Development and evaluation of methodologies for monitoring droughts and their impacts on agriculture in data-scarce areas

Presented by **Ronnie Javier Araneda Cabrera**

Doctoral Thesis

2022

Supervisors:

Jerónimo Puertas Agudo

María Bermúdez Pita

Civil Engineering Ph.D. Program

UNIVERSIDADE DA CORUÑA

The undersigned hereby certifies that they are supervisors of the Doctoral Thesis entitled “Development and evaluation of methodologies for monitoring droughts and their impacts on agriculture in data-scarce areas” developed by Ronnie Javier Araneda Cabrera, whose signature is also included, in the framework of the Civil Engineering Ph.D. Program at the University of A Coruña, consenting to its presentation and subsequent defence.

Los abajo firmantes hacen constar que son los directores de la Tesis Doctoral titulada “Development and evaluation of methodologies for monitoring droughts and their impacts on agriculture in data-scarce areas” realizada por Ronnie Javier Araneda Cabrera, cuya firma también se incluye, en el marco del Programa de Doctorado en Ingeniería Civil de la Universidade da Coruña, dando consentimiento para su presentación y posterior defensa.

Supervisors / Directores:

Jerónimo Puertas Agudo

María Bermúdez Pita

Doctoral candidate / Doctorando:

Ronnie Javier Araneda Cabrera

UNIVERSIDADE DA CORUÑA

The current doctoral thesis is presented to the Department of Civil Engineering of the Universidade da Coruña in compliance with the requirements for obtaining the title of Doctor in Civil Engineering. The dissertation is presented in the form of a compendium of scientific articles including the following referenced works:

La presente tesis doctoral es presentada ante el departamento de Ingeniería Civil de la Universidade da Coruña en cumplimiento de los requerimientos para la obtención del título de Doctor en Íngeniería Civil. La disertación se presenta en la modalidad de compendio de artículos científicos incluyendo los siguientes trabajos referenciados:

Research article 1 / Artículo científico 1:

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2021. Revealing the spatio-temporal characteristics of drought in Mozambique and their relationship with large-scale climate variability. J. Hydrol. Reg. Stud. 38, 100938. https://doi.org/10.1016/j.ejrh.2021.100938.

Research article 2 / Artículo científico 2:

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2021. Benchmarking of drought and climate indices for agricultural drought monitoring in Argentina. D790, 148090. https://doi.org/10.1016/j.scitotenv.2021.148090.

Research article 3 / Artículo científico 3:

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2021. Assessment of the performance of drought indices for explaining crop yield variability at the national scale: Methodological framework and application to Mozambique. Agric. Water Manag. 246. https://doi.org/10.1016/j.agwat.2020.106692.

Research article 4 / Artículo científico 4:

Araneda-Cabrera, R.J., Bermudez, M., Puertas, J., 2021. Índices de precipitación y vegetación estandarizados bivariables para evaluar y monitorear sequías agrícolas. / Bivariate standardized precipitation and vegetation indices for assessing and monitoring agricultural drought. Rev. Hidrolatinoamericana 5, 27–30.

In addition, a paper at an international conference and a book (manual) complete the work presented in this thesis. The first peer-reviewed research has been published in the Proceedings of the 6th IAHR Europe Congress, 2020, Warsaw, Poland. The second contribution has been published by the Publication Service of the Universidade da Coruña in Spanish and Portuguese. The two works are cited below:

Además, una comunicación en un congreso internacional y un libro (manual) completan el trabajo presentado en esta tesis. La primera investigación revisada por pares ha sido publicada en los proceedings del 6º Congreso Europeo de la IAHR, 2020, Warsaw, Polonia. El segundo trabajo ha sido publicado por el Servicio de Publicación de la Universidade da Coruña en versión español y portugués. Los trabajos se citan a continuación:

Conference paper / Artículo de congreso:

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2021. A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation drought index, in: 6th IAHR Europe Congress; Rowiński, Paweł; Okruszko, Tomasz; Nones, Michael, Eds. pp. 513–514.

Book Spanish versión / Libro versión en español:

Araneda-Cabrera, R.J., Puertas, J., Alvarez, M., Penas, V., 2022a. Manual para la obtención y aplicación de variables hidrometeorológicas desde bases de datos globales para la gestión de sequías, Universida. ed, Universidade da Coruña, Servizo de Publicacións. Universidade da Coruña, Servizo de Publicacións, A Coruña. https://doi.org/https://doi.org/10.17979/spudc.9788497498333.

Book Portuguese versión / Libro versión en portugués

Araneda-Cabrera, R.J., Puertas, J., Alvarez, M., Penas, V., 2022b. Manual para a obtenção e aplicação de variaveis hidrometeorológicas a partir de bases de dados globais para a gestão das secas, Universida. ed, Universidade da Coruña, Servizo de Publicacións. A Coruña. https://doi.org/https://doi.org/10.17979/spudc.9788497498340.

Acknowledgments/ Agradecimientos

Desde pequeño tuve la suerte, al otro lado del mundo en mi querida Cuenca (Ecuador), de que familia y amigos compartieran sus experiencias y vivencias universitarias conmigo. Este hecho despertó en mi un interés en la investigación que hoy se ve plasmada en este documento. En estas líneas quisiera mostrar el agradecimiento a personas e instituciones que han hecho posible que cumpla una de mis grandes metas, como es realizar esta tesis doctoral.

En primer lugar, quiero agradecer a mis tutores Jerónimo Puertas y María Bermúdez por haberme dado la oportunidad de trabajar y aprender con ellos. Por haberme dado las libertades para aprender y desarrollar ideas, aunque no siempre acertara con ellas. Por ayudarme a formar nuevas aptitudes y conocimientos. Les agradezco el compromiso tan firme que han mantenido constantemente, con una pandemia incluida en el proceso. Gracias por todo el cuidado en esta mi segunda casa y por la calidad humana transmitida. Este trabajo es tan suyo como mío.

Mi agradecimiento a la Xunta de Galicia y a la Unión Europea por la ayuda para mi contrato predoctoral (referencia ED481A- 2018/162) de la que he sido beneficiario durante estos últimos tres años y medio. Agradezco Programa IACOBUS (ayudas 2019) y al Programa inMOTION para becas para estancias predoctorales INDITEX-UDC 2020, ya que con su financiación me han permitido realizar estancias de investigación en Oporto (Portugal) y en Delft (Países Bajos) respectivamente. De igual manera agradezco a Aguas de Galicia y a Cooperación Galega (Xunta de Galicia), ya que parte del trabajo de esta tesis proviene de colaboraciones que he podido realizar en los proyectos SECARA y Aqua-Moz relacionados con estas instituciones.

Durante mis estancias, tuve el placer de estar bajo la supervisión de Micha Werner en Delft y de Rodrigo Maia en Oporto, a quienes agradezco sinceramente todo el tiempo y ayuda que me ha dedicado. En sus departamentos conocí a gente muy valiosa a quienes agradezco la gran acogida y el trato recibido, hicieron de mis estancias grandes experiencias tanto profesionales como personales.

A todos los compañeros y compañeras del GEAMA, gracias por crear un ambiente de compañerismo, los buenos momentos, los cafés y aguantar mis charlas sobre sequías. Con miedo a dejar a alguno fuera de este texto (pero no de la cabeza y el corazón) me arriesgo a nombrarlos. A los que estuvieron primero Juan, Esteban y Manu que siempre, desde el día uno, están dispuestos a darme una mano. A Manuel Álvarez y Víctor Penas por su apoyo y consejo en el desarrollo de la tesis. A los que cruzamos y aún compartimos camino Gonzalo, Juan Farfán, Nacho, Gabriela, Emilio, Rocío, Paolo, Andrea, Jandro, Carlos, Omar y Amal. Como no agradecer a Luis Cea, Luis Pena, Quique, José, Andrés, Igor, Delia, Humberto y “a los del CITEEC” por las buenas charlas y momentos en los pasillos y cafeterías de la UDC.

La vida me ha sonreído con muchos amigos y una gran familia (a los que me haría falta otra tesis para nombrarlos a todos) que todos los días me han apoyado desde lejos para llegar donde estoy. Gracias a todos por los ánimos y siempre estar pendientes.

A mis padres, mi hermana y abuelos especial agradecimiento, la distancia no ha sido nada para sentir su contante apoyo y cariño. No sé cuántas horas los aburrí contando sobre algún tema paper o informe. Gracias.

A mi compañera de vida, Raquel, la que todos los días me apoya y acompaña en el camino. Realizar una tesis tiene altibajos, ella siempre estuvo ahí. Gracias.

Contents

List of figures

Chapter 1

Chapter 2

Supplementary information

Chapter 3

Supplementary information

Chapter 4

Supplementary information

Chapter 5

Chapter 6

Appendix A

List of tables

Chapter 1

Chapter 2

Supplementary information

Chapter 3

Chapter 4

Supplementary information

Chapter 5

Chapter 6

Appendix A

Abstract

Drought is one of the natural phenomena that causes the greatest socio-economic and environmental damage. Its impacts are of particular importance in agriculture, as this activity is closely linked to food security and quality of life in many territories. Droughts can occur in any climatic regime in the world, with arid and semi-arid areas being the most affected and prone to drought events. In regions particularly exposed and vulnerable to drought, specific drought studies are needed to help manage and mitigate its impacts. This thesis is a contribution to the management of drought and its impacts, specifically on agriculture.

Several novel and bespoke methodologies were developed with the aim of increasing knowledge of drought phenomena and providing solutions for water resources and drought management. Freely available global scale hydrometeorological data sources were used, so that the methodologies can be applied to any country or region of the world. The case studies were Mozambique and Argentina, both are developing countries with significant agricultural activity (in terms of cropland extension) and prone to drought events. Methodologies focused on defining and understanding the spatio-temporal characteristics of droughts; defining and relating drought events to their triggers; validating tools for monitoring droughts and their impacts on agricultural activity; and knowledge transfer to all beneficiaries and stakeholders involved in drought management in data-scarce regions. The methodologies are of general applicability and can be replicated worldwide, providing meaningful information to the scientific, technical and management community to develop, calibrate or validate existing and new formulations. In addition, they could contribute to the creation of drought mitigation and adaptation plans aimed at reducing impacts, especially in agriculture.

Resumen

La sequía es uno de los fenómenos naturales que mayores daños socioeconómicos y medioambientales causa. Sus impactos son de especial importancia en la agricultura, ya que esta actividad está ligada a la seguridad alimentaria y calidad de vida de muchos territorios. Las sequías pueden ocurrir en cualquier régimen climático del mundo, siendo las zonas áridas y semiáridas las más afectadas y propensas a eventos de sequía en el futuro. En las regiones particularmente expuestas y vulnerables a la sequía, se necesitan estudios específicos sobre la sequía para ayudar a controlar y mitigar sus impactos. Esta tesis es una contribución a la gestión de las sequías y sus impactos, específicamente en la agricultura.

Se desarrollaron varias metodologías específicas y novedosas con el objetivo de aumentar el conocimiento de los fenómenos de la sequía y aportar soluciones para la gestión de los recursos hídricos y de la sequía. Se hizo uso de fuentes de datos hidrometeorológicos alternativos de libre acceso, de manera que las metodologías pueden aplicarse a cualquier país o región del mundo y a cualquier escala espacial. Los casos de estudio fueron países en vías de desarrollo con una importante actividad agrícola (extensión de cultivos) y propensos a eventos de sequía. Se uso Mozambique y Argentina debido a su situación económica y compleja disponibilidad de datos. Las metodologías se centraron en definir y comprender las características espaciotemporales de las sequías; en definir y relacionar los eventos de sequía con sus desencadenantes; en la validación de herramientas para el seguimiento de las sequías y sus impactos en la actividad agrícola; y, en la transferencia de conocimientos a todos los beneficiarios e implicados en la gestión de la sequía en regiones con escases de datos. Las metodologías y los resultados obtenidos pueden ser replicados en cualquier parte del mundo, proporcionando información significativa a la comunidad científica, técnica y de gestión para desarrollar, calibrar o validar formulaciones existentes y nuevas. Además, son herramientas que podrían contribuir a la creación de planes de mitigación y adaptación a la sequía destinados a reducir los impactos, especialmente en la agricultura.

Resumo

A seca é un dos fenómenos naturais que provocan maiores danos socioeconómicos e ambientais. Os seus impactos son de especial importancia no agro, xa que esta actividade está moi ligada á seguridade alimentaria e á calidade de vida en moitos territorios. As secas poden ocorrer en calquera réxime climático do mundo, sendo as zonas áridas e semiáridas as máis afectadas e propensas a sufrir eventos de seca. Nas rexións especialmente expostas e vulnerables á seca, son necesarios estudos específicos sobre a seca para axudar a xestionar e mitigar os seus impactos. Esta tese é unha contribución á xestión da seca e os seus impactos, concretamente na agricultura.

Desenvolvéronse varias metodoloxías novedosas e a medida co obxectivo de aumentar o coñecemento dos fenómenos da seca e proporcionar solucións para os recursos hídricos e a xestión da seca. Utilizáronse fontes de datos hidrometeorolóxicos a escala global de libre dispoñibilidade, de xeito que as metodoloxías poden aplicarse a calquera país ou rexión do mundo. Os estudos de caso foron países en desenvolvemento con actividade agrícola importante (extensión de terras de cultivo) e propensos a eventos de seca. Utilizáronse Mozambique e Arxentina debido á súa situación económica e á complexa dispoñibilidade de datos. Metodoloxías enfocadas a definir e comprender as características espazo-temporais das secas; definir e relacionar os eventos de seca cos seus desencadenantes; validación de ferramentas para o seguimento das secas e os seus impactos na actividade agraria; e, transferencia de coñecemento a todos os beneficiarios e partes interesadas implicadas na xestión da seca en rexións con escaseza de datos. As metodoloxías son de aplicabilidade xeral e pódense replicar en todo o mundo, proporcionando información significativa á comunidade científica, técnica e de xestión para desenvolver, calibrar ou validar formulacións existentes e novas. Ademais, poderían contribuír á creación de plans de mitigación e adaptación á seca destinados a reducir os impactos, especialmente no agro.

Chapter 1

Introduction, main results, and conclusions

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Writing – original draft, Visualization. María Bermudez: Conceptualization, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Writing – review & editing, Supervision.

1. Overview and background

Droughts are among the most common natural phenomena worldwide, and can occur anywhere under any climate regime (Bryant et al., 2005; Sheffield and Wood, 2012). However, it is one of the least understood hazards due to its complexity and difficult quantification (Hagenlocher et al., 2019a). Droughts can be perceived by the chains of energy production, food, water supply, etc., being generally detected when the consequences of the phenomenon are difficult to mitigate, so it is the disaster that causes more socioeconomic losses worldwide (WMO, 2006). Therefore, specific drought studies at regional or local levels are needed in order to support drought management stakeholders.

Droughts must be differentiated from other related phenomena such as scarcity and aridity. Scarcity represents a permanent situation of deficit (calculated with the water balance) in relation to the demand for water in a regional resource system, characterised either by an arid climate or by a rapid growth in water demand. Aridity is a natural structural situation of a region and therefore permanent (Svodova et al., 2016). Even in these circumstances, there should not be a deficit if the exploitation systems are adequately designed and exploited and the demands are kept within reasonable limits, in accordance with the climatic characteristics of the region. This requires planned medium and long-term actions.

The classification of droughts types are typically four: meteorological, agricultural, hydrological and socioeconomic (Wilhite and Glantz, 1985). Meteorological drought is associated to a deficit of precipitation over a region for a period of time; agricultural drought is related to a period when crops fail, and vegetation condition is affected because of deficits of soil moisture; hydrological drought refers to a deficit of water resources in the surface and subsurface; and socioeconomic drought is associated with failure in the water resources systems to meet water demands for socioeconomic good. The sequence of occurrence of the type of drought and their main impacts are illustrated in Fig. 1.1.

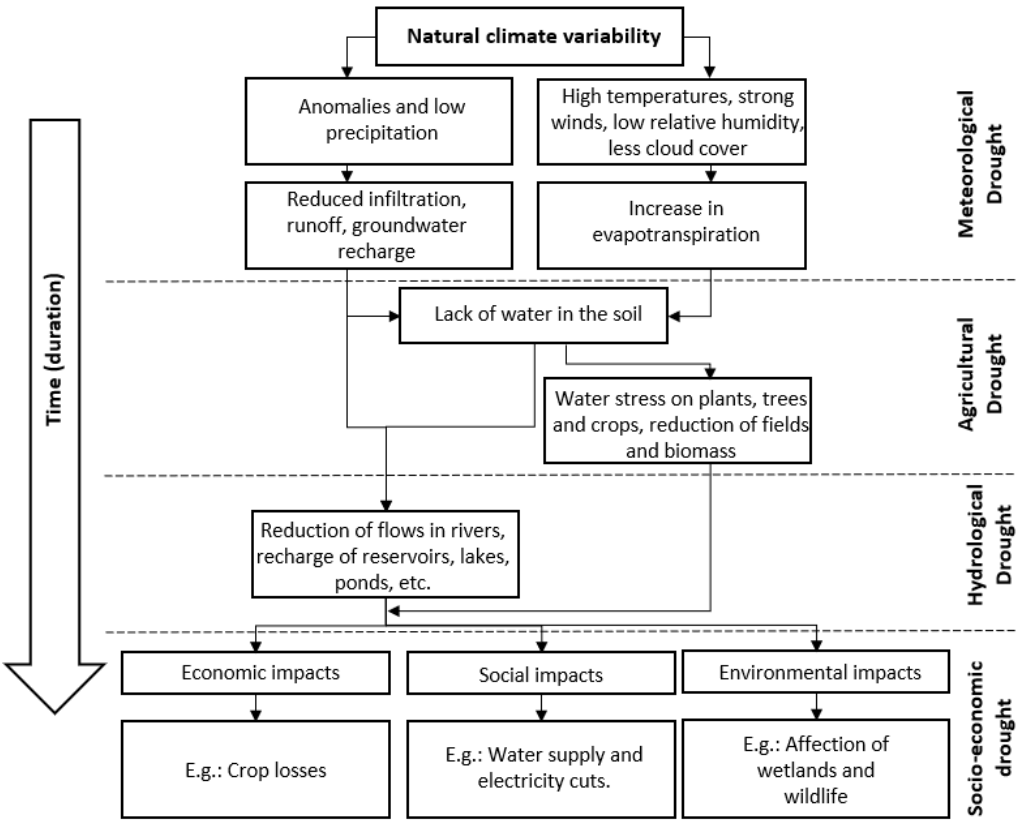


Fig. 1.1 Sequence of occurrence of drought types and their impacts. Source: National Drought Mitigation Center de la Universidad de Nebraska –Lincoln.

Droughts are commonly evaluated by events, which can be quantified and compared. An event can be defined according to the "run theory" (Yevjevich, 1969), which defines the duration, peak or maximum intensity, severity and probability or frequency of a drought event. It also makes it possible to delimit the affected areas and to calculate the extent of drought in a given area over a period of time. These characteristics can then be related to corresponding impacts, for example in agriculture (Yihdego et al., 2019).

The effects or impacts of droughts are expected to worsen in the next decades due to climate change. Changes in spatial-temporal patterns of precipitation and extreme temperatures could make droughts more recurrent (Mishra and Singh, 2011). As stated by the (IPCC, 2014) there are trends of increasing intensities and frequencies of droughts around the world, with arid and semi-arid areas possibly being the most affected. Recent studies have suggested that future drought events may affect larger areas, reaching a higher level of impact than any other climate-related hazard, especially on food supply activities (Romm, 2011). Ultimately, there is widespread acceptance that droughts in southern Africa and South America will intensify and extend in duration over the coming decades due to reduced precipitation and increased evapotranspiration (due to temperature increase effects). Countries such as Mozambique and Argentina, the case studies used in this thesis, are vulnerable to potential climate change impacts that may bring more severe extreme weather events affecting natural and socio-economic sectors.

One of these sectors is agriculture, which is the mainstay of food security in Mozambique y Argentina. At the same time, it is highly dependent on the availability of water from the environment, i.e. from rainfall, runoff and natural flows, since in these areas there are usually no major regulatory water works (e.g. for irrigation or dams). Temporal variability in crop yields can respond to many non-climatic influences, such as civil war and conflicts or epidemics (Ben-Ari and Makowski, 2014; Eriksen and Silva, 2009; Hellmuth et al., 2007; Schauberger et al., 2018), however, the main driver is climate variability (Lobell et al., 2011b). According to FAO, (2019), in developing countries (such as Argentina and Mozambique) the agricultural sector absorbs about 80% of the direct impacts caused by droughts, while a quarter of the global economic impacts come from crop losses caused by natural hazards, especially droughts. Therefore, in regions particularly exposed and vulnerable to drought, specific drought studies are needed to help mitigate its impacts.

The study of drought consists of different components, starting with the characterisation of drought in a given area, it encompasses the monitoring, analysis, prediction, visualisation and evaluation of the impacts of one or several drought events. According to Wilhite et al., (2000), the process of studying droughts in a given area (country or river basin) should start with a spatial division into regions according to drought characteristics, then a drought monitoring system should be developed to create early warnings of emerging drought conditions, and finally drought forecasting, which can be based on predictive models or known relationships between drought indices and indicators of drought triggers (e.g., teleconnections). In this last step, it is necessary to understand the climatic drivers that trigger drought events in the region and use this climate teleconnection information as a forecasting tool.

Several drought indices have been developed for the above purposes (World Meteorological Organization and Global Water Partnership, 2016). The indices are computed using as inputs climatic and hydrological variables. For instance, precipitation (McKee et al., 1993), soil moisture (Hao and AghaKouchak, 2013), vegetation condition (Kogan, 1995) and runoff (Mo, 2008). Some multivariate indices have been created as well, showing good results. For example, using precipitation and evapotranspiration (Vicente-Serrano, 2006a), vegetation condition and ground temperature (Kogan, 1995), and reservoir storage and water demands (Mehran et al., 2015). Drought indices are very powerful tools for identifying water deficits throughout time and monitoring the characteristics of past and present events. They also make it possible to monitor the characteristics of past and present drought events and to relate them to negative effects, such as reduced water volume in a reservoir or loss of vegetation and crops.

For the proper application of drought indices, time series of hydrometeorological data for each component of the hydrological cycle are required. The inputs used in the calculation of the drought indices require a high-level quality. According to the World Meteorological Organization, (2008), in situ measurements of the input variables are required with a length of at least 30 years and with a spatial distribution of at least one every 250km^2. These conditions are particularly difficult to meet in developing countries or certain regions of the world due to socio-economic, environmental and other conditions (Easterling, 2013). The main reason is often the high cost of set-up and maintenance of the instrumentation required to measure the variables needed to calculate drought indices over large territories, so that in many areas of the world the availability of data is extremely limited.

To obtain hydrological and climate data such as precipitation, soil moisture, streamflow, etc., several institutions have developed quasi-global databases which are shared openly and free of charge. These products are based on gauged (measured) point data that have been interpolated by statistical techniques; on results from global climate and hydrological models; and, on remotely sensed information from satellite radars. These global data sets offer products that meet the required quality for drought studies and have proved to be a valuable source of timely, spatially continuous data with improved information on climatic and hydrological monitoring (Funk et al., 2014; Harris et al., 2020). This opens up the possibility of widespread use of these data in drought studies.

The analyses and research conducted in this thesis are intended to address the knowledge gaps and the specific management needs in drought-prone areas. This study aimed to define methodologies for monitoring drought events and their impacts on crops in data-scarce regions. The methodologies sought to improve management systems and contingency plans by filling the gaps that exist in the different components of a proper drought study in some regions of the world. The methods were based on drought indices calculated with variables from alternative global databases (free and open access). The case studies were two large territories: Mozambique in the southern cone of Africa and Argentina in southern South America.

The following sub-sections of this introductory chapter present the justification and motivations for this thesis; the main characteristics of the case studies; the main and specific objectives; and an outline of the thesis including the main results obtained, the principal conclusions, and future work.

1.1 Motivation and justification of the thesis

Adequate drought management is conditioned by several prerequisites. These requirements can be framed in terms of the availability of meteorological and hydrological metrics, the understanding of the relationships between these two sets of parameters, and methodologies and approaches for their use and application (Mishra and Singh, 2010). In an area of interest such as a country or a river basin, these requirements will define the scope and effectiveness of drought preparedness and mitigation plans (among other aspects of management). In many regions of the world, these requirements are a constraint due to various socio-economic reasons that preclude optimal drought management (Easterling, 2013). Consequently, this fact results in significant impacts on socio-economic systems, especially in agriculture. The impact on this sector is directly related to food security and the quality of life of a large part of the population in developing territories. According to Cumani and Rojas, (2016), one third of the population in these regions is directly dependent on farming. This thesis is a contribution to the management of droughts and their impacts, specifically in agriculture.

Although global drought studies have increased in recent decades, leading to significant advances in understanding of the phenomena (Hagenlocher et al., 2019b), certain aspects are still unresolved, especially in developing regions of the world. In this framework, the thesis was motivated by the need to define feasible and reproducible methodologies to improve the understanding of droughts and their impacts on agriculture. Specifically, the thesis aimed to develop and validate methodologies to characterise the development of droughts in space and time, define drought indices that better describe

the local drought conditions, explain the impacts on agriculture and find the large-scale climate drivers of drought events. In order to be adaptable worldwide, these methodologies were based on the use of global, open access and free global hydrometeorological databases. As case studies, the work of this thesis was focused on Mozambique and Argentina.

Globally, Mozambique is involved in many development programmes led by the Food and Agriculture Organization of the United Nations (FAO) (Midgley et al., 2012) and the World Food Programme (WFP), among others, aimed at implementing climate change adaptation strategies to improve resilience to droughts and especially to the sustainability of agriculture. The University of A Coruña, working together with the Spanish Regional Government of Galicia (Xunta de Galicia), has carried out several collaboration and development projects with Mozambique. This thesis is related to the following projects:

- "Implementation of drought indices as monitoring and early warning tools for the development of drought preparedness and management plans in Northern Mozambique (SECARA Project)". Funded by Spanish Regional Government of Galicia (Xunta de Galicia) and Galician Cooperation (Cooperación Galega). Research for Development 2017.

  Original project title (in Spanish): "Proyecto Implementación de índices de sequías como herramientas de vigilancia y alerta temprana para el desarrollo de planes de preparación y gestión de sequías en el Norte de Mozambique (Proyecto SECARA)".

- "Strengthening technical and operational capacities for the improvement of water resources management in Mozambique (Aqua-Moz): Phase 1 and 2". Funded by Spanish Regional Government of Galicia (Xunta de Galicia) and Galician Cooperation (Cooperación Galega). Research for development 2019.

  Original project title (in Spanish): "Fortalecimiento de capacidades técnicas y operativas para la mejora de la gestión de los recursos hídricos en Mozambique (Aqua-Moz)": Fase 1 y 2.

Within the framework of the efforts to combat climate change, these projects aim to develop methodologies, at the technical and operational levels, for the improvement of water resources management in Mozambique. The actions identified focused on capacity building (in the administration and other relevant stakeholders) and the implementation of technical and management tools. The approach of the projects emphasises the replicability of the actions throughout the country with the objective that local beneficiaries will progressively become the protagonists of knowledge transfer to all Mozambican administrations.

Mozambique, at the scientific level, is a very complex scenario for conducting specific drought studies. The country does not have an operational measurement network that meets the optimal criteria established by Easterling, (2013). Furthermore, despite being a country highly dependent on rain-fed agricultural labour, data on agricultural yields are limited. In general, previous research on the impacts of drought is scarce, so the methodologies proposed during the development of this thesis had to be tested and validated in other settings where better-quality data is available, but where specific studies of similar droughts are required in order to contribute to water management. For this purpose, Argentina was chosen as a suitable complementary case study for this thesis.

In contrast to Mozambique, Argentina has a better and more reliable database of agricultural data. Annual crop yields at departmental level (second order administrative subdivision of the provinces) are documented by the Ministry of Agriculture, Livestock and Fisheries of Argentina (Ministerio de Agricultura, Ganadería y Pesca de Argentina). This database includes the sowed area, harvested area and total production of 30 different crops from 1961 to date. Nevertheless, Argentina, similar to Mozambique, also does not meet the optimal requirements of established climate and hydrological databases.

However, Argentina has made significant progress in terms of understanding droughts. Specific drought studies have been carried out in different areas of the country, many of them with a view to minimising the impact of the phenomenon on agricultural work. However, there are still many gaps, which coincide with those in Mozambique (explained in the following sections).

For these reasons, this country has been chosen as a second case study for the development of this thesis in order to compare and validate methodologies applied in Mozambique, while contributing and creating additional knowledge for water and drought managers in Argentina.

The following is a brief description of the characteristics of these two countries. The focus is on the description of the hydrological and climatic characteristics; the relevance of their agricultural activity; the drought risks they are prone to; and the specific information gaps they need to address to monitor and manage droughts in order to minimise the impact of droughts on crops.

1.2 Case studies

1.2.1 Mozambique

Mozambique (Fig. 1.2) is one of the poorest countries worldwide. It is located in southern Africa in one of the most prone to drought areas in the world (Eriksen and Silva, 2009; 2014: Climate Change IPCC, 2014; Osbahr et al., 2008; Patt and Schröter, 2008). The climate is tropical, with a hot and rainy summer season from October to March, and a cool and dry winter season from April to September (Midgley et al., 2012). Maximum and minimum averaged temperature varies from 19°C to 30°C, respectively. The national annual average precipitation is 1032 mm, which 75% of it falls during summertime. According to the International disaster Database (EM-DAT, 2021), Mozambique has suffered various drought events affecting more than 20 million people. The most important events were registered in 1982-1983, 1991-1993 and 2016.

The country's total population was more than 30 million inhabitants in December 2019 according to the National Institute of Statistics (“Instituto Nacional de Estatística”, www.ine.gov.mz), which more than 60% of them live in rural areas and have their main economic activity related to agriculture (Ministério da Agricultura e Segurança Alimentar, 2015).

According to Hellmuth et al., (2007) extreme hydrological events such as drought will be more frequent and more damaging in Mozambique due to climate change. Average rainfall trends generally indicate a slight drying, particularly in summer, while Global Circulation Models (GCMs) suggest that temperature will warm by between 3.1 °C and 3.4 °C, with warming of up to 4.8 °C possible towards the end of the twenty-first century (Midgley et al., 2012). The country is not properly equipped with the hydraulic structures it requires to regulate rainwater, nor does it have adequate national mitigation plans for this hazard in current and future conditions, so it is at very high risk of impacts such as food shortages and associated socio-economic issues. Indeed, historically Mozambican farmers have experienced recurrent losses due to water scarcity and climatic variability, with the most vulnerable people in rural areas particularly exposed to this climatic phenomenon. (Eriksen and Silva, 2009).

To the authors' knowledge, there is a lack of comprehensive national drought studies that could serve as a basis for the development of national drought monitoring and forecasting systems. Some studies have explored small parts of the country, focusing on the Limpopo basin (Dutra et al., 2013; Seibert et al., 2017; Trambauer et al., 2015). Therefore, there is a clear need to develop specific drought studies in this region, which can ultimately contribute to a better drought management.

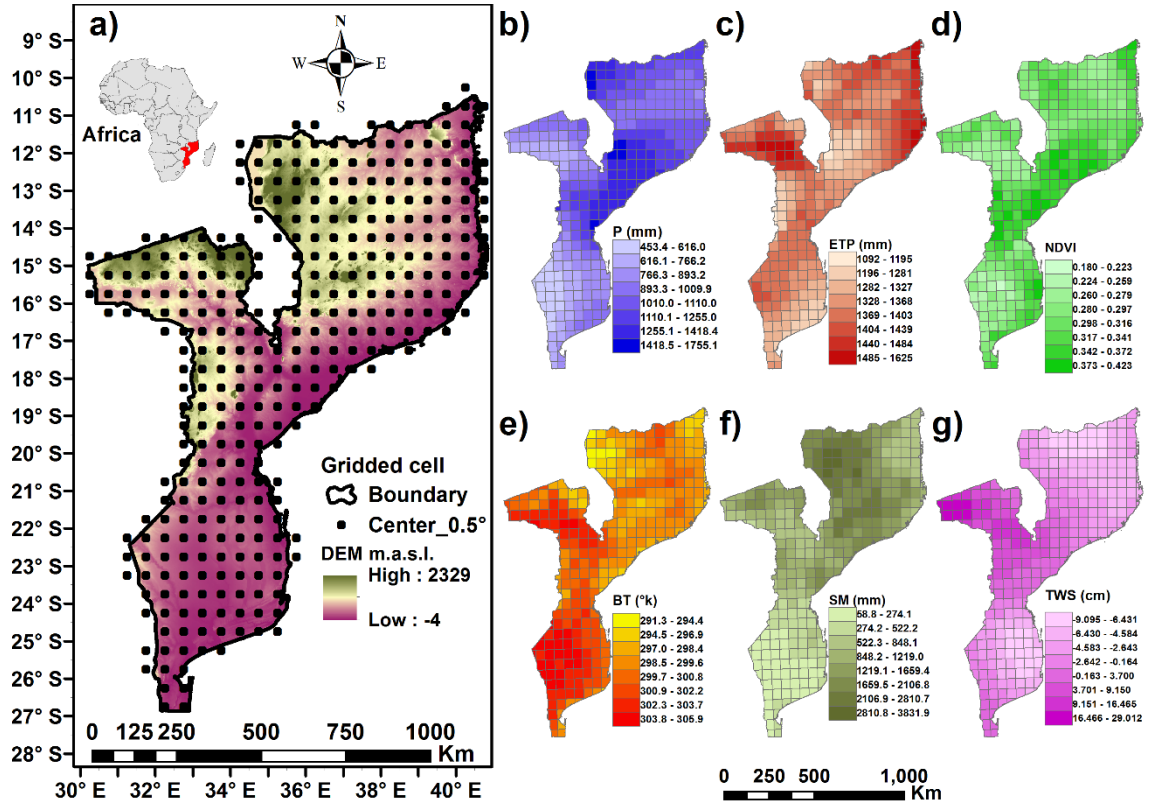


Fig. 1.2 (a) Location of Mozambique in Africa and its topography. Black dots illustrate the CRU grid points (0.5◦ × 0.5◦). Spatial distribution of annual mean values of: (b) precipitation (1973–2017), (c) potential evapotranspiration (ETP, 1973–2017), (d) Normalized Difference Vegetation Index (NDVI, 1983–2017), (e) Brightness Temperature (BT, 1983–2017), (f) Soil moisture (1973–2017), and (g) Terrestrial Water Storage (TWS, 2002–2017), across the country.

1.2.2 Argentina

Argentina covers 2791810 km^2 and is divided into five main administrative regions according to the National Institute of Statistics and Census of the Argentine Republic (INDEC: www.indec.gob.ar), 24 provinces (including the Autonomous City of Buenos Aires as a province), and 525 departments (Fig. 1.3). Due to its large surface area, the country has a wide climatic diversity, ranging from arid (south and centre-north) to fully humid (northeast) (Beck et al., 2018; Kottek et al., 2006). Argentina has the highest per capita crop production in the world (FAO, 2019) and at the same time has significant annual and interannual climate variability (Barros and Silvestri, 2002). These facts mean that the probability of crop yields in Argentina being reduced due to drought ranges from as high as 80% in some cases (Leng and Hall, 2019).

Argentina is one of the world's leading producers of cereals and oilseeds (FAO, 2017), which are largely cultivated in the Argentine Pampas and over half of which are rainfed (Cherlet et al., 2018). Average annual precipitation at the departmental scale varies between 70 and 1880 mm per year and, the average annual temperature ranges between 2 and 23 °C. Both precipitation and temperature increase from east to west and from south to north. Spring and summer are the wettest seasons, while autumn and winter are the coldest and driest. The value of Argentina's cereal production was USD 10.2 billion in 2013, representing 8.3% of its GDP (FAO, 2017).

The country is vulnerable to several natural phenomena (earthquakes, floods, etc.); however, droughts represent the greatest risk of agricultural losses (Cherlet et al., 2018). The triggers of drought events in Argentina have only been studied in specific areas (provinces or river basins), but not at the national or

regional level of agricultural production. Therefore, studies are needed to help decision-makers monitor the impacts of droughts on agriculture and minimise the associated risk.

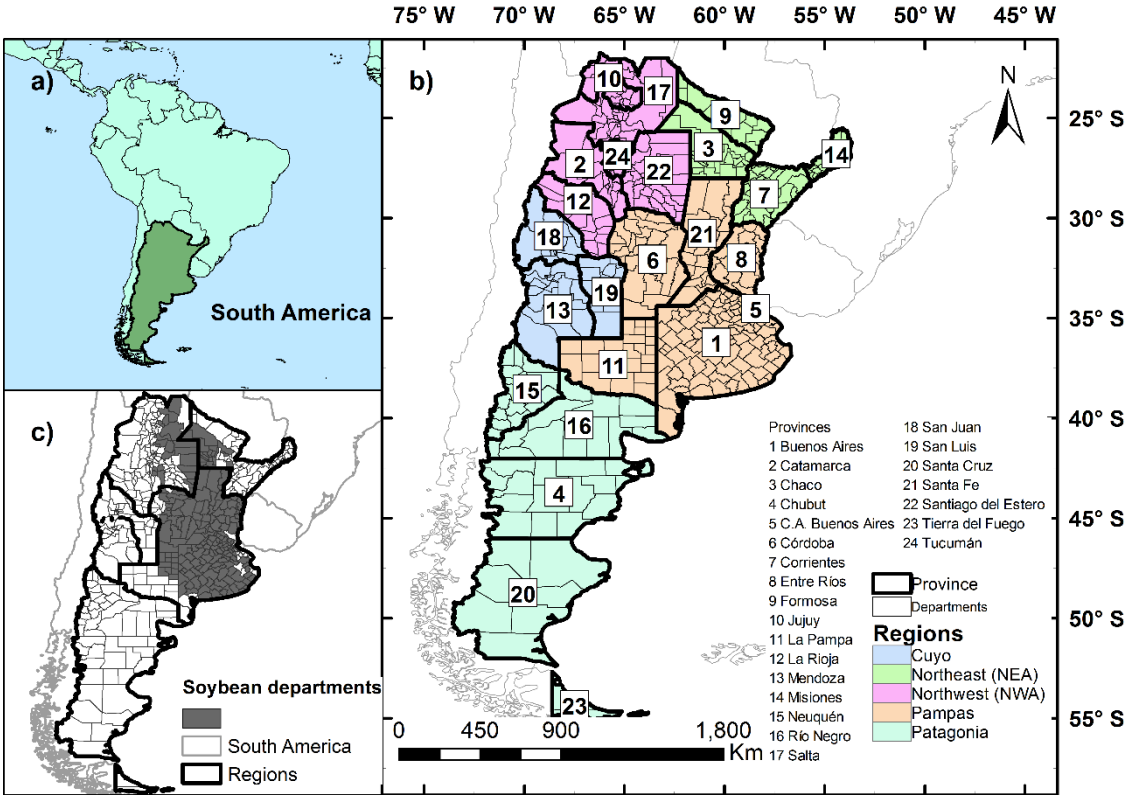


Fig. 1.3 Location of a) Argentina; b) regional, providence and department divisions; and c) departments with soybean production.

2. Objectives of the thesis

The main objective of the thesis was to investigate and develop methodologies, at a technical and procedural level, for monitoring drought and evaluating its impacts on agriculture in data-scarce areas. The methodologies consider the use of alternative hydrometeorological data sources that are freely available, so that they could be replicated and implemented in any region of the world at different spatial scales.

This general objective can be broken down into more specific objectives which are described here in terms of the main methodologies required in the study areas for appropriate drought management.

Understand the spatio-temporal characteristics of droughts and their triggers.

i. Characterise spatio-temporal conditions and drought trends.
ii. Identify the influence of large-scale climatic factors on drought events.

Validate the tools for monitoring droughts and their impacts with a focus on agriculture.

iii. Use global gauging-based and satellite-derived datasets to calculate several widely known drought indices. Compare and validate with historical drought records.
iv. Analyse the temporal variability between various drought indices to identify those that are strongly correlated and potentially provide redundant information.
v. Develop statistical crop models based on time series to explain and subsequently forecast national yields using drought indices as predictors.

vi. Propose a new index to improve the assessment and monitoring of agricultural droughts.

Transfer of knowledge to local beneficiaries and all those involved in drought management.

vii. Satisfy the need to obtain quickly and efficiently the data necessary for drought management and monitoring in any region.
viii. Create a tested tool for drought management.
ix. To transfer knowledge to Mozambican administrations and universities.

3. Thesis outline

The dissertation of this thesis is presented in the modality of a compendium of scientific papers. The results of this work have been published in four international journals, three of which are indexed by the Journal Citation Report: Journal of Hydrology: Regional Studies (Ranked Q1 with an Impact Factor of 5.437 according to the Journal Citation Reports 2021), Agricultural Water Management (Ranked Q1 with an Impact Factor of 6.611 according to the Journal Citation Reports 2021) and Science of the Total Environment (Ranked Q1 with an Impact Factor of 10.753 according to the Journal Citation Reports 2021); and, the fourth by the Regional Online Information System for Scientific Journals of Latin America, the Caribbean, Spain and Portugal, Latindex (Revista Hidrolatinoamericana). Furthermore, part of the results presented here have been published in the proceedings of an international congress and in an open access manual (book).

Chapter 1 summarises the main contents of the thesis and gives coherence and unity to the different studies that compose it, explaining how the objectives of the project are addressed. The remaining chapters are then organised in a way that each of them can be considered as an individual study, including its own specific state of the art, methodology, results, and conclusions.

The study presented in Chapter 2 aimed to characterise drought conditions and trends in Mozambique and to assess the influence of the main climatic factors as triggers of drought events. Drought conditions were studied using the Standardised Precipitation and Evapotranspiration Index (SPEI) and the "run theory". The SPEI was selected as a hydrometeorological drought index due to its flexibility as a multi-scalar index and because of its functionality under climate change conditions. Principal component analysis technique and k-means clustering method were applied to define homogeneous drought regions. The Mann-Kendall trend test and the Rescaled Range statistical analysis were used to define temporal characteristics of drought. The cross-correlation method, a spectral analysis based on the Fast Fourier Transform and a Cross-Wavelet analysis were used to identify possible climate indicators as drought triggers. Several large-scale climate indices were used to analyse their relationships with droughts.

Chapter 3 proposes a methodology for identifying the most appropriate drought indices and data sources for monitoring droughts and their impact on crops. Mozambique is used as a case study, as it represents a challenging example due to several characteristics that have already been discussed. A total of seven standardised multi-scale drought indices at different scales (1, 3, 6 and 12 months) were obtained from global databases: The Standardized Precipitation Index (SPI), the Standardized Precipitation and Evapotranspiration Index (SPEI), the Standardized Soil Moisture Index (SSI), the Standardized Vegetation Condition Index (SVCI), the Standardized Temperature Condition Index (STCI), the Standardized Vegetation Health Index (SVHI) and the Standardized Terrestrial Water Storage (STWS). These indices were compared and evaluated both as tools in drought management and as potential predictors of annual variability of agricultural yields at the national level. A statistical model of crop yields based on time series was used to measure the explanatory power of each index.

Subsequently, in Chapter 4, the methodologies used in chapters 2 and 3 are applied and validated in the Argentinean territory. An evaluation of a set of drought indices and climate indicators (similar to those used in Chapter 2 and 3) to monitor agricultural drought in Argentina is presented. First, the link between drought indices and climate indicators was investigated at the departmental administrative level and at

different time scales. Then, the effectiveness of drought indices in explaining crop yield variability, understood as impacts of agricultural droughts, was evaluated using statistical regression models (defined in Chapter 3). Complementarily, the results of drought indices and climate indicators were compared to explain crop yield variability.

According to the results obtained in the first chapters, and following the stated objectives, in Chapter 5 new indices are proposed to improve the assessment and monitoring of agricultural droughts through the explanation of annual crop variability. These are two multivariate composite drought indices that take into account meteorological and agricultural drought conditions by combining in a probabilistic framework the Standardized Precipitation Index (SPI), through precipitation, with the Vegetation Condition Index (VCI) and the Vegetation Health Index (VHI). The methodology to validate and compare the new indices was to use the statistical crop models developed in chapters 3 and 4. This chapter is divided into two parts, the first part studying these indices in Mozambique and part 2 in Argentina.

Chapter 6 responds to the specific objectives vi, vii and ix of the thesis related to collaboration, development, and knowledge transfer. A basic introductory tool for drought management is presented. It is a manual published in Spanish and Portuguese (original versions) that describes how to obtain the necessary data quickly and efficiently for drought management and monitoring in any region. These data are products of global databases that can be obtained free and open to everyone, so that the application of these data may be used in administrative (countries, cities, etc.) or natural (river basins, nature reserves, etc.) boundaries. In addition, the manual presents different uses and applications of these data as specific drought management tools. As part of the procedure, a brief introduction to the different aspects, concepts, and requirements necessary for the use of the manual is given. The case study for this manual was Mozambique.

