI: 10.1016/j.earscirev.2010.02.004.

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., Zhang, X. (2012) Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Sepulcre-Cantó, G., Horion, S., Singleton, A., Carrão, H., Vogt, J. (2012) Development of a combined drought indicator to detect agricultural drought in Europe. *Natural Hazards and Earth System Sciences*, 12:3519–3531.

Shafer, B.A., Dezman, L.E. (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.

Shahabfar, A., Eitzinger, J. (2013) Spatio-temporal analysis of droughts in semi-arid regions by using meteorological drought indices. *Atmosphere*. 4: 94-112, DOI:10.3390/atmos 4020094.

Sheffield, J., Wood, E.F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., Ogallo, L. (2014) A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bulletin of the American meteorological society*. 95:861-882, DOI:10.1175/BAMS-D-12-00124.1.

Shukla, S., Wood, A.W. (2008) Use of a standardized runoff index for characterizing hydrologic drought. *Geophysical Research Letters,* 35:L02405, DOI:10.1029/ 2007GL032487.

Shuttleworth, W.J., Gurney, R.J., Hsu, A.Y., Ormsby, J.P. (1989) FIFE: The variation in energy partition at surface flux sites, IAHS Publ., 186, 67–74.

Singh, J., Dhillon, S.S. (2006) Agricultural geography, Tata McGraw-Hill Publishing, 3rd edition, New Delhi.

Singh, R.P., Roy, S., Kogan, F. (2003) Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. *International Journal of Remote Sensing*, 24:4393–4402.

Smakhtin, V.U., Hughes, D.A. (2007) Automated estimation and analysis of meteorological drought characteristics from monthly data. *Environmental Modelling & Software*, 22: 880-890.

Soltani, S., Modarres, R. (2006) Classification of spatio-temporal pattern of rainfall in Iran using a hierarchical and divisive cluster analysis. *Journal of Spatial Hydrology*. 6:1-12.

Spearman, C. (1904) The proof and measurement of association between two things. *American Journal of Psychology*. 15: 72–101. DOI:10.2307/1412159

Steinemann, A., Hayes, M., Cavalcanti, L. (2005) Drought indicators and triggers. In: Wilhite, D. (Ed.), Drought and Water Crises: Science, Technology, and Management Issues. Dekker, New York, pp. 71–92.

Steinemann, A., Iacobellis, S.F., Cayan, D.R. (2015) Developing and evaluating drought indicators for decision-making. *Journal of Hydrometeorology*, 16:1793-1803, DOI: 10.1175/JHM-D-14-0234.1.

Stephens, M.A. (1974) EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*. 69:730–737.

Sur, C., Hur, J., Kim, K., Choi, W., Choi, M. (2015) An evaluation of satellite-based drought indices on a regional scale, *International Journal of Remote Sensing*, 36:5593-5612, DOI: 10.1080/01431161.2015.1101653.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D. (2002) The drought monitor. *Bulleting of the American Meteorological Society*, 83:1181–1190.

Svoboda, M.D., Fuchs, B.A., Poulsen, C.C., Nothwehr, J.R. (2015) The drought risk atlas: enhancing decision support for drought risk management in the United States. *Journal of Hydrology*, 526:274–286.

Tadesse, T., Wardlow, B.D., Brown, J., Hayes, M., Svoboda, M., Fuchs, B., Gutzmer, A.D. (2015) Assessing the vegetation condition impacts of the 2011 drought across the U.S. southern Great Plains using the Vegetation Drought Response Index (VegDRI). *Journal of Applied Meteorology and Climatology*, 54:153–169.

Tallaksen, L.M., van Lanen, H.A.J. (2004) Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Science, 48. Amsterdam, Elsevier Science B.V.

Thomas, A.C., Reager, J.T., Famiglietti, J.S., Rodell, M. (2014) A GRACE-based water storage deficit approach for hydrological drought characterization, *Geophysical Research Letter*, 41:1537–1545, DOI:10.1002/2014GL059323.

Tozer, C.R., Kiem, A.S., Verdon-Kidd, D.C. (2012) On the uncertainties associated with using gridded rainfall data as a proxy for observed. *Hydrology and Earth System Sciences*, 16, 1481-1499,doi:10.5194/hess-16-1481-2012.

Tsakiris, G., Nalbantis, I., Vangelis, H., Verbeiren, B., Huysmans, M., Tychon, B., Jacquemin, I., Canters, F., Vanderhaegen, S., Engelen, G., Poelmans, L., De Becker, P., Batelaan, O. (2013) A system-based paradigm of drought analysis for operational management. *Water Resources Management*. 27:5281-5297, DOI:10.1007/s11269-013-0471-4.