As part of the knowledge transfer to Mozambican administrations and universities, the manual (Chapter 6) was presented in an online course (Fig 1.14) given by the Ph.D. candidate to students of Civil Engineering, Geography and Environmental Engineering at the Universidade Eduardo Mondlane - ESUDER in Maputo (capital of Mozambique) and to officials of the Regional Water Administrations of the South (Administração Regional da Água do Sur) – ARA South of Mozambique.

Finally, this thesis includes two appendices with additional information. First, a summary of the results in Spanish is attached in Appendix A. Secondly, in Appendix B, the poster resulting from the communication at the 6th IAHR Europe Congress, 2020, Warsaw, Poland (Chapter 5 Part 1) is shown.

4. Main results and discussions

In the development of this PhD research, several methodologies required in for appropriate drought management were studied. The complete publications and results are presented in the following chapters, while this subsection summarises the main results of the thesis.

In Mozambican territory Principal Component Analysis with Varimax rotation method was used to define 3 homogeneous drought regions located in the North, South, and Centre of the country (Fig. 1.4). This regionalisation was validated with hierarchical and non-hierarchical clustering methods and with several SPEI accumulations. Monthly precipitation and potential evapotranspiration data were downloaded from the Climate Research Unit (CRU) (http://www.cru.uea.ac.uk/data) to calculate SPEI-12 for the 70-year period from January 1950 to December 2019. The CRU provides monthly climate time series with a resolution of 0.5˚ (≈ 55 km at the equator) worldwide (Harris et al., 2014). At this resolution, a total of 343 SPEI time series were calculated across Mozambique.

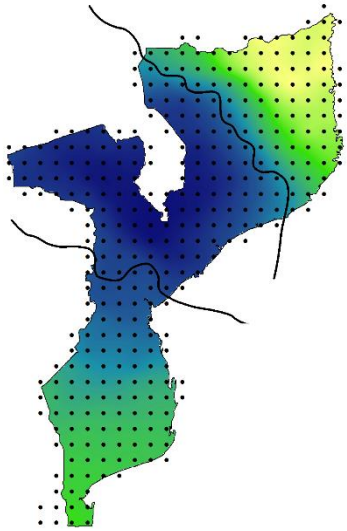


Fig. 1.4 Homogeneous drought regions defined in Mozambique. Dots are the centroids of the cells according to the CRU grid.

In the three defined regions the most important drought events were identified and coincided with the major ones at the national level, validating the use of SPEI and this regionalisation for future drought management. The Southern and Central regions having the most intense and severe drought events in the past. These results were obtained using the "run theory" method. The SPEI trends were negative, although in the Northern region they were statistically non-significant at the 5% level for the annual index, according with the Mann-Kendall trend test. In terms of persistence, the Hurst index (H) suggested that long-term negative trends will persist for the near future across the country, especially in the Southern and Central regions.

A number of large-scale climate indices were selected to analyse their correlations with the variability of SPEI time series averaged over homogeneous drought regions. The climate indices are based on fluctuations in sea level atmospheric pressure (SLP) from different locations around the globe (Darwin SLP, Tahiti SLP, SOI and NAO indices), and sea surface temperature (SST) from the Atlantic (TNA, TSA, NAT, SAT and TASI indices), the Pacific (ENSO Niño 1 +2, Niño 3, Niño 4, Niño 3.4 and PDO indices) and the Indian Ocean (SWIO, WTIO, SETIO and DMI indices). According to the cross-correlations, the time series anomalies of the climate indices were negative with Darwin, Niño 3.4, Niño 3, Niño 4, WTIO and SETIO, and positive with the SOI, showing peaks around 2-3 months before the negative peak of the SPEI. The Northern region showed persistently low correlations with the climate indices analysed, while the strongest correlations were obtained in the Southern and Central regions.

The climate indices with the best correlation (Darwin SLP and El Niño 4) and the SPEI series have a periodicity associated with high energies between 40 and 120 months (3.5 and 10 years) according to the spectral analysis using the Fast Fourier Transform technique and the Cross-Wavelet analysis (Fig. 1.5). In the Northern region, periods between 35 and 60 months (3 and 5 years) were found. These periods are consistent and similar to those reported in other studies in Africa (Oguntunde et al., 2018). Both Darwin SLP and El Niño 4 events showed strong impacts on the monthly SPEI series, especially in the Southern and Central regions, indicating that they play a relevant role in the characteristics of drought evolution in these regions (something that had already been seen with correlations).

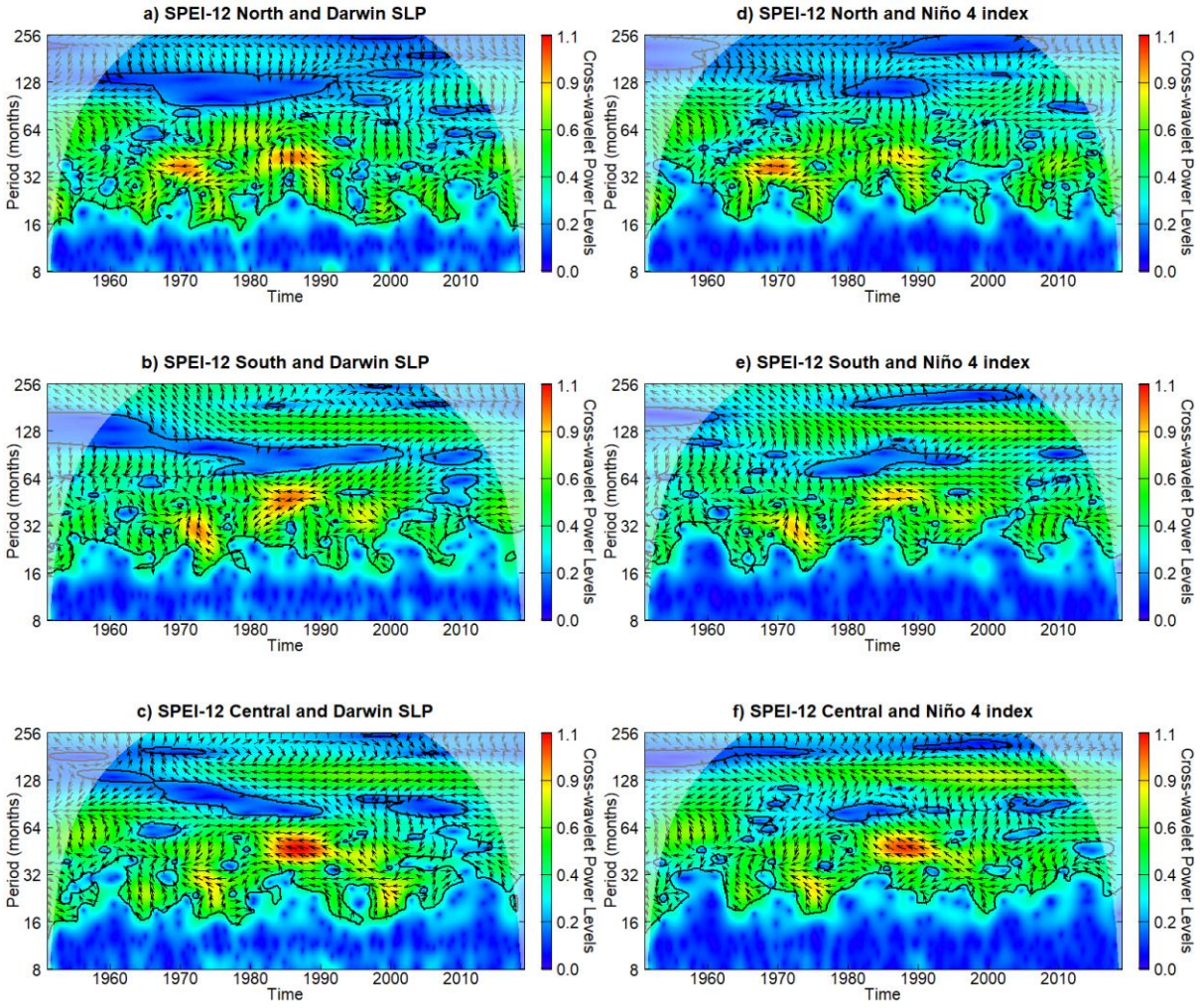


Fig. 1.5 Cross-wavelet transform between SPEI of a) North, b) South, c) Centre regions and Darwin PSL climatic index; and of d) North, e) South, f) Centre regions and the Niño 4 climatic index. The y-axis is equivalent to the periods defined with the FTT (Period = 1/Frequency); the coloured bar denotes the energy density (red plus high energy density); the 5% confidence level against red noise is shown in an outline with the thick black line; and the relative phase relationship is represented with arrows (with the anti-phase pointing to the left, the in-phase pointing to the right).

Once a framework for understanding the spatio-temporal conditions of droughts and triggers has been defined and applied, a methodology for identifying the most appropriate drought indices and data sources for monitoring droughts and their impact on crops in Mozambique was investigated. A total of seven standardised drought indicators were used: Standardized Precipitation Index (SPI), Standardized Precipitation and Evapotranspiration Index (SPEI), Standardized Soil Moisture Index (SSI), Standardized Vegetation Condition Index (SVCI), Standardized Temperature Condition Index (STCI), Standardized Vegetation Health Index (SVHI) and Standardized Terrestrial Water Storage (STWS) at different scales (1, 3, 6 and 12 months). Primary hydrological and vegetation variables were obtained from global databases based on measured and satellite data: the CRU, the Center for Satellite Applications and Research and the environmental satellites for the U.S. Oceanic and Atmospheric Administration (NOAA STAR), the TerraClimate dataset (Abatzoglou et al., 2018) and the GRACE satellites (Tapley et al., 2004).

Both time series of national average values of drought indices and percentage area affected by droughts series showed similar detection capability to the historical drought records from EMDAT (EM-DAT, 2021) and IRI (Hellmuth et al., 2007). In general, the meteorological indices (SPEI and SPI) were the

best matched to the historical records, followed by the hydrological index SSI (Fig. 1.6). These facts were resolved using performance metrics based on probability of detection (POD) or hit rate, and probability of false detection (POFD) or false alarm rate (Wilks, 2006).

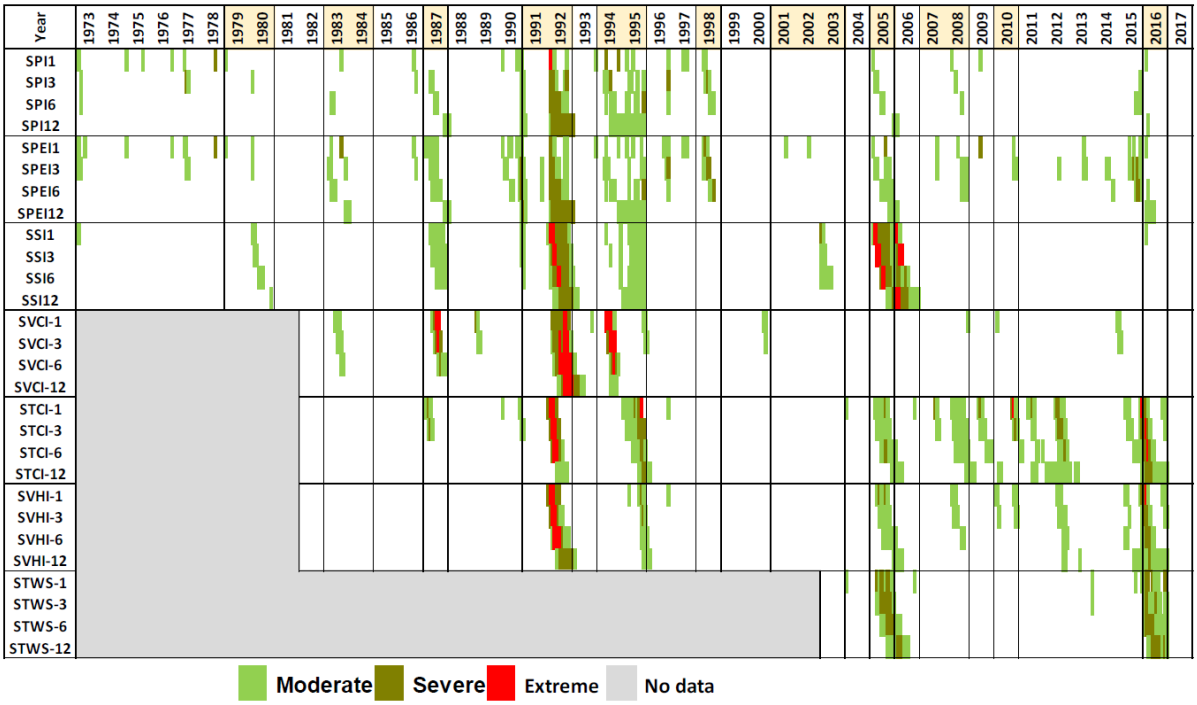


Fig. 1.6 Monthly temporal evolution of SPI, SPEI, SSI, SVCI, STCI, SVHI, and STWS at the national scale (−1, −3, −6, and −12-month aggregations). Intensity levels can be interpreted in conjunction with McKee et al., (1993). Historical drought years according to the records are highlighted in yellow.

Subsequently, the drought indices showed consistency with the physical significance of the different indices according a correlation analysis performed between them. Although there is a time lag between meteorological forcing and hydrological responses, indices based on meteorological variables with 6- and 12-month accumulations showed strong correlations with indices based on vegetation/hydrological data. The results suggest that only one of the SPI and SPEI meteorological indices with 6- and 12-month accumulations could be used alone in any further analysis. SCVI can be used but considering that it has no significant correlation with the hydroclimatic indicators, it should be used together with another index. The results observed in this section validated in the first instance the use of the SPEI in Chapter 1 as a representative drought index of hydro-meteorological conditions in Mozambique.

Then, using a national crop yield statistical model, the annual yield of 12 crops (maize, millet, sorghum, wheat, cashew nuts, cassava, potatoes, sugar cane, tea, tobacco and vegetables) was assumed to be the response of a function of independent variables, which in this case were the drought indices and the percentages of area affected by drought according to the different categories defined in McKee et al., (1993). The results indicated that the best candidate predictors were different for each crop (Fig. 1.7). This is because not all crops are equally sensitive to drought, nor do they have the same water harvesting or storage capacity. For most crops, variability was explained by two generic indicators: SPEI-3 for cereals (maize, millet and sorghum) and SSI-12 for other crops (cashew nuts, cassava, potatoes, tea, tobacco and vegetables). SPEI-6 best explained variability for wheat and SVCI-6 for sugar cane. The best predictors of agricultural yield variability were predictors incorporating spatial information as the national indices lost spatial information in their computation.

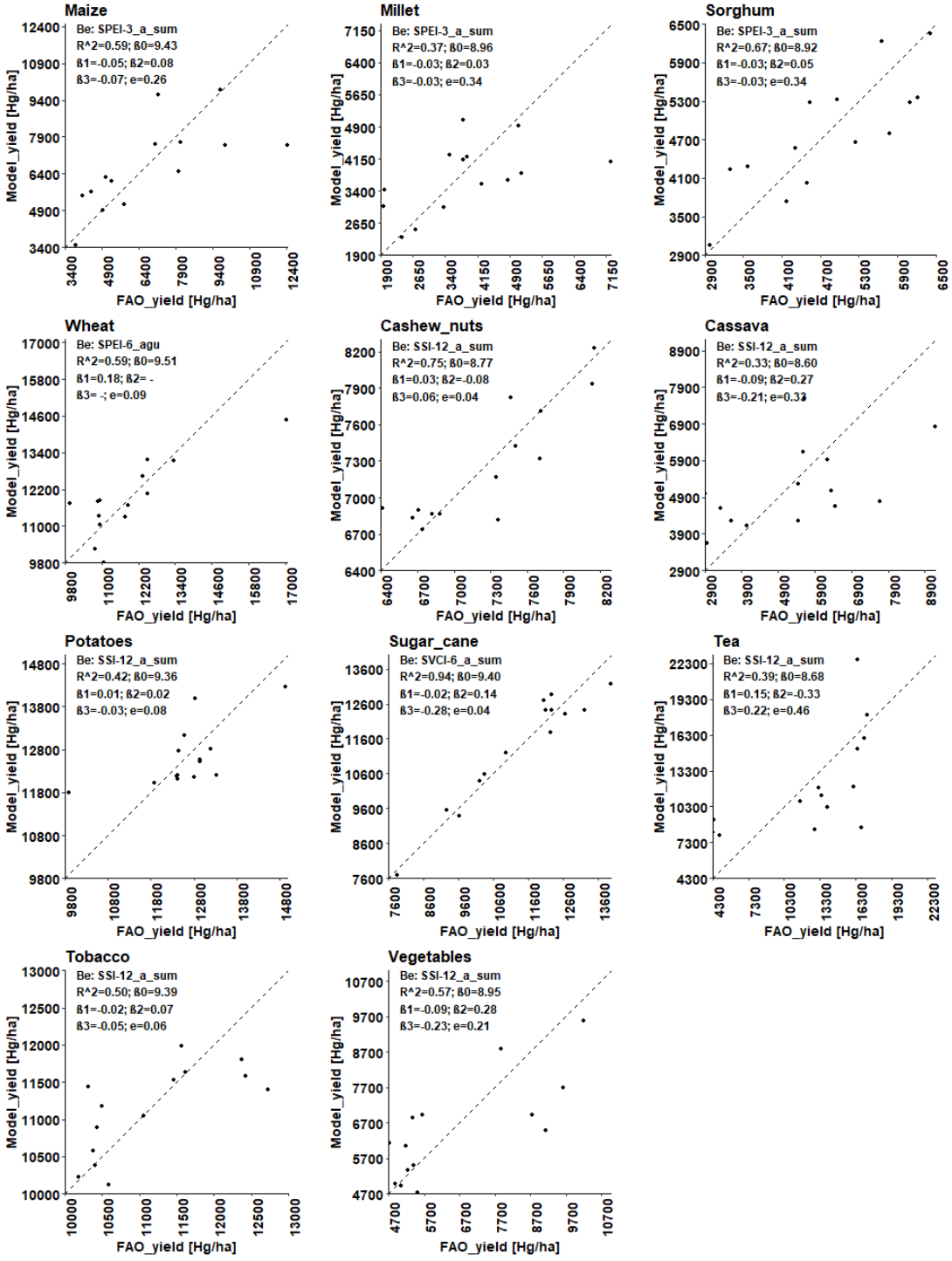


Fig. 1.7 Crop yields as measured by FAO versus crop yields as calculated using the best explanatory variable candidate (indicated as Be). Fitted parameters are also shown. The dashed line corresponds to the 1:1 line.

Based on these results presented so far, we identified some limitations in the methodology that required further work to complement this research and fill in the gaps. On the one hand, due to the large variety of climate indicators, the relationship of these with droughts (with drought indices) should be taken with caution due to the complexity of the climate system. This topic is covered in Chapters 2 and 3. On the other hand, the availability of agricultural data in Mozambique, including special disaggregation,

requires that the developed methodologies be tested in other settings. These factors motivated the research presented in Chapter 4.

First, correlations between drought indices and climate indices were tested in other scenarios at smaller spatial scales and compared as possible predictors of crop variability. And second, since it was generally observed that the introduction of spatial disaggregation improved the results, it was found necessary to validate the statistical model with a disaggregation of crop evolution by area, assuming a priori that the results would be much more accurate. The study was applied in Argentina and evaluated eight drought indices, seven already explained in previous chapters (SPI, SPEI, SSI, SVCI, SVCI, STCI, SVHI, STWS) and additionally an eighth, the Standardised Palmer Drought Severity Index (SPDSI) (Ma et al., 2014), all computed at 3-, 6- and 12-month scales. They were related to 19 climate indices (similar to those discussed above in Chapter 2) and their performance was compared to explain variability in soybean production. Soybeans were chosen as a representative rainfed crop of the impacts of the drought on agriculture.

Pearson's correlations (r) between climate indices and drought indices aggregated 12 months at departmental level across the country were analysed. A seasonal analysis (summer, autumn, winter, and spring) was then performed with the 6-month aggregated drought indices and a selected number of climate indices selected in the previous step. The seasonal analysis was carried out for the whole country and for all soybean-producing departments, prioritising the summer analysis because it is the soybean sowing and growing period. Climate indices located in the equatorial Pacific Ocean such as Tahiti, SOI, El Niño 3.4 (Fig. 1.8) and El Niño 4 showed stronger correlations with drought indices in Argentina. These indices have been found to indicate the triggering of droughts in other parts of the world (Gupta and Jain, 2021; Vicente-Serrano et al., 2017), demonstrating the global importance of their variability in extreme hydrological events.

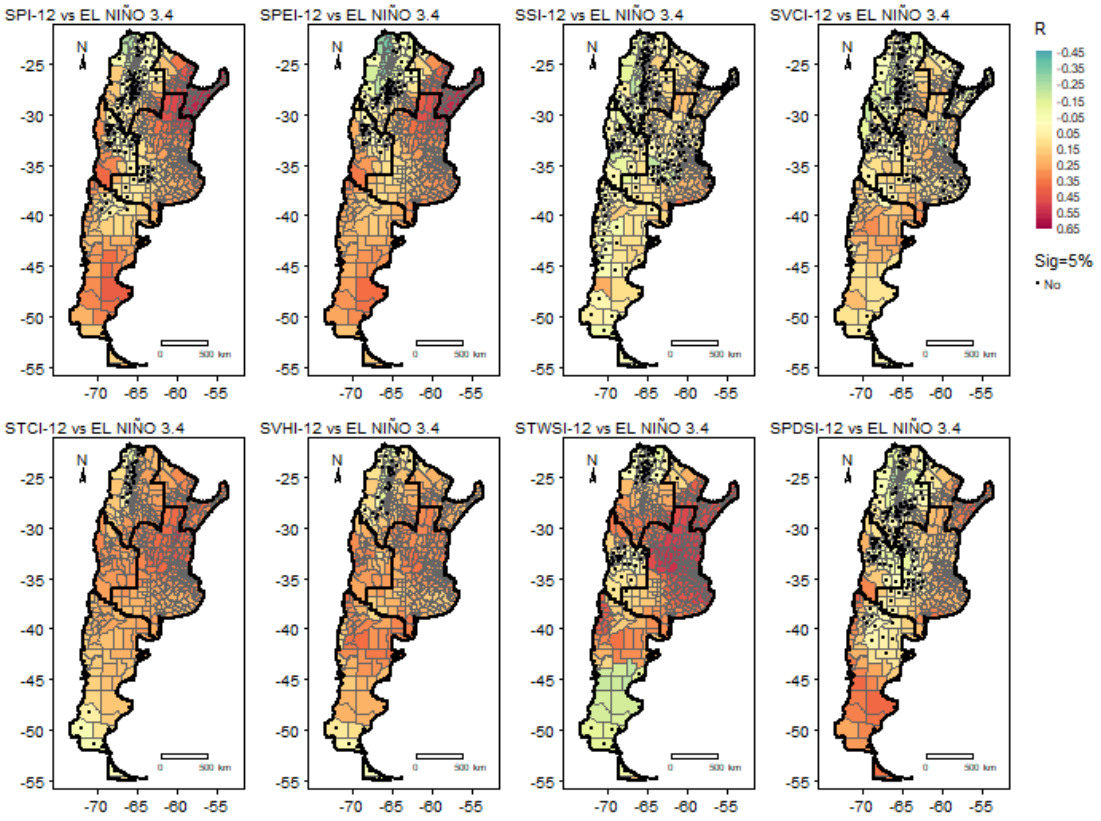


Fig. 1.8 Spatial pattern of correlations between the 8 drought indices with a 12-month aggregation period and El Niño 3.4 climate index based on 1982–2019 data (STWSI 2003–2019). Black dots indicate where correlations were not statistically significant.

Persistent negative correlations (with very few exceptions) during spring and summer and positive correlations during autumn and winter were observed between El Niño 3.4 and the drought indices. Similar patterns were observed for the other three selected indices (Tahiti, SOI, El Niño 3.4 and El Niño 4). Spatially, regions with weak correlations had the highest number of departments with non-significant correlations, and this number increased in the winter and spring seasons. For example, Fig. 1.9 illustrates the spatial patterns of correlations between SPEI-6 and El Niño 3.4. Considering only soybean-producing departments and the summer summation season, when soybean planting and growth occurs, SPEI and STWSI had the strongest correlations.

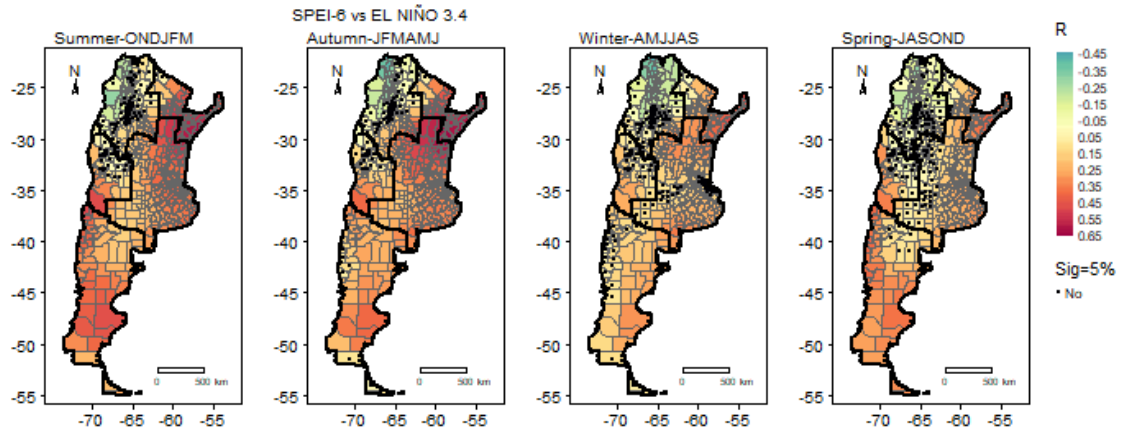


Fig. 1.9 Spatial pattern of correlations between the seasonal time series SPEI-6 and El Niño 3.4 based on 1982–2019 data.

The effectiveness of drought indices and climate indices in explaining variability in soybean yields, understood as impacts of agricultural droughts, was then evaluated using statistical regression models. The SVHI, SVCI, STCI, SPEI and SPI, which are based on both meteorological and vegetation variables, best explained the variability of soybean yields in the three models applied (time-series, panel model and cross-section models). The spatial distributions of the coefficients of determination using the time-series model with 6-month aggregate predictors are shown in Fig. 1.10. Soybean yield variability (impacts associated with agricultural droughts) responded better to the drought indices than to the meteorological indices. The SVHI and SPEI aggregated for 6 months and corresponding to the month of March (soybean growing season) were found to best explain the state of soybean production in selected regions.

Then, based on the results obtained in chapters 3 and 4, the investigations described in Chapter 5 were generated. Specifically, based on the results of the sections related to the analysis of drought indices as predictors of annual crop variability in Mozambique and Argentina. In this chapter, using precipitation and VCI/VHI, mixed indices for monitoring agricultural droughts were proposed and compared with the traditional primary indices (SPI and SVCI/SVHI) in the two territories through the statistical time-series crop yield model. The method developed to characterise agricultural droughts was a multivariate approach that depends on two individual variables (Mehran et al., 2015): 3-month cumulative precipitation and the VCI or VHI. With the formulation of Gringorten, (1963) the non-exceedance probabilities of both variables are calculated and combined using the multivariate framework explained in Yue et al., (1999). The empirical probabilities are transformed into a standardised index and the Standardised Multivariate (or Bivariate) Precipitation and Vegetation Condition, and Health index (MSPVI and MSPHI) are obtained. These indices are interpreted similarly to the original SPI (McKee et al., 1993).

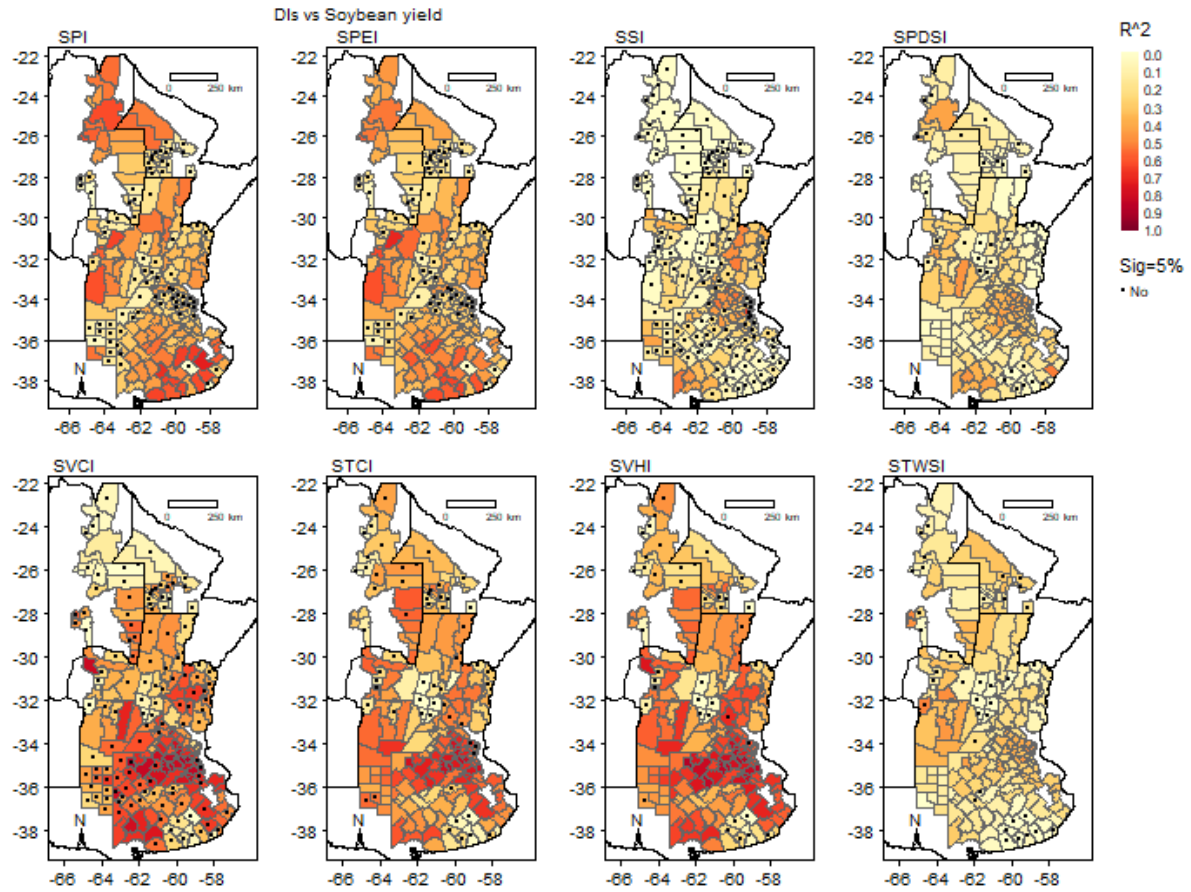


Fig. 1.10 Spatial patterns of the determination coefficient results of the time-series models between drought indices aggregated 6 months (ONDJFM) and soybean yield based on soybean-producing departments and 2004–2019 data.

The research was conducted in Mozambique and Argentina. In the first case, agricultural data for maize, soybean and wheat from FAO and precipitation from CRU were used; in the second, data for soybean and maize crops from the Ministry of Agriculture, Livestock and Fisheries of Argentina (Ministerío de Agricultura, Ganadería y Pesca de Argentina) and precipitation from TerraClimate were employed. In both cases the VCI and VHI were obtained from NOAA STAR.

The temporal patterns of the proposed indices (the case of Mozambique for the SMPVCI is shown in the Fig. 1.11) show that the MSPVCI displays critical information on the onset and recovery of drought events. The new indices detect major drought events according to the historical records of the EM-DAT International Disaster Database, where for example the period 1991-1992 was one of the most damaging drought events (in Mozambique). Within this period, the MSPVCI reported almost 50% of the Mozambican territory under extreme drought [index < -2.0 according to McKee et al., (1993)] while the SVCI reported 22% and the SPI 11% (the spatial distribution of the 1992 drought event in May is represented in Fig. 1.12). The MSPVHI showed very similar spatio-temporal patterns to the MSPVCI, indicating that the indices showed a higher sensitivity to these events and that using them to monitor droughts would mean being on the safe side (Monteleone et al., 2020).

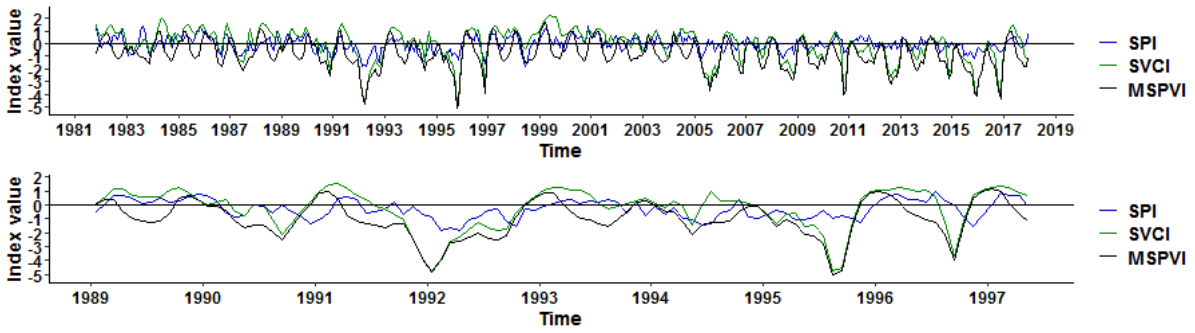


Fig. 1.11 The SPI-3, SVCI and the MSPVI series in 1982-2017 (above) and 1989-1997 (below) in Mozambique.

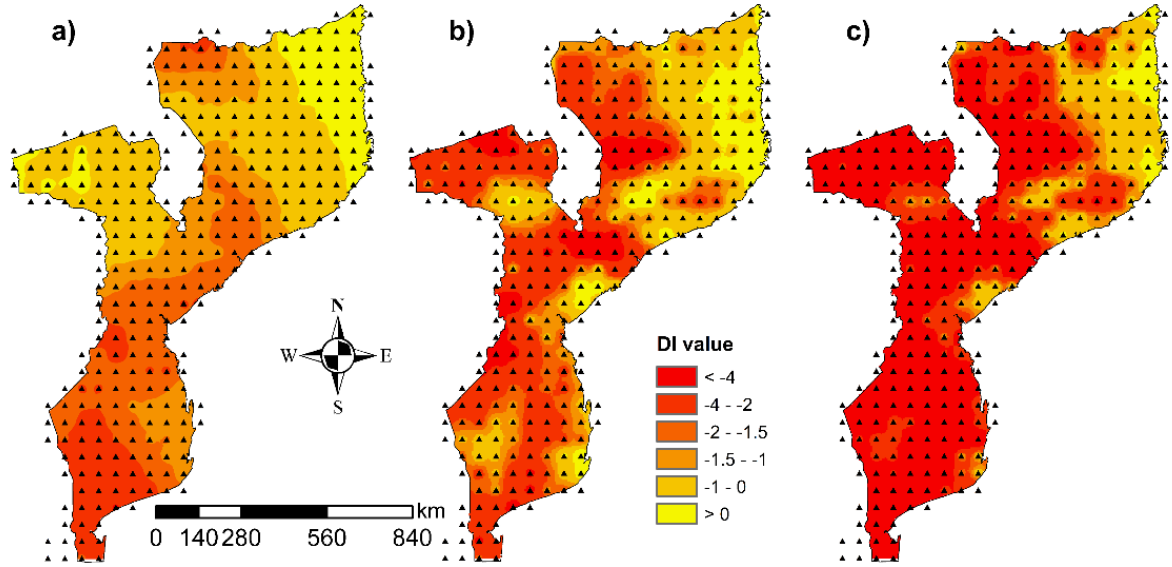


Fig. 1.12 Spatial distribution of the a) SPI-3, b) SVCI and c) MSPVCI during May/1992 drought event.

In the comparison of the proposed indices with the conventional ones to explain the annual variability of crops in both territories, the results were similar (Argentina and Mozambique). In the Argentinean case, the results of applying the statistical model are presented in Table 1.1 and plotted in Fig. 1.13. Compared to the classical SPI and SVCI/SVHI, the multivariate indices reported to explain better the variability of annual yields of maize and soybean (in Mozambique sorghum, maize, and wheat).

Table 1.1 Coefficients of determination (R^2) between soybean and maize yields and the classical drought indices 2004-2019.

	SPI	SVCI	SVHI
Soybean	0.476	0.572	0.686
Maize	0.306	0.407	0.670

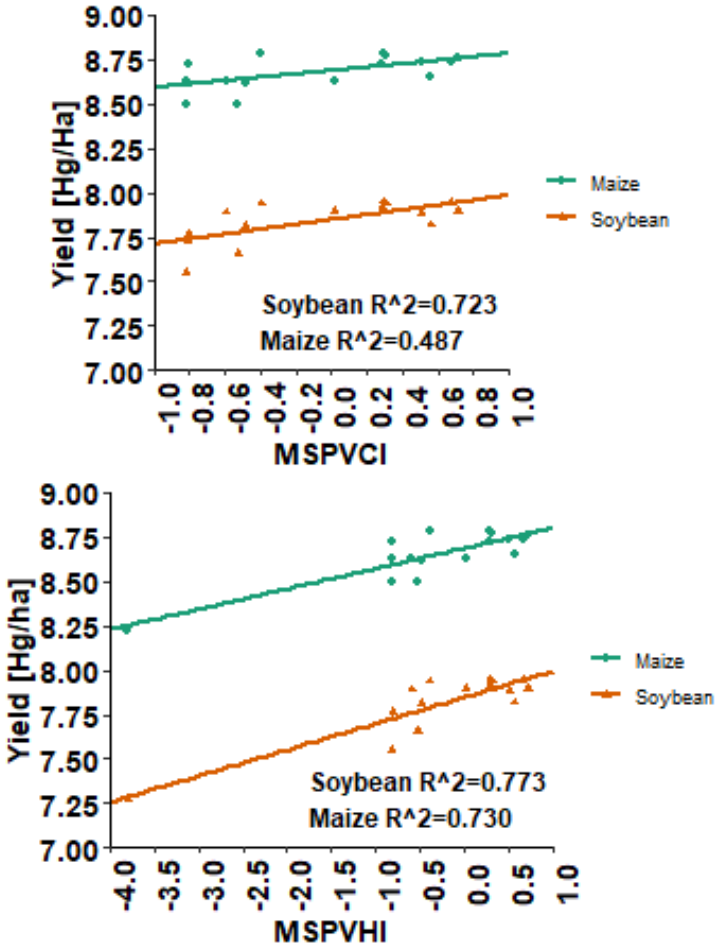


Fig. 1.13 Linear regression and coefficient of determination (R^2) between annual soybean and maize yields with MSPVCI (top) and MSPVHI (bottom) in Argentina.

Finally, to conclude this thesis, as a final material of the performed research, Chapter 6 provides a manual entitled "Manual for obtaining and applying hydrometeorological variables from global databases in drought management". This work gathers a series of concepts and techniques related to drought management that have been developed during the thesis that are applied through several scripts (programming codes) coded in the R Software (RStudio Team, 2016). These scripts divided into 3 sections show the step-by-step for i) downloading, extracting, manipulating, and saving hydrometeorological variables from global databases; ii) using and applying these variables for drought management; and iii) presenting results through graphs in different formats. Mozambique was used as an example case study.

As an additional outcome of this work, a course was conducted by the PhD candidate for students and workers in the water sector in Mozambique. The main objectives of the course were to teach concepts and tools related to drought monitoring; to obtain and manipulate hydrometeorological data in Mozambique; and to apply water balances and drought indices in Mozambique. The course was conducted online for 12 hours in November 2021 (Fig 1.14).

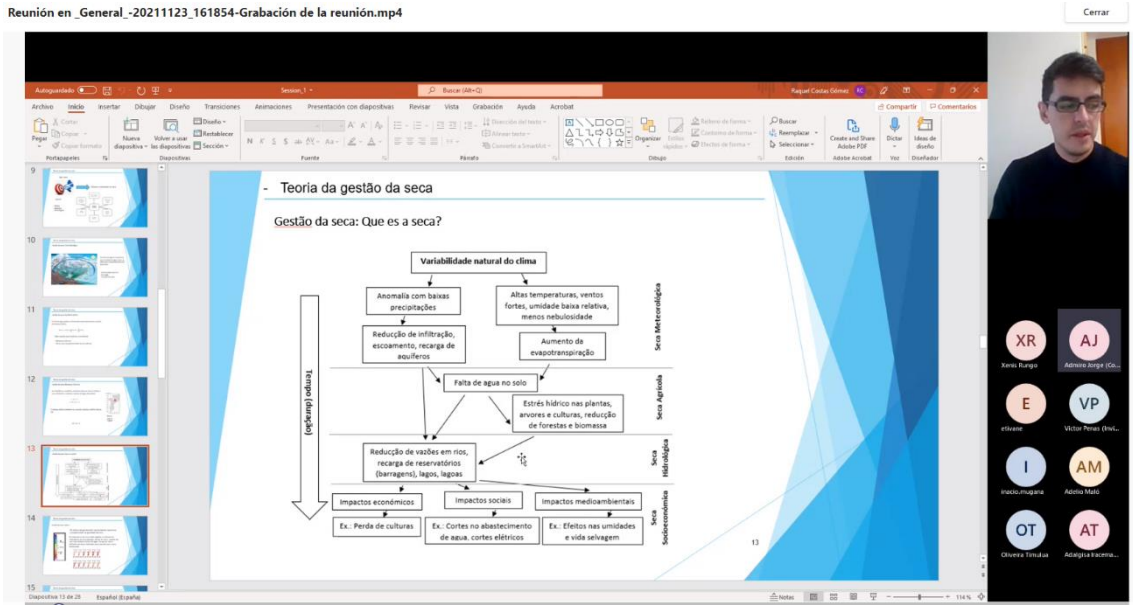


Fig. 1.14 Screenshot of the second session of the online course "Obtenção e aplicação de variáveis hidrometeorológicas de bancos de dados globais para a monitororia de secas" ("Obtaining and applying hydro-meteorological variables from global databases for drought monitoring") to students of the Universidade Eduardo Mondlane - ESUDER and public employees of the ARA South of Mozambique, taught by the PhD candidate.

5. General conclusions

In this thesis, several methodologies were developed with the aim of increasing knowledge of drought phenomena and providing solutions for water resources and drought management. The use of freely available alternative hydrometeorological data sources was considered, so that the methodologies can be applied to any country or regions of the world and at (almost) any spatial scale. The case studies were Mozambique and Argentina as they are very drought-prone territories, suffer continuously from its effects especially on agriculture, and do not have the necessary management tools to control and manage the phenomenon. The methodologies responded to the need to understand the spatio-temporal characteristics of droughts and their triggers; to the validation of tools for monitoring droughts and their impacts with a special focus on agricultural activity; and, in accordance with the need for local beneficiaries to progressively become the protagonists of knowledge transfer to all those involved in drought management. The thesis was carried out in 5 independent studies from which the following main conclusions have been drawn:

- In the first study, a five-step methodology consisting of several coherently organised methods was proposed for use in Mozambique or any other region (country or river basin) that requires a first assessment of the spatio-temporal characteristics of droughts and where local meteorological and hydrological monitoring data are extremely limited. With the monthly SPEI-12 as the drought index from 1950 to 2019 at a high resolution (0.5˚) 3 homogeneous drought regions were defined located in the north, south and centre of the country. The southern and central regions have had the most intense and severe drought events in the past. In the three regions the trends are statistically significant towards a higher incidence of droughts and the results suggest that this trend could persist in the near future. Strong correlations were found between two climate indices (El Niño 4 (ENSO) and Darwin SLP) and droughts in the southern and central regions, with a lag of 2-3 months. The periods of SPEI and these two climate indices have a similar periodicity of between 3 and 8 years and are strongly correlated in antiphase for periods between 1.4 and 10.4 years. These statements are particularly novel in Mozambique. Thus, these climate indices could be used to develop a drought forecasting system, providing sufficient lead time to establish prevention strategies.

- In the following research, the main objective was to test a methodology that could evaluate various drought indices as drought monitoring tools and test their ability to explain annual variability in crop yields. The case study was Mozambique, a particularly difficult country due to its poor hydrological monitoring system and lack of local/regional crop yield data. The proposed indicators successfully detected major drought events from 1973 to 2017 according to historical records, and accurately captured their duration and intensity. The SPEI and SSI indices had the best ability to detect historical droughts using nationally averaged time series and percentage of area affected by drought. Crop yield variability was best explained by indices associated with agricultural and hydrological variables and when using the percentage of area affected due to the spatial decomposition involved in their computation. This variability was explained using SPEI-3 for cereals (maize, millet and sorghum); SSI-12 with other crops such as cashew nuts, cassava, potatoes, tea, tobacco and vegetables; SPEI-6 for wheat and SVCI-6 for sugar cane. The proposed methodology allowed us to confirm the use of these drought indices in their different temporal accumulations as a tool for monitoring and characterising droughts and modelling annual yields of specific crops in Mozambique.

- Linking the findings observed in the first two studies, 8 drought indices and 19 climate indices were evaluated for agricultural drought monitoring in Argentina. The relationship between these two groups of indices and their ability to explain the impacts of agricultural drought was explored, using the variability of annual soybean yield as a proxy for the impacts of drought on agricultural activity. The drought indices are particularly related to climate indices located in the Pacific Ocean, including El Niño 3.4 and El Niño 4, with the drought indices that include temperature in their calculation (STCI, SPEI and SVHI) being the best correlated with them. For soybean-producing areas, SPEI was the drought index that responded best to variations in the climate indices. Correlations were positive and strong in the hot and wet season (summer), while in the cold and dry season (winter) they were negative and less strong. Soybean yield variability responded better to drought indices than to climate indices. The SVHI and SPEI aggregated for 6 months and corresponding to the month of March (soybean growing season) were found to best explain the state of soybean production in selected regions. The results provide useful tools to understand drought in various stages of the water cycle and its association with variability in soybean production in Argentina. Therefore, this research could be of interest to water managers and especially to soybean producers at the national and regional level in Argentina. However, the methodologies could be replicated anywhere in the world.

- Subsequently, following the findings of the previous studies (chapters 3 and 4), a methodology for calculating the Multivariate Standardised Precipitation and Vegetation Condition Index (MSPVCI) and the Multivariate Standardised Precipitation and Vegetation Health Index (MSPVHI) was presented and validated. These indices were calculated by combining precipitation (Standardised Precipitation Index, SPI) with the Vegetation Condition Index (VCI) and Vegetation Health Index (VHI) in a probabilistic framework. The multivariate (bivariate) indices were compared with the original indices both in their ability to detect historical drought events and to explain the annual variability of the main crops in the study territories. They were applied in Mozambique and Argentina where the results suggested similar conclusions. Both indices are shown to be more sensitive to the onset and recovery from drought events, show a more significant decreasing trend, report a higher spatial coverage of extreme droughts, and better explain the variability of annual yields compared to the conventional SPI and VCI/VHI. The proposed drought indices could be useful for water managers in the studied territories, as well as in other countries or regions, especially for farmers and drought managers.

- To complete the objectives of the thesis, specifically the one related to the transfer of knowledge through tools and procedures, the last chapter of the thesis presented an introductory manual to

obtain hydrometeorological variables in any region and apply them to drought management. The work gathers a series of concepts and techniques related to drought management studied in this thesis that are applied through several scripts (programming codes) developed in the R software. These scripts divided in 3 sections show the step-by-step to a) download, extract, manipulate and save hydrometeorological variables from global databases; b) use and apply these variables for drought management; and c) present results through graphs in different formats. A great advantage of this manual is the flexibility of its tools, which can be used in different places and allow for the creation of new scripts and the generation of new ways of using the information downloaded by users. The methodologies discussed in this document can be expanded in future work, such as obtaining data from climate change projection models, trend analysis, representation of results for regions within a country (cities or provinces) or river basins (sub-basins), among others. The manual was presented to students and public water sector workers in Mozambique.

Overall, research has sought to contribute tools for understanding and modelling droughts, their components, impacts, and triggers. The methodologies and results obtained can be replicated anywhere in the world, providing meaningful information for the scientific, technical and management community to develop, calibrate or validate existing and new formulations. In addition, they are tools that could contribute to the creation of drought mitigation and adaptation plans aimed at reducing impacts, especially in agriculture.

6. Future research

The spatio-temporal characteristics of droughts and their relationship with possible climatic predictors of the phenomenon were investigated. The work has provided insight into the spatial and temporal distribution of droughts in the case studies. The results are of great use to regional water administrations for the development of drought contingency plans. In Mozambique, simplified management regions have been defined and characterised; in addition, potential predictors of drought have been significantly identified. In Argentina, these relationships were defined at the departmental administrative level and taking into account areas of agricultural interest and time of year. Future studies related to drought prediction at the regional level could be very useful and complementary to these investigations. Although the proposed methodologies can be used anywhere in the world, given their restrictions and the large number of topics covered, certain limitations will have to be taken into account in future studies. For example, the use of SPEI-12 in Mozambique derives certain conditions, so it would be advisable to consider additional time scales in future work to obtain more information on the temporal patterns of drought in each region and to strengthen possible predictive models.

Several multi-scale drought indices were investigated and compared according to their performance in explaining the annual variability of rainfed crop yields.In Mozambique with its particular conditions and later in Argentina - where agricultural data were more widely available and spatially disaggregated - the results were very promising. However, in Mozambique, further analysis of agricultural droughts and their impacts at the regional level is needed to provide an improved basis for drought management at the local level. While in Argentina, it would be particularly interesting to explore the predictive, rather than explanatory, capacity of selected drought indices and climate indices to predict the agricultural impacts of droughts in the country. These future studies could follow the prospect of being replicated anywhere in the world (of general application) using global databases.

Linking previous research and possible future studies, two agricultural drought indices were proposed to monitor their impacts by explaining annual crop variability (MSPVCI and MSPVHI). In general, the results showed better performance according to the statistical models applied and provide an opportunity for future studies to improve the use of these drought indices, such as assessing the predictive capacity of agricultural yields. However, following the line of improving and increasing monitoring and

predictive tools, different methodologies could be chosen to associate meteorological and agricultural (including hydrological) variables in a single index or in a single process based on several indices.

Overall, future studies proposed on the basis of the results and conclusions generated in this thesis are mainly associated with drought prediction and its impacts. These studies could be complemented by considering several aspects that are a current challenge for scientists, technicians, and decision-makers in drought management. For example, climate change, the wide variety of new databases and the different institutional and governmental strengths. This wide range of possible future research could continue to generate courses and documents such as the manual presented in this thesis, so that the knowledge and tools are accessible, used and shared by the different actors involved not only at the national but also at the local level.

Chapter 2

Revealing the spatio-temporal characteristics of drought in Mozambique and their relationship with large-scale climate variability

Ronnie J. Araneda-Cabrera, María Bermúdez and Jerónimo Puertas.

Journal: Journal of Hydrology: Regional Studies

Date: 19 October 2021

DOI: https://doi.org/10.1016/j.ejrh.2021.100938

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – original draft, Visualization. María Bermudez: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Resources, Writing – review & editing, Supervision.

Abstract

Study Region: Mozambique. Study Focus: Mozambique does not currently have the necessary tools for systematic monitoring and forecasting of drought at a subnational scale. The purpose of this study was to characterize drought conditions and trends throughout the country and to evaluate the influence of major climatic drivers on drought events (period 1950-2019). Drought conditions were studied by means of the Standardized Precipitation and Evapotranspiration Index (SPEI) and run theory. The principal component analysis technique and the k-means clustering method were applied for defining homogenous drought regions. The Mann-Kendall trend test and Rescaled Range statistical analysis were used for defining the temporal characteristics of drought. The cross-correlation method, a spectral analysis based on the Fast Fourier Transform and a cross-wavelet analysis, were used to identify possible climate drivers. The results are ultimately intended to contribute to the development of a drought monitoring system in this country. New Hydrological Insights for the Region: Three homogeneous drought regions can be defined in Mozambique. The South and Centre regions showed more intense and severe drought events. In all regions, a significant trend towards a higher incidence of droughts and long-term a persistence was found. El Niño-Southern Oscillation and Darwin Sea Level Pressure anomalies were identified as significant drivers of drought variability, especially in the southern regions. These climate indices can be used as predictors in drought forecasting models.

Keywords

Drought, Standardized Precipitation and Evapotranspiration Index (SPEI), regionalization, trend, persistence, teleconnections, climatic indices.

Highlights

- A stepwise method for spatio-temporal drought assessment and prediction is proposed.
- The method is applied to Mozambique, which lacks effective drought monitoring tools.
- Three regions with distinctive drought characteristics were identified in Mozambique.
- A significant trend towards a higher incidence of droughts was found in all regions.
- The El Niño 4 SST and Darwin SLP climatic indices can be used to forecast drought.

1. Introduction

Droughts are among the most common natural phenomena worldwide, and can occur anywhere under any climate regime (Bryant et al., 2005; Sheffield and Wood, 2012). However, they constitute one of the least understood natural hazards due to their complexity and difficulties in quantification (Hagenlocher et al., 2019a). The effects of droughts can be felt in chains of energy production, food, water supply, etc., yet are generally detected when the consequences of the phenomenon are difficult to mitigate, and thus droughts constitute the disaster that causes the greatest socioeconomic losses worldwide (WMO, 2006). In addition, the effects of droughts are expected to worsen in the coming decades as a result of climate change. Changes in spatial-temporal patterns of precipitation and extreme temperatures are likely to make droughts more recurrent (Mishra and Singh, 2011). As stated by the IPCC (2014), there are trends of increasing intensities and frequencies of droughts around the world, with arid and semi-arid areas possibly being the most affected. This might invalidate traditional methods for examining the impact of environmental factors on drought, based as they are on the assumption of stationarity (Jehanzaib et al., 2020). Specific spatiotemporal-drought assessments at regional or local levels are needed for reliable decision-making in the context of adaptation planning to future climate conditions (see, e.g., Jehanzaib and Kim, 2020; Kim and Jehanzaib, 2020; and the references therein).

According to Wilhite et al. (2000), one of the first steps for drought assessment and management in a given zone (country or river basin) should be a spatial division into regions according to drought characteristics. The techniques of hierarchical and non-hierarchical clustering (Santos et al., 2010; Vicente-Serrano, 2006b) and Principal Component Analysis (Agutu et al., 2017; Lovino et al., 2014; Vicente-Serrano, 2006a) can be applied directly to drought indicators for identifying homogenous drought regions. The next steps should be aimed at developing a drought monitoring system to create early warnings of emerging drought conditions. The temporal variability of droughts in the area under analysis must be studied to accomplish this objective. The historical drought variability is usually analysed by means of a technique such as run theory (Yevjevich, 1969), accompanied by trend and persistence tests of the drought characteristics (Ayantobo et al., 2017; Huang et al., 2016; Zambreski et al., 2018). The subsequent step is then to forecast droughts. Thus, it is necessary to understand the climate drivers that trigger drought events in the region and to use this teleconnection information as a forecasting tool. To achieve this goal, methods such as cross-correlation (Araneda-Cabrera et al., 2021a; Hair et al., 1998) and various spectral analysis applied to climate indices and drought indicators have become popular as a means of identifying appropriate drought predictors (El Kenawy et al., 2016; Espinosa et al., 2019; Fleming et al., 2002; Zeleke et al., 2017).

A large number of drought indicators serve as a basis for these analyses (Svodova et al., 2016). The Palmer Drought Severity Index (PDSI) (Palmer, 1965) is widely used due to its versatility and effectiveness (Alley, 1984; Nam et al., 2015; Quiring and Papakryiakou, 2003). The Standardized Precipitation Index (SPI) (McKee et al., 1993), currently recommended by the World Meteorological Organization, uses rainfall series to define drought periods and has also been used widely in studies around the world (Ayantobo et al., 2017; Stagge et al., 2017). One of its main advantages, compared to the PDSI, is that it can be computed for multiple time scales, which allows the assessment of water availability according to the process under consideration (e.g., 3-month SPI provides a seasonal estimation of precipitation, whereas 12-month SPI reflects long-term precipitation patterns) (Guttman, 1999, 1998). More recently, the Standardized Precipitation and Evaporation Index (SPEI) was introduced by Vicente-Serrano et al. (2010). Its versatility is similar to that of the SPI and it has the advantage of considering both precipitation and evapotranspiration. Vicente-Serrano et al. (2010) found that SPEI had a better performance than SPI and PDSI under global warming scenarios since it could reflect the increase in drought severity associated with higher water demand due to evapotranspiration. However, the need for long-term and high-quality input data is often a problem for the application of these indicators, especially in poorly monitored regions (Easterling, 2013). In recent years, global databases such as Climate Research Unit (CRU) (Harris et al., 2014), TerraClimate (Abatzoglou et al., 2018) and the Global Precipitation Climatology Centre (GPCC) (Rudolf et al., 2011) have emerged as

alternative data sources, showing good performance in drought studies (Araneda-Cabrera et al., 2021b, 2020; Lovino et al., 2014).

In this study, we evaluate the spatiotemporal distribution of drought in Mozambique, and explore its relationships with large-scale climate variability. This country is one of the poorest in the world, highly dependent on rain-fed agriculture, and very prone to droughts. It has very little water infrastructure and a lack of monitoring systems, so its resilience to extreme hydrological events is very low (Osbahr et al., 2008). To the best of the authors' knowledge, there is a lack of comprehensive drought studies at the national level that might provide the basis for the development of national drought monitoring and forecasting systems. The possibility of drought forecasting has only been explored in small parts of the country, with a focus on the Limpopo Basin (Dutra et al., 2013; Seibert et al., 2017; Trambauer et al., 2015, 2014). Climate variability has also been studied in specific regions as part of vulnerability and adaptation assessments (Eriksen and Silva, 2009; Macarringue et al., 2017; Osbahr et al., 2008; Uele et al., 2017). Similarly, the relationship between droughts and climate indices has not been widely analysed in the country. Manhique et al. (2011) note that El Niño-Southern Oscillation (ENSO) appears to play a significant role in the inter-annual frequency of the main summer rainfall over Mozambique. However, other climate indices, such as Darwin sea level pressure, have been shown to influence the climate in neighbouring countries (Manatsa et al., 2008a, 2008b).

The main objective of the current study is to characterize drought conditions and trends over Mozambique between 1950 to 2019, and to identify the influence of large-scale climatic drivers on drought events. We follow a five-step methodology (described in section 2) of general applicability (i.e., one that could be used in any other country or region worldwide) comprising drought regionalization, characterization, trend analysis, long-term dependence, and cross-dependence with climatic factors. Given that Mozambique is heavily reliant on agriculture, the analysis focuses on persistent drought, which can affect agricultural production and food security. In this way, the results are intended to be of practical value to water managers and users. The ultimate aim is to support drought management planning with tools that enable better monitoring and prediction of risk at the regional scale.

2. Materials and methodology

The methodology developed in this study involves the following steps: (1) the calculation of the SPEI drought index with a 12-month time scale (SPEI-12) for the period 1950-2019; (2) the application of Principal Component Analysis (PCA) analysis and the k-means clustering method to define homogeneous drought regions that follow the time patterns of the drought index series; (3) the characterization of the drought events according to run theory; (4) the application of the modified Mann-Kendall (MMK) trend test method and the Rescaled Range (R/S) analysis to determine the temporal variability of droughts; and (5) the exploration of the relationships between the SPEI time series and several large-scale climate indices to find appropriate drought predictors. The cross-correlation method, the Fast Fourier Transform (FFT) and cross-wavelet analysis were used for this latter purpose. The methods employed at each step were complementary and not exclusive. Fig. 2.1 summarizes the steps in a methodological flow-chart.

The entire methodology was developed and computed using the R Software (RStudio Team, 2016). Specifically, we used the R package "SPEI" (Begueria and Vicente-Serrano, 2017), "stat" (Bolar, 2019), "lmomRFA" , "trend" (Pohlert, 2020) , "pracma" (Borchers, 2019), "tseries" (Trapletti et al., 2020), "stats" (developed by R Core Team and contributors worldwide) and "biwavelet" (Gouhier et al., 2016).

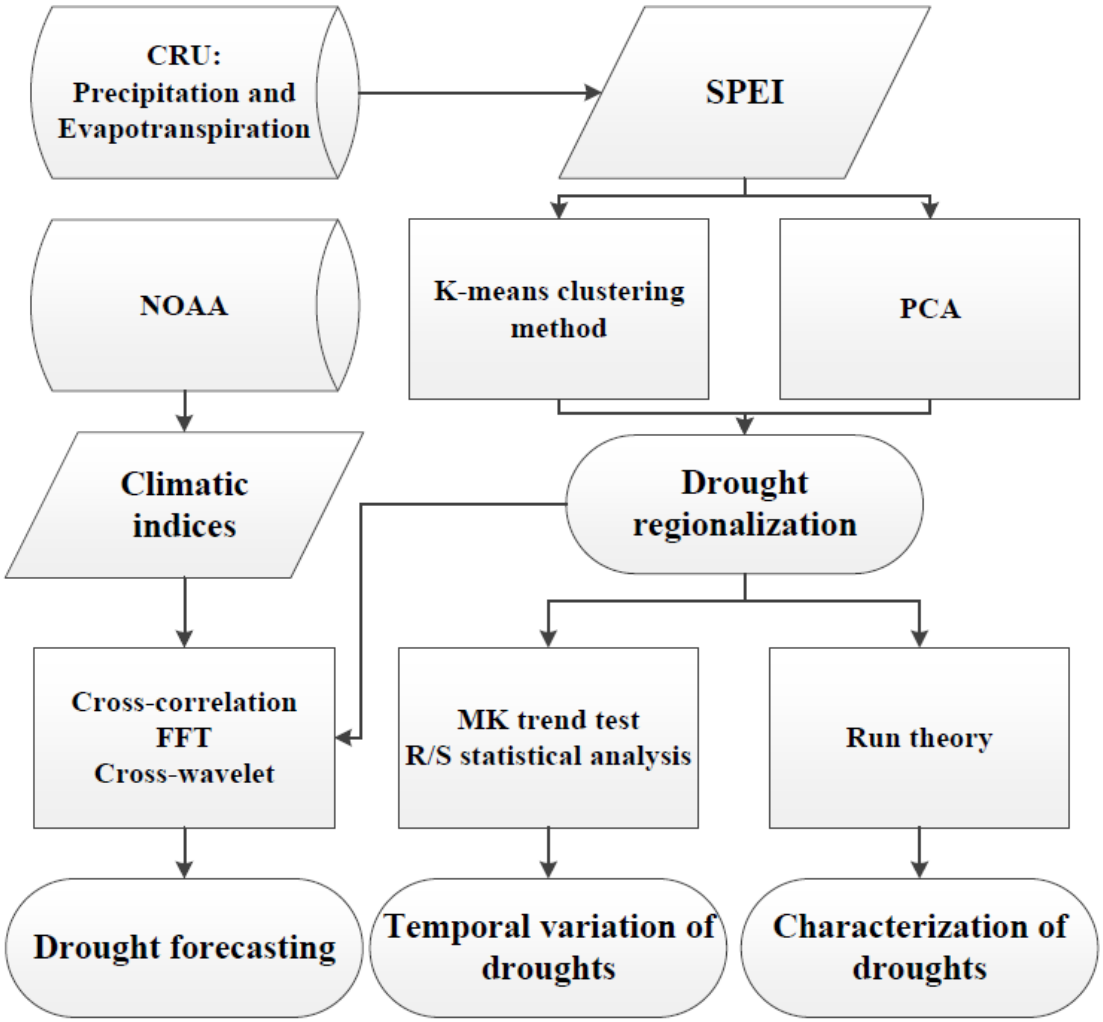


Fig. 2.1 Workflow of the overall methodology.

2.1. Study area

Mozambique covers an area of 801,000 km^2 (Fig. 2.2a). It is in the southern cone of Africa, one of the most drought-prone areas worldwide, where extreme hydrological events are expected to become more frequent and damaging due to climate change (Eriksen and Silva, 2009; Osbahr et al., 2008; Patt and Schröter, 2008). The climate is tropical, with a hot and rainy summer season from November to March (80% of the year's precipitation falls during this period), and a cool and dry winter season from April to October. The national average annual precipitation varies from 683 mm to 1,276 mm, with the south and central west being the drier regions. The mean annual temperature varies from 23°C to 25°C, with the coastal regions of the centre and north of the country, plus the centre-west, being the warmest. The national average annual precipitation level has high interannual variability and has been below the average for most years in the past two decades (Fig. 2.2b). The average annual temperature has increased considerably over the last 20 years, where its anomalies have reached +1°C in relation to the period 1950-2019 (Fig. 2.2c). According to the International Disaster Database (EM-DAT, 2021), the drought events of 1987, 1991-1992, 1995 and 2016 were among the severest in the country, causing losses of 650 million dollars and affecting at least 24 million people (MunichRE, 2018). In these years, low precipitation and high temperature anomalies were observed.

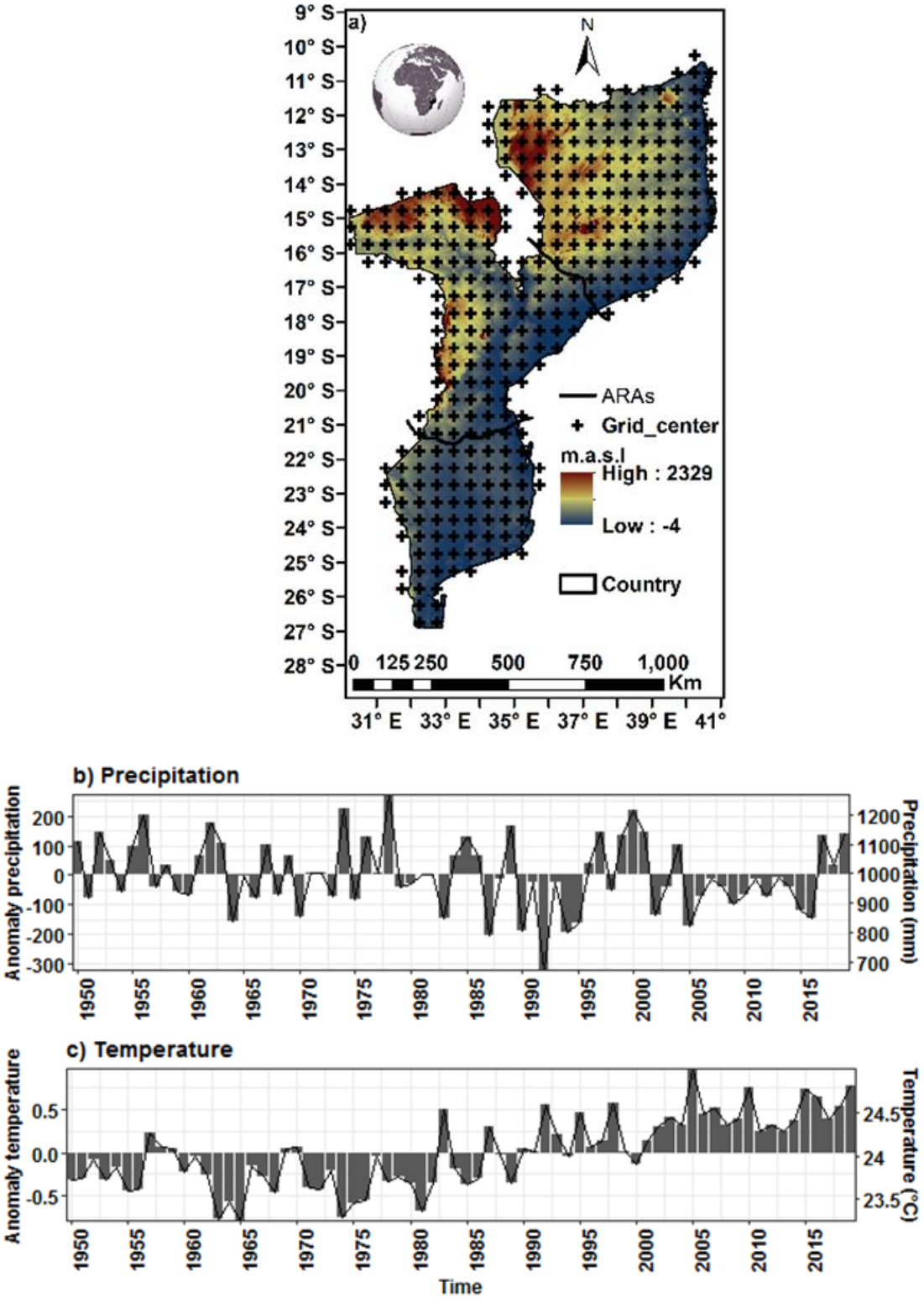


Fig. 2.2 a) Location of Mozambique and its topography with black marks that illustrate the Climate Research Unit (CRU) grid points (0.5° resolution) (Harris et al., 2014). Annual national average series of b) precipitation and c) temperature (solid lines) and anomalies with respect to the mean from 1950 to 2019 (bars). Dashed lines indicate linear trends.

Although rainfall volumes are higher in the north than in the south, droughts are a recurrent problem throughout the country, affecting the socio-economic activities and negatively impacting the quality of life of the population.

The National Directorate of Water (DNA) divides the country into three Regional Water Administrations—ARAs (Administração Regional de Águas)—which are responsible for managing extreme events such as droughts (Conselho de Ministros, 2020): South ARA, from the country's southern border to the Save river; Central ARA, running from the Save river to the Licungo river; and North ARA, comprising the territory from the Licungo river to the country's northern border with Tanzania (Fig. 2a). It should be noted that this division into 3 ARAs is very recent. Five ARAs were initially created in Mozambique as part of the 1991 Water Law, based on geographic and infrastructural conditions specific to their jurisdictional areas, prioritizing their institutional capacity (Inguane et al., 2014).

2.2. The Standardized Precipitation and Evapotranspiration Index (SPEI)

The SPEI was selected as a hydro-meteorological drought index due to its flexibility as a multi-scalar index (unlike the PDSI), and because of its functionality under climate change conditions, in that it can account for the role of temperature increase in future drought conditions (unlike the SPI). The SPEI has been used successfully for drought monitoring in various studies around the world (Lovino et al., 2014; Meresa et al., 2016), and has been found to be more effective than other indices in capturing drought responses for ecological, agricultural, and hydrological applications (Vicente-Serrano et al., 2012).

The SPEI is based on the probability distribution of a long-term climatic water balance ($CWB = P - ETP$) time series, where P is the precipitation and ETP the potential evapotranspiration. It is typically computed by summing CWB over k months (similar to SPI), termed accumulation periods, and fitting these accumulated values to a parametric statistical distribution from which probabilities are standardized ($\mu = 0$, $\sigma = 1$). Given that the CWB series can have values below zero, a three-parameter distribution is needed to model them (Vicente-Serrano et al., 2010a). The three-parameter log-logistic distribution has been found to fit the CWB series very well across most of the world for most time scales (Vicente-Serrano et al., 2010b), so we have used this distribution here.

We utilized a 12-months accumulation ($k = 12$) to analyse the interannual variability of drought conditions. The choice of time scale is driven by the objectives of the study, which ultimately are to provide useful information for water managers and users for monitoring and forecasting drought. Drought at this scale can cause yield reductions for both rainfed and irrigated crops, posing a severe threat to food security in this country. It has also been seen in Mozambique that SPEI-12 is strongly correlated to SPEI at other scales, and that it is effective at detecting the country's historical drought records.(Araneda-Cabrera et al., 2021b). However, it should be noted that using this timescale presents some challenges, such as generating time series of independent measurements and capturing the whole drought cycle in a watershed or region.

Monthly P and ETP data were downloaded from the Climate Research Unit (CRU TS v. 4.04) (http://www.cru.uea.ac.uk/data) to compute the SPEI. CRU offers monthly climatic time series at a 0.5° resolution (≈ 55 km in the Equator) worldwide (Harris et al., 2014). At this resolution, a total of 343 time series of SPEI were computed in Mozambique (Fig. 2.2a). The time span of the analysis was from January 1950 to December 2019 (70 years).

For the purposes of this study, a drought event started when the SPEI took values lower than -1 and ended when its value returned to values higher than this threshold, which corresponds to moderate droughts according to the categories in McKee et al. (1993). In addition, to ensure that drought events were independent of each other, and to group mutually dependent droughts, we used the inter-event time method introduced by Zelenhasić and Salvai (1987), which is still widely applied in the recent literature (e. g. Liu et al., 2020; Rivera et al., 2021). Drought events were designated as independent if the inter-event time lasted more than 2 months (i.e., 2 consecutive months above the proposed drought threshold); while a drought qualified as an event if it lasted more than 2 months. The events were characterized according to run theory, explained in detail in Yevjevich (1969). The intensity (Int) is the minimum

monthly value that is reached by the index throughout the event, the duration (Du) is the number of months that the event lasts, and the severity (Sev) is computed as the sum of monthly SPEI values throughout the event.

2.3. Drought regionalization

Principal component analysis (PCA) was used to identify drought patterns in the SPEI series, and hence define homogeneous regions with similar drought variability and characteristics. It has been widely used for similar regionalization purposes in other parts of the world (Espinosa et al., 2019; Lovino et al., 2014; Santos et al., 2010; Vicente-Serrano, 2006a).

The method consisted of calculating the covariance matrix of the data (SPEI series) with the corresponding eigenvalues and eigenvectors. The principal components (PC) are given by linear combinations of the time series (SPEI) with maximum variance (Rencher, 2002). The number of regions were defined by the number of chosen PCs. There are several methods for finding the right number of PCs (Cangelosi and Goriely, 2007). Here the criterion selected was that they explain at least 75% of the accumulated variance, while the following PC represents less than 5%. Then, the main components were rotated (rotated principal components, RPC) using the Varimax technique (Espinosa et al., 2019) to locate more accurately the spatial patterns of drought variability, to improve their interpretation, and to redistribute the final explained variance (Vicente-Serrano, 2006a). To identify the spatial patterns of the SPEI, Pearson's correlation coefficients (r) were calculated between each RPC and the SPEI series of each CRU, resulting in smooth and gradual patterns of the SPI-12 field (Espinosa et al., 2019). When a group of centroids (CRU cells) had the high correlations r with an RPC, we delimited a new region.

In parallel, hierarchical clustering analysis was applied through the k-means method (Santos et al., 2010; Vicente-Serrano, 2006b; Wilks, 2006). The goal was to compare the number of optimal clusters with the number of PCs obtained according to the above criterion to validate the regionalization defined by the PCA method. To choose the optimal number of clusters, we used the Euclidean distances between the created clusters, which yields the lowest possible number with the greatest possible homogeneity.

Euclidean distances ensured heterogeneity between clusters, so in order to guarantee the homogeneity within clusters, the regional heterogeneity measure Hn proposed by Hosking and Wallis, (1993) was used, this as a means of assessing whether the resulting regions were statistically homogeneous. A region is considered "acceptably homogeneous" if $Hn < 1$, "possibly heterogeneous" if $1 < Hn < 2$ and "definitely heterogeneous" if $Hn \geq 2$.

From this point onwards, all subsequent analyses were performed with SPEI-12 series representative of the resulting homogeneous regions, obtained by averaging all the time series contained in each of them.

2.4. Trend and persistence analysis

This step sought to analyse the temporal variability of droughts by exploring trends and their long-term persistence. On the one hand, the Mann-Kendall (MK) trend test was used to analyse whether the SPEI time series presented a significant trend, either positive or negative. The MK trend test is a rank-based non-parametric method that analyses the difference in signs between the previous and subsequent data points, using the standard normal variant (Z) (Hipel and McLeod, 1994). Although the MK trend method requires the measurements to be independent, for simplicity of analysis in the present paper we have assumed that monthly SPEI-12 meet this condition, as other studies have also done (García-Garizábal, 2017; Yao et al., 2018).

On the other hand, the Rescaled Range (R/S) statistical analysis was applied to the SPEI time series in order to quantify the long-term persistence of trends. We ascertained whether the drought trends observed to be statistically significant in the study period (past) persist in time (future), since this is

related to the predictability of droughts and climate change (Koutsoyiannis, 2005, 2003). This analysis was introduced by Hurst (1956). The long-term persistence of trends in the time series is analysed by estimating the autocorrelation properties of the time series. For instance, this allows us to see whether humid years cluster in multiannual humid periods or if drought years cluster in multiannual dry periods. Such an estimation is made by means of the Hurst index (H), which is a measure of long-term persistence. The H index classifies the time series into 3 types according to their value. When $H = 0.5$, the series is completely uncorrelated and its future trend is different or equal to the past one; when $H < 0.5$, the future trend of the series will the opposite of the past series; and when $H > 0.5$, the future trend of the series will be the same as the past trend. In the latter two cases, with the lowest and highest values of H, respectively, the strength of the persistence is greatest. The steps of the computation can be seen in Gao et al. (2020).

These two analyses were applied to the monthly and annual SPEI time series averaged over the homogeneous drought regions obtained in the previous step. The annual series was assessed in this section to strengthen the limitations of using SPEI-12, as these measures can be considered non-dependent.

2.5. Relationships with large-scale climate indices

A series of large-scale climate indices were selected to analyse their possible relationships with the variability of the SPEI time series averaged over the homogeneous drought regions. The climate indices are based on the fluctuations of atmospheric pressure at sea level (SLP) of different points around the globe (Darwin SLP, Tahiti SLP, SOI and NAO indices), and the sea surface temperature (SST) of the Atlantic (TNA, TSA, NAT, SAT and TASI indices), Pacific (ENSO indices Niño 1+2, Niño 3, Niño 4, Niño3.4 and PDO) and Indian (SWIO, WTIO, SETIO and DMI indices) oceans. The climate indices used are listed in Table 2.1 together with the data sources and the available period of data. For consistency with the SPEI time series, a 12-month moving average was applied to the climatic indices from 1950 to 2019, with a monthly resolution, except for PDO, SWIO, TASI, NAT and SAT, which were not available for the whole time period.

Table 2.1 Climate indices considered in the correlation analysis.

Variable/data set	Period available	Data availability
Darwin sea level pressure (Darwin SLP)*	Jan 1882-now	http://cpc.ncep.noaa.gov/data/indices/darwin
Tahiti sea level pressure (Tahiti SLP)*	Jan 1882-now	http://cpc.ncep.noaa.gov/data/indices/tahiti
Southern Oscillation Index (SOI)**	Jan 1866-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/SOI/
ENSO indices (ERSSTv5): El Niño 1+2, El Niño 3, El Niño 4, and El Niño 3.4*	Jan 1950-now	https://cpc.ncep.noaa.gov/data/indices/ersst5.nino.mth.81-10.ascii
Pacific Decadal Oscillation (PDO)**	Jan 1948-Dec 2018	https://psl.noaa.gov/data/correlation/pdo.data
South Western Indian Ocean (SWIO)	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/ind/swio.php
Western Tropical Indian Ocean (WTIO)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmiwest.had.long.data
South-eastern Tropical Indian Ocean (SETIO)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmieast.had.long.data
Indian Ocean dipole mode index (DMI)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmi.had.long.data
Tropical Northern Atlantic Index (TNA)**	Jan 1948-now	https://psl.noaa.gov/data/correlation/tna.data
Tropical Southern Atlantic Index (TSA)**	Jan 1948-now	https://psl.noaa.gov/data/correlation/tsa.data
North Atlantic Tropical (NAT)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/nat.php
South Atlantic Tropical (SAT)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/sat.php
Tropical Atlantic (TASI)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/tasi.php
North Atlantic Oscillation (NAO)**	Jan 1950-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/nao.long.data

*; **; *** specifies source:
* Climate Prediction Centre of NOAA
** Physical Sciences Laboratory of NOAA
*** Ocean Observations Panels for Climate of NOAA

The cross-correlation method (Hair et al., 1998) was applied between the monthly and annual SPEI series and the climatic indices to quantify the strength of the link between them. Since the relationships between climatic and drought indices necessarily could not occur at the same time, we first analysed the correlation with zero lag time and then looked for the time lag (on a monthly scale) in which the correlation is greater between the two series. The time lag is associated with the early prediction of one series using the other, while the correlation coefficient indicates how strong that relationship would be. Based on the cross-correlation results, a reduced set of climate indices was selected for the next steps of the methodology.

A spectral analysis based on Fast Fourier Transform (FFT) and a cross-wavelet analysis was then performed between the monthly SPEI series and the climatic indices that showed the best correlations in the previous analysis. The idea was to further explore the relationship between these two types of indicators in each homogeneous drought region of the country.

The FFT is a well-known mathematical procedure that allows us to convert signals (time series) from the time domain to the frequency domain. This process is very useful for decomposing a time series comprising various pure frequencies (sinuses and cosines) in only a few recurring periods (Period=1/Frequency) of different lengths. Here, we looked for the periods in the SPEI and their highly correlated climatic indices. For details of the mathematical process, see Fleming et al. (2002).

Cross-wavelet analysis was initially introduced by Hudgins et al. (1993), and explores the relationships between two associated time series (in this case, the SPEI series and the climatic indices). It combines wavelet transformation with cross-spectrum analysis and can notionally capture the characteristic changes and associated oscillations of these two time series in both the time and frequency fields (Grinsted et al., 2004). A detailed description of the calculation method and applications can be found in Torrence and Compo (1997).

These methods have been applied in other similar studies. For example, the cross correlation method was used by Lima and AghaKouchak (2017) in Amazonia to correlate the PDSI with climate indices; Santos et al. (2010) used FFT in Portugal to determine the periodicity of droughts according to the SPI-6; and Räsänen et al. (2016) applied cross-wavelet analysis in mainland Southeast Asia to analyse the relationship between the ENSO and the Palmer drought Severity Index (PDSI, Palmer, 1965).

3. Results

3.1. Spatial distribution of droughts

Following the procedure described in Section 2.2, a total of 343 time series of SPEI values were calculated, these corresponding to the CRU coordinates presented in Fig. 2.2a. Each SPEI-12 time series had a length of 840 months (from 1950 to 2019).

PCA analysis was applied to the matrix that contained the times series of SPEI (with 840 rows corresponding to the length of the time series, and 343 columns corresponding to the coordinates of CRU), to transform its variables into principal components (PC) by simple linear transformations. The first PC explained a large percentage of the total variance (46.83 %). The variance retained by the PC2 and PC3 were 25.17% and 7.51%, respectively. Following the proposed criteria, these three PCs were chosen for the Varimax orthogonal rotation, since together they explained 79.5% of the variance, and the part of the total variance retained by the next PC (PC4) was below 5% (4.26%). Inverse Distance Weighting (IDW) interpolation was used to plot Pearson's correlations between the three chosen RPCs and each of the 343 SPEI series (Fig. 2.3). This technique was applied only for plotting purposes; regionalization was based on the values at the CRU grid points. Three regions were clearly defined: North (Reg1), South (Reg2) and Centre (Reg3) following the classification obtained by the best correlations ($r > 0.60$). The coordinates of each SPEI series belonging to each RPC are listed in the supplementary materials (Table S2.1).

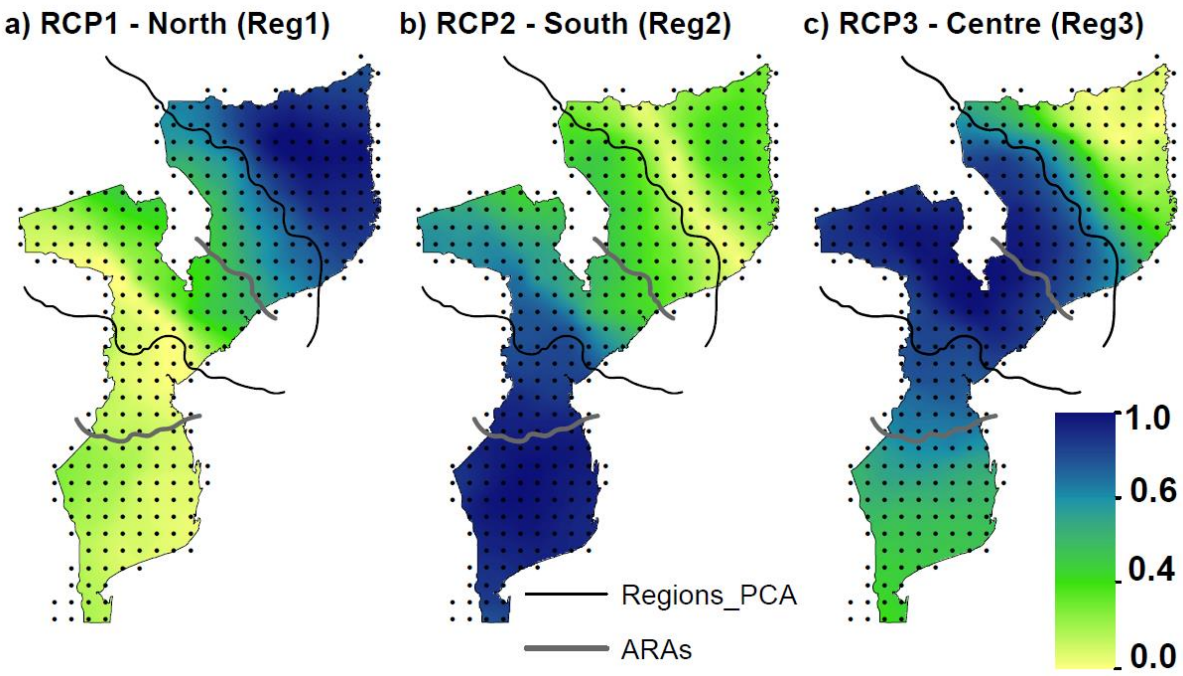


Fig. 2.3 Spatial distribution of the correlation coefficients between RPCs and SPEI. The continuous black lines define the regions obtained with correlations greater than 0.60. The grey lines show the three Regional Water Administrations (ARAs).

It should be noted that other temporal aggregations (SPEI-3 and SPEI-6) were also tested in this step, obtaining very similar homogeneous drought regions, as can be seen in Fig. S2.1 of the Supplementary materials. In what follows we used the regionalization obtained with PCA on the SPEI-12 series.

The non-hierarchical k-means clustering method was applied to the 343 SPEI time series to validate the regionalization obtained by PCA. According to PCA, a successful regionalization would be a classification into 3 groups, so here we analysed a clustering of 2, 3 and 4 regions (i.e., a variation of ± 1 with respect to the PCA result). The spatial extent of the resulting clusters is shown in Fig. 2.4, and the coordinates of the SPEI series belonging to each cluster are detailed in the supplementary materials (Table S2.1). Using the Euclidean distance between clusters method (Table 2.2), the classification into 2 groups (Fig. 2.4a) is considered inadequate, since the distance between both clusters in this configuration (26.76) is less than the distance between them (37.36) when grouped into 3 clusters (Fig. 2.4b). When clustering into 4 groups (Fig. 2.4c), the area representing cluster 1 in the 3-group division (Fig. 2.4b) is divided into clusters 1 and 4, while cluster 2 remains invariant. To be accepted as a better classification, the two distances (cluster 1 and 4) with respect to cluster 2 should be greater than the distance between cluster 1 and 2 when grouped into 3 clusters. Since this is not true for the distance between clusters 2 and 4 (35.41 < 37.36), the division into 4 clusters is rejected. Although not shown in this paper, regionalization was also performed with Ward's hierarchical clustering method (Wilks, 2006), obtaining similar results.