Tsakiris, G., Pangalou, D., Vangelis, H. (2007) Regional drought assessment based on the Reconnaissance Drought Index (RDI), *Water Resources Management*, 21:821–833.

USDA (1994) Major world crop areas and climatic profiles. World Agricultural Outlook Board, US Department of Agriculture. *Agricultural Handbook*, 664:157–170.

Van Dijk, A.I.J.M. (2010) AWRA Technical Report 3: Landscape Model (version 0.5) Technical Description. WIRADA/CSIRO report.

van Dijk, A.I.J.M., Beck, H.E., Crosbie, R.S., Jeu, R.A., Liu, Y.Y., Podger, G.M., Timbal, B., Viney, N.R. (2013) The millennium drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, *Water Resources Research*, 49:1040–1057, DOI:10.1002/wrcr.20123.

Van Loon, A.F. (2015) Hydrological drought explained. *WIREs Water*, 2:359–392, doi:310.1002/wat1002.1085.

Van Loon, A.F. et al. (2016a) Drought in the Anthropocene. *Nature Geoscience*, 9:89-91, doi:10.1038/ngeo2646.

Van Loon, A.F., et al. (2016b) Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches. *Hydrol. Earth Syst. Sci.*,20(9), 3631-3650.

Van Loon, A.F., Van Lanen, H. (2012) A process-based typology of hydrological drought. *Hydrological and Earth System Sciences*, 16:1915-1946.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I. (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate,* 23:1696–1718.

Victorian Resources Online (VRO) (2015) Victorian Land Use Information System (VLUIS). http://vro.depi.vic.gov.au/dpi/vro/vrosite.nsf/pages/landuse-home, based on the updated on 20/2/2016.

Wang, A., Bohn, T.J., Mahanama, S.P., Koster, R.D., Lettenmaier, D.P. (2009) Multimodel ensemble reconstruction of drought over the continental United States, *Journal of Climate*, 22: 2694–2712.

Wang, K., Dickinson, R.E. (2012) A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Review of Geophysics*, 50: RG2005, DOI:10.1029/2011RG000373.

Wang, L., Qu, J.J. (2009) Satellite remote sensing applications for surface soil moisture monitoring: A review, *Frontiers of Earth Science in China*, 3: 237–247.

Wells, N., Goddard, S., Hayes, M.J. (2004) A self-calibrating Palmer drought severity index. *Journal of Climate*, 17:2335–2351.

Wilhite, D.A. (2000) Drought as a natural hazard: concepts and definitions. drought: a Global assessment, D. Wilhite, Ed., 1:3–18.

Wilhite, D.A. (2005) Drought and water crises: science, technology, and management issues, pp. 432, vol. 86, CRC Press.

Wilks, D.S. (1995) Statistical methods in the atmospheric sciences: an introduction, International Geophysics Series, vol. 59, 467 pp., Academic, San Diego, Calif.

Willmott, C.J. (1984) On the evaluation of model performance in physical geography. Spatial Statistics and Models, G. L. Gaile and C. J. Willmott, eds., 443-460.

Wood, E.F., Schubert, S.D., Wood, A.W., Peters-Lidard, C.D., Mo, K.C., Mariotti, A., Pulwarty, R.S. (2015) Prospects for advancing drought understanding, monitoring, and prediction. *Journal of Hydrometeorology*, 16:1636-1657, DOI: 10.1175/JHM-D-14-0164.1.

Wu, H., Hayes, M.J., Wilhite, D.A., Svoboda, M.D. (2005) The effect of the length of record on the Standardized Precipitation Index calculation. *International Journal of Climatology*, 25: 505–520.

Xia, Y., Ek, M.B., Peters-Lidard, C.D., Mocko, D., Svoboda, M., Sheffield, J., Wood, E.F., (2014a) Application of USDM statistics in NLDAS-2: optimal blended NLDAS drought index over the continental United States. *Journal of Geophysical Research: Atmospheres*, 119:2947-2965, DOI:2013JD020994.

Xia, Y., Ek, M.B., Mocko, D., Peters-Lidard, C.D., Sheffield, J., Dong, J., Wood, E.F. (2014b) Uncertainties, correlations, and optimal blends of drought indices from the NLDAS multiple land surface model ensemble. *Journal of Hydrometeorology*, 15:636–1650.

Yirdaw, S.Z., Snelgrove, K.R., Agboma, C.O. (2008) GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian prairie, *Journal of Hydrology*, 356: 84–92.

Zaitchik, B.F., Rodell, M., Reichle, R.H. (2008) Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, 9: 535–548.

Zhang, A., Jia, G. (2013) Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sensing of Environment*, 134:12–23.