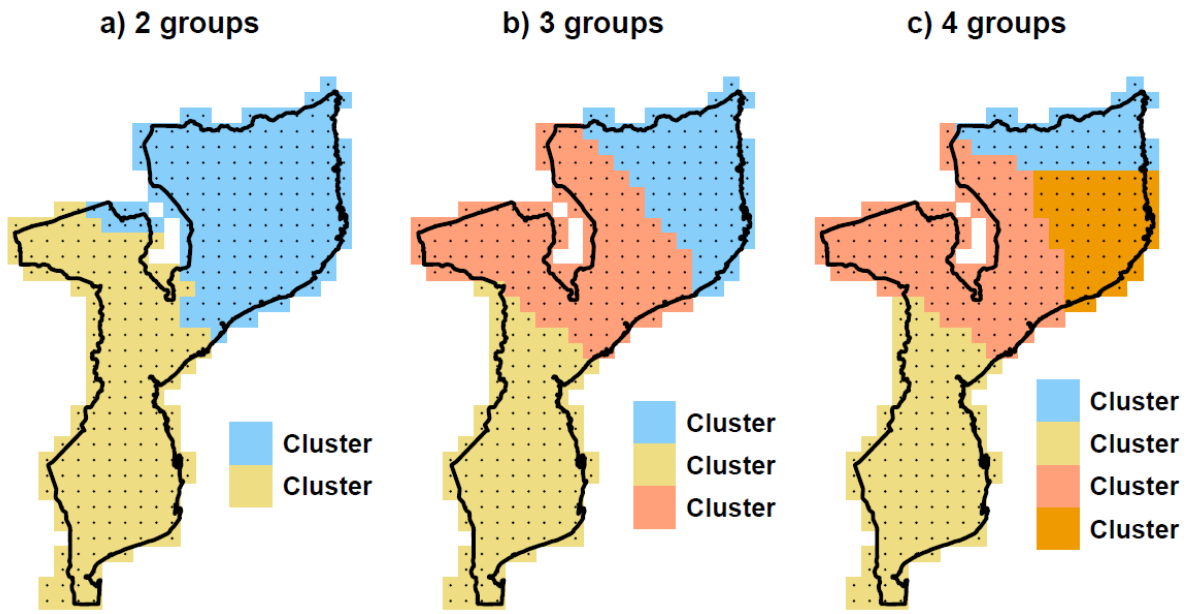
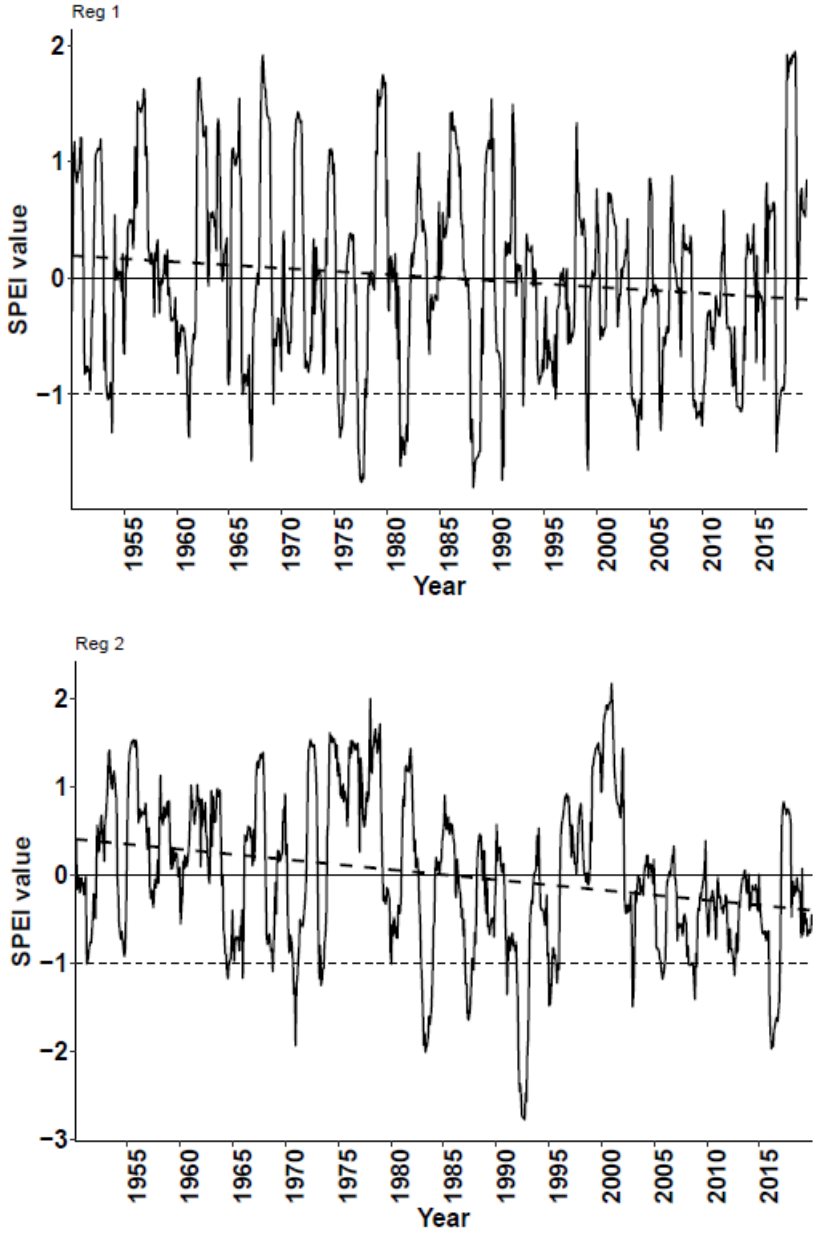


Fig. 2.4 Comparison of clustering in (a) two, (b) three and (c) four groups.

Based on these results, in what follows we used the regionalization obtained with PCA, which divides Mozambique into 3 homogeneous drought regions (Fig. 2.3). In these 3 regions, the Hn index was -0.111 for the northern region, -0.004 for the central region and -0.024 for the southern region. Thus, both heterogeneity between clusters and homogeneity within clusters was guaranteed. The resulting regionalisation differs from other divisions created for purposes other than drought management (FEWS NET Moçambique, 2014; INGC, 2009). However, as shown in Fig. 3, the regions as defined do not entirely coincide with those used by the Mozambican government (ARAs), although the spatial positioning patterns (North, Centre, and South) are maintained. Therefore, these regions could be used for drought assessment, monitoring, and management. The SPEI series were averaged in each region (93 SPEI series on the North, 110 on the South and 140 on the Centre regions) to use in the subsequent analysis (Fig. 2.5).

Table 2.2 Euclidean distances between clusters for the analysis with two, three and four classification groups.

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**
	Two classification groups			
Cluster 1	0	0	0	0
Cluster 2	26.76	0	0	0
	Three classification groups			
Cluster 1	0	0	0	0
Cluster 2	22.88	0	0	0
Cluster 3	37.36	26.16	0	0
	Four classification groups			
Cluster 1	0	0	0	0
Cluster 2	38.32	0	0	0
Cluster 3	30.21	22.20	0	0
Cluster 4	20.44	35.41	23.80	0


Reg 1
SPEI value
2
1
0
−1
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
Year
Reg 2
SPEI value
2
1
0
−1
−2
−3
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
Year

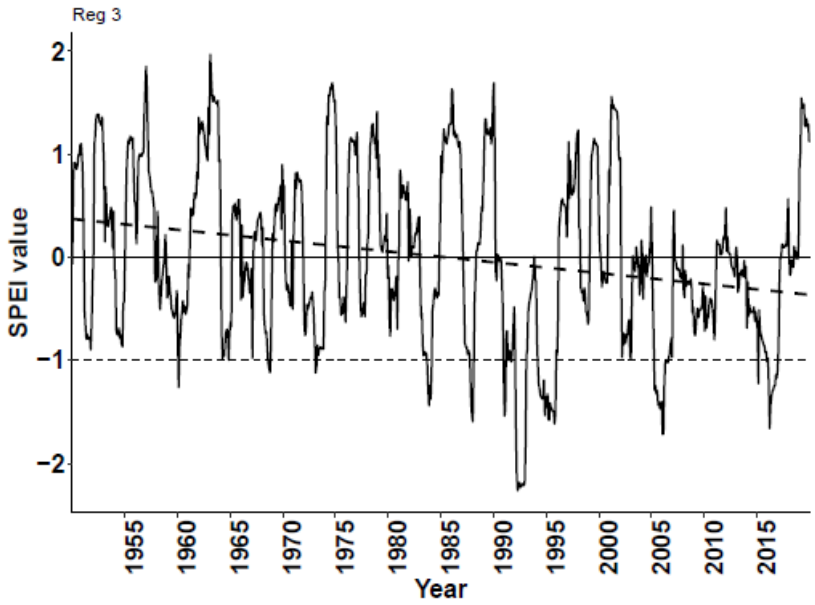


Fig. 2.5 SPEI time series averaged over the three regions: North (Reg1), South (Reg2) and Centre (Reg3). The thin horizontal dashed line represents the threshold considered to define a drought event while the thick dashed line represents the linear trend of the time series.

3.2. Characterization of drought events

Run theory was adopted to characterize the drought events in the 3 regions previously determined (Table 2.3). A lower number of drought events, but of longer duration and higher severity, were found in the Centre region. On average, drought events were less intense and severe in the North. They occurred on average every 4.9 years in the North, every 4 years in the South, and every 5.9 years in the Centre region. The longest, most intense, and severest events began in 1982, 1987, 1991, 1994, 2005, 2009, and 2016. These drought events have been listed as the major ones at the national level (EM-DAT, 2021; Masih et al., 2014), and most of them affected all three regions.

Table 2.3 Identification of drought events and their characteristics with run theory in the three homogeneous drought regions.

North				South				Centre			
Begin	**Du**	**Int**	**Sev**	**Begin**	**Du**	**Int**	**Sev**	**Begin**	**Du**	**Int**	**Sev**
06/1953	3	-1.04	-3.11	06/1964	3	-1.17	-3.31	08/1968	3	-1.12	-3.22
01/1961	3	-1.37	-3.70	10/1970	5	-1.93	-6.97	02/1973	2	-1.12	-2.19
12/1966	4	-1.57	-5.06	05/1973	4	-1.25	-4.56	09/1983	6	-1.44	-7.44
05/1975	8	-1.37	-9.50	11/1982	14	-2.01	-22.70	10/1987	5	-1.59	-6.66
04/1977	8	-1.75	-12.91	01/1987	8	-1.64	-11.07	12/1990	3	-1.54	-4.10
04/1981	9	-1.62	-12.94	10/1987	2	-1.07	-2.10	02/1992	12	-2.26	-25.54
12/1987	12	-1.80	-18.27	12/1991	15	-2.77	-32.81	03/1994	21	-1.62	-29.40
12/1990	3	-1.74	-4.72	01/1995	4	-1.48	-5.40	04/2005	16	-1.72	-21.02
12/1998	3	-1.65	-4.22	09/1995	4	-1.22	-4.35	02/2016	11	-1.66	-14.17
04/2003	12	-1.48	-13.64	12/2002	2	-1.49	-2.80				
12/2005	3	-1.31	-3.61	09/2005	4	-1.18	-4.49				
02/2009	12	-1.27	-13.62	04/2008	7	-1.41	-7.77				
04/2013	7	-1.15	-7.84	12/2015	12	-1.97	-20.47				
01/2017	5	-1.49	-6.21								
Average:	6.57	-1.47	-8.52	Average:	6.46	-1.58	-9.91	Average:	8.78	-1.56	-12.64

3.3. Temporal variability and persistence of droughts

The representative SPEI time series for each region are shown in Fig. 2.5. They are clearly non-monotonic, non-stationary according to ACF analysis (autocorrelation function, not shown here), and trend-stationary (statistical significance level of 1%) according to the KPSS test (the Kwiatkowski–Phillips–Schmidt–Shin test).

The MK test (Z index) and the R/S analysis (H index) applied to the annual and monthly SPEI time series of the three homogeneous drought regions are presented in Table 2.4. In all cases and regions, the trends were negative (illustrated in Fig. 2.5), although in the North region they were statistically non-significant at a level of 5% for the annual SPEI.

Table 2.4 Trends (Z) and Hurst index (H) of monthly and annual SPEI in the period 1950–2019 in the homogenous drought regions.

Region		Annual	Monthly
North	**Z**	-1.25*	-3.86
	H	-	0.69
South	**Z**	-2.32	-7.88
	H	0.61	0.74
Centre	**Z**	-2.10	-6.83
	H	0.61	0.74

* Trend statistically non-significant ($\rho > 0.05$)

In all cases, the Hurst index (H) was greater than 0.5, which suggests that the negative long-term trends will persist in the near future. In the South and Centre regions, the H values were higher, suggesting that trends will persist with greater strength, while long-term trend persistence strength will be weak in the North region.

3.4. Identification of large-scale climate drivers

The cross-correlations and lagged cross-correlations between the proposed climate indices (Table 2.1) and the SPEI of the homogeneous drought regions are shown in Table 2.5. The best correlations are shown in the top rows of the table. The anomalies in the time series of the best-correlated climate indices are shown in Fig. 2.6. Most of the correlations were negative (Darwin, Niño 3.4, Niño 3, Niño 4, WTIO and SETIO), and only one was positive (SOI), indicating they are anti-phase or in-phase, respectively, relative to the SPEI.

Table 2.5 Cross correlations between the SPEI time series of each region and the climatic indices. r is the correlation coefficient, r_lag is the lagged correlation coefficient obtained when lag = lag_months (greater correlation) and r_annual is the correlation coefficient between the series aggregated annually.

Climatic indices	North				South				Centre			
	r	r_lag	lag months	r annual	r	r_lag	lag months	r annual	r	r_lag	lag months	r annual
Niño 4	-0.20	-0.24	4	-0.24	-0.45	-0.48	3	-0.49	-0.53	-0.56	3	-0.59
Niño 3.4	-0.17	-0.30	6	-0.23	-0.45	-0.46	2	-0.50	-0.51	-0.57	3	-0.59
Darwin	-0.14	-0.25	7	-0.21	-0.43	-0.44	2	-0.48	-0.50	-0.54	3	-0.58
Niño 3	-0.15	-0.34	7	-0.23	-0.42	-0.43	1	-0.47	-0.47	-0.53	4	-0.55
SOI	0.14	0.21	6	0.17	0.37	0.38	2	0.42	0.45	0.48	3	0.52
WTIO	-0.19	-0.20	4	-0.20	-0.38	-0.38	0	-0.40	-0.41	-0.41	0	-0.43
SETIO	-0.28	-0.29	2	-0.24	-0.32	-0.40	-6	-0.35	-0.38	-0.41	-3	-0.39
Niño 1+2	-0.12	-0.29	8	-0.19	-0.31	-0.31	-1	-0.36	-0.35	-0.40	4	-0.42
Tahiti	0.10	-0.18	20	0.09	0.27	0.27	1	0.29	0.33	0.35	2	0.36
PDO**	-0.20	-0.21	3	-0.21	-0.22	-0.28	-9	-0.25	-0.30	-0.30	-1	-0.32
SWIO*	-0.13	-0.13	-2	-0.10	-0.17	-0.20	-5	-0.19	-0.24	-0.26	-3	-0.23
NAO	0.17	0.18	1	0.16	-0.22	-0.29	4	-0.26	-0.22	-0.33	5	-0.32
TASI*	-0.14	-0.24	-10	-0.17	0.30	0.30	-1	0.31	0.18	0.36	11	0.22
TNA	-0.24	-0.24	-2	-0.22	-0.10	-0.20	17	-0.12	-0.17	-0.18	-2	-0.15
DMI	0.05	-0.21	-11	-0.02	-0.17	-0.30	5	-0.21	-0.16	-0.26	5	-0.22
NAT*	-0.27	-0.28	-1	-0.26	0.08	0.18	7	0.09	-0.13	0.21	11	-0.09
SAT*	-0.12	-0.21	-10	-0.16	0.24	0.27	-3	0.24	0.12	0.29	11	0.13
TSA	-0.12	-0.13	3	-0.14	-0.07	-0.25	-19	-0.09	-0.07	-0.22	-16	-0.08

* Period Jan1982-Dec2019, ** Period Jan1950-Dec2018

The North region persistently showed a poor correlation with the climate indices analysed, while the strongest correlations were obtained in the South and Centre regions. In these two regions, strong correlations with the different El Niño indices were noticeable (r up to -0.59), with the higher correlations being found for the El Niño 4 and El Niño 3.4 SST indices, which showed crests around 2-3 months earlier than the negative peak of the SPEI. Another strong correlation (r = -0.58) with a similar time lag was found with the Darwin SLP index. Based on the above observations, El Niño 4 and Darwin indices were chosen for the spectral analysis using the FFT technique and cross-wavelet analysis.

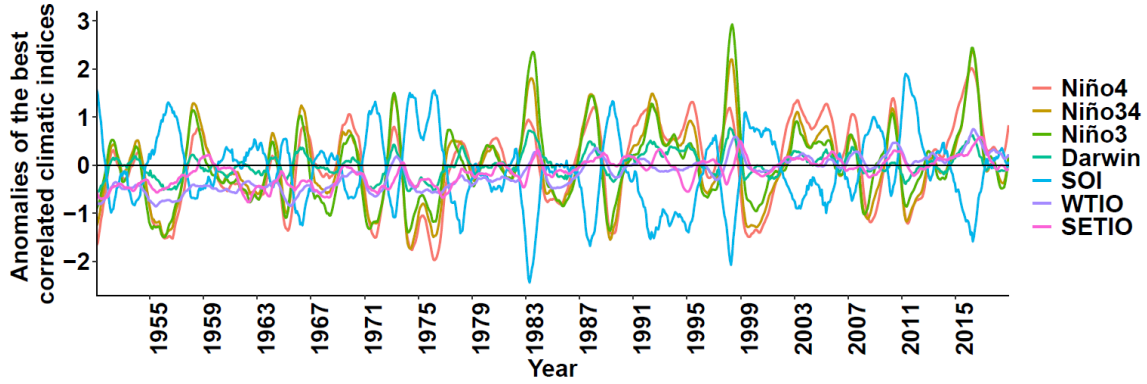


Fig. 2.6 Anomalies of the best correlated climatic indices with the SPEI time series in the period 1950-2019.

FFT allowed us to appreciate the periodic behaviour of the monthly SPEI patterns in each homogeneous drought region, and of Darwin SLP and Niño 4 climatic indices. The periodograms are set out in Fig. 2.7. The results showed that climatic indices and SPEI series have a periodicity associated with high energies between 40 and 120 months (3.5 and 10 years). In the North region, periods of between 35 and 60 months (3 and 5 years) were found. These periods are consistent and similar to those reported in other studies in Africa (Oguntunde et al., 2018, 2017).

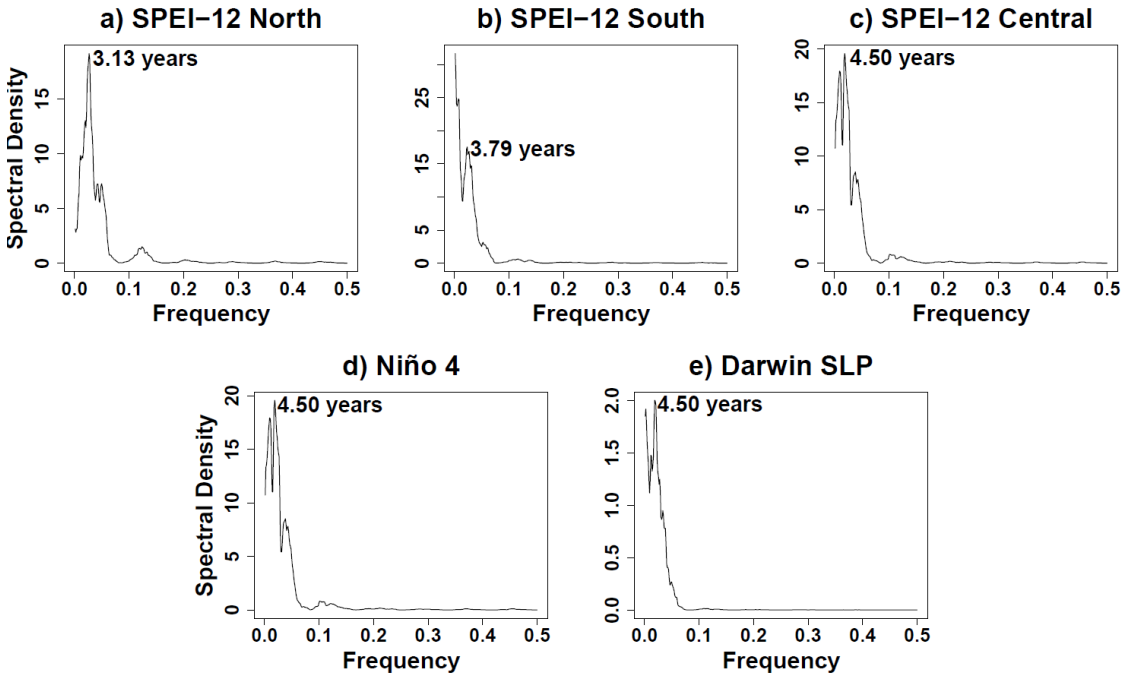


Fig. 2.7 Fast Fourier Transform of monthly SPEI time series of a) North, b) South, c) Centre regions, d) the climatic index Niño 4, and e) the climatic index Darwin SLP. Peaks have been transformed to years.

The spectral analysis of the SPEI, Darwin SLP, and El Niño 4 patterns was expanded using cross-wavelet transform (Fig. 2.8). Both Darwin SLP and El Niño 4 events showed strong impacts on the monthly series of SPEI, especially in the South and Centre regions, indicating that they play a relevant role in the characteristics of the evolution of droughts in Mozambique. Specifically, the positive events of Niño 4 show statistically significant negative links (confidence level 95%) with the monthly SPEI series of the South and Centre regions, with a signal of 16 to 128 months (1.3–10.7 years). In the North region, climatic indices did not show strong effects on the evolution of droughts (something previously seen in the low correlations); however, the statistically significant strongest signals were found for 16 to 64 months (1.4–5.4 years) over the entire study period. In the three regions, the energy density is higher in the periods where drought events were detected (e.g., the major drought event of 1991).

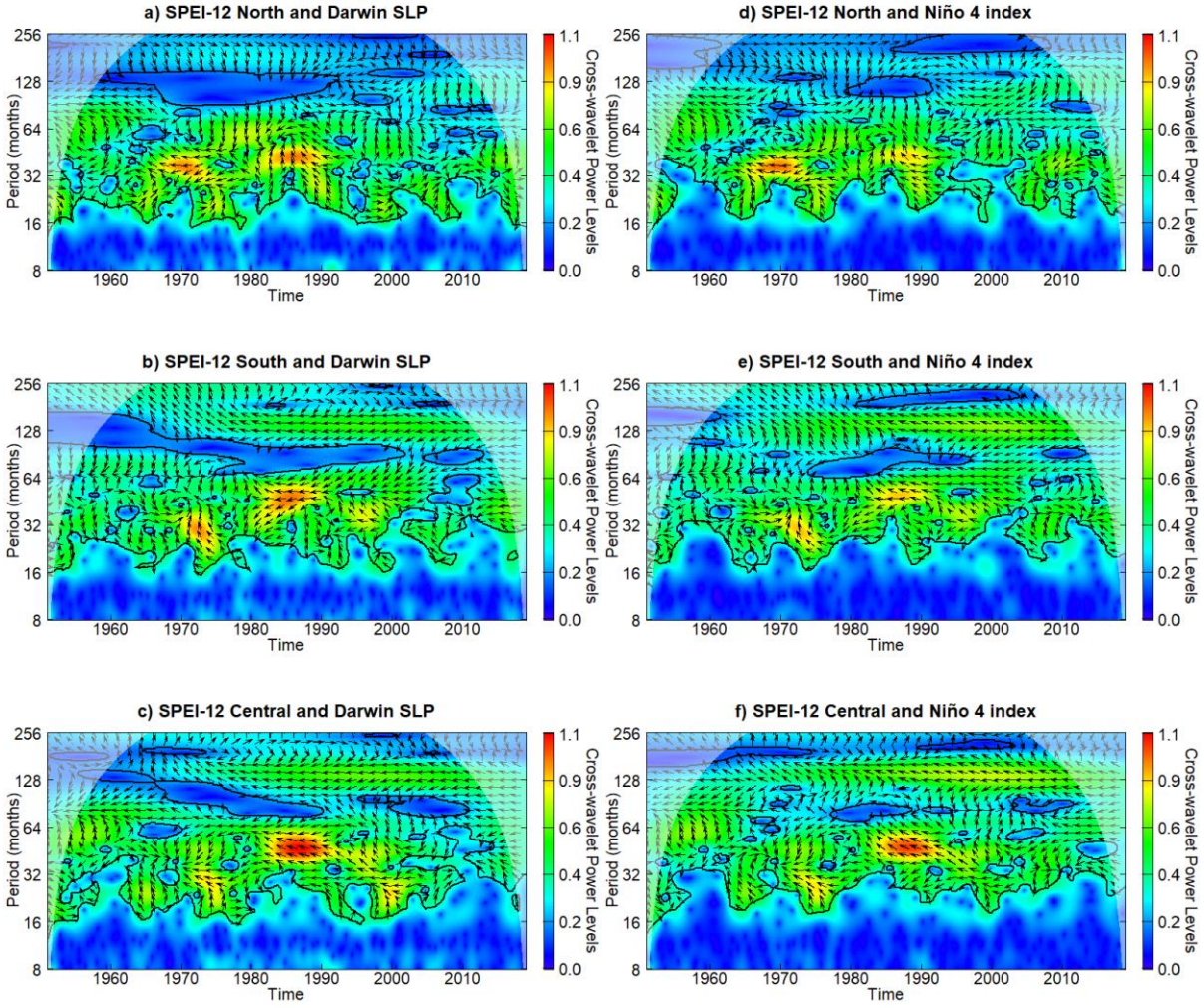


Fig. 2.8 Cross-wavelet transform between SPEI of a) North, b) South, c) Centre regions and Darwin PSL climatic index; and of d) North, e) South, f) Centre regions and the Niño 4 climatic index. The y-axis is equivalent to the periods defined with the FTT (Period = 1/Frequency); the coloured bar denotes the energy density (red plus high energy density); the 5% confidence level against red noise is shown in an outline with the thick black line; and the relative phase relationship is represented with arrows (with the anti-phase pointing to the left, the in-phase pointing to the right).

4. Discussion

The above results provide new insights into the spatial and temporal patterns of drought in Mozambique, and their relationship with the large-scale climate variability.

The resulting drought regionalisation differs from other divisions created for other management purposes (FEWS NET Moçambique, 2014; INGC, 2009). As shown in Fig. 2.3, the regions as defined also do not entirely coincide with those used by the Mozambican government for water resources management (ARAs), although the spatial positioning patterns (North, Centre, and South) are maintained. The use of the proposed divisions as drought management areas would allow for more appropriate regional strategies for assessing, monitoring, and responding to drought (Wilhite et al., 2000). The most important drought events identified in these regions have been listed and coincided with the major ones at the national level.

The trends observed in the SPEI series were consistent with the trends found for the SPEI input variables (i.e., precipitation and temperature) shown in Fig. 2.2b and Fig. 2.2c., where trends were positive for temperature and negative for precipitation. In addition, these findings are in line with similar ones reported by Jury (2013), who used satellite sensor data to analyse climate trends in southern Africa and found that temperature and precipitation trends were positive and negative, respectively, over the period 1980-2010 in Mozambique as a whole. On the other hand, the persistence analysis points to an increase in the incidence of droughts throughout the country. Although these results should be taken with caution due to the climate system complexity, they further highlight the need for the development of drought forecasting tools and more specific, in-depth studies on drought variability in the homogeneous regions. These results are consistent with those found at the continental level and in other regions of Africa (Masih et al., 2014; Rouault and Richard, 2005), and with the variations seen in recent years on precipitation and temperature in Mozambique (Jury, 2013; Uele et al., 2017).

Relationships found between droughts and El Niño 4 and El Niño 3.4 SST indices agree with those described in Manhique et al. (2011) for southern Africa and Mozambique as a whole, and in Manatsa et al. (2008a) for the neighbouring tropical country of Zimbabwe. The strong climatic influence of the Darwin SLP index was also found by Manatsa et al., (2008b) in Zimbabwe. Other tropical regions of the world have also found El Niño 4 and Darwin to be drought triggers (D'Arrigo and Smerdon, 2008; Gu et al., 2020; Lyon and Barnston, 2005), so these (and following) results are also of hydrological interest for countries located in these climatic regions.

The periodic behaviour of the monthly SPEI patterns according to FTT match those reported in section 3.2., and coincide with the historical drought records documented in EM-DAT, (2020) and Masih et al. (2014). Thus, these results are of great importance, as well as being novel for the region. The periodicity found in the El Niño 4 (the same oscillation period was observed for Darwin SPL) index is consistent with the widely accepted 3-7 year period (McPhaden et al. 1998).

The results of the spectral analysis consistently point to an anti-phase relationship between the drought events detected by SPEI and the climatic indices Darwin SLP and El Niño 4 in the South and Centre regions. The North region showed a poor relationship with the climatic indicators. In this region of the country, drought events are less intense and severe, although the number of events is higher. This may be explained by the relationship between the Temperate Tropical Depressions and ENSO. This relationship affects the precipitation patterns in Southeast Africa, making ENSO less influential in the North of Mozambique and meaning that it is a wetter area than the rest of the country (Manhique et al. 2011).

These climate indices could be used in drought forecasting models as predictors of drought in Mozambique, with a lead time of 2-3 months based on the lagged correlations. Such a lead time would enable the establishment of preventive measures against possible upcoming droughts (e.g., accumulating water in reservoirs, prioritizing water use for different uses, etc.). In this way, a forecasting model could

be employed to infer the probability and intensity of drought events in the short-term future, relying on past values of the climate indices (Hao et al., 2018), and allowing actions to be implemented when a drought is expected.

5. Conclusions

The main objective of this study was to investigate the spatio-temporal characteristics of droughts and their relationships with possible predictors of the phenomenon in Mozambique. Although Mozambique is very prone to droughts and suffers continuously from their effects, it does not have the necessary management tools to monitor and predict the phenomenon. The proposed five-step methodology consists of several methods organised in a coherent way for use in Mozambique or in any other region (country or river basin) that requires a first assessment of the spatio-temporal characteristics of droughts. Here, efforts have been made to adapt the methodology specifically to Mozambique, where local meteorological and hydrological monitoring data are extremely limited.

The monthly SPEI-12 was calculated as an indicator of drought from 1950 to 2019 at a high resolution (0.5°). Principal Component Analysis with the Varimax rotation method was used to define 3 homogeneous drought regions located in the North, South, and Centre of the country. This regionalization was validated with hierarchical and non-hierarchical clustering methods. The regions as delimited do not coincide entirely with those identified by the Mozambican National Directorate of Water but are preferable for drought monitoring and management.

Based on run theory, the South and Centre regions are the ones that have presented the most intense and severe drought events in the past. A statistically significant trend towards a higher incidence of droughts was found in the three regions and Rescale Range analysis suggests that this trend might persist in the near future. This section presented valuable information for Mozambique on the temporal variability of droughts. However, given the limitations derived from the use of SPEI-12, it would be advisable to consider additional time scales in future studies to gain further insights into the temporal patterns of drought in each region.

Strong correlations between two climatic indices—El Niño 4 (ENSO) and Darwin SLP—and droughts were found in the South and Centre regions, with a time lag of 2-3 months. These climate indices are representative drought triggers in tropical regions of the world such as Mozambique. With the FFT technique, it was found that the periods of SPEI and these two climate indices have a similar periodicity of between 3 and 8 years, this being a novel statement with reference to Mozambique. Spectral analysis by means of cross-wavelet transform confirmed that SPEI and El Niño 4 and Darwin SLP are strongly related in anti-phase for periods between 1.4 and 10.4 years. These climate indices could thus be used to develop a drought forecasting system, providing sufficient lead time to establish prevention strategies.

In summary, this study has provided an understanding of the spatial and temporal distribution of droughts in Mozambique. The results are of great potential use for Mozambique's regional water administrations towards developing drought contingency plans. Simplified management regions have been defined, characterised, and strongly related to potential drought predictors. The proposed methodology can be used elsewhere in the world; however, given its limitations and the large number of topics it covers, certain limitations will need to be considered in future studies.

Acknowledgments

Ronnie Araneda gratefully acknowledges financial support from the Spanish Regional Government of Galicia (Xunta de Galicia) and the European Union through the predoctoral grant reference ED481A-2018/162. María Bermúdez was supported by the European Union H2020 research and innovation

program under the Marie Skłodowska-Curie Grant Agreement No. 754446 and the Research and Transfer Fund of the University of Granada - Athenea3i. The authors would like to thank Manuel Alvarez and Victor Penas for their comments and support during the development of the study.

Supplementary materials

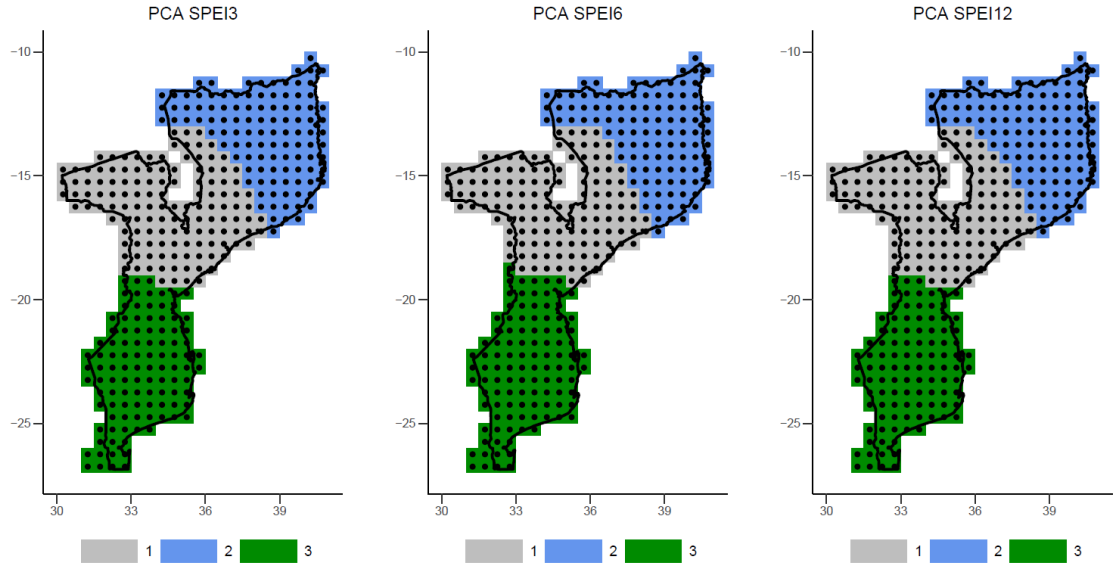


Fig. S2.1 Homogeneous drought regions in Mozambique according to the PCA using the SPEI-3, -6 and -12. It can be seen that the regions are roughly equal.

Table S2.1 The grouping of the 343 CRU coordinates (Lat for latitude and Long for longitude) for the SPEI-12 time series by the non-hierarchical k-means clustering method with 2, 3 and 4 clusters (cl) (Fig. 2.4) and PCA analysis (Fig. 2.3).

Coordinates		Grouping method				Coordinates		Grouping method			
Lat	Long	2 cl	3 cl	4 cl	PCA	Lat	Long	2 cl	3 cl	4 cl	PCA
40.3	-14.8	1	2	2	1	35.8	-11.3	1	2	3	1
40.8	-14.8	1	2	2	1	36.3	-11.3	1	2	3	1
40.3	-14.3	1	2	2	1	37.8	-11.3	1	2	3	1
40.8	-14.3	1	2	2	1	38.3	-11.3	1	2	3	1
40.3	-13.8	1	2	2	1	38.8	-11.3	1	2	3	1
40.8	-13.8	1	2	2	1	39.3	-11.3	1	2	3	1
40.3	-13.3	1	2	2	1	39.8	-11.3	1	2	3	1
40.8	-13.3	1	2	2	1	39.8	-10.8	1	2	3	1
40.3	-12.8	1	2	3	1	30.3	-14.8	2	3	1	3
40.8	-12.8	1	2	3	1	30.8	-14.8	2	3	1	3
40.3	-12.3	1	2	3	1	31.3	-14.8	2	3	1	3
40.8	-12.3	1	2	3	1	31.8	-14.8	2	3	1	3
40.3	-11.8	1	2	3	1	32.3	-14.8	2	3	1	3
40.3	-11.3	1	2	3	1	32.8	-14.8	2	3	1	3
40.3	-10.8	1	2	3	1	33.3	-14.8	1	3	1	3
40.8	-10.8	1	2	3	1	33.8	-14.8	1	3	1	3
40.3	-10.3	1	2	3	1	34.3	-14.8	1	3	1	3
35.8	-14.8	1	3	1	3	34.8	-14.8	1	3	1	3
36.3	-14.8	1	3	1	3	31.8	-14.3	2	3	1	3
36.8	-14.8	1	3	1	3	32.3	-14.3	2	3	1	3
37.3	-14.8	1	3	1	3	32.8	-14.3	1	3	1	3

37.8	-14.8	1	3	2	3	33.3	-14.3	1	3	1	3
38.3	-14.8	1	2	2	1	33.8	-14.3	1	3	1	3
38.8	-14.8	1	2	2	1	34.3	-14.3	1	3	1	3
39.3	-14.8	1	2	2	1	34.8	-13.8	1	3	1	3
39.8	-14.8	1	2	2	1	34.8	-13.3	1	3	1	3
35.3	-14.3	1	3	1	3	34.3	-12.8	1	3	1	3
35.8	-14.3	1	3	1	3	34.8	-12.8	1	3	1	3
36.3	-14.3	1	3	1	3	34.3	-12.3	1	3	1	3
36.8	-14.3	1	3	1	3	34.8	-12.3	1	3	1	3
37.3	-14.3	1	3	2	3	34.3	-11.8	1	3	1	3
37.8	-14.3	1	2	2	1	34.8	-11.8	1	3	3	1
38.3	-14.3	1	2	2	1	40.3	-16.3	1	2	2	1
38.8	-14.3	1	2	2	1	40.3	-15.8	1	2	2	1
39.3	-14.3	1	2	2	1	40.3	-15.3	1	2	2	1
39.8	-14.3	1	2	2	1	40.8	-15.3	1	2	2	1
35.3	-13.8	1	3	1	3	35.3	-19.8	2	1	4	2
35.8	-13.8	1	3	1	3	35.3	-19.3	2	1	4	3
36.3	-13.8	1	3	1	3	35.8	-19.3	2	1	4	3
36.8	-13.8	1	3	1	3	35.3	-18.8	2	1	4	3
37.3	-13.8	1	3	2	1	35.8	-18.8	2	3	1	3
37.8	-13.8	1	2	2	1	36.3	-18.8	2	3	1	3
38.3	-13.8	1	2	2	1	35.3	-18.3	2	3	1	3
38.8	-13.8	1	2	2	1	35.8	-18.3	2	3	1	3
39.3	-13.8	1	2	2	1	36.3	-18.3	2	3	1	3
39.8	-13.8	1	2	2	1	36.8	-18.3	1	3	1	3
35.3	-13.3	1	3	1	3	35.3	-17.8	2	3	1	3
35.8	-13.3	1	3	1	3	35.8	-17.8	1	3	1	3
36.3	-13.3	1	3	1	3	36.3	-17.8	1	3	1	3
36.8	-13.3	1	3	1	3	36.8	-17.8	1	3	1	3
37.3	-13.3	1	2	2	1	37.3	-17.8	1	3	1	3
37.8	-13.3	1	2	2	1	37.8	-17.8	1	3	1	3
38.3	-13.3	1	2	2	1	35.3	-17.3	2	3	1	3
38.8	-13.3	1	2	2	1	35.8	-17.3	1	3	1	3
39.3	-13.3	1	2	2	1	36.3	-17.3	1	3	1	3
39.8	-13.3	1	2	2	1	36.8	-17.3	1	3	1	3
35.3	-12.8	1	3	1	3	37.3	-17.3	1	3	1	3
35.8	-12.8	1	3	1	3	37.8	-17.3	1	3	1	3
36.3	-12.8	1	3	1	1	38.3	-17.3	1	3	2	3
36.8	-12.8	1	2	3	1	38.8	-17.3	1	3	2	1
37.3	-12.8	1	2	3	1	35.3	-16.8	2	3	1	3
37.8	-12.8	1	2	3	1	35.8	-16.8	2	3	1	3
38.3	-12.8	1	2	3	1	36.3	-16.8	1	3	1	3
38.8	-12.8	1	2	3	1	36.8	-16.8	1	3	1	3
39.3	-12.8	1	2	3	1	37.3	-16.8	1	3	1	3
39.8	-12.8	1	2	3	1	37.8	-16.8	1	3	1	3
35.3	-12.3	1	3	3	1	38.3	-16.8	1	3	2	3
35.8	-12.3	1	3	3	1	38.8	-16.8	1	3	2	1
36.3	-12.3	1	2	3	1	39.3	-16.8	1	2	2	1
36.8	-12.3	1	2	3	1	39.8	-16.8	1	2	2	1
37.3	-12.3	1	2	3	1	35.3	-16.3	2	3	1	3
37.8	-12.3	1	2	3	1	35.8	-16.3	1	3	1	3
38.3	-12.3	1	2	3	1	36.3	-16.3	1	3	1	3
38.8	-12.3	1	2	3	1	36.8	-16.3	1	3	1	3

39.3	-12.3	1	2	3	1	37.3	-16.3	1	3	1	3
39.8	-12.3	1	2	3	1	37.8	-16.3	1	3	1	3
35.3	-11.8	1	3	3	1	38.3	-16.3	1	3	2	3
35.8	-11.8	1	2	3	1	38.8	-16.3	1	3	2	1
36.3	-11.8	1	2	3	1	39.3	-16.3	1	2	2	1
36.8	-11.8	1	2	3	1	39.8	-16.3	1	2	2	1
37.3	-11.8	1	2	3	1	35.8	-15.8	1	3	1	3
37.8	-11.8	1	2	3	1	36.3	-15.8	1	3	1	3
38.3	-11.8	1	2	3	1	36.8	-15.8	1	3	1	3
38.8	-11.8	1	2	3	1	37.3	-15.8	1	3	1	3
39.3	-11.8	1	2	3	1	37.8	-15.8	1	3	1	3
39.8	-11.8	1	2	3	1	38.3	-15.8	1	3	2	3
38.8	-15.8	1	3	2	1	35.3	-20.8	2	1	4	2
39.3	-15.8	1	2	2	1	32.3	-24.8	2	1	4	2
39.8	-15.8	1	2	2	1	32.8	-24.8	2	1	4	2
35.8	-15.3	1	3	1	3	33.3	-24.8	2	1	4	2
36.3	-15.3	1	3	1	3	33.8	-24.8	2	1	4	2
36.8	-15.3	1	3	1	3	34.3	-24.8	2	1	4	2
37.3	-15.3	1	3	1	3	34.8	-24.8	2	1	4	2
37.8	-15.3	1	3	2	3	31.8	-24.3	2	1	4	2
38.3	-15.3	1	3	2	1	32.3	-24.3	2	1	4	2
38.8	-15.3	1	2	2	1	32.8	-24.3	2	1	4	2
39.3	-15.3	1	2	2	1	33.3	-24.3	2	1	4	2
39.8	-15.3	1	2	2	1	33.8	-24.3	2	1	4	2
32.8	-19.8	2	1	4	2	34.3	-24.3	2	1	4	2
33.3	-19.8	2	1	4	2	34.8	-24.3	2	1	4	2
33.8	-19.8	2	1	4	2	31.8	-23.8	2	1	4	2
34.3	-19.8	2	1	4	2	32.3	-23.8	2	1	4	2
34.8	-19.8	2	1	4	2	32.8	-23.8	2	1	4	2
32.8	-19.3	2	1	4	2	33.3	-23.8	2	1	4	2
33.3	-19.3	2	1	4	2	33.8	-23.8	2	1	4	2
33.8	-19.3	2	1	4	2	34.3	-23.8	2	1	4	2
34.3	-19.3	2	1	4	2	34.8	-23.8	2	1	4	2
34.8	-19.3	2	1	4	2	31.3	-23.3	2	1	4	2
32.8	-18.8	2	1	4	2	31.8	-23.3	2	1	4	2
33.3	-18.8	2	1	4	3	32.3	-23.3	2	1	4	2
33.8	-18.8	2	1	4	3	32.8	-23.3	2	1	4	2
34.3	-18.8	2	1	4	2	33.3	-23.3	2	1	4	2
34.8	-18.8	2	1	4	2	33.8	-23.3	2	1	4	2
32.8	-18.3	2	1	4	2	34.3	-23.3	2	1	4	2
33.3	-18.3	2	1	4	3	34.8	-23.3	2	1	4	2
33.8	-18.3	2	1	4	3	31.3	-22.8	2	1	4	2
34.3	-18.3	2	1	4	3	31.8	-22.8	2	1	4	2
34.8	-18.3	2	1	4	3	32.3	-22.8	2	1	4	2
32.8	-17.8	2	1	4	3	32.8	-22.8	2	1	4	2
33.3	-17.8	2	1	4	3	33.3	-22.8	2	1	4	2

33.8	-17.8	2	1	4	3	33.8	-22.8	2	1	4	2
34.3	-17.8	2	3	1	3	34.3	-22.8	2	1	4	2
34.8	-17.8	2	3	1	3	34.8	-22.8	2	1	4	2
32.8	-17.3	2	1	4	3	31.3	-22.3	2	1	4	2
33.3	-17.3	2	1	4	3	31.8	-22.3	2	1	4	2
33.8	-17.3	2	3	1	3	32.3	-22.3	2	1	4	2
34.3	-17.3	2	3	1	3	32.8	-22.3	2	1	4	2
34.8	-17.3	2	3	1	3	33.3	-22.3	2	1	4	2
32.3	-16.8	2	1	1	3	33.8	-22.3	2	1	4	2
32.8	-16.8	2	1	1	3	34.3	-22.3	2	1	4	2
33.3	-16.8	2	3	1	3	34.8	-22.3	2	1	4	2
33.8	-16.8	2	3	1	3	31.8	-21.8	2	1	4	2
34.3	-16.8	2	3	1	3	32.3	-21.8	2	1	4	2
34.8	-16.8	2	3	1	3	32.8	-21.8	2	1	4	2
30.8	-16.3	2	3	1	3	33.3	-21.8	2	1	4	2
31.3	-16.3	2	3	1	3	33.8	-21.8	2	1	4	2
31.8	-16.3	2	3	1	3	34.3	-21.8	2	1	4	2
32.3	-16.3	2	3	1	3	34.8	-21.8	2	1	4	2
32.8	-16.3	2	3	1	3	32.3	-21.3	2	1	4	2
33.3	-16.3	2	3	1	3	32.8	-21.3	2	1	4	2
33.8	-16.3	2	3	1	3	33.3	-21.3	2	1	4	2
34.3	-16.3	2	3	1	3	33.8	-21.3	2	1	4	2
34.8	-16.3	2	3	1	3	34.3	-21.3	2	1	4	2
30.3	-15.8	2	3	1	3	34.8	-21.3	2	1	4	2
30.8	-15.8	2	3	1	3	32.3	-20.8	2	1	4	2
31.3	-15.8	2	3	1	3	32.8	-20.8	2	1	4	2
31.8	-15.8	2	3	1	3	33.3	-20.8	2	1	4	2
32.3	-15.8	2	3	1	3	33.8	-20.8	2	1	4	2
32.8	-15.8	2	3	1	3	34.3	-20.8	2	1	4	2
33.3	-15.8	2	3	1	3	34.8	-20.8	2	1	4	2
33.8	-15.8	2	3	1	3	32.8	-20.3	2	1	4	2
34.3	-15.8	2	3	1	3	33.3	-20.3	2	1	4	2
30.3	-15.3	2	3	1	3	33.8	-20.3	2	1	4	2
30.8	-15.3	2	3	1	3	34.3	-20.3	2	1	4	2
31.3	-15.3	2	3	1	3	34.8	-20.3	2	1	4	2
31.8	-15.3	2	3	1	3	32.3	-26.8	2	1	4	2
32.3	-15.3	2	3	1	3	32.8	-26.8	2	1	4	2
32.8	-15.3	2	3	1	3	32.3	-26.3	2	1	4	2
33.3	-15.3	2	3	1	3	32.8	-26.3	2	1	4	2
33.8	-15.3	2	3	1	3	31.8	-25.8	2	1	4	2
34.3	-15.3	2	3	1	3	32.3	-25.8	2	1	4	2
34.8	-15.3	2	3	1	3	32.8	-25.8	2	1	4	2
35.3	-24.8	2	1	4	2	31.8	-25.3	2	1	4	2
35.3	-24.3	2	1	4	2	32.3	-25.3	2	1	4	2

35.3	-23.8	2	1	4	2	32.8	-25.3	2	1	4	2
35.3	-23.3	2	1	4	2	33.3	-25.3	2	1	4	2
35.3	-22.8	2	1	4	2	33.8	-25.3	2	1	4	2
35.8	-22.8	2	1	4	2	31.8	-26.8	2	1	4	2
35.3	-22.3	2	1	4	2	31.8	-26.3	2	1	4	2
35.8	-22.3	2	1	4	2	31.3	-26.8	2	1	4	2
35.3	-21.8	2	1	4	2	31.3	-26.3	2	1	4	2
35.3	-21.3	2	1	4	2						

Chapter 3

Assessment of the performance of drought indices for explaining crop yield variability at the national scale: Methodological framework and application to Mozambique

Ronnie J. Araneda-Cabrera, María Bermúdez and Jerónimo Puertas.

Journal: Agricultural Water Management

Date: 16 December 2020

DOI: https://doi.org/10.1016/j.agwat.2020.106692

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – original draft, Visualization. María Bermúdez: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Resources, Writing – review & editing, Supervision.

Abstract

Droughts are one of the most damaging and complex natural disasters in the world, and they frequently affect agricultural production. Drought monitoring is essential for decision-makers seeking to minimize the socio-economic impacts related to drought events. In this study, we propose a methodology to identify the most suitable drought indices and data sources for monitoring the impact of drought on crops. Mozambique is used as a case study, as it represents a challenging example because of its poor hydroclimatic monitoring network and a lack of disaggregated data for agricultural production. A total of seven standardized drought indicators (SPI, SPEI, SSI, SVCI, STCI, SVHI, and STWS) at different scales (1, 3, 6, and 12 months) were obtained from global databases and evaluated as possible predictors of the annual variability of agricultural yields at the national level. A statistical model of crop yields based on time series was used to measure the explanatory capacity of each index. SPEI and SSI were the most effective at detecting the country's historical drought records regardless of whether nationally averaged values or the percentages of area affected by drought (PAA) were used. However, PAA was found to be a more accurate predictor of variability in crop yields. The variability of most cereals (maize, millet and sorghum) was adequately explained by the PAA of SPEI-3, with that of other crops (cashew nuts, cassava, potatoes, tea, tobacco and vegetables) being explained by the PAA of SSI-12. Specific indicators were proposed for monitoring wheat and sugar cane. These results can directly support managers and decision makers in developing drought contingency plans in Mozambique. To further demonstrate the potential of this methodology, it should be tested in other regions with a greater availability of agricultural data, including spatial disaggregation.

Keywords

Drought index, drought impacts, crop yield, statistical model, Mozambique.

Highlights

- A method to identify indices (DIs) for monitoring drought impact on crops is proposed.
- SPEI and SSI were the DIs that best detected historical droughts in Mozambique.
- Percentage area affected by drought (PAA) better indicates crop yield variability than nationally averaged DIs values.
- PAA of SPEI-3 and SSI-12 explained the variability of most crop yields.
- The usefulness of regional disaggregated data for such impact analysis is recognized.

1. Introduction

Droughts represent one of the most extensive, costly, recurrent, and complex types of natural disasters worldwide (Bryant et al., 2005; Mishra and Singh, 2010). Given that it is related to water availability, the agricultural sector is especially sensitive to this natural phenomenon. Negative impacts, like a decrease in quantity and/or quality of crops, directly affect food security and consequently, the quality of life within a region or country (Backlund et al., 2008). Such impacts are more perceptible in high drought risk regions (e.g., Southern Africa) and rainfed agriculture systems (Tigkas et al., 2019).

Droughts are classified into four widely accepted types: Meteorological, agricultural, hydrological, and socio-economic (Wilhite and Glantz, 1985). Each type is typically characterized and described through drought indices (DIs) (Hayes et al., 2011); several DIs have been developed over the last century (World Meteorological Organization and Global Water Partnership, 2016). Most DIs require climatic and/or hydrological data with at least 30 years of observation as inputs for a reliable temporal drought analysis, whereas at least one data station per 5000 km2 is recommended for spatial analyses (AghaKouchak et al., 2015).

In recent decades, different institutions have constructed and updated various large-scale climate and hydrological data sets (e.g. Abatzoglou et al., 2018; Beck et al., 2017a; Harris et al., 2020; Thomas et al., 2014). These products provide gauge-based, satellite-derived, or reanalysis-based estimates, and can constitute a suitable alternative for calculating DIs in data-scarce regions (e.g., Nashwan et al., 2020). Though these global products are not without limitations (Beck et al., 2017b; Sun et al., 2018), they have been used in several drought and agriculture-related studies worldwide and constitute a reliable data source (Agutu et al., 2017; Champagne et al., 2019; Du et al., 2018; García-León et al., 2019; Jayanthi et al., 2013; Lawal et al., 2019; Potopová et al., 2020; Rojas et al., 2011).

Drought studies related to agriculture usually use correlation tests and statistical models to explain the relationship between farming yields and drought indicators (Shi et al., 2013). These models can be based on a single point or area (time series methods), spatial and temporal variations (panel methods), or solely spatial variations (cross-section methods) (Lobell and Burke, 2010). The type of model can be chosen depending on the spatial and temporal detail of the crop yield series (Shi et al., 2013). García-León et al. (2019) and Peña-Gallardo et al. (2019a) researched the relationships between drought indices and crop yields at the provincial and regional scale in Spain using panel methods. For the same purpose, Jayanthi et al. (2013) used the time series method in Malawi because they only had data on agricultural yields at the national level. Regardless of the method chosen, crop yields are frequently subject to a detrending process to extract the yield trend and remove the variability in productivity caused by non-climate factors (e.g., improvements in farming techniques, seed hybrid development, and irrigation optimization) (Champagne et al., 2019; Peña-Gallardo et al., 2019b), before developing the statistical model.

The above studies showed that site- and crop-specific studies are required to identify the most suitable DI and data sources for monitoring the impact of drought on crops. However, there is currently no standard methodology to assess the performance of DIs in explaining crop yield variability. To fill this research gap, in this study we propose a methodological framework to be applied at the national scale, comprising three steps. First, we use global gauge-based and satellite-derived datasets to calculate several well-known DIs and validate them with historical drought records. In addition to nationally average drought indicator values, new aggregated descriptors based on the areas affected by drought are considered to better capture regional variability when working at this scale. Second, we analyse the time variability among DIs to identify those that are strongly correlated and potentially provide redundant information. Then, we develop a time-series based statistical crop model to predict national yields using DIs as predictors. Building on the above analysis, we identify a DI or set of DIs, as well as the data sources to compute them, that can be used to monitor crop yields.

The proposed methodology is applied to Mozambique as a case study, with an emphasis on determining a climatic drought-related explanation of the national crop yield variability. The case study is challenging because this country does not currently have an operational measurement network that meets optimum criteria (Easterling, 2013), data on agricultural yields are limited, and previous research on drought impacts is scarce. However, it is also of considerable interest because of the role played by agriculture in the sustainable development in this region and the concerns raised by climate change. Mozambique is one of the poorest and least developed countries in the world. Approximately 70% of the population works in agriculture, representing 24% of the GDP (Ministério da Agricultura e Segurança Alimentar, 2015). It is also located in one of the most drought- and climate change-prone areas (Eriksen and Silva, 2009; 2014: Climate Change IPCC, 2014; Osbahr et al., 2008; Patt and Schröter, 2008), which increases the vulnerability of its agricultural sector. For these reasons, this country is immersed in several development programs led by the Food and Agricultural Organization (FAO) (Midgley et al., 2012) and the World Food Programme (WFP) (WFP, 2007), among others, aimed at implementing climate change adaptation strategies to enhance the resilience and sustainability of agriculture. It is in this context that this work explores the impact of drought on agricultural production in Mozambique.

The overall aim of this study is to develop a methodology of general applicability to identify the most suitable DIs and data sources for monitoring the impact of drought on crops at national scale. Two methodological aspects are worth noting: (1) the reliance on freely global-scale datasets to obtain a comprehensive set of potential DIs, and (2) the exploration of alternatives to averaging DI values over the entire country, that are intended to better capture the local and regional drought conditions when working at such scale. The proposed methodology may thus be of special interest in countries with data scarcity, where ground data observations are neither sufficient nor timely available for drought monitoring, and countries with significant regional variability in drought occurrence, where nationally average drought indicator values may conceal regional differences. The methodology is tested in Mozambique. To the knowledge of the authors, this is the first study that compares drought indices for an agricultural drought risk assessment in this country. This study endeavours to act as a tool for supporting decision-makers, focusing on the performance of DIs in explaining yields and yield variability at the national scale. The results can help assess drought-related risks to crop production and are ultimately intended to contribute to developing an agricultural drought monitoring system in this country. As such, we acknowledge the need to reconcile the demands for highly detailed analysis with the extent of the resource requirements (infrastructure, operational needs, etc.) and data availability.

2. Materials and methods

2.1.Study area

Mozambique lies in southeast Africa (Fig. 3.1a) and covers a continental area of 801.590 km². The weather system is dominated in the north by the Inter-Tropical Convergence Zone and in the south by Antarctic Polar Fronts and Tropical Temperate Troughs (Manhique et al., 2011). The climate is tropical, with a hot and rainy summer season from November to March, and a cool and dry winter season from April to October (Midgley et al., 2012). The annual average temperature varies from 17.8°C to 32.8°C (Fig. 3.1e), and has increased by 1.25–2.0°C in the last 60 years (Ragab and Prudhomme, 2002). The annual average precipitation is 1032 mm, 75% of which occurs during summertime.

Owing to these conditions, the sowing and harvesting season generally extends from November to April (Rojas et al., 2011). Farming is one of the main activities in the country (70% of the population depending on subsistence farming), with over 80% of the total cultivated area used to produce staple food crops. Because of the lack of hydraulic infrastructure for irrigation, over 95% of this agricultural production is mainly rainfed and without fertilizer consumption (FAO, 2016a).

Fig. 1. a) Location of Mozambique in Africa and its topography. Black dots illustrate the CRU grid points (0.5°×0.5°). Spatial distribution of annual mean values of: b) precipitation (1973–2017), c) potential evapotranspiration (ETP, 1973–2017), d) Normalized Difference Vegetation Index (NDVI, 1983–2017), e) Brightness Temperature (BT, 1983–2017), f) Soil moisture (1973-2017), and g) Terrestrial Water Storage (TWS, 2002–2017), across the country.

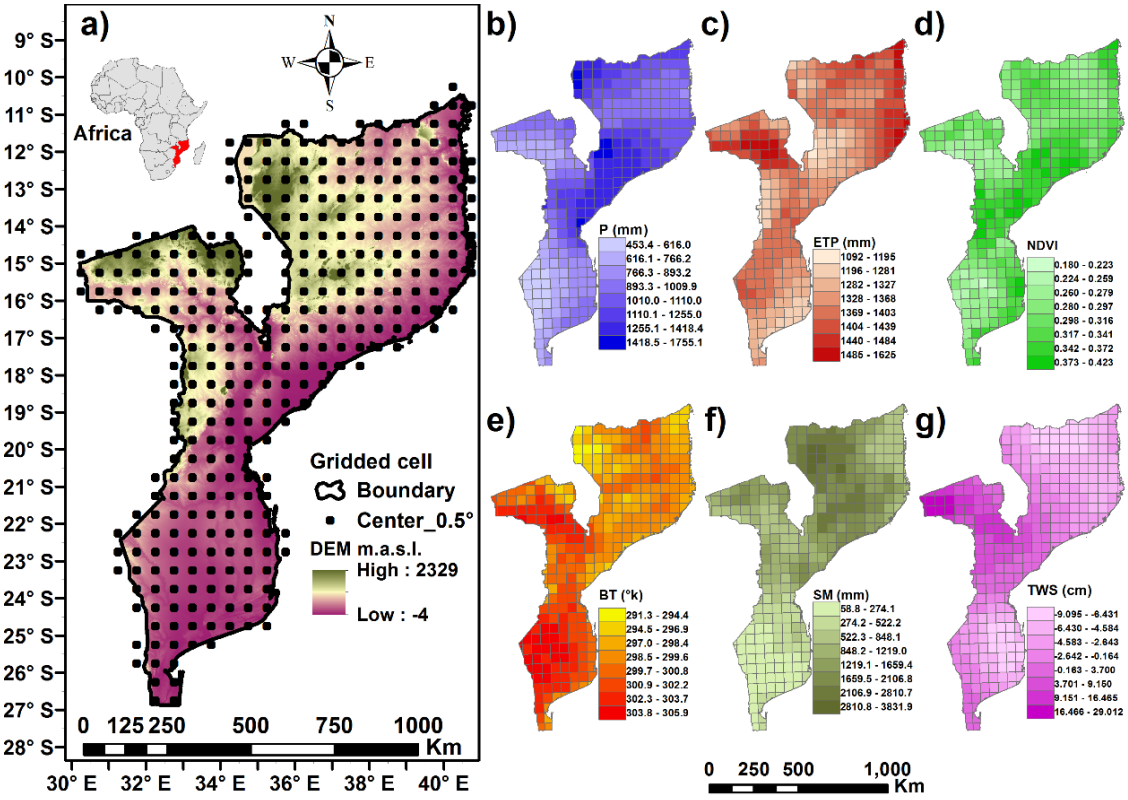


Fig. 3.1 The country is prone to drought, which has caused temporary food insecurity in the past (FAO, 2016a). According to the International Disaster Database (EM-DAT, 2021) and the International Research Institute for Climate and Society (IRI) (Hellmuth et al., 2007), Mozambique has experienced various annual and interannual drought episodes in recent decades. In terms of socio-economic impacts, the most important were the droughts that occurred in 1979–1980, 1983–1984, 1987, 1991–1992, 1994–1995, 1998, 2001–2003, 2005, 2007–2008, 2010, and 2016.

2.2.Meteorological, hydrological and vegetation data

The data used for calculating the DIs (Table 3.1) comprises Precipitation (P) and Potential Evapotranspiration (ETP) as meteorological information, the Normalized Difference Vegetation Index (NDVI) and Brightness Temperature (BT) related to vegetation conditions, and the Soil Moisture (SM) and Terrestrial Water Storage (TWS) as hydrological measures.

Monthly P and ETP were obtained from Climatic Research Unit CRU TS3.10 (CRU) at the University of East Anglia (Harris et al., 2020) for the period between 1973 and 2017 (https://crudata.uea.ac.uk/cru/data/hrg/) at a 0.5° resolution. A total of 343 CRU grid points covering the entire Mozambican territory were used for the study (Fig. 3.1).

NDVI and BT are derived from spectral reflectance at the blue, red, and near-infrared (NIR) wavelengths observed from space by orbiting satellites (Deering, 1978). These data were obtained from the Center for Satellite Applications and Research (STAR) and the environmental satellites for the U.S. Oceanic and Atmospheric Administration (NOAA). Datasets consist of 7-day value composites at 8 km resolution from 1982 to 2017 (https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php).

Monthly SM time series from 1973 to 2017 and a 1/24° resolution were obtained from the TerraClimate dataset (https://climate.northwestknowledge.net/TERRACLIMATE/index_directDownloads.php) (Abatzoglou et al., 2018). These values have been derived using a one-dimensional soil water balance model based on the primary climate variables of this dataset.

The GRACE satellites can accurately observe and measure the TWS changes over global land areas (Tapley et al., 2004). Monthly TWS data at a 1° resolution (https://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/) from 2003 to 2016 were used in this study. The gridded TWS data were scaled following the process explained by Landerer and Swenson (2012).

To establish consistency with the CRU spatial and temporal resolution, the NDVI, BT and SM data were resampled to a 0.5° resolution and monthly scale. Four-point bilinear resampling was applied to TWS data for the same reason. The spatial distribution of the annual mean values of each variable for the study period is shown in Fig. 3.1b-g.

2.3.Drought indices and area affected by droughts

The drought indices (DIs) used in this study (Table 3.1) rely on the meteorological, vegetation condition, and hydrological data described in the previous section. Although several DIs exist (Dai, 2011; World Meteorological Organization and Global Water Partnership, 2016), we selected seven widely known DIs that have been successfully used in drought and agriculture-related investigations (Agutu et al., 2017; Sun et al., 2012).

Table 3.1 Summary of drought indices and data sources used in this study. The temporal accumulations (n) were 1, 3, 6, and 12 months.

Index	Input	Data source	Original temporal resolution	Original spatial resolution	Time span
SPI-n	P	CRU	Monthly	0.5º	1973–2017
SPEI-n	P, ETP	CRU	Monthly	0.5º	1973–2017
SSI-n	SM	TerraClimate	Monthly	1/24º	1973–2017
SVCI-n	NDVI	NOAA STAR	Weekly	1/24º	1982–2017
STCI-n	BT	NOAA STAR	Weekly	1/24º	1982–2017
SVHI-n	VCI, TCI	NOAA STAR	Monthly	0.5º	1982–2017
STWS-n	TWS	GRACE	Monthly	1º	2002–2017

The Standardized Precipitation Index (SPI-n) (McKee et al., 1993) shows precipitation anomalies with respect to the long average of P, considering a window period of n months. Its computation consists of fitting P time series accumulated at a chosen n-months period with a two-parameter gamma probability distribution and later standardizing this result. The Standardized Precipitation and Evapotranspiration Index (SPEI-n) (Vicente-Serrano et al., 2010), the Standardized Soil Moisture Index (SSI-n) (Hao and AghaKouchak, 2013), the Standardized Vegetation Condition Index (SVCI-n), the Standardized Temperature Condition Index (STCI-n), the Standardized Vegetation Health Index (SVHI-n) and the Standardized Terrestrial Water Storage (STWS-n) (Agutu et al., 2017) were calculated following the same mathematical procedure as the SPI, but used the corresponding variables as inputs rather than P. The SPEI-n uses the difference between P and ETP, which is understood as the climatic water balance. SSI, SVCI-n, STCI-n, and SVHI-n used the SM, VCI, TCI, and VHI, respectively, as previously calculated according to Kogan (1995), whereas the STWS-n used the TWS. The SPEI-n and STWS-n were fitted to a three-parameter log-logistic probability distribution following Vicente-Serrano et al. (2010). The seven indices were calculated for four temporal accumulations (n) of 1, 3, 6 and 12 months, considering that farming periods for selected crops are less than one year. The indices were evaluated

according to the categories explained in Table 3.2 for standardized indexes, following the recommendations of McKee et al. (1993).

We computed the DIs time series for each CRU cell and then aggregated them into a single national time series. For this purpose, we averaged the DIs values over the entire territory (343 cells), weighted by cell area in the country (a), to obtain a single national monthly time series for each DI. We also calculated the percentage area affected by droughts (PAA) by considering the intensity thresholds indicated in Table 3.2. We determined the percentage of cells under the different intensity categories using the DIs time series for each cell on a monthly scale. The annual PAA was obtained through aggregating from a monthly to annual series over the calendar year. The idea behind the calculation of the PAA is to evaluate alternatives to national averaged DI values that can better capture, in a single value, subnational differences in drought conditions. The spatial averaging performed for computing the national DI values can, in some cases, hide extreme drought conditions occurring at a more regional scale, which is not the case with the PAA computation. The time span covered by the DIs and PAA series depended on the availability of main data, as indicated in Table 3.1.

Table 3.2 Intensity categories of droughts used in this study for standardized indices.

Standardized Indices	Category
< -2.00	Extreme drought
< -1.50	Severe drought
< -1.00	Moderate drought

2.4.Crop yield data

The country-level annual crop yield data (Y) for Mozambique was obtained from the FAO data portal (http://www.fao.org/faostat/en/#data/QC) for the period of 2002–2017 (Fig. 3.2). Crops included maize, millet, sorghum, wheat, cashew nuts, cassava, potatoes, sugar cane, tea, tobacco, and vegetables. According to Kasnakoglu and Mayo, (2004), this data source represents one of the most credible, readily available Y dataset because of its monitoring data quality and statistical process. More recently, Agutu et al. (2017) used it successfully in a nearby region to characterize agricultural drought.

Crop statistics are not routinely compiled in Mozambique at the sub-national level (e.g., by agroecological zones or provincial levels). The availability of sub-national production statistics is, in fact, very limited in the majority of sub-Saharan African countries (You et al., 2009). At present, regional yields in Mozambique are available for around 8 non-consecutive years (depending on the specific crop and region) from sources like the FAO Agro-MAPS database (http://kids.fao.org/agromaps) (George, 2006) or the Agricultural Statistics Yearbooks (Ministério da Agricultura e Segurança Alimentar, 2015). However, these data are insufficient to support the type of analysis proposed in this paper and were, therefore, not considered in this study.

Mozambique has only recently started to improve its farming techniques. This could be attributed to the social, economic and political issues Mozambique has faced in recent decades, causing limited developments in agriculture (FAO, 2016a). However, we decided that it might be necessary to apply a detrending process to the crop data in order to eliminate variations that may result from abiotic factors (market prices, government policy, etc.). Y trends were calculated using the Mann-Kendall (MK) trend test method (Mann, 1945): Millet, cashew nuts, cassava, potatoes, sugar cane, tobacco, and vegetables had positive and significant trends; tea showed a negative and non-significant trend, whereas the other crops had positive and non-significant trends. The crop yield series were detrended by fitting a linear regression model and extracting the residuals. The average crop yield was added to the residual series to produce the detrended yield data. The analyses shown in this paper were conducted on the detrended

data series, and the detrending process was applied only for Y with significant trends (hereafter detrended crops yield are simply referred as Y).

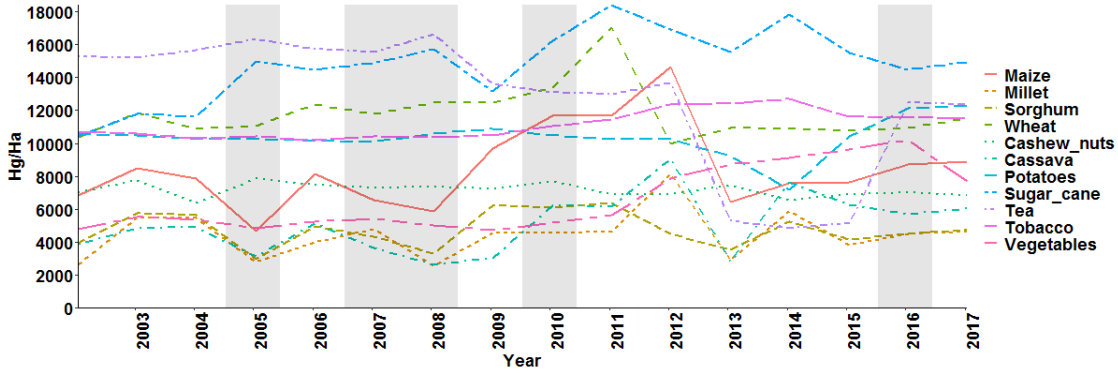


Fig. 3.2 Time series of the main yield crops in Mozambique after applying a detrending process (Source: FAO data portal http://www.fao.org/faostat/en/#data/QC). Drought years according to the records are shaded in grey. NOTE: Yields of cassava, potatoes and vegetables are divided by 10, and that of sugar cane is divided by 40 for visualization.

2.5.Skill assessment and benchmarking of drought indices

In the first step of the proposed methodology, the performance of the different drought indices for drought detection was analysed. For this purpose, the historical drought record from EMDAT and IRI was compared with the time series of the average national DIs values and the annual PAA for each DI. The EMDAT is part of the Centre for Research on the Epidemiology of Disasters (CRED), which initiated its active disaster data collection in 1973 (Guha-Sapir et al., 2015); hence, skill assessment was conducted from 1973 to 2017 on a yearly basis. In this work, a year qualified as a drought year if at least two consecutive months were under moderate drought intensity (Table 3.2) according to the DIs series. This two-month criterion allows the exclusion of short droughts, which are presumably of minor importance, as done in previous studies (Spinoni et al., 2019). In the case of the PAA series, a drought year was designated if the annual PAA value reached 30%. Although this is an arbitrary threshold, the percentage of land impacted by drought is directly related to the number of agricultural households affected, and therefore with the drought impact on productivity (Rojas et al., 2011). The performance metrics used were the probability of detection (POD) or hit rate, and the probability of false detection (POFD) or false alarm rate, computed according to Wilks (2006):

$$POD = 100\frac{H}{H+M} \quad (1)$$

$$POFD = 100\frac{FA}{CN+FA} \quad (2)$$

where H are hits, M misses, FA are false alarms, and CN are correct negatives. H represents the coincidence of drought of both series (historical records and DIs/PAA series), and M corresponds to the presence of a drought in the records and the absence of this event in the DIs/PAA series. FA occurs if there is no drought in the records, but one occurs in the DIs/PAA series, and finally, CN represents the years in which there is no drought in both series. The Euclidean distance between the point (x_1=POFD, y_1=POD) of each DI and PAA series and the point with the best possible performance (x_2=POFD=0, y_2=POD=100) was used to benchmark performance. Shorter distances thus indicated better performance.

A correlation analysis between the DIs series was performed in the second step of the methodology. We calculated the Pearson correlation coefficients between all monthly DIs series averaged at the country scale. The correlation coefficients were analysed to assess if the information provided by the different DIs was redundant.

2.6.Statistical crop yield model

In the third step of the methodology, we developed a statistical crop yield model for Mozambique by assuming that Y was the response of a function of k independent variables X which, in this context, included the DIs and PAA as possible predictors. Given $Y \in [0, \infty)$, and following Shi et al. (2013), an exponential (time series) model was adopted as follows:

$$\ln(Y_t) = f(X_{t1}, X_{t2}, ..., X_{tk}) = \beta_0 + \sum_{j=1}^{k} \beta_j (X_{tj}) + \varepsilon_t \quad (3)$$

where Y is the vector of annual crop yields (Fig. 3.2), *t* is the year, X represents the vector with the candidate predictors, β_j are the constant coefficients and ε the error.

Both single and multiple candidate predictors are considered to develop the models. Equation (3) transforms in a simple linear regression model with one explanatory variable or in a multiple linear regression, respectively. The times series considered as candidate predictors for the models are: 1) The national average DIs of all months of the year (January to December), 2) the annual national average DIs, 3) the annual average PAA under a certain intensity category (moderate, severe, or extreme). The multiple linear regression models use the annual average PAA under the three aforementioned intensity categories as candidate predictors (i.e., three independent variables). Fig. 3.3 summarizes the steps in a methodological flow-chart.

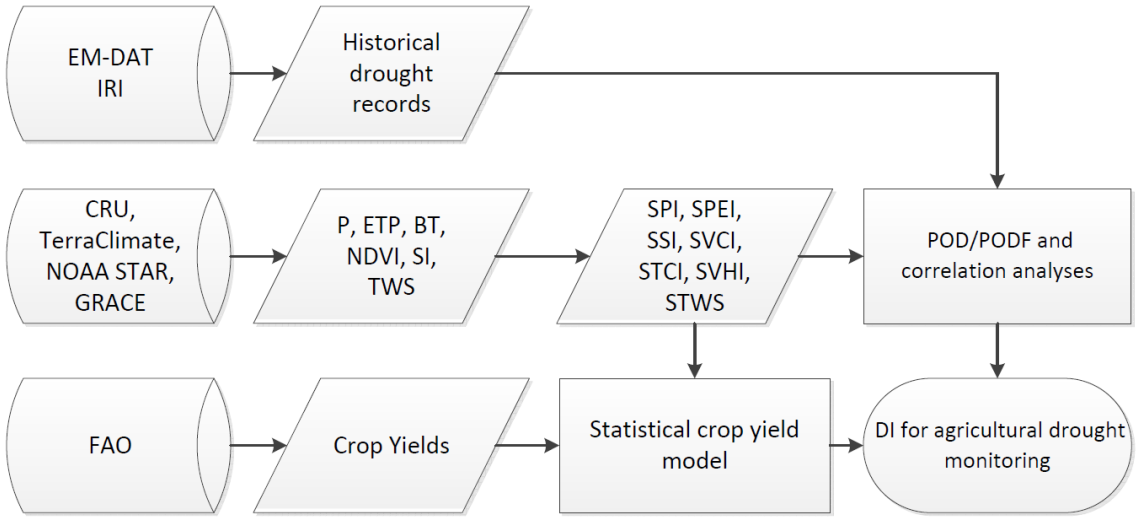


Fig. 3.3 Methodological flow-chart of the study.

3. Results and Discussion

3.1. Comparison with historical drought records

The temporal patterns of the national averaged DIs series are plotted in Fig. 3.4. Each DI detected several droughts of different intensity categories: Moderate (green), severe (orange), and extreme (red). The longer the temporal accumulation (n) for each DI is, the later the dry periods are detected, and they are also less frequent. This result can be explained by the time scale of the DI and the drought propagation through the hydrological cycle. The historical droughts of 1987, 1991–1992, 1994–1995, 2005, and 2016 were the main drought events detected by the majority of DIs. According to the DI linked to soil moisture (SSI), an extreme intensity was reached in up to four of these events. The DIs related to vegetation conditions (SVCI, STCI, and SVHI) also classified the 1991–1992 event as extreme. The remaining DIs (SPI, SPEI, and STWS) distinguished several drought events (including the five named above), but the intensities were lower (moderate and severe). In these DIs, extreme intensities were completely smoothed out by the spatial averaging process.

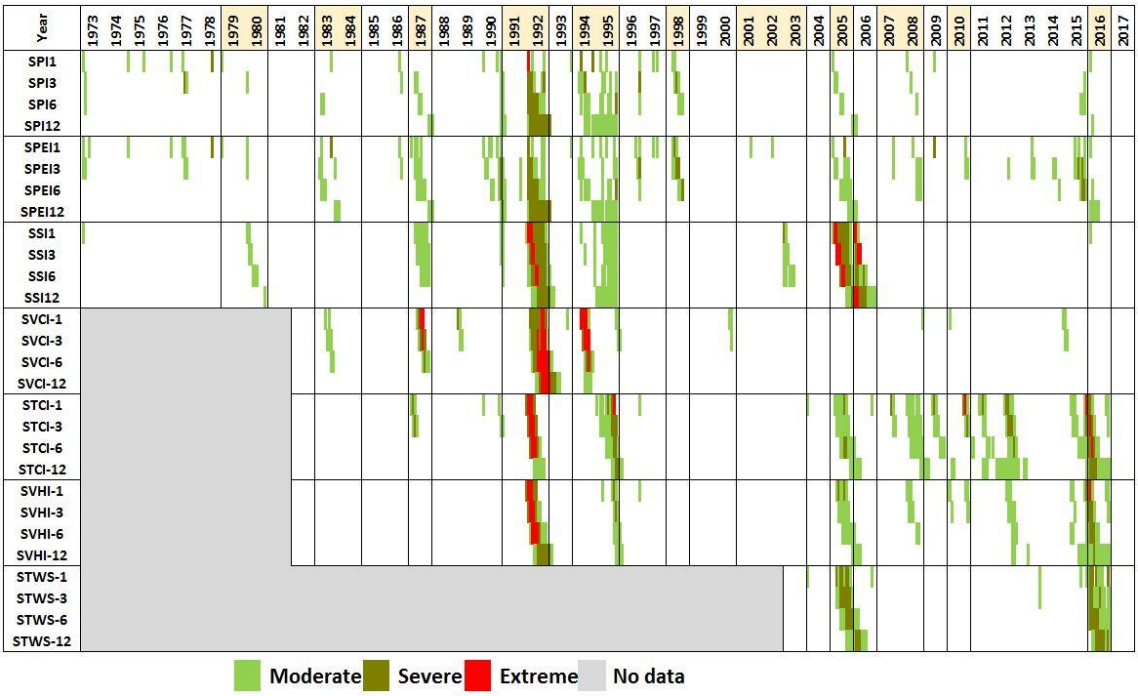


Fig. 3.4 Monthly temporal evolution of SPI, SPEI, SSI, SVCI, STCI, SVHI, and STWS at the national scale (-1, -3, -6, and -12-month aggregations). Intensity levels can be interpreted in conjunction with Table 2. Historical drought years according to the records are highlighted in yellow.

STCI-n and STWS-n indicated drought conditions in Mozambique between 12.0% and 14.1% during the analysed period. SPI-n and SVCI-n reported the shortest time under drought conditions (less than 7.2%). When the extreme drought threshold was considered, SCVI-n indicated that such conditions were present less than 2.1% of the time, whereas less than 0.4% of the time was classified in this intensity category, according to SPI-n.

Similar to the DIs time series observations, the historical droughts in 1987, 1991–1992, 1994–1995, 2005, and 2016 were the events that were most clearly detected by majority of the PAA series. These drought periods have also been highlighted in previous studies (Brida and Owiyo, 2013; Jayanthi et al., 2013; Trambauer et al., 2014). In these significant events, the total area affected by drought and its distribution in intensity categories differed between DIs. In 1992, up to 74 % of the Mozambican

territory was below the moderate drought threshold, and 53 % reached the extreme intensity according to SPEI-12. In 2016, STWS-6 showed that 90 % of the country was facing a moderate drought, 5% of which corresponded to an extreme intensity; however, SPEI-12 indicated a PAA greater than 20% in the extreme category.

To benchmark the capacity of detection of historical drought events by the different indices, the POD and POFD metrics were computed, as described in Section 2.5. Table 3.3 shows the skill assessment of the national averaged DIs series and PAA series according to the distance (d) to the point with the best possible performance. The rankings of DIs according to their performances were quite similar for both series (DIs and PAA). In both cases, the best performance was obtained by SPEI-3 and SPEI-6, with more than 50% of POD and with POFD ranging from 7% to 26%. Similar performances were obtained with these two indices when considering thresholds for drought detection between 20 and 35 % of the annual PAA value, demonstrating that the indexes are robust with respect to slight variations in the definition of a drought year. These results are consistent with previous studies that found that EMDAT drought disasters were best matched with severe droughts identified using meteorological DIs for mid-high temporal accumulations (United Nations, 2009). The worst detection performance was provided by the DIs related to vegetation conditions (SVHI and SVCI). Nonetheless, all the DIs detected the most important drought events because they affected the entire hydrological cycle.

Table 3.3 Skill assessment results according to POD and POFD. The DIs not included in the table have distances (d) to the point with the best possible performance larger than 62.

DIs	H	M	FA	CN	POD	POFD	d	PAA	H	M	FA	CN	POD	POFD	d
SPEI-6	9	9	2	25	50.00	7.41	50.55	**SPEI-3**	9	9	2	25	50.00	7.41	50.55
SPEI-3	10	8	7	20	55.56	25.93	51.45	**SPEI-6**	9	9	2	25	50.00	7.41	50.55
STCI-3	8	8	4	16	50.00	20.00	53.85	**SPEI-12**	8	10	2	25	44.44	7.41	56.05
SPEI-12	8	10	1	26	44.44	3.70	55.68	**SSI-3**	8	10	2	25	44.44	7.41	56.05
STCI-1	8	8	5	15	50.00	25.00	55.90	**SSI-6**	8	10	2	25	44.44	7.41	56.05
SPI-6	7	11	1	26	38.89	3.70	61.22	**SPI-6**	7	11	0	27	38.89	0.00	61.11
SSI-1	7	11	1	26	38.89	3.70	61.22	**SPI-12**	7	11	1	26	38.89	3.70	61.22
SSI-3	7	11	1	26	38.89	3.70	61.22	**SSI-1**	7	11	1	26	38.89	3.70	61.22

The national averaged DIs series are likely missing some climatic, hydrological, or vegetative information of specific regions within the country. For example, drought conditions that affect an area of the country might be concealed by wet conditions in another region. This is not the case for the PAA series, which a priori makes them a better option for working at a national scale. However, both the DIs and the PAA series showed a similar detection capability of historical records. Overall, the meteorological indices (SPEI and SPI) most closely matched the historical records. However, none of the DIs were able to perfectly capture all the drought periods collected from historical records (i.e., POD lower than 100), and they identified drought conditions outside the drought years recorded in the historical disaster databases (i.e., POFD greater than 0).This may be attributed to the following reasons: On the one hand, unrelated circumstances might have worsened the consequences of what would otherwise be considered as a mild drought, being recorded as drought year in the records. For example, a) other natural disasters occurring between dry periods, such as the heavy floods in 1981 and 1985 and between 2000 and 2001 (Brida and Owiyo, 2013; Midgley et al., 2012; Patt and Schröter, 2008), b) civil war and conflicts, such as the conflict from 1982–1984 against Zimbabwe (Hellmuth et al., 2007), c) epidemics, such as the cholera outbreak in 1983–1984 (Eriksen and Silva, 2009), or d) food security crises, as in the 2001–2005 period (FAO, 2006). On the other hand, the drought events in Mozambique were reported at the national level, even if they occurred only in a specific region of the country. Therefore, the national DIs series obtained after averaging at national level and the PAA may not adequately reflect regional information.

3.2.Correlation between drought indices

The occurrence, intensity, and duration of drought periods in Mozambique showed some variability between the different DIs. In this step, we compared the drought variation patterns between each DI. Fig. 3.5 shows the Pearson coefficients between the DIs time series; the strongest correlations are painted blue, and non-significant (ρ>0.05) correlations are denoted as struck through.

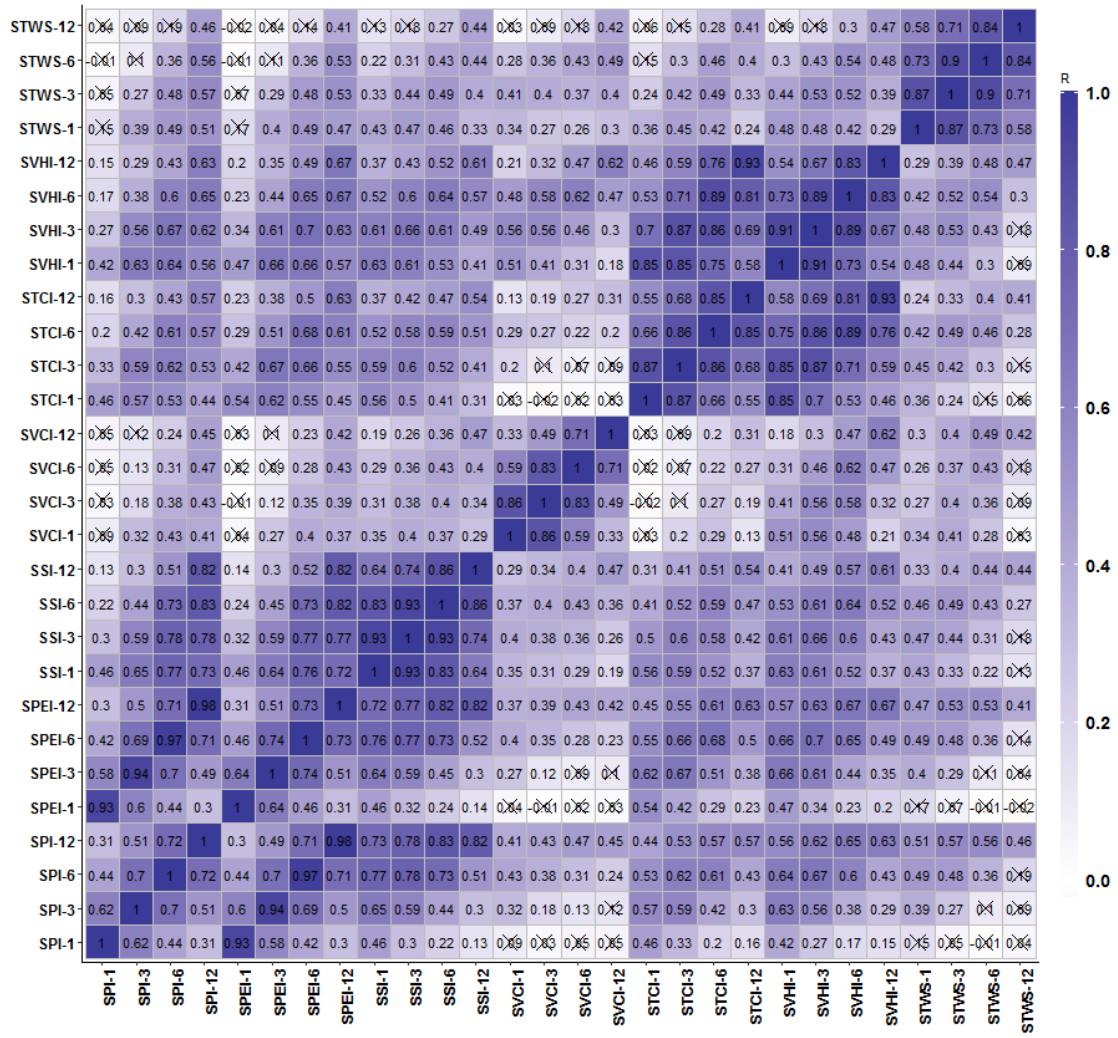


Fig. 3.5 Correlation coefficients (R) between the DIs, where the non-significance level (ρ>0.05) is indicated by strikethrough.

Overall, the strongest correlations were found between the indices associated with meteorological variables (SPI and SPEI) and between those related to soil moisture (SSI). These were followed by those associated with soil conditions and temperature (STCI and SVHI). In general, a DI with a temporal accumulation of 1 month had a very high or nearly perfect correlation with the same DI at a 3-month scale. The same was observed between the temporal aggregations of 3 and 6 months, and between 6 and 12 months.

The indexes associated with the vegetation condition, albeit without meteorological variables such as temperature in their calculation (SVCI), as well as STWS exhibited considerably different variabilities

from the rest of DIs, with low correlation values. These two DIs also had the most non-significant correlations with other DIs. The SVCI is an index that indicates the condition of the vegetation, and thus can reflect other factors external to the climate (e.g., irrigation, forestation, and afforestation). STWS, on the other hand, characterizes the total availability of surface and underground water, so its variation depends on the complete hydrological process of the area and not solely on one or two climatological parameters. The lowest correlations were found for the highest temporal accumulation (SWT-12), which could be related to multiannual droughts (not analysed in this paper).

The results of the correlation analyses are consistent with the physical meaning of the different DIs. Although there is a time lag between the meteorological forcing and the hydrological responses, DIs based on meteorological variables with 6- and 12-month accumulations showed strong correlations with DIs based on vegetation/hydrological data. The results suggest that only one of these meteorological DIs (SPI and SPEI with n=6 and n=12 months) could be used alone in any subsequent analysis, as they provided very similar information and had high correlations with other DIs. SCVI may be used but considering that it does not have an important correlation with hydroclimatic indicators, it should be used in conjunction with another DI.

3.3.Explanation of crop yield variability

Because of the close relationship between Y and drought conditions, especially in areas with little hydraulic infrastructure such as Mozambique (non-irrigated agriculture), a time-series statistical model was proposed to find the DI that better explains the annual yield variability of the 11 selected crops, as explained in Section 2.6. The candidate predictors were the averaged national DIs of each month (Jan to Dec), the annual averaged national DIs (annual), the annual PAA at moderate (a_mod), severe (a_sev) and extreme (a_ext) intensities, and the sum of the annual PAA at moderate, severe, and extreme intensities (a_sum), totalling 17 time estimates for each DI. Because 28 DIs were analysed, this created a total of 17x28=476 candidate predictors for each type of crop. Fig. S3.1 (supplementary materials) shows all R^2 values of the statistical model results for each crop. The crop types are represented on the y-axis and the candidate predictors are on the x axis. The strongest positive correlations are plotted as red. Many of the estimates were non-significant ($p < 0.05$).

The main reason for such low coefficients has already been mentioned: there are no regional data on crops. Droughts can, however, affect specific regions of the country that contribute little to the overall national yield. The use of national aggregated estimates in these cases means that we are searching for correlations between droughts affecting one area of the country and yields produced in others. It is also worth noting that this study is focused on rainfed agriculture, but some crops (e.g., vegetables or sugar cane) might be irrigated in some areas, adding noise to the yield data. Mozambique is thus a highly complex scenario. Area-based analysis in countries where disaggregated data are available will certainly improve the results.

However, if a certain number of indicators containing reasonable correlations with crops are determined, this can act as a valuable tool for decision-makers, who currently do not have objective data to guide their policies. As more data become available, the same methodology can be applied to improve drought monitoring.

The best candidate predictors of Y were different for each crop. This is because not all crops are equally sensitive to drought, nor do they have the same water harvesting or storage capacities. In general, the best predictors were those based on the PAA and related to agricultural and hydrological droughts. Being predictors that incorporate spatial information below the national level, they showed a better relationship with Y than national indicators that lost spatial information in their computation. Again, we note that introducing spatial disaggregation improved the results. If, in addition to incorporating spatial

information from the DIs, data for the evolution of crops by area were available, the results would be much more accurate.

Considering that some of the indicators had high correlations between themselves, which meant that their spatial-temporal drought patterns were similar, the determination coefficients resulting from the proposed model were understandably similar between themselves. For example, SPI-6 and SPEI-3 best explained the variability of the sorghum, with R^2 values of 0.70 and 0.67, respectively (with a correlation of 0.70 between them). To limit the number of indicators, the predictive capacities between these two DIs can be considered highly similar; SPEI-3 can ultimately be chosen, as it also provides reasonable results in other crops such as millet and maize. Thus, the proposal of indicators should be compacted to a minimum number that can explain the variability of the 11 crops. Fig. S3.1 (supplementary materials) allows for an analysis of which DI reasonably explains the evolution of each of the crops.

Of the 11 crops analysed, 8 of them showed a high correlation with 2 predictor candidates: SPEI-3_a_sum reasonably explained the evolution of cereals (maize, millet, and sorghum) and SSI-12_a_sum explained the rest of the crops (cashew nuts, cassava, potatoes, tea, tobacco and vegetables) (Fig. 6). This suggests that DIs that consider not solely precipitation but soil moisture conditions, either directly or indirectly, can provide a better assessment of the potential impacts on agricultural production. The differences in the phenological characteristics and the cultivation period of the studied crops can justify the need for considering different DIs (SPEI and SSI) at both short and long timescales (3 months and 12 months). The variability of the cereal yields is typically best explained by short term meteorological DI (Chen et al., 2016; Peña-Gallardo et al., 2019a), whereas other crops such as tubers and vegetables respond to soil moisture drought conditions at a longer timescale (Daryanto et al., 2016; Sorensen, 2005). It should also be noted that the choice of a limited set of indices, for the sake of simplicity and operational use, comes at the cost of more difficult physical interpretation of the results.

There were some crops that did not neatly fit into these general trends. Wheat also responded well to SPEI, as did all other cereals, but SPEI-6_Aug presented a much better fit than SPEI-3_a_sum, used for all other cereals; thus, the former predictor stands out. One possible reason may be a widespread or uniform distribution of wheat crops, where the zoning provided by PAA-based indicators does not offer an added value. This is however a hypothesis, as disaggregated data are not available.

The last peculiarity is that of sugar cane. This crop had acceptable correlations with other indicators, so it could have been included within the general block. However, sugar cane also had an excellent correlation with the SVCI-6_a_sum indicator, as demonstrated in Fig. 3.6, so the best-fit DI was selected in this case. It is difficult to estimate the reason for the adjustment, although Lisboa et al. (2018) highlighted the good strong predictive capacity of sugarcane yields using NDVI data. The modelled and measured Y are compared for all crops in Fig. 3.6.

The lack of spatially detailed agricultural data made Mozambique a difficult place to find relationships between Y and DIs. Nevertheless, the method applied here discovered relationships that adequately explained the variance of Y for some crops, proving that it can be used in other regions or countries. However, as these are national approximations for a considerable territory, these results should be considered with care.

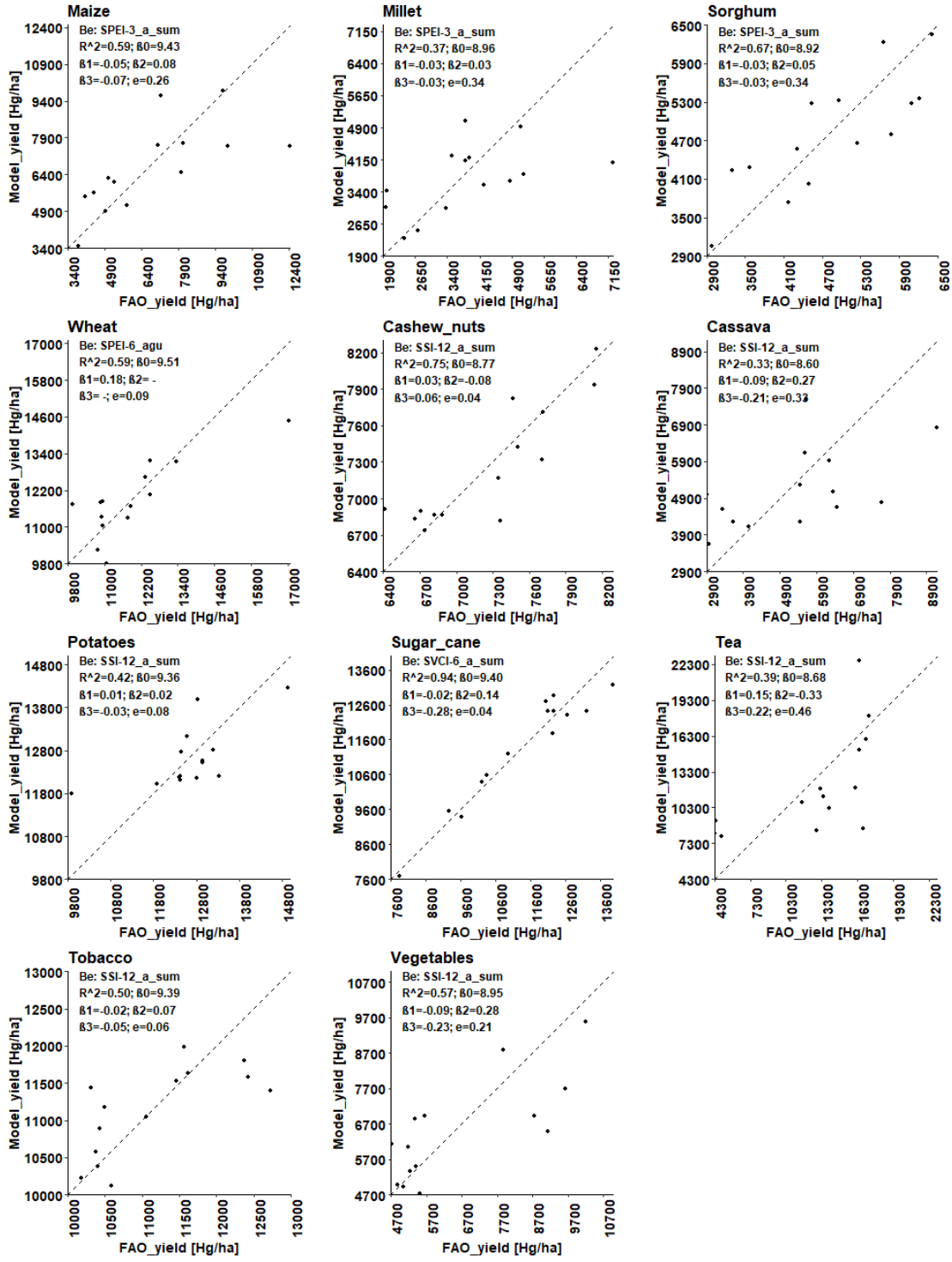


Fig. 3.6 Crop yields as measured by FAO versus crop yields as calculated using the best explanatory variable candidate (indicated as Be). Fitted parameters are also shown. The dashed line corresponds to the 1:1 line.

4. Concluding remarks

The aim of this research was to test a methodology that could evaluate various drought indices as tools for monitoring drought and to test their ability to explain the annual variability of crop yields. The case

study was carried out for Mozambique, which was especially challenging given its poor monitoring system and lack of local/regional data on crop yields; thus, data from global databases and national agricultural data were used.

The proposed indicators successfully detected the main drought events from 1973 to 2017 according to historical records, and accurately noted their duration and intensity. The SPEI and SSI indicators had the best capacity to detect historical droughts through using both nationally averaged time series and the PAA.

Variability in crop yields was associated with agricultural and hydrological droughts. This variability was explained for the majority of crops using two generic indicators: SPEI-3_a_sum explained the performance of various cereal crops (maize, millet, and sorghum), whereas SSI-12_a_sum correlated well with several other crops (cashew nuts, cassava, potatoes, tea, tobacco, and vegetables). Some specific indicators were proposed for two specific crops: SPEI-6 of August (wheat) and SVCI-6_sum (sugar cane). SPEI and SSI also offered the best results in terms of their ability to explain historical drought events. In addition, because the state level agricultural data were used, the annual area affected by drought (PAA used in both SPEI-3_a_sum and SSI-12_a_sum) explained the variance of agricultural yields more effectively than the national level drought indicators.

In summary, the proposed methodology allowed us to confirm the use of these drought indicators - in their different temporal accumulations - as a tool to monitor and characterize droughts and model the annual yields of specific crops in Mozambique. This methodology should be tested in other regions with a greater availability of agricultural data, including spatial disaggregation. These results can be used as a support mechanism by managers and decision makers for drought contingency plans in Mozambique. However, there is a need to deepen the analysis of droughts in this country at the regional level, to provide an improved basis for drought management at the local level.

Acknowledgements

Ronnie Araneda gratefully acknowledges financial support from the Spanish Regional Government of Galicia (Xunta de Galicia) and the European Union through the predoctoral grant (reference ED481A-2018/162). María Bermúdez was supported by the European Union H2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 754446 and the Research and Transfer Fund of the University of Granada - Athenea3i.

Supplementary materials

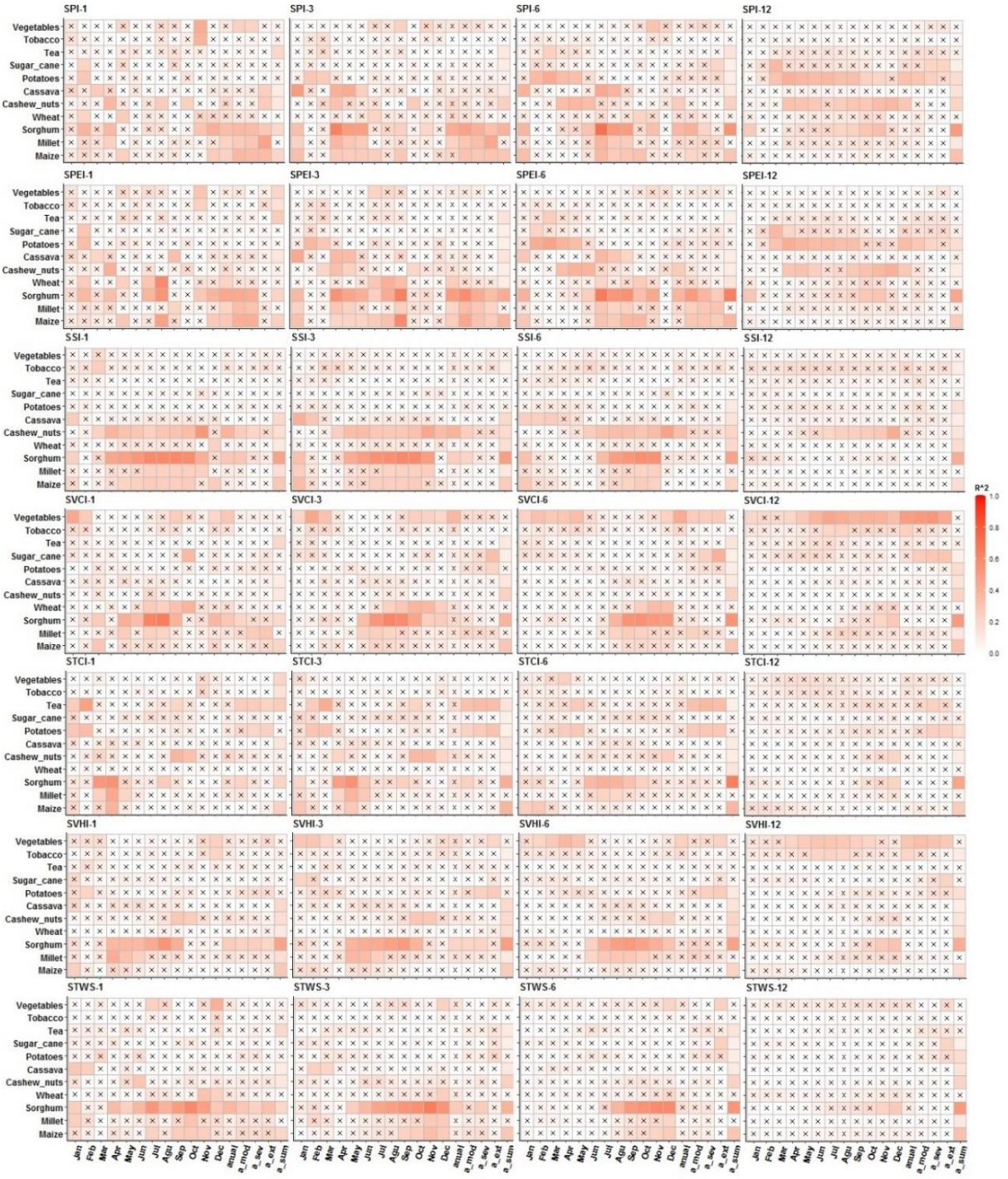


Fig. S3.1 Determination coefficients (R^2) of the statistical crop models. The different candidate predictors (x-axis) represent the monthly DI values (Jan to Dec), the average annual values (annual), and the PAA for the multilinear regression model (a_sum) and each intensity category: moderate (a_mod), severe (a_sev), and extreme (a_ext). The strongest positive correlations are shown in red. Non-significance values ($p > 0.05$) are marked with a cross.

Chapter 4

Benchmarking of drought and climate indices for agricultural drought monitoring in Argentina

Ronnie J. Araneda-Cabrera, María Bermúdez and Jerónimo Puertas.

Journal: Science of the Total Environment

Date: 28 May 2021

DOI: https://doi.org/10.1016/j.scitotenv.2021.148090

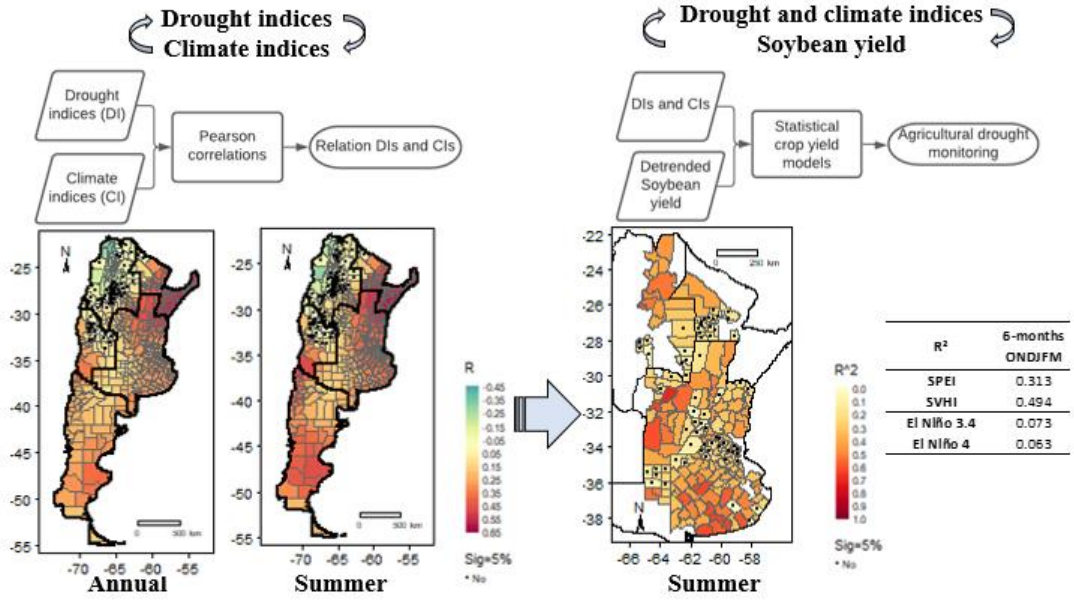
Author contributions:

Ronnie J. Araneda-Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. María Bermúdez: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Resources, Writing – review & editing, Supervision.

Abstract

Site-specific studies are required to identify suitable drought indices (DIs) for assessing and predicting drought-related impacts. This study presents a benchmark of eight DIs and 19 large-scale climate indices (CIs) to monitor agricultural drought in Argentina. First, the link between the CIs and DIs was investigated at the departmental-administrative level and at different temporal scales. Then, the effectiveness of the DIs in explaining the variability of crop yields, understood as impacts of agricultural droughts, was evaluated using statistical regression models. Soybeans were used as the reference crop. Additionally, the performances of DIs and CIs in explaining the variability of crop yields were compared. The CIs located in the Pacific Ocean (El Niño 3.4 and El Niño 4) were found to have the best correlations with the DIs (R values up to 0.49). These relationships were stronger with longer temporal aggregations and during the wet and hot seasons (summer), showing a significant role in the triggering of droughts in Argentina. The DIs that best corelated with CIs were those that included temperature in their calculations (STCI, SVHI, and SPEI). The impacts of droughts on soybean production were better explained using DIs than with CIs (up to 89% vs 8% of variability explained) as predictors of the statistical models. SVHI-6 and SPEI-6, depending on the area of interest, were, during the phenological period of crop growth (summer), the most effective DIs in explaining annual variations in soybean yields. The results may be of interest in water resource management, drought risk management, and the Argentinean soybean production sector. Furthermore, they provide a foundation for future studies aimed at forecasting agricultural droughts and their impacts.

Graphical Abstract



Keywords

Drought indices, agricultural drought, teleconnections, statistical models, soybean yield.

Highlights

- El Niño 3.4 and 4 climate indices (CIs) are best correlated with drought indices (DIs).
- The spatial patterns of CI-DI correlations can be used to define homogenous drought regions.
- DIs were better predictors of soybean yield variability than CIs.
- SPEI-6 and SVHI-6 indices were the best DIs in explaining the variability of soybean yields.
- A foundation is laid for finding suitable indicators for the forecasting of drought.

1. Introduction

Drought is a natural stochastic hazard that causes substantial socio-economic and environmental losses worldwide (Golnaraghi et al., 2014). It is a complex phenomenon, usually initiated when precipitation presents with volumes below normal in a particular place (meteorological drought). Such anomalies go on to affect agriculture and hydrology (agricultural and hydrological drought, respectively), an issue reviewed in Mishra and Singh (2010). To date, several drought indices (DIs) have been developed with the intention of characterising and monitoring the phenomenon (World Meteorological Organization and Global Water Partnership, 2016; Zargar et al., 2011). These include the Standardised Precipitation Index (SPI; McKee et al., 1993) and the Standardised Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano et al., 2010) for meteorological droughts, the Vegetation Condition Index (VCI) and the Vegetation Health Index (VHI; Kogan, 1995) for agricultural droughts, the Standardised Streamflow Index (SSI; Hao and AghaKouchak, 2013), plus the Palmer Hydrological Severity Index (PHSI; Palmer, 1965; Zargar et al., 2011) for hydrological droughts. Each DI has advantages and disadvantages, which have been discussed in previous studies (Keyantash and Dracup, 2002; Mishra and Singh, 2010).

Since the introduction of the concept of a drought timescale by McKee et al. (1993), DIs have been used to quantify drought events in each component of the terrestrial water cycle (e.g. precipitation, soil moisture, and groundwater). In turn, these DIs can be associated with certain water uses (e.g. agriculture and electric generation) (Guttman, 1998). Agricultural activity, on which food security and much of the global economy depends, has increased significantly in recent years due to the growing demand for food by an increasing population (Tester et al., 2010). However, growth in farming has not been linear in time, as annual variations have been characterised by significant decreases in production. Although crop yields can be affected by a variety of factors, including wars, social crises, and plagues, drought is a key factor in yield variability, especially for rainfed crops (Leng and Hall, 2019; Lobell et al., 2011a, 2011b; Zampieri et al., 2017). Several studies have successfully correlated DIs with variability in crop yields worldwide (Araneda-Cabrera et al., 2021b; García-León et al., 2019; Peña-Gallardo et al., 2019a; Quiring and Papakryiakou, 2003; Vicente-serrano et al., 2012). Generally, these relationships are assessed using statistical models (Shi et al., 2013). DIs can thus be used as predictors of crop yields in such models and may explain the impact of agricultural drought.

Drought variability can be linked to large-scale climate oscillations (Hassan and Nayak, 2020; Singh, 2012), which are quantified by climatic indices (CIs) that rely on sea surface temperature (SST) and sea pressure level (SPL). Some well-known CIs based on SST are the ENSO indices (in the Pacific Ocean: The El Niño 3.4, for instance), the Caribbean Index (CAR; in the Caribbean Sea), and the south-eastern tropical Indian Ocean (SETIO; in the Atlantic Ocean). On the other hand, CIs based on SPL include Darwin and Tahiti (in the Pacific Ocean) and North Atlantic Oscillation (NAO; in the Atlantic Ocean) in the Atlantic Ocean. Several studies have linked climate indices to DIs (Huang et al., 2016; Manatsa et al., 2008b; Oñate-Valdivieso et al., 2020; Santos et al., 2019). Other studies have used climate indices to forecast droughts (Dutra et al., 2013; Tan and Perkowski, 2015).

Because variability in crop yields can be related to DIs, which in turn can be linked to CIs, some studies have directly connected crop yield variability with CIs (Anderson et al., 2017; Iizumi et al., 2014; Leng and Hall, 2019; Wang et al., 2020). Correlations between crop yields and CIs are usually lower than those with DIs, although no research which specifically addresses and supports this assertion currently exists. For instance, in the United States, Anderson et al. (2017) found correlations between soybean yield and the Oceanic Niño Index (ONI) of up to $r = 0.30$, while Peña-Gallardo et al. (2019b) highlighted correlations between the same crop yield and the SPEI of as much as $r = 0.70$.

There is no single DI that can explain variability in crop yields (i.e., drought-related impacts on agricultural production). Similarly, there is no single CI (teleconnection) that can represent all climate variability and can thus be used to predict drought conditions over large regions (Stenseth et al., 2003).

This is due to the very large differences in environmental physical factors (climate, soil composition, topography, etc.) that make both climate and crop development respond differently in each location. Hence, the present study compares the performance of eight DIs and 19 CIs to determine which are the most appropriate for use in agricultural drought monitoring throughout Argentina.

Argentina has the highest per capita crop production in the world (FAO, 2019) and is also the third largest soybean producer (Food and Agriculture Organization, FAO; http://faostat.fao.org). The country has substantial annual and inter-annual climate variability (Barros and Silvestri, 2002), and the likelihood of soybean yield reduction due to droughts ranges from 70 to 81% (when experiencing moderate to exceptional droughts, respectively) (Leng and Hall, 2019). In Argentina, DIs have been related to crop variability (e.g. D'Ambrosio et al., 2013; Seiler et al., 2007) and to CIs (e.g. Díaz et al., 2018; Rivera et al., 2018; Vicario et al., 2015) for specific regions, such as certain provinces or river basins. Similarly, relationships have been established between some crop yields and CIs, but only for specific regions (e.g. Anderson et al., 2017; Iizumi et al., 2014; Podestá et al., 1999). However, to the best of the authors' knowledge, there are no studies that benchmark different DIs and large-scale CIs for explaining agricultural drought and associated crop variability.

The goals of this study are: a) to determine which climate index (or indices) is best associated with droughts and with which drought index; b) to establish a drought index (or a set of indices) that can explain the annual variability in crop yield using soybeans as a benchmark crop; and c) to compare drought and climate indices as predictors of crop yield variability through three statistical models. The study was conducted throughout the country at the level of administrative departments. The ultimate aim is to support decision makers, farmers, and agricultural drought managers in Argentina. However, the methodology can be applied to other countries or regions and at any spatial scale.

2. Materials and methods

2.1.Study area

Continental Argentina was defined as the case study (Fig. 4.1a). It covers 2,791,810 km^2 and is divided into five main administrative regions according to the National Institute of Statistics and Censuses Argentine Republic (Spanish acronym INDEC: www.indec.gob.ar), 24 provinces (including the Autonomous City of Buenos Aires as a province), and 525 departments (m number of departments) (Fig. 4.1b). Due to its extensive area, the country sees wide climatic diversity, from arid (south and centre-north) to fully humid (northeast) (Beck et al., 2018; Kottek et al., 2006). However, 55% of the country has drylands (Cherlet et al., 2018). Argentina is one of the major worldwide producers of cereals (FAO, 2017), which are cultivated largely in the Argentine Pampas. The departmental average annual precipitation varies between 70 and 1880 mm per year and, the average annual temperature ranges from 2 to 23 °C. Both precipitation and temperature increase from east to west and from south to north. The spring and summer seasons are the most humid, while autumn and winter are the coldest and driest. The value of Argentina's cereal production was $10.2 billion in 2013, representing 8.3% of its GDP (FAO, 2017). The country is vulnerable to several natural phenomena (earthquakes, floods, etc.); however, droughts represent the greatest risk for agricultural losses (Cherlet et al., 2018). Argentina's agricultural year is defined as July to June (https://www.argentina.gob.ar/agricultura-ganaderia-y-pesca), while the hydrological year varies across the country according to regional precipitation patterns. At the national level, the precipitation data used in this study (see Section 2.2) show that the driest month is June; therefore, in this study, the hydrological year was taken to coincide with the agricultural year.

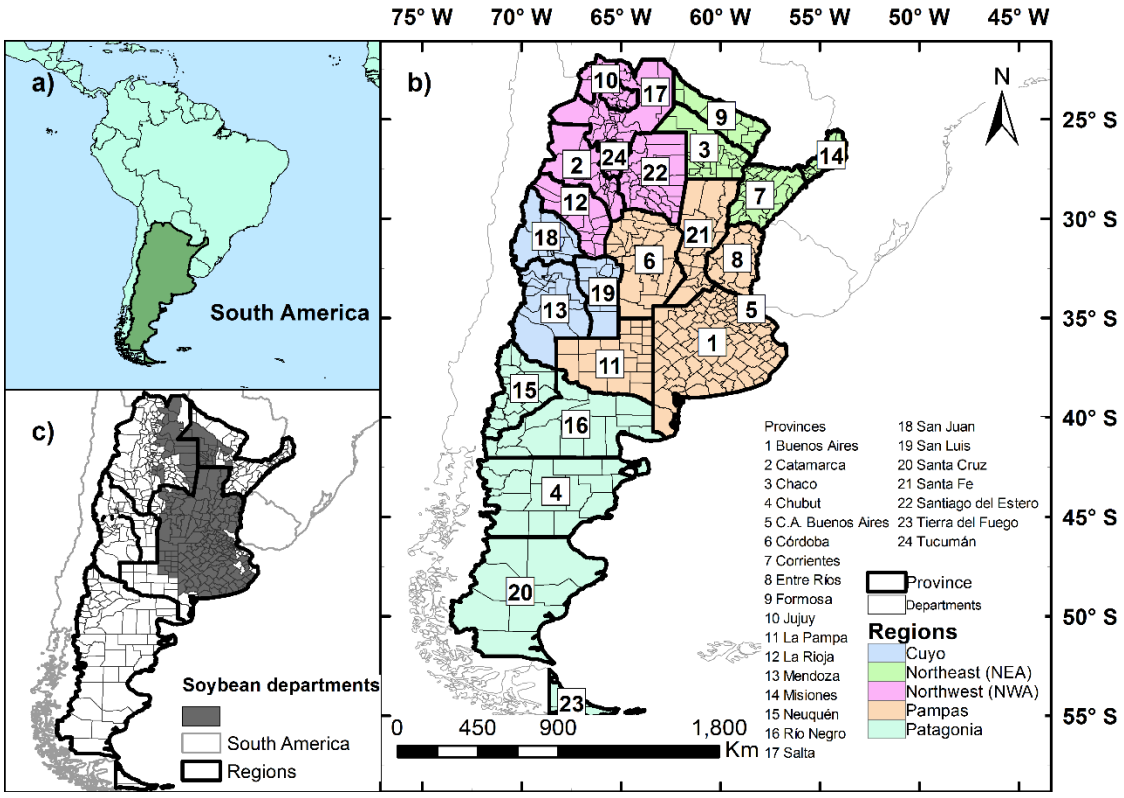


Fig. 4.1 Location of a) Argentina; b) regional, providence and department divisions; and c) departments with soybean production.

2.2.Drought indices

Various indices based on meteorological, vegetation condition, and hydrological variables were calculated for each department on a monthly scale: SPI, SPEI, SSI, SPDSI, SVCI, STCI, SVHI, and STWSI. Each index, together with its data source, is briefly described, as follows:

The Standardised Precipitation Index (SPI) is one of the best-known DIs and is recommended by the World Meteorological Organization (WMO). It was introduced by McKee et al. (1993). It is based on transforming precipitation into a normal function ($\overline{x} = 0$ and $\sigma = 1$) by means of a probability distribution function of two gamma parameters. It is very versatile because it can be calculated for any time scale n. The detailed procedure for its computation can be found in Kumar et al. (2009). Monthly precipitation data were downloaded from the TerraClimate database (Abatzoglou et al., 2018) at http://www.climatologylab.org/terraclimate.html at a spatial resolution of 1/24° (≈4 km at the equator) and were averaged for each department before calculating the SPI.

The Standardised Precipitation and Evapotranspiration Index (SPEI) was introduced by Vicente-Serrano et al. (2010). It is similar to the SPI but is computed by standardising the water deficit ($D = \text{Precipitation} - \text{Evapotranspiration}$). Here, D is fitted to a three-parameter log-logistic function: precipitation and potential evapotranspiration series data were downloaded from the TerraClimate database and calculated for each department from the average prior to the computation of the SPEI.

The Standardised Soil Moisture Index (SSI) (Hao and AghaKouchak, 2014) is based on the standardisation of soil moisture following the mathematical steps of the SPI. Here, we used soil moisture obtained from the Climate Change Initiative (CCI) program of the European Space Agency (ESA) at

https://www.esa-soilmoisture-cci.org, version v04.7, at a 0.25° spatial grid (Dorigo et al., 2017). These data were downscaled to the departmental level using a bilinear resampling of their centroids.

The Standardised Vegetation Condition Index (SVCI), Standardised Temperature Condition Index (STCI), and Standardised Vegetation Health Index (SVHI) (Agutu et al., 2017) were calculated following the standardisation procedure of the SPI to the VCI, TCI, and VHI time series, as defined by Kogan (1995). These data were downloaded from the Center for Satellite Applications and Research (STAR) and the Environmental Satellites for the U.S. Oceanic and Atmospheric Administration (NOAA) at https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. Datasets consist of 7-day value composites at 8 km resolution that were averaged to monthly and departmental scales before being standardised.

The Standardised Palmer Drought Severity Index (SPDSI) (Ma et al., 2014) is the result of standardising the widely-used drought index PDSI (Palmer, 1965). The PDSI was downloaded from the TerraClimate database and averaged at the departmental scale. Then, the time series were processed according to the SPEI computing steps.

The Standardised Total Water Storage Index (STWSI) (Agutu et al., 2017) was computed using the SPEI procedure. Instead of D, the input is the total water storage anomaly (TWS) – surface water and groundwater – derived from the Gravity Recovery and Climate Experiment (GRACE) developed by the National Aeronautics and Space Administration (NASA) and the German Aerospace Centre (Landerer and Swenson, 2012). In this study, we used data (level water thickness in cm) provided by the Jet Propulsion Laboratory (JPL) as part of the GRACE Follow On mission JPL RL06_v02 (Landerer et al., 2020) at https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RL06_V2. Because the data are represented on a 0.5° grid, the time series were downscaled to the departmental scale using bilinear resampling towards their centroids prior to the standardisation procedure.

The datasets used in this study have been validated and used successfully in other drought-related studies (Araneda-Cabrera et al., 2020; Rojas et al., 2011; Thomas et al., 2014; Y. Wang et al., 2019). The TerraClimate database offers data from January 1958; however, only the data since August 1981 were downloaded, so as to have the same time span as that offered by NOAA STAR. GRACE began its mission in April 2002; thus, we have data from that date. All variables were obtained up to December 2019. The DIs were calculated for seasonal and annual time scales (n) of 3, 6, and 12 months, since crop data that reflect agricultural droughts have sub-annual cycles (cf Section 2.4, below). Because of the aggregation of n months as part of the computation of the DIs, the time length of the DIs is n months less than their primary variables. Therefore, the final common timespan covered by the DIs was from August 1982 and April 2003 (STWSI) to December 2019. Some data was missing for certain months due to technical problems (VCI, TCI, VHI, and TWS), and here a linear interpolation method was adopted to complete the data based on the neighbouring months. This method is effective and widely-used for handling missing data (Noor et al., 2015). As previously noted, the VCI, TCI, VHI, PDSI, and TWS variables were standardised according to the recommendations proposed for the calculation of the SPI and SPEI. In this way, all DI values could be interpreted in the same, with values below -0.5 indicating droughts with variable intensities (McKee et al., 1993).

2.3.Large-scale climatic indices

A wide variety of climate indices were considered in this study. These are based on the SST from the Atlantic (TNA, TSA, NAT, SAT, and TASI), Pacific (ENSO indices ERSSTv5: Niño 1+2, Niño 3, Niño 4, Niño3.4, and PDO), and Indian (SWIO, WTIO, SETIO, and DMI) Oceans, and SPL from various locations across the world (Darwin, Tahiti, SOI, and NAO). Additionally, we used the CAR associated with the SST from the Caribbean Sea. Thus, we used a total of 19 monthly aggregated climate indices that can be freely obtained in near real-time, the details of which are shown in Table 4.1. For this study,

we downloaded the climate indices for the same time span as for the Dis, from August 1981 and April 2002 (STWSI) to December 2019. To establish consistency in the DIs, 3-, 6-, and 12-month running means were applied to the CIs.

Table 4.1 Selected climatic indices and associated free sources.

Variable/Data set	Period available	Data availability
Darwin Sea Level Pressure (Darwin SLP)*	Jan 1882-now	http://cpc.ncep.noaa.gov/data/indices/darwin
Tahiti Sea Level Pressure (Tahiti SLP)*	Jan 1882-now	http://cpc.ncep.noaa.gov/data/indices/tahiti
Southern Oscillation Index (SOI)**	Jan 1866-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/SOI/
ENSO indices (ERSSTv5): El Niño 1+2, El Niño 3, El Niño 4, and El Niño 3.4*	Jan 1950-now	https://cpc.ncep.noaa.gov/data/indices/ersst5.nino.mth.81-10.ascii
Pacific Decadal Oscillation (PDO)**	Jan 1948-Dec 2018	https://psl.noaa.gov/data/correlation/pdo.data
Caribbean Index (CAR)**	Jan 1950-now	https://psl.noaa.gov/data/correlation/CAR_ersst.data
South Western Indian Ocean (SWIO)	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/ind/swio.php
Western Tropical Indian Ocean (WTIO)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmiwest.had.long.data
Southeastern Tropical Indian Ocean (SETIO)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmieast.had.long.data
Indian Ocean Dipole Mode Index (DMI)**	Jan 1870-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmi.had.long.data
Tropical Northern Atlantic Index (TNA)**	Jan 1948-now	https://psl.noaa.gov/data/correlation/tna.data
Tropical Southern Atlantic Index (TSA)**	Jan 1948-now	https://psl.noaa.gov/data/correlation/tsa.data
North Atlantic Tropical (NAT)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/nat.php
South Atlantic Tropical (SAT)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/sat.php
Tropical Atlantic (TASI)***	Nov 1981-now	https://stateoftheocean.osmc.noaa.gov/sur/atl/tasi.php
North Atlantic Oscillation (NAO)**	Jan 1950-now	https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/nao.long.data

" * ";" ** ";" *** ";" **** " specifies Source

*Climate Prediction Center of NOAA

**Physical Sciences Laboratory of NOAA

***Ocean Observations Panels for Climate

2.4.Crop yield data

Annual crop yields at the departmental level were obtained from the Ministry of Agriculture, Livestock, and Fisheries of Argentina (Ministerio de Agricultura, Ganadería y Pesca de Argentina) at https://datos.agroindustria.gob.ar/dataset/estimaciones-agricolas. This database includes the sowed area, harvested area, and total production of 30 different crops from 1961 to 2019. Each year was measured from July to June (agricultural year). Soybeans were chosen as a representative rainfed crop because since 2000 it has been the cereal with the highest growth both in farmed areas and total production in the country. It is the most important crop in Argentina (Anderson et al., 2017; FAO, 2016b; Leng and Hall, 2019; Magrin et al., 2005). The annual soybean yield (kg/Ha) was calculated by dividing production by the sowed area. All departments showing continuous series over time were considered, representing a total of 193 soybean-producing departments (Fig. 4.1c), where the annual yield data timespan was 16 years, from July 2004 to June 2019 (Fig. 4.2). In the case of soybeans, sowing is carried out from October to December, and harvesting from April to June. Notably, lower median and mean soybean yields were obtained in 2009 and 2018, years affected by significant drought episodes (EM-DAT, 2021).

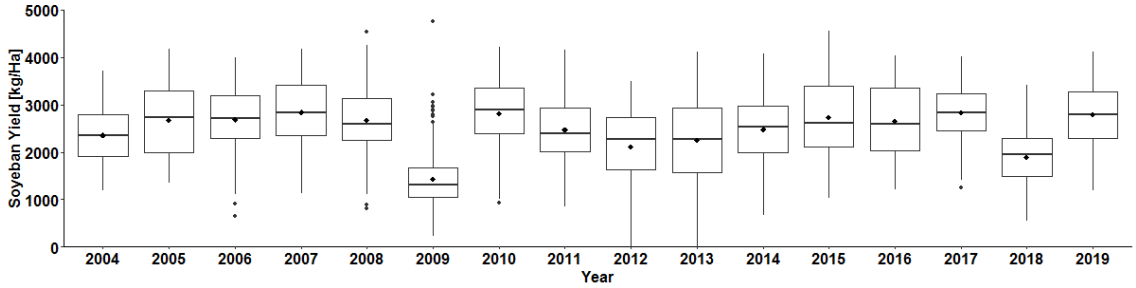


Fig. 4.2 Temporal series of detrended soybean yields for the 193 departments for the period 2004–2019. The solid black line shows the median, and the black dot shows the mean.

Because crop yields are affected by factors other than climate, including agricultural innovations, technological improvements in sowing practices, and seed selection, crop yields generally have a positive trend (Peña-Gallardo et al., 2019b; Tian et al., 2018). This is also evident in Argentina; therefore, the yield series were detrended to remove the variability in productivity caused by non-climate factors using a linear regression model adjusted to the soybean yield series of each department. The average crop yield of each series was added to the residual model series to produce non-trend yield data in Kg/Ha (hereafter Y) following the procedure explained in detail in Lobell et al., (2011b) and used in other studies (e.g. Tian et al., 2018).

2.5.Relationship between drought indices and climatic indices

Pearson correlations (r) were calculated between CIs and DIs aggregated for 12 months at the departmental level throughout the country. Because the sign of the correlations is important (positive or negative indicate in-phase or anti-phase relation), r values were used instead of other metrics as the coefficient of determination (R^2). Then, a seasonal analysis was also performed with DIs and a selected number of CIs (in the previous step) with 6-month aggregations (for summer: March, for autumn: June, for winter: August, and for spring: December). Seasonal computing was performed for the entire country (m=525) and for all soybean producing departments (m=193). In the latter, we prioritised the summer analysis because this is the sowing and growing period for soybeans. The STWSI, with a shorter data length, may show better correlations with the CIs than the other DIs; thus, statistical significance will be a determining factor in the comparison of results.

2.6.Statistical crop yield models

To compare DIs as explanations for the variability of crop yields, we trained three statistical models (Lobell and Burke, 2010; Shi et al., 2013) over the 193 soybean producing departments: time-series, panel, and cross-section models. In all three cases, we assumed Y (in each department) was the response of a function of k independent variables X, which, in this context, are the DIs as possible predictors.

Given $Y \in [0, \infty)$, the time series model was implemented in each department as follows:

$$\ln\left(Y_t\right) = f\left(X_{t1}, X_{t2}, ..., X_{tk}\right) = \beta_0 + \sum_{j=1}^{k} \beta_j \left(X_{tj}\right) + \varepsilon_t , \quad (4)$$

where Y is the departmental vector of annual soybean yields, t is the year, X represents the vector of the candidate predictors, β_0 and β_j are the parameters (intercept and constant coefficients) to be fit, and ε is the error.

A panel regression model was executed combining the 193 departments:

$$\ln\left(Y_{i,t}\right) = \beta_{i,0} + \sum_{j=1}^{k} \beta_{i,j} \left(X_{i,tj}\right) + \varepsilon_{i,t} , \quad (5)$$

where i represents each soybean-producing department, and $\beta_{i,0}$ is an intercept.

Finally, the average departmental yields and DIs were computed to estimate the cross-sectional model.

$$\ln\left(Y_{i,avg}\right) = \beta_0 + \sum_{j=1}^{k} \beta_j \left(X_{i,avgj}\right) + \varepsilon_i \quad (6)$$

The predictor candidates were: the June DIs with a 12-month aggregation period, this for consistency with the annual crop yield data; the March DIs with a 6-month aggregation period, coinciding with the sowing and growing periods of soybeans; and the December and March DIs with a 3-month aggregation period, this to capture the sowing and growing periods of soybeans, respectively. The coefficients of determination of the models were used as the comparative statistics.

Moreover, to compare the performance of DIs and CIs as crop yield predictors, those CIs that showed the highest correlations with the DIs in the analysis described in Section 2.5 were evaluated as candidate predictors in the time-series model, following equation (1). A 3-, 6-, or 12-month running mean from June, March, or December was applied to the CIs for consistency with the DIs.

The overall procedure for identifying suitable DIs and CIs for monitoring agricultural drought in Argentina is summarized in Fig. 4.3.

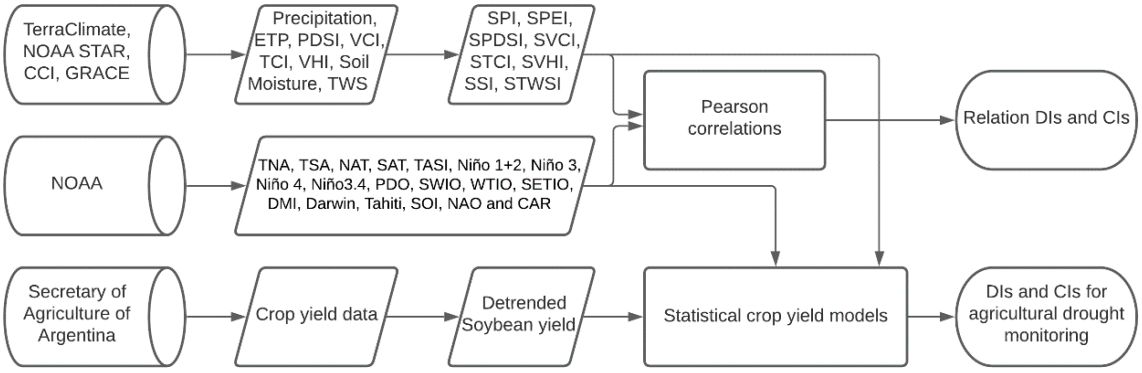


Fig. 4.3 Methodological flow-chart of the study.

3. Results

3.1.Identification of the large-scale climate drivers of drought

The spatiotemporal patterns of the correlations between CIs and DIs in Argentina were seen to vary significantly. The greater the temporal aggregation, the greater the correlations and the higher the percentage of departments with significant correlations (Table S4.1). Table 4.2 shows the correlations between the DIs and CIs with a 12-month aggregation. The STWSI showed the strongest correlation with all the CIs. However, some DIs (SPI, SPEI, STCI, SVHI, and SPDSI) showed a higher percentage of departments with significant correlations. The CIs that best correlated with all the DIs were Tahiti and SOI, with negative correlations, and El Niño 3.4 and El Niño 4, with positive correlations. These four CIs showed percentages of departments with statistically significant correlations of at least 74% ($\rho<0.05$) and were selected for further analysis. The spatial patterns of the correlations between the DIs and selected CIs were similar (Fig. 4.4 and Fig. S4.1). DIs that included temperature in their calculus (STCI, SPEI, and SVHI) had the highest percentage of departments with significant correlations. The Pampas and NEA regions had the strongest negative correlations, followed by Cuyo and NWA with negative correlations, and finally Patagonia, where correlations were low and negative (Fig. 4.4). Table 4.3 shows the results of the correlations by region, which support the spatial patterns shown in Fig. 4.4 and Fig. S4.1.

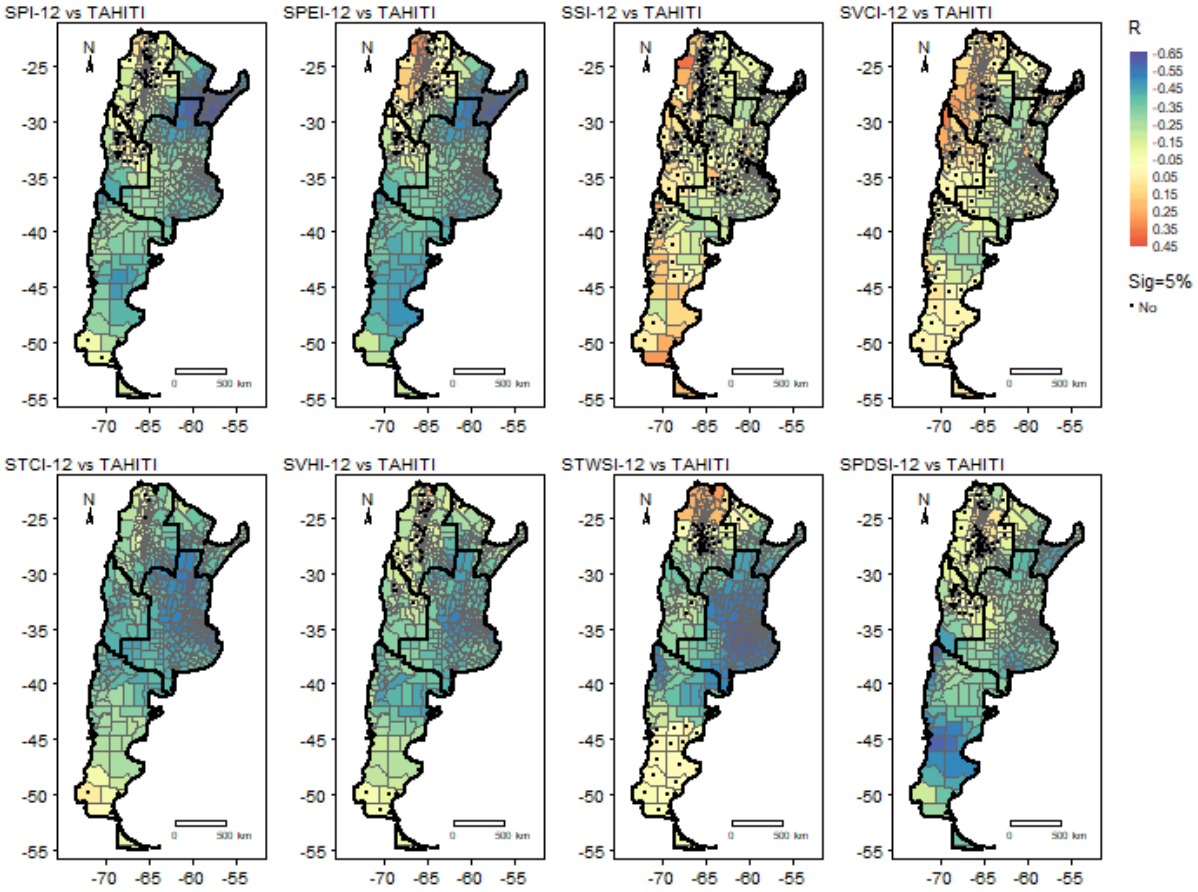


Fig. 4.4 Spatial pattern of correlations between the 8 DIs with a 12-month aggregation period and Tahiti SLP CI based on 1982–2019 data (STWSI 2003–2019). Black dots indicate where correlations were not statistically significant.

Table 4.2 Median correlations between DIs and CIs aggregated for 12 months based on all departments (m=525) and 1982-2019 (STWSI 2003-2019). Percentage of departments with significant correlation at 5% level (ρ ≤ 0.05) are showed.

DI		Darwin	Tahiti	Niño3.4	Niño3	Niño4	Niño1+2	SOI	PDO	CAR	SWIO	WTIO	SETIO	DMI	TNA	TSA	NAT	SAT	TASI	NAO
SPI-12	*	0.18	*-0.33*	*0.24*	0.19	*0.23*	0.18	*-0.27*	-0.01	-0.03	0.09	-0.02	0.08	-0.05	-0.13	-0.16	-0.07	-0.09	-0.13	0.11
	**	88.2	*85.9*	*77.0*	89.1	*74.1*	75.6	*87.2*	55.2	65.9	48.6	65.7	61.3	49.1	75.0	82.5	42.9	49.3	70.1	57.1
SPEI-12	*	0.21	*-0.36*	*0.24*	0.21	*0.20*	0.21	*-0.29*	0.04	-0.03	0.12	-0.01	0.09	-0.07	-0.15	-0.11	-0.07	-0.05	-0.08	0.10
	**	90.1	*86.3*	*84.0*	90.1	*79.8*	79.6	*86.9*	64.6	53.3	55.8	57.9	65.9	48.6	77.3	63.0	43.4	31.6	41.9	55.4
SSI-12	*	0.07	*-0.06*	*0.07*	0.13	*0.04*	0.22	*-0.08*	-0.10	0.32	0.06	0.13	-0.04	0.23	0.15	0.04	0.07	0.00	-0.02	-0.16
	**	40.8	*48.8*	*43.6*	74.7	*39.4*	86.9	*49.7*	61.5	94.3	50.5	82.5	38.9	89.7	71.2	31.0	57.3	14.9	18.1	83.8
SVCI-12	*	0.02	*-0.09*	*0.11*	0.10	*0.14*	0.04	*-0.06*	-0.03	0.21	0.02	0.11	0.11	0.01	0.14	-0.01	0.08	-0.01	-0.01	-0.07
	**	51.8	*73.7*	*60.8*	60.4	*67.2*	38.5	*68.2*	39.4	81.5	41.9	59.0	61.9	44.6	69.5	42.3	53.9	23.8	30.7	54.5
STCI-12	*	0.34	*-0.43*	*0.28*	0.29	*0.17*	0.26	*-0.42*	0.20	-0.13	0.04	-0.07	0.02	-0.10	-0.28	-0.24	-0.18	-0.13	-0.16	0.27
	**	97.1	*98.9*	*93.9*	95.6	*89.3*	97.3	*99.0*	85.9	64.8	25.7	50.1	25.9	55.6	94.7	95.6	81.1	78.3	86.3	95.6
SVHI-12	*	0.24	*-0.34*	*0.25*	0.25	*0.20*	0.20	*-0.32*	0.11	0.03	0.05	0.03	0.09	-0.06	-0.12	-0.17	-0.08	-0.08	-0.10	0.16
	**	91.6	*91.0*	*93.3*	93.5	*86.5*	90.3	*92.8*	59.8	53.1	41.5	49.5	52.4	47.2	68.4	76.4	55.8	49.0	56.8	75.6
STWSI-12	*	0.40	*-0.37*	*0.38*	0.30	*0.48*	0.25	*-0.42*	0.54	0.49	0.02	0.40	0.21	0.08	-0.17	0.25	-0.03	0.11	0.16	0.31
	**	86.9	*87.6*	*86.1*	78.9	*88.6*	83.4	*92.4*	90.9	87.2	14.7	89.0	69.3	60.8	59.4	81.0	21.0	36.0	62.7	85.0
SPDSI-12	*	0.15	*-0.30*	*0.18*	0.12	*0.16*	0.15	*-0.27*	0.06	0.01	0.06	-0.03	0.08	-0.06	-0.19	-0.10	-0.15	-0.06	-0.08	0.07
	**	73.5	*83.8*	*73.0*	71.2	*78.5*	70.9	*81.7*	69.1	73.0	37.1	69.0	60.4	47.0	68.2	60.2	62.7	33.3	46.7	45.0

* r; ** Percentage of departments with 5% significance level

A seasonal analysis was conducted correlating the seasonal series between the 6-month-aggregated DIs at the national level (Table S4.2 in the supplementary materials). The El Niño 3.4 and El Niño 4 had stronger correlations than did the other two indices. The SSI and STWSI showed stronger correlations for the cold seasons (winter and spring), while the rest of the DIs did so during the warmer seasons (summer and autumn). Persistent (very few exceptions) negative correlations during spring and summer and positive correlations during autumn and winter between El Niño 3.4 and the DIs were observed (Table S4.3). Similar patterns were observed for the other three selected CIs (Table S4.4 in the supplementary materials). Spatially, regions with weak correlations had the most departments with non-significant correlations, and this number increased in the winter and spring seasons. For example, Fig. 4.5 illustrates the spatial patterns of correlations between SPEI-6 and El Niño 3.4.

Table 4.3 Median correlations between DIs and CIs aggregated for 12 months based on the departments of Cuyo (m=44) Patagonia (m=53), Pampas (m=233), NEA (m=76), and NWA (m=119) and 1982–2019 data (STWSI 2003–2019).

DI	CI	Region				
		Cuyo	**Patagonia**	**Pampas**	**NEA**	**NWA**
SPI-12	Tahiti SLP	-0.009	-0.319	-0.376	*-0.533*	-0.094
SPEI-12		-0.083	-0.397	-0.375	*-0.505*	-0.010
SSI-12		-0.004	0.067	-0.085	*-0.116*	-0.056
SVCI-12		0.112	-0.075	-0.189	*-0.027*	0.147
STCI-12		-0.433	-0.369	-0.481	*-0.399*	-0.360
SVHI-12		-0.247	-0.350	-0.407	*-0.263*	-0.130
STWSI-12		-0.263	-0.330	-0.545	*-0.346*	-0.076
SPDSI-12		-0.263	-0.330	-0.545	*-0.346*	-0.076
SPI-12	El Niño 3.4	0.045	0.235	0.250	*0.501*	0.049
SPEI-12		0.075	0.250	0.262	*0.478*	-0.021
SSI-12		0.015	0.064	0.092	*0.141*	0.050
SVCI-12		0.032	0.185	0.157	*0.061*	-0.044
STCI-12		0.295	0.204	0.301	*0.330*	0.239
SVHI-12		0.274	0.299	0.273	*0.255*	0.137
STWSI-12		0.087	0.225	0.502	*0.438*	0.228
SPDSI-12		0.087	0.225	0.502	*0.438*	0.228
SPI-12	El Niño 4	-0.106	0.181	0.326	*0.390*	-0.050
SPEI-12		-0.055	0.169	0.303	*0.363*	-0.128
SSI-12		-0.007	0.105	0.059	*0.083*	-0.014
SVCI-12		0.089	0.194	0.172	*0.121*	0.050
STCI-12		0.143	0.134	0.217	*0.251*	0.115
SVHI-12		0.176	0.236	0.230	*0.214*	0.081
STWSI-12		0.148	0.338	0.610	*0.519*	0.246
SPDSI-12		0.148	0.338	0.610	*0.519*	0.246
SPI-12	SOI	-0.092	-0.328	-0.283	*-0.499*	-0.109
SPEI-12		-0.152	-0.387	-0.303	*-0.474*	-0.027
SSI-12		-0.014	0.006	-0.094	*-0.132*	-0.068
SVCI-12		0.069	-0.121	-0.148	*0.032*	0.146
STCI-12		-0.466	-0.358	-0.452	*-0.377*	-0.355
SVHI-12		-0.336	-0.373	-0.362	*-0.221*	-0.153
STWSI-12		-0.275	-0.332	-0.567	*-0.411*	-0.196
SPDSI-12		-0.275	-0.332	-0.567	*-0.411*	-0.196

Table 4.4 Median correlations between DIs and the selected CIs (Tahiti SLP, El Niño 3.4, El Niño 4, and SOI) aggregated for 6 months based on soybean producing departments (m=193) and 1982–2019 data (STWSI 2003–2019).

DI		Tahiti	El Niño 3.4	El Niño 4	SOI
SPI-6	Summer-ONDJFM	-0.29	0.35	0.38	-0.34
SPEI-6		*-0.29*	*0.37*	*0.38*	*-0.35*
SSI-6		-0.12	0.03	-0.03	-0.15
SPDSI-6		-0.16	0.10	0.13	-0.14
SVCI-6		-0.20	0.22	0.24	-0.23
STCI-6		-0.27	0.29	0.33	-0.35
SVHI-6		-0.26	0.28	0.32	-0.32
STWSI-6		*-0.32*	*0.28*	*0.40*	*-0.41*

Considering only the soybean-producing departments and the summer season, when soybean sowing and growth occurs, the SPEI and STWSI had the strongest correlations with the four selected CIs (Table 4.4). For these specific departments and seasons, the r medians were persistently higher than the correlation found when all departments were considered. However, similar patterns were found in other seasons (Supplementary Table S4.5).

Table 4.5 Determination coefficient results of the a) time-series model (medians), b) panel model, and c) cross-section model between the DIs and soybean yield based on soybean-producing departments (m=193) and 2004–2019 data.

	a)					b)					c)			
R^2 - Time-series model	**3-months JFM**	**3-months OND**	**6-months ONDJFM**	**12-months Year**	**R^2 - Panel model**	**3-months JFM**	**3-months OND**	**6-months ONDJFM**	**12-months Year**	**R^2 - Cross-section model**	**3-months JFM**	**3-months OND**	**6-months ONDJFM**	**12-months Year**
SPI	0.188	0.148	*0.298*	0.268	**SPI**	0.128	0.071	*0.175*	0.154	**SPI**	0.380	0.196	*0.504*	0.504
SPEI	0.233	0.160	*0.313*	0.253	**SPEI**	0.160	0.072	*0.174*	0.148	**SPEI**	0.437	0.220	*0.503*	0.482
SSI	0.195	0.024	*0.142*	0.247	**SSI**	0.207	0.021	*0.062*	0.157	**SSI**	0.255	0.003	*0.119*	0.346
SPDSI	0.335	0.016	*0.152*	0.173	**SPDSI**	0.176	0.000	*0.065*	0.087	**SPDSI**	0.608	0.007	*0.228*	0.262
SVCI	0.575	0.134	*0.467*	0.333	**SVCI**	0.268	0.073	*0.234*	0.185	**SVCI**	0.847	0.470	*0.771*	0.504
STCI	0.556	0.215	*0.452*	0.345	**STCI**	0.319	0.119	*0.275*	0.262	**STCI**	0.743	0.475	*0.631*	0.507
SVHI	0.625	0.199	*0.494*	0.416	**SVHI**	0.324	0.106	*0.276*	0.274	**SVHI**	0.821	0.511	*0.711*	0.549
STWSI	0.249	0.094	*0.189*	0.204	**STWSI**	0.195	0.040	*0.143*	0.161	**STWSI**	0.392	0.139	*0.288*	0.286

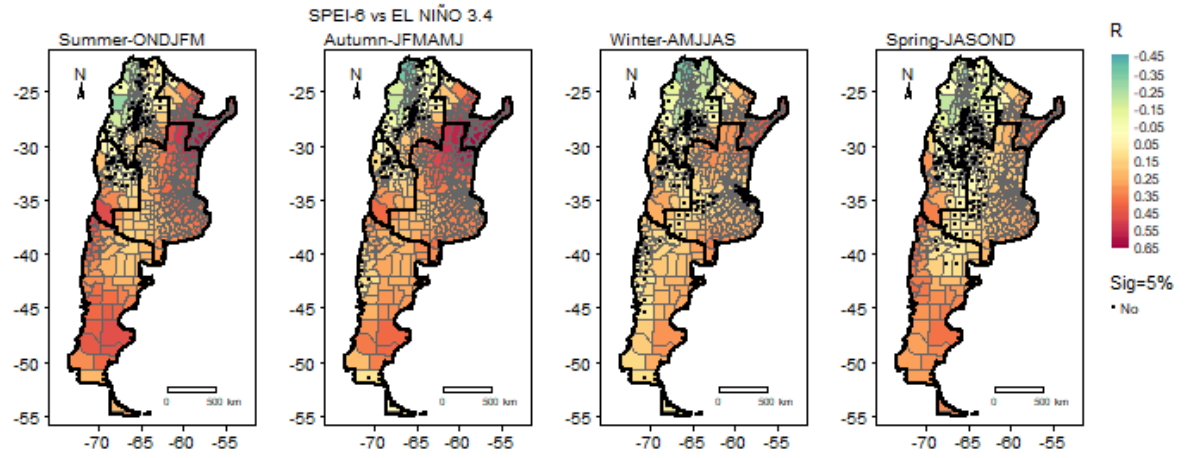


Fig. 4.5 Spatial pattern of correlations between the seasonal time series SPEI-6 and El Niño 3.4 based on 1982–2019 data.

3.2. Comparison of the performance of DIs in explaining soybean yield variability

The results of the three statistical models with the soybean yield data and DIs are presented in Table 4.5. The determination coefficients (R^2) of the cross-section model were consistently higher than those of the other models. The SVHI, SVCI, STCI, SPEI, and SPI, which are based on meteorological and vegetation variables, better explained the variability of soybean yields in all three models and performed better using their 3- and 6-month aggregations for the summer season (in March) as a predictor. The other DIs (SSI, SPDSI, and STWSI) had a lower explanatory power and performed better with the 12-month aggregation in June.

The spatial distributions of the determination coefficients using the time-series model with predictors aggregated for six months are shown in Fig. 4.6. For all DIs, the models produced positive coefficients. This was not surprising, given that when DI values are lower (indicating more intense/severe drought events), reductions in crop (soybean) yields are expected. Specifically, the time-series model estimated positive coefficients in more than 96% of the departments (except SSI, which did so in 84%). The SVHI, which has the highest median determination coefficients, provided the best explanation of the variability of soybean yields in the central areas of the Pampas region. This zone obtained the highest R^2 values using any of the DIs as predictors. The SPEI was the best predictor of soybean yield in departments located in the southern and northern parts of the soybean belt.

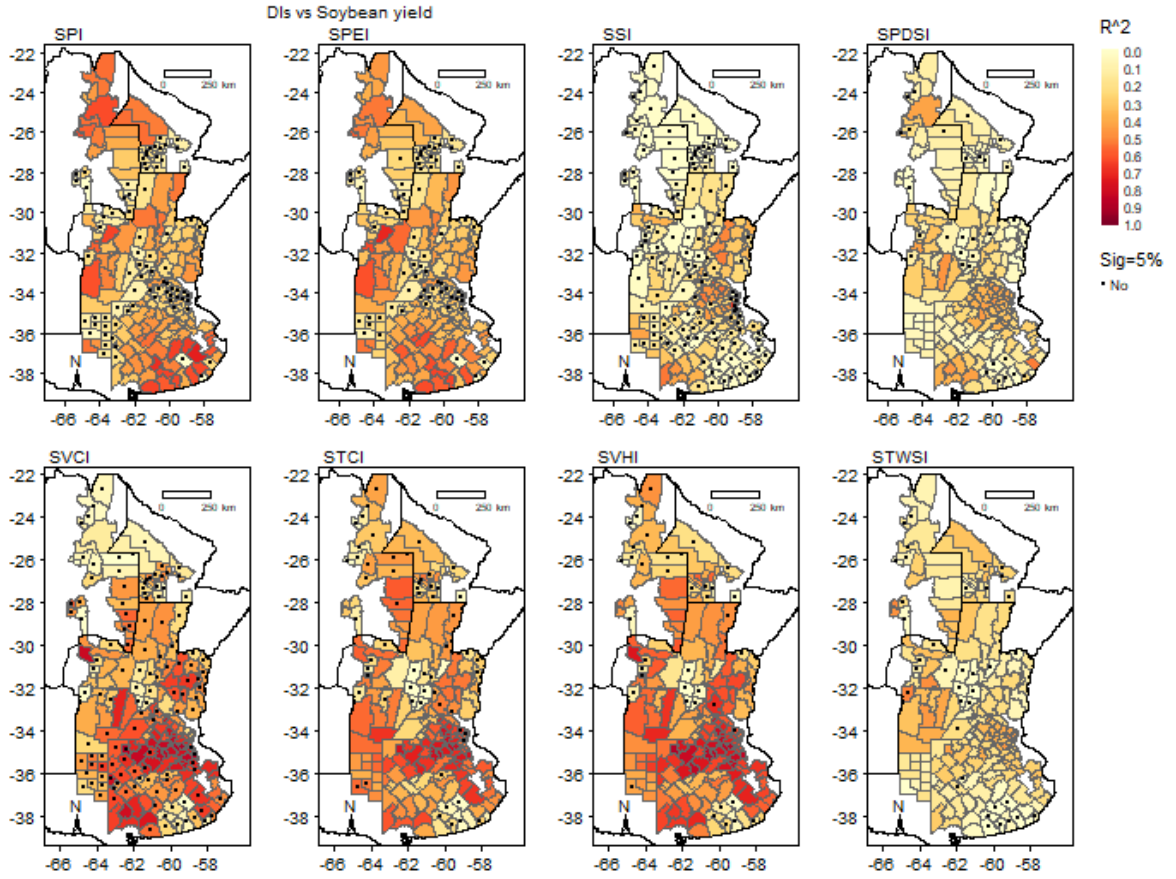


Fig. 4.6 Spatial patterns of the determination coefficient results of the time-series models between DIs aggregated 6 months (ONDJFM) and soybean yield based on soybean-producing departments (m=193) and 2004–2019 data.

3.3 Comparison of the performance of CIs in explaining soybean yield variability

Table 4.6 highlights the results of the time series models computed using the soybean yield data and four selected CIs (Tahiti, SOI, El Niño 3.4, and El Niño 4). Similar to the results with the DIs as predictors, the CI 3- and 6-months running-mean in summertime (March) explained the variability of soybean yields better than the CI 12-months running mean. The highest R^2 was found using El Niño 3.4 and El Niño 4. However, this maximum goodness of fit explained only 11.90% of the variability in soybean yield. In other words, the performance of CIs as possible predictors of soybean yield was very poor in Argentina for these specific timescales (6 and 12 months) and months (March and June).

Table 4.6 Median correlations between the selected CIs (Tahiti SLP, El Niño 3.4, El Niño 4 and SOI) aggregated for 3,6, and 12 months and the soybean yield based on soybean-producing departments (m=193) and 2004–2019 data.

R^2 - Time-series model	3-months JFM	3-months OND	6-months ONDJFM	12-months Year
El Niño 3.4	*0.087*	0.074	0.073	0.004
SOI	0.129	0.025	*0.044*	0.000
TAHITI	0.110	0.110	*0.027*	0.015
El Niño 4	0.119	0.082	*0.063*	0.002

4. Discussion

4.1.Drought indices and teleconnections

CIs located in the equatorial Pacific Ocean (Tahiti, SOI, El Niño 3.4, and El Niño 4) showed stronger correlations with the DIs in Argentina. Although relatively low r values were found for this type of study, they can be considered acceptable (Lovino et al., 2018; Robledo et al., 2013). These CIs have been found to indicate the triggering of droughts in other parts of the world (Gupta and Jain, 2021; Vicente-Serrano et al., 2017), demonstrating the global importance of their variability in extreme hydrological events. On a larger time scale (12 months), correlations were persistently higher, something that has also been found in other studies (Lovino et al., 2018; Singh and Shukla, 2020). However, it might be possible to find stronger relationships on other time scales, although this was not investigated in the present study. Some DIs (i.e., SSI and SVCI) showed strong correlations with other CIs, which suggests that the explanations for droughts in Argentina are highly complex and cannot be covered by a single climate driver.

Stronger negative correlations between CIs and DIs were found in warm seasons, whereas in cold seasons they were less strong yet positive. These findings coincide with the results in Hurtado and Agosta (2020) and De La Casa and Ovando (2006), in which the climatic drivers were better correlated with the characteristic patterns of the summer than with the winter rainfall regime. Such results, then, suggest that extremely hot and humid summers and extremely dry and cold winters are associated with the variability of these CIs, and may be used for monitoring and forecasting droughts in Argentina, similar to other regions (Dikshit et al., 2021; Seibert et al., 2017).

Spatially, the DIs and CIs showed a stronger correlation in regions with high rainfall regimes and high temperature variability (Pampas and NEA), where the climate classification is semi-arid (temperate climate) according to the Köppen-Geiger climate classification (Kottek et al., 2006). The weakest correlations were found in the arid regions (NEA and north of Cuyo). In the Patagonia region, where the climate is cold, correlations were persistently weaker throughout the year. Similar results have been obtained by Robledo et al. (2013), who found stronger relations between drought conditions and climate drivers in the northeast and central regions of Argentina than in the south. In these regions, CIs might be used in conjunction with other variables (e.g., measurements of streamflow in rivers, water levels in lakes or reservoirs, or snow cover) to better monitor and predict droughts. The reader is referred to Hao et al., 2018 for a review of commonly used predictors for statistical drought prediction.

The spatial patterns of correlations between DIs and CIs indicated that the five administrative regions used were not the most suitable for drought monitoring based on climate drivers (Fig. 4.5). Therefore, at the national level, we believe that it is necessary to define the climate drivers of droughts in homogeneous drought regions. As an example, Fig. 4.7 sets out how four regions, based on correlations, can be defined: i) northern Patagonia + eastern Pampas + southern NEA; ii) northern NWA; iii) western Cuyo and Patagonia; and iv) the central part of the country from central NWA to northern Patagonia across eastern Cuyo and western Pampas. The precise definition of homogeneous drought regions is, however, beyond the scope of this work. Specific clustering methodologies, such as principal component analysis and hierarchical and non-hierarchical clustering methods, should be used to establish this regionalisation, ensuring a robust definition of drought regions. These techniques could be applied to identify patterns in the DI and CI series, hence defining regions with similar drought variability and characteristics, as done for example in Espinosa et al., (2019).

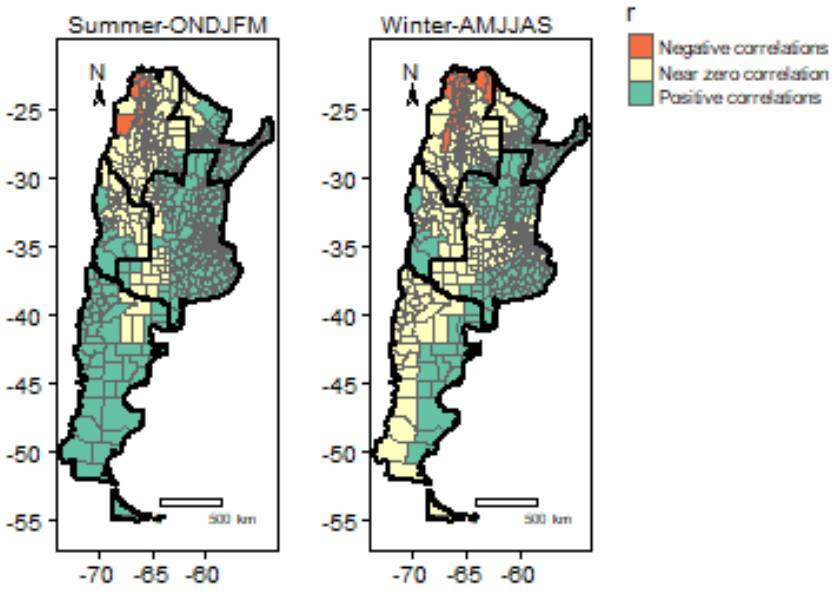


Fig. 4.7 Spatial distribution of positive (r > 0.2), negative (r < -0.2), and near zero (-0.2 ≤ r ≤ 0.2) correlations between SPEI-6 and the El Niño 3.4. The black lines show the regional divisions.

Because each DI is associated with a specific part of the water cycle, and it is not common for the entire water cycle also to be under stress at the time due to the drought development process (from meteorological to hydrological) (Huang et al., 2017; Mishra and Singh, 2010), the DIs showed different relationship levels with the CIs. Nevertheless, the STWSI (associated with the total amount of water in the environment, both ground and surface) was notable, showing strong correlations with all CIs. It was also interesting that STWS showed stronger correlations with CIs in the cold than in the hot seasons. This is possibly due to the time lag between hydrological droughts and agricultural and meteorological droughts (Huang et al., 2017; Van Loon, 2015). However, although STWSI primary variables (GRACE data) have been used to successfully detect some significant drought events in Argentina (Aragón et al., 2011; Chen et al., 2010), to claim that the STWSI is the DI that best correlates with the CIs may imply a bias, due to the short time that its data have been available. Thus, a longer time series needs to be used to verify these results.

The STCI, SPEI, and SVHI showed high percentages of departments with significant correlations with CIs (better than the percentage obtained by STWSI, this perhaps due to the length of the time series), indicating that they could be used throughout the country, despite having a lower correlation than the STWSI. These DIs all include temperature in their calculus, which is a very important variable in the relationship between droughts and climate drivers, especially in Argentina (Carcedo and Gambin, 2019). Furthermore, the SPEI showed the strongest correlations with the CIs in the analysis based on soybean-producing departments (Fig. 4.1c), particularly with El Niño 3.4, and specifically for warm seasons (the growing season for crops such as soybeans). These findings might be of particular interest to drought managers and farmers in Argentina. Further research could define CIs, such as El Niño 3.4, as predictors of agricultural drought in this area.

4.2.Drought indices and soybean yield variability

This study has indicated that the utilisation of a cross-section model resulted in the highest R^2, followed by the time-series model. Such a pattern of results between statistical models is consistent with the results shown in Lobell and Burke (2010). However, statistical models based on a time-series show a

better spatial understanding of the relationship between drought and crop yield, which was expected, since having subnational data would a priori provide better approximations (Lobell and Burke, 2010).

The time-series model allowed us to identify areas (or departments) in which soybeans were more sensitive to drought. Soybeans responded better to the SVHI in departments where yields were higher (explaining up to 88.8% of agricultural variability). These areas are located in the north of the Buenos Aires Province (Fig. 4.8). DIs based on vegetation condition variables (e.g., SVHI) have been associated with crop variability with other crop yields. For instance, in Spain, García-León et al. (2019) found better responses in wheat, barley, oat, rye and maize to the VHI (non-standardised SVHI), and in Argentina, Seiler et al. (2007) explained corn yield variability with the VCI and TCI (non-standardised SVHI SVCI and STCI). The SPEI, on the other hand, accurately explains areas with low yields in the south and north of the soybean belt (up to 71.5% of variability). The SPEI has also been widely associated with the response of crop yields to droughts (Chen et al., 2016; Peña-Gallardo et al., 2019a). Interestingly, the STWSI, which has a strong relationship with the CIs, performed poorly in explaining soybean responses to droughts, which can be explained by the fact that this index is associated with longer drought timescales, more representative of hydrological droughts (reservoirs, aquifers, etc.) than meteorological or agricultural droughts (Zhou et al., 2020).

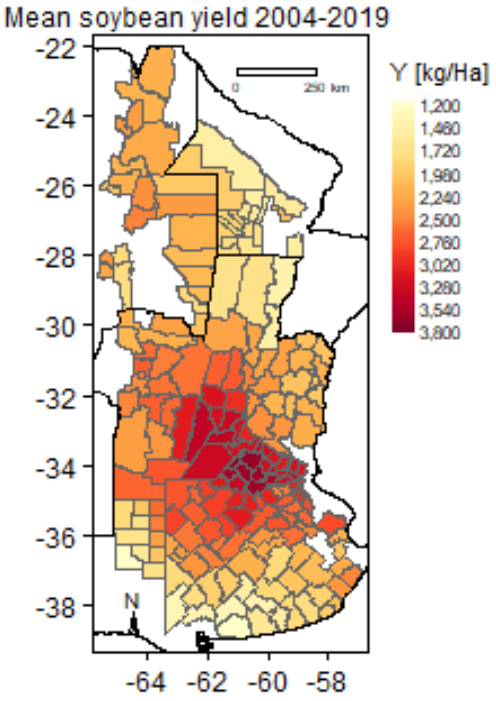


Fig. 4.8 Spatial distribution of average annual soybean yield (period 2004–2019).

The findings also indicate that CIs explain variability in soybean yield very poorly. This does not necessarily mean that climate drivers and crop (soybean) production are not associated. Relations between ENSO and crop yields have been established by Anderson et al. (2018) and Podestá et al. (1999). Furthermore, in this study, we analysed three timescales (3, 6, and 12 months) for specific months (December, March, and June); therefore, an extension of this study using a variety of different time scales may be necessary to verify the results. However, for the purposes of annual/seasonal monitoring of soybean production in Argentina, based on our results, we recommend using a DI rather than a CI.

One of the limitations of this study was the lack of data associated with irrigation. In areas with purely rainfed crops, meteorological variables are more important for crop development (Kuwayama et al.,

2018). Therefore, the disparity in the results using the SPEI and SVHI could be attributed to the fact that in areas with higher yields, there might have been some type of additional irrigation. According to (FAO et al., 2015), there are some small-scale irrigated areas irrigated fields in the northern Buenos Aires Province. This should be studied in detail in each departmental unit. In other locations, crops (including soybeans) usually have short time dependencies (1–4 months) (Peña-Gallardo et al., 2018). In future studies, it will be necessary to refine the analysis of the link between soybean yields and droughts using various other time scales, for instance, following the methodology used by Peña-Gallardo et al. (2019b) in the United States. However, at the national level, but with a focus on specific regions, the use of SPEI and SVHI aggregated for 3 and 6 months during phenological growth to monitor the state of soybean production is recommended. This information might prove useful for local farmers.

5. Conclusions

In this study, 8 DIs, and 19 CIs were benchmarked for agricultural drought monitoring in Argentina. First, the relationship between DIs and CIs was explored. Then, DIs were evaluated based on their capacity to explain the impacts of agricultural drought (annual soybean yield variability). Finally, the best response of crop production to DIs, rather than to CIs, was presented.

DIs were particularly related to the CIs located in the Pacific Ocean, including El Niño 3.4 and El Niño 4. DIs that include temperature in their computation (STCI, SPEI, and SVHI) correlated best with CIs across the country. For soybean production areas, SPEI was the DI that best responded to variations in CIs. Correlations were positive and strong in the warm and wet season (summer), while in the cold and dry season (winter), they were negative and less strong. Droughts were strongly linked to the CIs defined in some Argentinian regions.

The time series model showed a sound spatial characterization of the relationship between drought and crop yield. Soybean yield variability (impacts associated with agricultural droughts) responded better to DIs than to CIs. The SVHI and SPEI aggregated for 6 months and corresponding to the month of March (soybean growth season) were found to best explain the state of soybean production in certain regions.

The results provide useful drought insight tools in various parts of the water cycle and their association with variability in soybean production in Argentina Therefore, this research might be of interest to water managers and especially soybean producers, at national and regional level in Argentina. It may also serve as a foundation for future studies on drought in Argentina. It would be particularly interesting to explore the predictive rather than the explanatory capacity of the selected DIs and CIs for forecasting droughts in the country. The methodology is also of general applicability and relies on freely global-scale datasets, so it could be replicated in other regions of the world.

Acknowledgements

Ronnie Araneda gratefully acknowledges financial support from the Spanish Regional Government of Galicia (Xunta de Galicia) and the European Union through the predoctoral grant reference ED481A-2018/162. María Bermúdez was supported by the European Union H2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant Agreement No. 754446 and the Research and Transfer Fund of the University of Granada - Athenea3i.

Supplementary materials

Table S4.1 Median correlations between DIs and CIs aggregated for 6 and 12 months based on all departments (m=525) and 1982–2019 data (STWSI 2003–2019). The percentage of departments with significant correlations at 5% level (ρ ≤ 0.05) are shown.

DI		Darwin	Tahiti	Niño3.4	Niño3	Niño4	Niño1+2	SOI	PDO	CAR	SWIO	WTIO	SETIO	DMI	TNA	TSA	NAT	SAT	TASI	NAO
SPI-3	*	0.03	-0.12	0.18	0.13	0.16	0.06	-0.19	-0.02	-0.05	0.09	0.00	0.00	0.01	-0.07	-0.09	-0.06	-0.05	-0.10	-0.09
	**	3.6	65.0	72.4	66.5	62.7	19.2	75.2	39.2	48.0	48.6	50.3	23.8	44.6	38.5	47.6	26.7	26.3	56.8	47.2
SPEI-3	*	0.04	-0.13	0.20	0.14	0.16	0.07	-0.20	0.02	-0.05	0.14	0.01	0.03	-0.02	-0.08	-0.06	-0.06	-0.01	-0.06	-0.10
	**	5.9	71.0	80.2	77.9	71.2	23.4	83.8	49.5	43.8	73.0	49.5	33.9	41.0	42.9	27.8	22.5	18.9	35.4	52.0
SSI-3	*	0.02	-0.03	0.06	0.07	0.04	0.08	-0.07	-0.08	0.24	0.04	0.09	-0.02	0.13	0.11	0.01	0.05	-0.03	-0.06	0.14
	**	0.0	6.1	29.1	25.0	28.8	27.2	38.1	46.9	91.0	34.3	50.9	16.6	72.2	63.8	11.2	43.4	4.0	18.7	68.0
SVCI-3	*	0.01	-0.06	0.13	0.10	0.15	0.03	-0.07	-0.01	0.24	0.04	0.11	0.12	0.00	0.15	0.04	0.09	0.00	0.01	0.19
	**	1.3	38.7	63.8	53.5	76.4	9.3	55.0	38.3	85.3	43.0	61.1	64.6	37.3	70.9	35.4	53.7	18.3	25.9	77.3
STCI-3	*	0.10	-0.18	0.19	0.14	0.15	0.08	-0.32	0.23	-0.12	0.04	-0.04	0.02	-0.07	-0.19	-0.09	-0.15	-0.09	-0.11	-0.21
	**	60.0	86.5	90.9	85.1	82.9	29.7	94.9	90.1	60.4	28.0	42.1	27.4	29.5	82.9	49.9	72.4	42.9	65.9	90.9
SVHI-3	*	0.07	-0.15	0.19	0.14	0.18	0.06	-0.25	0.15	0.05	0.06	0.03	0.10	-0.07	-0.07	-0.03	-0.06	-0.05	-0.06	-0.05
	**	25.5	70.5	89.3	83.0	84.2	28.2	88.8	74.1	58.1	46.7	47.0	58.1	43.4	54.5	32.8	43.4	30.3	39.8	53.9
STWSI-3	*	*0.11*	*-0.15*	*0.27*	*0.15*	*0.39*	*0.06*	*-0.33*	*0.47*	*0.41*	*0.08*	*0.37*	*0.14*	*0.12*	*-0.03*	*0.28*	*0.00*	*0.12*	*0.16*	*0.19*
	**	14.9	61.3	79.6	54.5	86.7	1.0	84.6	90.7	81.3	19.4	86.7	57.9	60.2	13.7	81.5	18.7	45.3	62.7	79.8
SPDSI-3	*	0.03	-0.15	0.16	0.10	0.14	0.08	-0.20	0.06	0.00	0.07	0.00	0.08	-0.04	-0.12	-0.05	-0.07	-0.03	-0.05	-0.08
	**	12.0	68.8	70.3	53.9	72.8	42.7	76.4	57.1	70.5	41.9	60.6	58.3	38.5	59.6	32.0	41.7	22.7	34.7	53.7
SPI-6	*	0.07	-0.20	0.22	0.15	0.20	0.09	-0.24	-0.02	-0.04	0.09	-0.01	0.01	0.00	-0.09	-0.12	-0.04	-0.06	-0.09	0.07
	**	32.6	73.5	76.4	79.2	70.1	47.0	79.2	49.9	60.4	50.5	59.6	37.9	47.6	50.5	65.5	24.0	31.8	48.2	40.4
SPEI-6	*	0.08	-0.22	0.24	0.17	0.20	0.11	-0.26	0.03	-0.04	0.12	0.01	0.04	-0.03	-0.10	-0.07	-0.04	-0.01	-0.04	0.08
	**	46.1	78.3	83.4	84.6	73.5	64.0	84.6	59.4	51.0	56.6	57.5	44.0	43.2	57.0	37.5	19.6	27.6	30.1	43.6
SSI-6	*	0.02	-0.04	0.05	0.08	0.02	0.11	-0.06	-0.06	0.31	0.05	0.13	-0.04	0.19	0.16	0.08	0.07	0.02	0.00	-0.14
	**	1.5	17.0	28.8	42.7	33.0	61.5	38.1	43.0	94.1	40.8	82.9	25.0	87.0	71.4	45.5	54.1	13.5	16.2	79.8
SVCI-6	*	0.01	-0.07	0.12	0.09	0.16	0.04	-0.07	-0.02	0.23	0.04	0.11	0.11	0.01	0.13	0.02	0.09	0.00	0.01	-0.06

	**	6.3	52.8	62.5	51.2	73.7	13.1	63.2	38.3	86.1	43.0	58.7	59.4	39.6	67.0	34.9	53.7	18.3	25.9	47.0
STCI-6	*	0.14	-0.26	0.25	0.20	0.17	0.13	-0.38	0.22	-0.12	0.04	-0.04	0.01	-0.07	-0.22	-0.14	-0.15	-0.09	-0.11	0.19
	**	86.5	93.9	93.1	91.6	88.8	79.8	96.8	90.3	61.7	28.0	45.5	27.4	31.8	87.0	73.7	72.4	42.9	65.9	88.0
SVHI-6	*	0.10	-0.21	0.23	0.19	0.20	0.11	-0.30	0.14	0.04	0.06	0.04	0.09	-0.06	-0.09	-0.07	-0.06	-0.05	-0.06	0.10
	**	61.3	82.3	92.0	91.4	87.8	64.8	91.2	72.6	57.1	46.7	49.7	55.0	41.9	61.1	48.0	43.4	30.3	39.8	57.7
STWSI-6	*	*0.13*	*-0.22*	*0.32*	*0.21*	*0.43*	*0.13*	*-0.37*	*0.51*	*0.45*	*0.05*	*0.39*	*0.18*	*0.12*	*-0.07*	*0.30*	*0.01*	*0.14*	*0.19*	*0.15*
	**	28.8	79.0	82.7	73.9	87.4	27.0	86.3	90.9	83.2	9.3	87.0	62.7	63.8	24.2	84.0	21.5	51.0	82.7	56.4
SPDSI-6	*	0.05	-0.18	0.16	0.10	0.15	0.09	-0.23	0.06	0.00	0.04	-0.01	0.08	-0.04	-0.14	-0.06	-0.09	-0.05	-0.06	0.02
	**	21.5	73.5	71.4	58.3	76.6	50.7	79.6	61.9	72.0	28.4	64.0	59.2	40.6	62.5	38.7	48.0	26.7	36.6	33.7
SPI-12	*	0.18	-0.33	0.24	0.19	0.23	0.18	-0.27	-0.01	-0.03	0.09	-0.02	0.08	-0.05	-0.13	-0.16	-0.07	-0.09	-0.13	0.11
	**	88.2	85.9	77.0	89.1	74.1	75.6	87.2	55.2	65.9	48.6	65.7	61.3	49.1	75.0	82.5	42.9	49.3	70.1	57.1
SPEI-12	*	0.21	-0.36	0.24	0.21	0.20	0.21	-0.29	0.04	-0.03	0.12	-0.01	0.09	-0.07	-0.15	-0.11	-0.07	-0.05	-0.08	0.10
	**	90.1	86.3	84.0	90.1	79.8	79.6	86.9	64.6	53.3	55.8	57.9	65.9	48.6	77.3	63.0	43.4	31.6	41.9	55.4
SSI-12	*	0.07	-0.06	0.07	0.13	0.04	0.22	-0.08	-0.10	0.32	0.06	0.13	-0.04	0.23	0.15	0.04	0.07	0.00	-0.02	-0.16
	**	40.8	48.8	43.6	74.7	39.4	86.9	49.7	61.5	94.3	50.5	82.5	38.9	89.7	71.2	31.0	57.3	14.9	18.1	83.8
SVCI-12	*	0.02	-0.09	0.11	0.10	0.14	0.04	-0.06	-0.03	0.21	0.02	0.11	0.11	0.01	0.14	-0.01	0.08	-0.01	-0.01	-0.07
	**	51.8	73.7	60.8	60.4	67.2	38.5	68.2	39.4	81.5	41.9	59.0	61.9	44.6	69.5	42.3	53.9	23.8	30.7	54.5
STCI-12	*	0.34	-0.43	0.28	0.29	0.17	0.26	-0.42	0.20	-0.13	0.04	-0.07	0.02	-0.10	-0.28	-0.24	-0.18	-0.13	-0.16	0.27
	**	97.1	98.9	93.9	95.6	89.3	97.3	99.0	85.9	64.8	25.7	50.1	25.9	55.6	94.7	95.6	81.1	78.3	86.3	95.6
SVHI-12	*	0.24	-0.34	0.25	0.25	0.20	0.20	-0.32	0.11	0.03	0.05	0.03	0.09	-0.06	-0.12	-0.17	-0.08	-0.08	-0.10	0.16
	**	91.6	91.0	93.3	93.5	86.5	90.3	92.8	59.8	53.1	41.5	49.5	52.4	47.2	68.4	76.4	55.8	49.0	56.8	75.6
STWSI-12	*	*0.40*	*-0.37*	*0.38*	*0.30*	*0.48*	*0.25*	*-0.42*	*0.54*	*0.49*	*0.02*	*0.40*	*0.21*	*0.08*	*-0.17*	*0.25*	*-0.03*	*0.11*	*0.16*	*0.31*
	**	86.9	87.6	86.1	78.9	88.6	83.4	92.4	90.9	87.2	14.7	89.0	69.3	60.8	59.4	81.0	21.0	36.0	62.7	85.0
SPDSI-12	*	0.15	-0.30	0.18	0.12	0.16	0.15	-0.27	0.06	0.01	0.06	-0.03	0.08	-0.06	-0.19	-0.10	-0.15	-0.06	-0.08	0.07
	**	73.5	83.8	73.0	71.2	78.5	70.9	81.7	69.1	73.0	37.1	69.0	60.4	47.0	68.2	60.2	62.7	33.3	46.7	45.0

* |R|

** Percentage of departments with 5% significance level

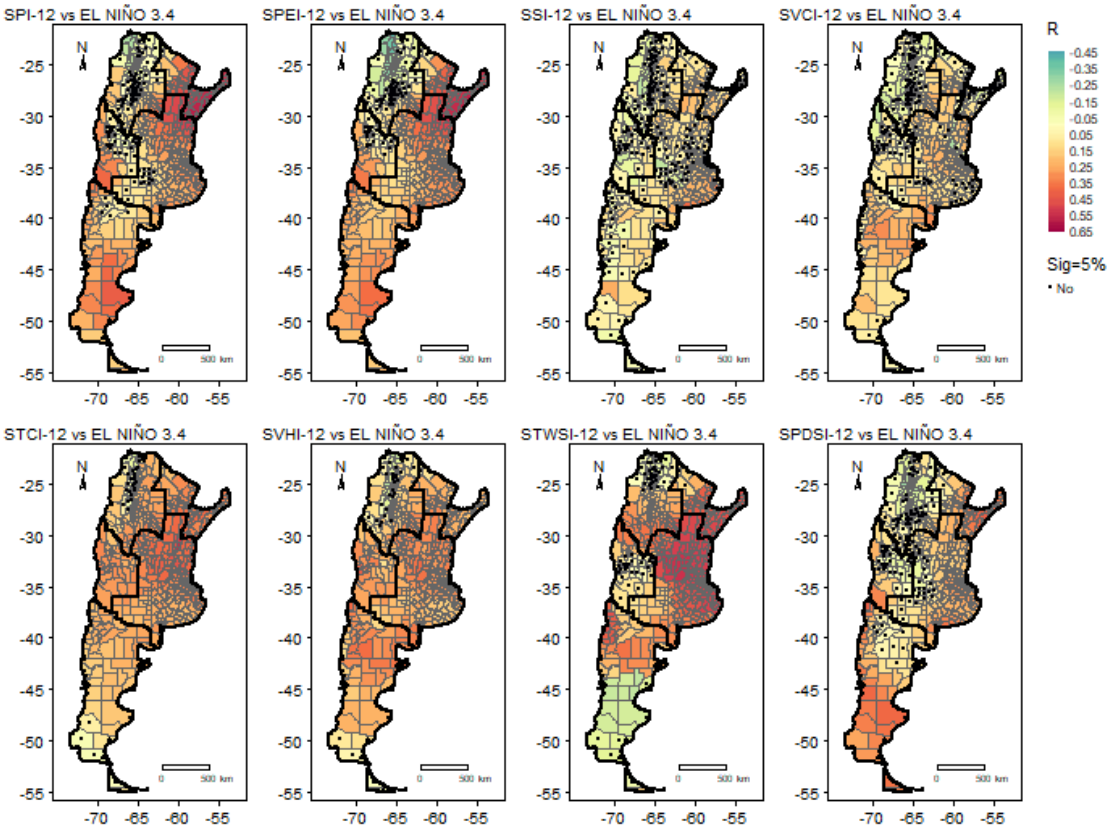


Fig. S4.1 Spatial pattern of correlations between the 8 DIs with a 12-month aggregation period and El Niño 3.4 CI based on 1982–2019 data (STWSI 2003–2019). Black dots indicate where correlations were not statistically significant.

Table S4.2 Median correlations between the seasonal time series of the DIs and the selected CIs (Thaiti SLP, El Niño 3.4, El Niño 4, and SOI) aggregated for 6 months based on all departments (m=525) and 1982–2019 data (STWSI 2003–2019).

Medians		Tahiti	El Niño 3.4	El Niño 4	SOI
SPI-6	Summer-ONDJFM	-0.24	0.30	0.30	*-0.31*
	Autumn-JFMAMJ	-0.23	0.22	0.27	-0.23
	Winter-AMJJAS	-0.13	0.13	0.15	-0.15
	Spring-JASOND	-0.23	0.22	0.21	-0.21
SPEI-6	Summer-ONDJFM	-0.26	*0.33*	*0.32*	-0.33
	Autumn-JFMAMJ	-0.26	0.27	0.30	-0.27
	Winter-AMJJAS	-0.14	0.15	0.16	-0.15
	Spring-JASOND	-0.21	0.19	0.18	-0.18
SSI-6	Summer-ONDJFM	-0.10	0.02	-0.06	-0.11
	Autumn-JFMAMJ	-0.12	0.04	-0.06	-0.11
	Winter-AMJJAS	-0.11	0.10	-0.03	*-0.18*
	Spring-JASOND	-0.13	0.07	-0.02	-0.18
SPDSI-6	Summer-ONDJFM	-0.15	0.11	0.12	-0.15
	Autumn-JFMAMJ	-0.27	0.24	0.24	*-0.29*
	Winter-AMJJAS	-0.18	0.22	0.22	-0.27
	Spring-JASOND	-0.14	0.08	0.09	-0.13
SVCI-6	Summer-ONDJFM	-0.14	0.17	*0.19*	-0.19
	Autumn-JFMAMJ	-0.17	0.17	0.17	-0.19
	Winter-AMJJAS	-0.07	0.12	0.15	-0.06
	Spring-JASOND	-0.08	0.13	0.18	-0.06
STCI-6	Summer-ONDJFM	-0.21	0.25	0.24	-0.30
	Autumn-JFMAMJ	-0.32	0.32	0.30	*-0.37*
	Winter-AMJJAS	-0.19	0.26	0.22	-0.31
	Spring-JASOND	-0.18	0.13	0.16	-0.25
SVHI-6	Summer-ONDJFM	-0.20	0.25	0.27	-0.28
	Autumn-JFMAMJ	-0.31	0.30	0.29	*-0.34*
	Winter-AMJJAS	-0.17	0.24	0.23	-0.24
	Spring-JASOND	-0.16	0.15	0.20	-0.19
STWSI-6	Summer-ONDJFM	-0.25	0.25	*0.36*	-0.34
	Autumn-JFMAMJ	-0.38	0.39	0.48	-0.45
	Winter-AMJJAS	-0.26	0.40	*0.50*	-0.47
	Spring-JASOND	-0.25	0.17	0.34	-0.33

Table S4.3 Median correlation coefficients between the seasonal time series of the DIs and El Niño 3.4 aggregated for 6 months based on the departments of Cuyo (m=44), Patagonia (m=53), Pampas (m=233), NEA (m=76), and NWA (m=119) and 1982-2019 data (STWSI 2003–2019).

Region	DI	Summer - ONDJFM	Autumn - JFMAMJ	Winter - AMJJAS	Spring - JASOND
Cuyo	SPI-6	*0.08*	*0.02*	-0.07	-0.02
	SPEI-6	*0.00*	*0.11*	0.02	-0.12
	SSI-6	*-0.07*	*-0.05*	-0.11	-0.06
	SPDSI-6	*0.00*	*-0.04*	-0.10	-0.03
	SVCI-6	*-0.04*	*-0.01*	-0.02	-0.06
	STCI-6	*-0.26*	*0.30*	0.21	-0.35
	SVHI-6	*-0.18*	*0.22*	0.12	-0.27
	STWSI-6	*-0.27*	*0.21*	0.26	-0.30
Pampas	SPI-6	*-0.26*	*0.26*	0.32	-0.25
	SPEI-6	*-0.27*	*0.31*	0.34	-0.29
	SSI-6	*-0.15*	*0.04*	-0.05	-0.14
	SPDSI-6	*-0.33*	*0.27*	0.34	-0.32
	SVCI-6	*-0.24*	*0.22*	0.27	-0.25
	STCI-6	*-0.37*	*0.34*	0.35	-0.40
	SVHI-6	*-0.35*	*0.33*	0.37	-0.37
	STWSI-6	*-0.54*	*0.54*	0.62	-0.62
Patagonia	SPI-6	*-0.17*	*0.12*	0.10	-0.19
	SPEI-6	*-0.25*	*0.26*	0.21	-0.30
	SSI-6	*-0.09*	*0.05*	0.04	-0.10
	SPDSI-6	*-0.28*	*0.26*	0.20	-0.29
	SVCI-6	*-0.23*	*0.17*	0.14	-0.25
	STCI-6	*-0.34*	*0.24*	0.19	-0.34
	SVHI-6	*-0.36*	*0.24*	0.18	-0.36
	STWSI-6	*-0.30*	*0.32*	0.42	-0.33
NEA	SPI-6	*-0.45*	*0.53*	0.47	-0.49
	SPEI-6	*-0.43*	*0.53*	0.46	-0.48
	SSI-6	*-0.17*	*0.10*	0.00	-0.18
	SPDSI-6	*-0.36*	*0.39*	0.32	-0.39
	SVCI-6	*-0.17*	*0.19*	0.20	-0.17
	STCI-6	*-0.32*	*0.40*	0.39	-0.37
	SVHI-6	*-0.29*	*0.36*	0.35	-0.33
	STWSI-6	*-0.36*	*0.51*	0.53	-0.46
NWE	SPI-6	*-0.02*	*0.10*	0.05	-0.07
	SPEI-6	*0.05*	*0.04*	-0.01	-0.01
	SSI-6	*-0.10*	*0.02*	-0.09	-0.10
	SPDSI-6	*0.03*	*0.01*	-0.06	0.00
	SVCI-6	*-0.04*	*0.05*	0.03	-0.05
	STCI-6	*-0.19*	*0.24*	0.19	-0.27
	SVHI-6	*-0.14*	*0.16*	0.11	-0.19
	STWSI-6	*0.04*	*0.17*	0.13	-0.05

Table S4.4 Median correlation coefficients between the seasonal time series of the DIs and the selected CIs (Thaiti SLP, El Niño 3.4, El Niño 4, and SOI) aggregated for 6 months based on the departments of Cuyo (m=44), Patagonia (m=53), Pampas (m=233), NEA (m=76), and NWA (m=119) and 1982–2019 data (STWSI 2003–2019).

Region	DI	Tahiti				El Niño 3.4				El Niño 4				SOI			
		a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
Cuyo	SPI-6	*0.03*	0.02	-0.07	-0.02	*0.08*	0.02	-0.07	-0.02	*0.00*	0.08	-0.04	-0.12	*-0.03*	0.06	-0.06	-0.09
	SPEI-6	*-0.01*	0.07	-0.01	-0.08	*0.00*	0.11	0.02	-0.12	*-0.03*	0.11	0.00	-0.12	*-0.01*	0.03	-0.07	-0.02
	SSI-6	*-0.07*	-0.04	-0.08	-0.06	*-0.07*	-0.05	-0.11	-0.06	*-0.10*	0.03	-0.08	-0.14	*-0.13*	0.03	-0.05	-0.15
	SPDSI-6	*-0.01*	-0.04	-0.08	0.01	*0.00*	-0.04	-0.10	-0.03	*0.00*	0.01	-0.09	-0.08	*-0.03*	0.02	-0.06	-0.04
	SVCI-6	*-0.03*	0.02	0.03	-0.05	*-0.04*	-0.01	-0.02	-0.06	*0.01*	0.04	0.02	0.00	*0.03*	0.07	0.09	0.00
	STCI-6	*-0.16*	0.23	0.20	-0.28	*-0.26*	0.30	0.21	-0.35	*-0.18*	0.27	0.18	-0.34	*-0.19*	0.14	0.14	-0.28
	SVHI-6	*-0.10*	0.19	0.17	-0.23	*-0.18*	0.22	0.12	-0.27	*-0.12*	0.20	0.12	-0.26	*-0.10*	0.15	0.15	-0.19
	STWSI-6	*-0.19*	0.16	0.25	-0.25	*-0.27*	0.21	0.26	-0.30	*-0.24*	0.19	0.26	-0.36	*-0.25*	0.12	0.25	-0.30
Patagonia	SPI-6	*-0.21*	0.27	0.24	-0.32	*-0.17*	0.12	0.10	-0.19	*-0.11*	0.13	0.09	-0.23	*-0.27*	0.33	0.29	-0.42
	SPEI-6	*-0.26*	0.33	0.30	-0.36	*-0.25*	0.26	0.21	-0.30	*-0.15*	0.16	0.10	-0.27	*-0.26*	0.30	0.25	-0.40
	SSI-6	*-0.06*	0.00	0.01	-0.09	*-0.09*	0.05	0.04	-0.10	*-0.08*	0.11	0.02	-0.16	*-0.09*	0.05	0.00	-0.14
	SPDSI-6	*-0.21*	0.24	0.25	-0.33	*-0.28*	0.26	0.20	-0.29	*-0.19*	0.22	0.16	-0.29	*-0.25*	0.21	0.21	-0.34
	SVCI-6	*-0.15*	0.20	0.20	-0.25	*-0.23*	0.17	0.14	-0.25	*-0.10*	0.13	0.12	-0.12	*-0.08*	0.16	0.20	-0.13
	STCI-6	*-0.23*	0.23	0.22	-0.31	*-0.34*	0.24	0.19	-0.34	*-0.20*	0.21	0.17	-0.31	*-0.19*	0.18	0.23	-0.29
	SVHI-6	*-0.23*	0.26	0.27	-0.36	*-0.36*	0.24	0.18	-0.36	*-0.21*	0.23	0.20	-0.30	*-0.21*	0.27	0.33	-0.33
	STWSI-6	*-0.14*	0.27	0.40	-0.34	*-0.30*	0.32	0.42	-0.33	*-0.19*	0.28	0.39	-0.30	*-0.17*	0.20	0.38	-0.31
Pampas	SPI-6	*-0.31*	0.35	0.40	-0.36	*-0.26*	0.26	0.32	-0.25	*-0.14*	0.16	0.17	-0.17	*-0.30*	0.23	0.26	-0.26
	SPEI-6	*-0.31*	0.36	0.40	-0.37	*-0.27*	0.31	0.34	-0.29	*-0.15*	0.18	0.18	-0.18	*-0.28*	0.22	0.24	-0.24
	SSI-6	*-0.10*	0.01	-0.06	-0.11	*-0.15*	0.04	-0.05	-0.14	*-0.14*	0.12	-0.01	-0.20	*-0.14*	0.07	-0.02	-0.18
	SPDSI-6	*-0.21*	0.15	0.22	-0.22	*-0.33*	0.27	0.34	-0.32	*-0.22*	0.24	0.30	-0.28	*-0.21*	0.08	0.17	-0.19
	SVCI-6	*-0.20*	0.21	0.25	-0.24	*-0.24*	0.22	0.27	-0.25	*-0.10*	0.17	0.23	-0.09	*-0.12*	0.16	0.23	-0.08
	STCI-6	*-0.27*	0.28	0.32	-0.35	*-0.37*	0.34	0.35	-0.40	*-0.22*	0.27	0.25	-0.35	*-0.22*	0.15	0.19	-0.28
	SVHI-6	*-0.27*	0.29	0.34	-0.34	*-0.35*	0.33	0.37	-0.37	*-0.20*	0.26	0.27	-0.28	*-0.22*	0.18	0.24	-0.24
	STWSI-6	*-0.33*	0.29	0.41	-0.43	*-0.54*	0.54	0.62	-0.62	*-0.39*	0.50	0.61	-0.61	*-0.37*	0.21	0.41	-0.40
NEA	SPI-6	*-0.33*	0.53	0.47	-0.42	*-0.45*	0.53	0.47	-0.49	*-0.28*	0.34	0.28	-0.29	*-0.25*	0.36	0.28	-0.22
	SPEI-6	*-0.30*	0.50	0.43	-0.39	*-0.43*	0.53	0.46	-0.48	*-0.25*	0.33	0.25	-0.26	*-0.22*	0.29	0.23	-0.17

NWE	SSI-6	*-0.12*	0.06	-0.01	-0.14	*-0.17*	0.10	0.00	-0.18	*-0.13*	0.17	0.03	-0.20	*-0.14*	0.11	0.03	-0.19
	SPDSI-6	*-0.16*	0.18	0.14	-0.16	*-0.36*	0.39	0.32	-0.39	*-0.27*	0.39	0.32	-0.35	*-0.15*	0.16	0.14	-0.11
	SVCI-6	*-0.11*	0.18	0.19	-0.11	*-0.17*	0.19	0.20	-0.17	*-0.06*	0.11	0.15	-0.02	*-0.03*	0.10	0.14	0.03
	STCI-6	*-0.18*	0.25	0.23	-0.25	*-0.32*	0.40	0.39	-0.37	*-0.14*	0.29	0.28	-0.23	*-0.06*	0.06	0.12	-0.07
	SVHI-6	*-0.16*	0.25	0.26	-0.20	*-0.29*	0.36	0.35	-0.33	*-0.12*	0.24	0.26	-0.15	*-0.04*	0.09	0.15	-0.01
	STWSI-6	*-0.18*	0.27	0.35	-0.32	*-0.36*	0.51	0.53	-0.46	*-0.23*	0.46	0.54	-0.45	*-0.15*	0.16	0.33	-0.25
	SPI-6	*-0.12*	0.13	0.11	-0.12	*-0.02*	0.10	0.05	-0.07	*0.08*	-0.05	-0.11	0.11	*-0.03*	0.00	-0.03	0.03
	SPEI-6	*-0.07*	0.06	0.03	-0.04	*0.05*	0.04	-0.01	-0.01	*0.12*	-0.08	-0.13	0.16	*0.02*	-0.08	-0.11	0.12
	SSI-6	*-0.10*	0.02	-0.07	-0.11	*-0.10*	0.02	-0.09	-0.10	*-0.10*	0.09	-0.06	-0.16	*-0.14*	0.07	-0.04	-0.18
	SPDSI-6	*0.01*	-0.03	-0.09	0.03	*0.03*	0.01	-0.06	0.00	*0.06*	-0.06	-0.12	0.04	*0.06*	-0.07	-0.15	0.05
	SVCI-6	*-0.07*	0.13	0.11	-0.12	*-0.04*	0.05	0.03	-0.05	*-0.01*	-0.01	0.01	0.03	*-0.05*	0.10	0.10	-0.05
	STCI-6	*-0.12*	0.10	0.12	-0.17	*-0.19*	0.24	0.19	-0.27	*-0.09*	0.23	0.16	-0.23	*-0.06*	0.05	0.07	-0.13
	SVHI-6	*-0.10*	0.12	0.13	-0.16	*-0.14*	0.16	0.11	-0.19	*-0.05*	0.12	0.09	-0.15	*-0.06*	0.08	0.11	-0.13
	STWSI-6	*0.10*	0.05	0.05	-0.01	*0.04*	0.17	0.13	-0.05	*-0.05*	0.25	0.27	-0.28	*0.01*	0.08	0.15	-0.18

a: Summer-ONDJFM c: Winter-AMJJAS
b: Autumn-JFMAMJ d: Spring-JASOND

Table S4.5 Median seasonal correlations between DIs and the selected CIs (Thaiti SLP, El Niño 3.4, El Niño 4, and SOI) aggregated for 6 months based on soybean-producing departments (m=193) and 1982–2019 data (STWSI 2003–2019).

DIs		Tahiti	El Niño 3.4	El Niño 4	SOI
SPI-6	Summer-ONDJFM	-0.29	0.35	0.38	-0.34
	Autumn-JFMAMJ	-0.26	0.28	0.33	-0.27
	Winter-AMJJAS	-0.15	0.19	0.19	-0.18
	Spring-JASOND	-0.26	0.22	0.22	-0.22
SPEI-6	Summer-ONDJFM	*-0.29*	*0.37*	*0.38*	*-0.35*
	Autumn-JFMAMJ	-0.28	0.33	0.35	-0.31
	Winter-AMJJAS	-0.15	0.19	0.19	-0.18
	Spring-JASOND	-0.24	0.19	0.20	-0.20
SSI-6	Summer-ONDJFM	-0.12	0.03	-0.03	-0.15
	Autumn-JFMAMJ	-0.17	0.06	-0.03	-0.16
	Winter-AMJJAS	-0.16	0.14	0.01	-0.23
	Spring-JASOND	-0.17	0.09	0.01	-0.22
SPDSI-6	Summer-ONDJFM	-0.16	0.10	0.13	-0.14
	Autumn-JFMAMJ	-0.29	0.26	0.31	-0.30
	Winter-AMJJAS	-0.20	0.25	0.29	-0.28
	Spring-JASOND	-0.15	0.08	0.11	-0.12
SVCI-6	Summer-ONDJFM	-0.20	0.22	0.24	-0.23
	Autumn-JFMAMJ	-0.23	0.22	0.27	-0.24
	Winter-AMJJAS	-0.09	0.15	0.23	-0.08
	Spring-JASOND	-0.12	0.16	0.22	-0.07
STCI-6	Summer-ONDJFM	-0.27	0.29	0.33	-0.35
	Autumn-JFMAMJ	-0.37	0.36	0.38	-0.40
	Winter-AMJJAS	-0.22	0.28	0.28	-0.35
	Spring-JASOND	-0.23	0.15	0.20	-0.28
SVHI-6	Summer-ONDJFM	-0.26	0.28	0.32	-0.32
	Autumn-JFMAMJ	-0.34	0.33	0.37	-0.36
	Winter-AMJJAS	-0.20	0.26	0.29	-0.27
	Spring-JASOND	-0.22	0.17	0.23	-0.23
STWSI-6	Summer-ONDJFM	-0.32	0.28	0.40	-0.41
	Autumn-JFMAMJ	-0.50	0.49	0.58	-0.59
	Winter-AMJJAS	*-0.37*	*0.48*	*0.59*	*-0.59*
	Spring-JASOND	-0.35	0.20	0.40	-0.38

Chapter 5

Part 1: A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation drought index

Ronnie J. Araneda-Cabrera, María Bermúdez and Jerónimo Puertas.

Communication: 6th IAHR Europe Congress, 2020, Warsaw, Poland

Date: July 2020

ISBN: 9781787284395

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – original draft, Visualization. María Bermúdez: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Resources, Writing – review & editing, Supervision.

Part 2: Bivariate standardized precipitation and vegetation indices for assessing and monitoring agricultural drought

Ronnie J. Araneda-Cabrera, María Bermúdez and Jerónimo Puertas.

Journal: Revista Hidrolatinoamericana, Vol. 5 (2021)

Date: 07 September 2021

ISSN: 2520-2960

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – original draft, Visualization. María Bermúdez: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Jerónimo Puertas: Conceptualization, Resources, Writing – review & editing, Supervision.

Part 1: A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation drought index

Abstract

Droughts are one of the most costly and harmful hazards worldwide. Several drought indices have been developed in order to characterize and monitor this phenomenon and help decision-makers to reduce the associated socio-economic risks. In this study, we present a methodology for calculating a multivariate standardized precipitation and vegetation index (MSPVI). It uses as input the precipitation and the satellite-based vegetation condition index. We show the statistical procedure for calculating it and then applied it in Mozambique as a case of study. We found the MSPVI is more sensitive to the onset and recovery of agricultural droughts than the SPI and SVCI. The MSPVI series has significantly negative trends with Mann-Kendall Z values higher than SPI and SVCI. The MSPVI can be a useful tool for decision-makers.

Keywords

Agriculture drought; multivariate; crop yield; drought index; monitoring.

1. Introduction

Many areas of the world face droughts events and according to IPCC, 2014 these events will increase in frequency and intensity because of climate change. Among others, one of the most important fields affected by droughts is agriculture, which is related to food security and socio-economic crisis. For this reason, various drought indices have been developed in order to characterize and monitor agricultural droughts and reduce the associated socioeconomic risk.

The drought indices are an integration of one or more hydrological variables (Hao et al., 2015). For instance, the Standardized Precipitation Index (SPI) (McKee et al., 1993) accumulated to three months is extended and used in agricultural drought monitoring. Nevertheless, nowadays satellite observations let us obtain new indicators for this purpose, related to vegetation condition. The Vegetation Condition Index (VCI) has been used in various agricultural droughts studies showing good results (Kogan, 1995).

In this contribution, we propose a methodology that integrates both meteorological and satellite-based vegetation variables. Under the assumption that the bivariate extreme value distribution with the Gumbel marginal distributions can be used to model the monthly precipitation data (P) and the monthly VCI, precipitation (P) and VCI are joined into a new multivariate standardized index. This index could characterize agricultural drought with the advantage of explaining the condition of vegetation and precipitation variations at the same time showing a complete status of a region or river basin. Then, we apply this index over Mozambique (case of study) and compare its performance with the SPI and standardized VCI (SVCI).

2. The multivariate standardized precipitation and vegetation index (MSPVI)

The developed method for characterizing agricultural droughts is a multivariate approach that depends on two individual variables (Mehran et al., 2015): P and VCI. First, P is averaging in three months periods. Then the non-exceedance probabilities of both variables are calculated following the formula (Gringorten, 1963):

$$P_t = \frac{I-0.44}{N+0.12} \tag{7}$$

where P_t is the cumulative frequency at month t $(1 \ldots N)$, I is the position from the smallest to largest and N denotes the sample size. Later, we transform the empirical probabilities into a standardized index (SI):

$$SI_t = \varphi^{-1}(P_t) \tag{2}$$

where φ denotes the standard normal distribution function. Next, the two univariate indicators are combined using the multivariate framework explained in Yue et al., 1999. Changing in (1) the position I for the number of occurrences of the pair $(SI(P), SI(VCI))$ for $SI(P) \leq SI(P_t)$ and $SI(VCI) \leq SI(VCI_t)$ we obtained the joint bivariate empirical probability at month t. Finally, using (2) we calculate the multivariate standardized precipitation and vegetation index (MSPVI). This index is interpreted similarly to the original SPI and can be computed for various time spans accumulating the raw variables. Based on our methodology, SPI and SVCI have different temporal patterns than the

originals due to the probability correspondent to any given quantile of a joint distribution of two variables is not identical to that of the univariate distribution of each individual variable.

3. Case of study: Mozambique

Mozambique is one of the poorest and least developed countries in the world. About 60% of the population of this country works in agriculture, representing about 24% of the GDP. It is located in one of the most drought and climate change prone areas and has a very high risk of food scarcity. Using monthly country-averaged P data obtained from the Climatic Research Unit CRU TS3.10 (CRU) from 1981 to 2017 and monthly VCI obtained from NOAA STAR (same period), we calculated the MSPVI. Additionally, the SPI-3 and SVCI were calculated in order to compare them with the new index.

The temporal patterns are shown in Fig. 5.1. It can be seen that MSPVI shows critical information about the onsets and recoveries of drought events. Furthermore, it detects the main historical drought record according to the International Disaster Database EM-DAT, where the 1991-1992 was one of the most harmful drought events. The MSPVI shows a significantly decreasing trend of -3.53 (z-value), while SPI and SVCI -1.72 and -3.28 respectively. Spatial distribution of the 1992 (May) drought event are plotted in Fig. 5.2. It can be seen the MSPVCI identified a major area affected by droughts. At national scale correlation between annual benchmark crops yields and DI of July are sowed in Table 5.1. Improvements using MSPVCI instead SPI or SVCI are noted.

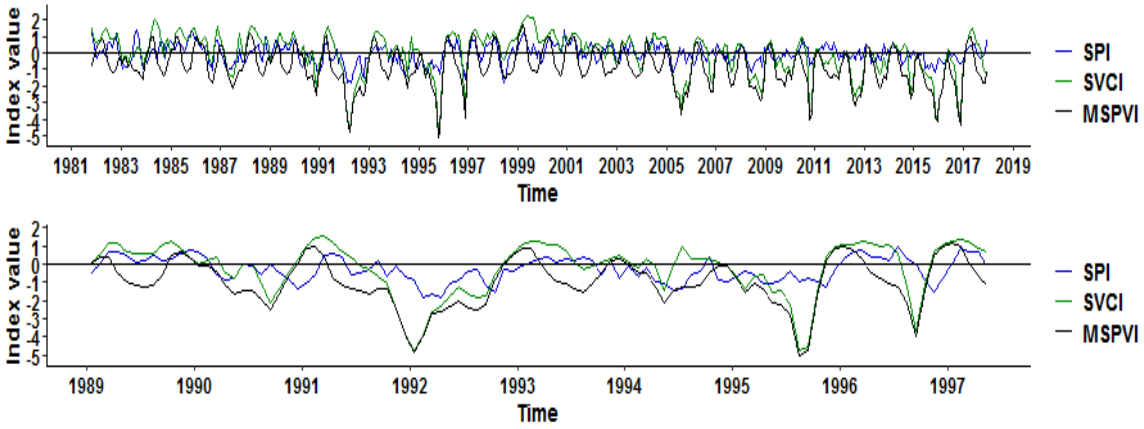


Fig. 5.1 The SPI-3, SVCI and the MSPVI series in 1982-2017 (above) and 1989-1997 (below).

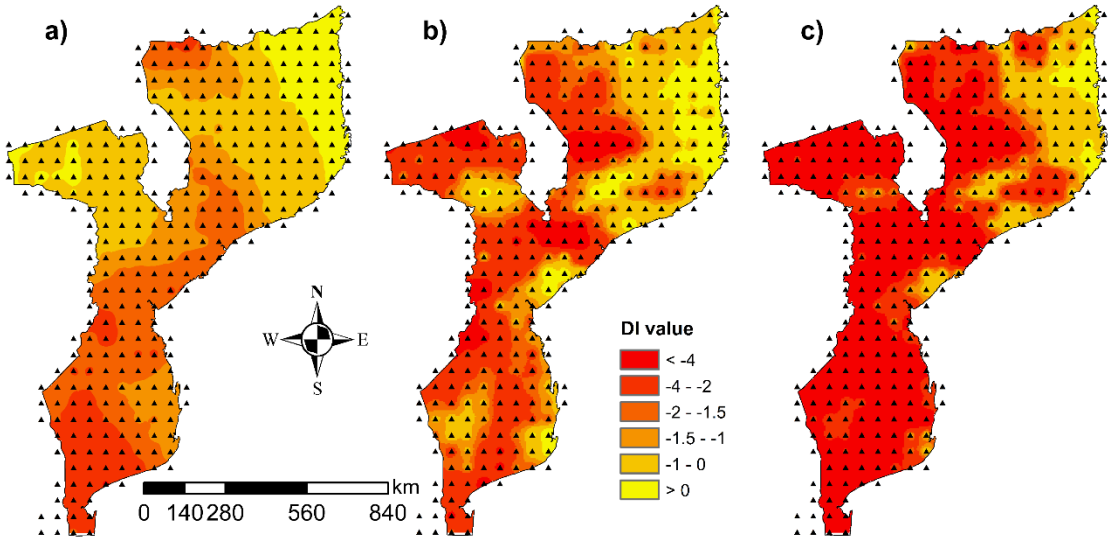


Fig. 5.2 Spatial distribution of the a) SPI-3, b) SVCI and c) MSPVI during May/1992 drought event

Table 5.1 Pearson correlation coefficients between Mozambican annual crop yields and drought indices from July.

Crop	MSPVCI	SPI-3	SVCI
Maize	0.593	0.515	0.172
Sorghum	0.833	0.3735	0.247
Wheat	0.538	0.363	0.003

4. Conclusions

In this study, we present a methodology for calculating the multivariate standardized precipitation and vegetation index (MSPVI). We calculated the MSPVI from open global datasets over Mozambique as a case of study. The proposed agricultural drought index is more sensitive to the onset and recovery of droughts, while it detects a more conservative spatial distribution of droughts than the SPI and SVCI. The MSPVI series has significantly negative trends higher than SPI and SVCI. Further it explains better the annual variability of maize and sorghum yield. The MSPVCI may be used for evaluating and monitoring agricultural drought being a useful tool for decision-makers.

Supplementary materials

Appendix B shows the complementary poster exposed in the 6th IAHR Europe Congress, 2020.

Acknowledgments

Ronnie Araneda gratefully acknowledges financial support from the Spanish Regional Government of Galicia (Xunta de Galicia) and the European Union through the predoctoral grant reference ED481A-2018/162, and the IACOBUS Program. María Bermúdez was supported by the European Union H2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 754446 and the Research and Transfer Fund of the University of Granada - Athenea3

Part 2: Bivariate standardized precipitation and vegetation indices for assessing and monitoring agricultural drought

Abstract

In this study, we assess two bivariate standardized precipitation and vegetation indices (BSPVCI and BSPVHI) as a tool for monitoring agricultural droughts, which are associated with losses in crop production and related to food security. These indices combine in a probabilistic framework the Standardized Precipitation Index (SPI) with the Vegetation Condition Index (VCI) and Vegetation Health Index (VHI). We found that both indices are more sensitive to drought onset and recovery, and explain crop variability considerably better than SPI, VCI or VHI alone in Argentina (case study).

Keywords

Agriculture drought; bivariate; crop yield; drought index; monitoring.

1. Introduction

Droughts are among the natural phenomena that cause the greatest socio-economic and ecological damage (Golnaraghi et al., 2014). They are classified into meteorological, agricultural, hydrological and socio-economic droughts. The first are understood as a deficit in the normal precipitation volumes in a given region from which the others are triggered (Wilhite and Glantz, 1985). Agricultural droughts are of particular interest as they can influence the variability of agricultural production, potentially compromising food security.

To assess and monitor droughts, several drought indices (DI) have been developed in recent decades (Svodova et al., 2016). Among the most widely used DIs for monitoring agricultural droughts is the Standardised Precipitation Index (SPI) (McKee et al., 1993), which is based on precipitation and recommended by the World Meteorological Organisation. Other DIs are the Vegetation Condition Index (VCI) and Vegetation Health Index (VHI) (Kogan, 1995) based on remotely sensed data that detect and classify vegetation stress caused by drought.

In this research we propose two new agricultural drought indices for assessment and monitoring. These are two composite DIs that take into account meteorological and agricultural drought conditions by combining precipitation (the SPI) with VCI and VHI in a probabilistic framework. These new DIs are evaluated and compared with the traditional SPI, VCI and VHI.

2. Methodology

2.1 Case study

Continental Argentina (Fig. 5.3) has a wide climatic variability. The average annual precipitation per department varies between 70 and 1880 mm per year, and the average annual temperature ranges between 2 and 23 °C. Both precipitation and temperature increase from east to west and from south to north. Argentina is one of the largest grain producers in the world, with soybean and maize being the most important crops (FAO, 2017). The South American country has faced important drought events in the past, with the most important ones occurring in 1988-1989, 1994-1996, 2009 and 2018 (EM-DAT, 2021), with strong links to the El Niño Southern Oscillation (ENSO) (Araneda-Cabrera et al., 2021a). In these years, a significant decrease in agricultural production was detected, which is concentrated in the northwest of the country (Fig. 5.3) and is mostly rainfed (> 95%). For these reasons, Argentina was taken as a case study to evaluate the new agricultural drought indices.

2.2 Data

Monthly precipitation series were downloaded from the TerraClimate database (Abatzoglou et al., 2018), while weekly VCI and VHI series were downloaded from NOAA STAR. The two databases provide their products at a spatial scale of 1/24° (≈ 4.62km), which were obtained for the period 1981-2019 over the whole country. The 3 variables were aggregated at departmental level (525 departments) and monthly over the whole of Argentina.

In each department, the 3-month cumulative SPI was computed following the steps described in McKee et al., (1993). The 3-month temporal accumulation was chosen since agricultural droughts are a seasonal phenomenon and have been found in other studies to be the best option to explain agricultural variability (McKee et al., 1993). To be comparable with the SPI, the same standardisation process was applied to the VCI and VHI series, obtaining the SVCI and SVHI. Annual agricultural yield data for soybean and maize were downloaded from the Ministry of Agriculture, Livestock and Fisheries of Argentina for the period 2004-2019.

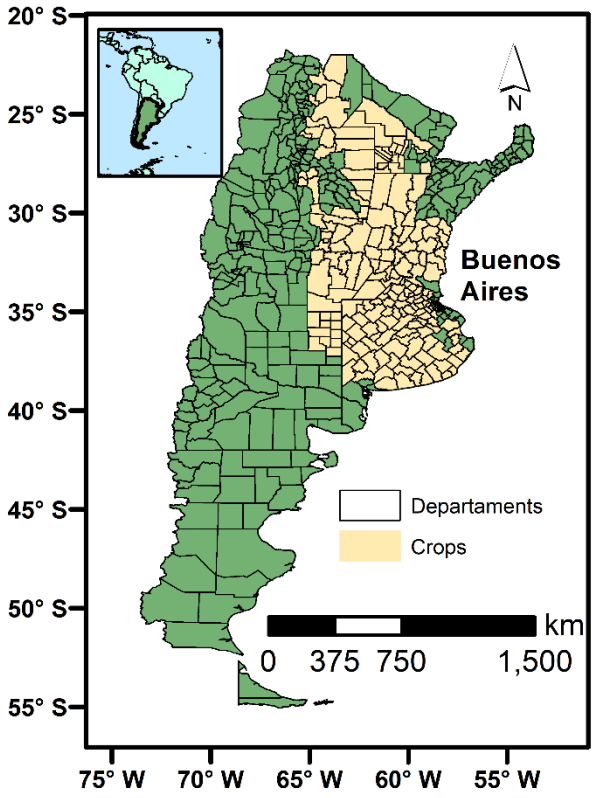


Fig. 5.3 Location of Argentina and its departments. The departments with soybean and maize crops are detailed.

2.3. Bivariate Standardised Precipitation and Vegetation Indices (BSPVCI and BSPVHI)

The approach of each index depends on two individual variables, precipitation and VCI for the BSPVCI and VHI for the BSPVHI. In both cases the procedure starts by accumulating precipitation (Pr) in 3-month periods. Then, the non-exceedance probabilities of both variables are calculated following the formula of (Gringorten, 1963):

$$P_t = \frac{I-0.44}{N+0.12} \tag{8}$$

Where P_t is the cumulative frequency in month t $(1 \ldots N)$, I is the position from lowest to highest and N denotes the sample size. Subsequently, the empirical probabilities are transformed into a standardised index (SI):

$$SI_t = \varphi^{-1}(P_t) \tag{2}$$

where φ denotes the standard normal distribution function. Next, the two univariate indices are combined using the multivariate framework explained in Yue et al. (1999). Changing in (1) the position I by the number of occurrences of the pair $(SI(\mathrm{Pr}), SI(VCI/VHI))$ by $SI(\mathrm{Pr}) \le SI(P_t)$ and $SI(VCI/VHI) \le SI(VCI/VHI_t)$ we obtain the joint bivariate empirical probability in month t. Finally, using (2) we calculate the standardised bivariate precipitation and vegetation indices (BSPVCI and BSPVHI). These indices are interpreted similarly to the original SPI. Based on our methodology,

the BSPVCI and BSPVHI have different temporal patterns from the original variables, because the probability corresponding to any quantile in a joint distribution is not identical to that of the distribution of each individual variable.

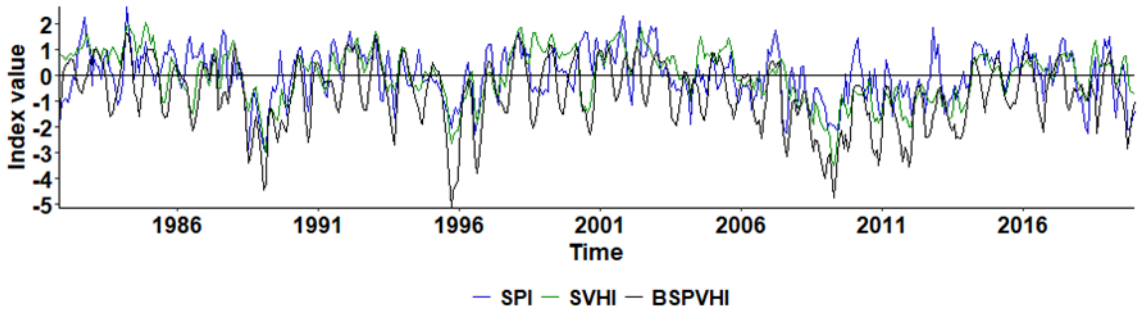


Fig. 5.4 SPI, SVHI and BSPVHI national time series.

2.4. Statistical crop model

In order to compare DI as predictors of variability in maize and soybean yields (associated with agricultural droughts), the time-series model used in Araneda-Cabrera et al., (2021):

$$\ln(Y_t)=\beta_0+\beta_1(X)+\varepsilon_t \quad (3)$$

where Y is the vector of the annual yield of each crop in the whole country, X is the vector with the predictors, which in this case were the SPI, SVCI, SVHI, BSPVCI and BSPVHI for the month of March (maize and soybean harvest season), β_0 y β_1 are the parameters to be adjusted and ε_t is the error. Since variations in yields can be affected by factors external to the weather (e.g., improved technologies and/or seeds), the crop yield series was previously trend extracted using a fitted linear regression model

3. Results and discussion

3.1 Assessment of BSPVCI and BSPVHI

The SPI, SVHI and BSPVHI detected the main drought events that have been recorded in recent years (Fig. 5.4). However, the BSPVHI showed a higher sensitivity to these events. For example, in March 2009 (major drought of the last decades in Argentina) it reached an intensity of -4.79, much higher than -3.51 or -1.98 of the SVHI and SPI, respectively. Spatially (Fig. 5.5), during March 2009, the BSPVHI reported almost 50% of the Argentinean territory under extreme drought (IS < -2.0) while the SVHI reported 22% and SPI 11%. The BSPVCI showed very similar spatio-temporal patterns to the BSPVHI.

These results indicate that bivariate DI are more sensitive to drought events, so using them as an agricultural drought monitoring tool would imply being on the safe side, a result that Monteleone et al., (2020) agrees with. However, the use of these indicators needs to be further developed through specific drought validation and characterisation studies.

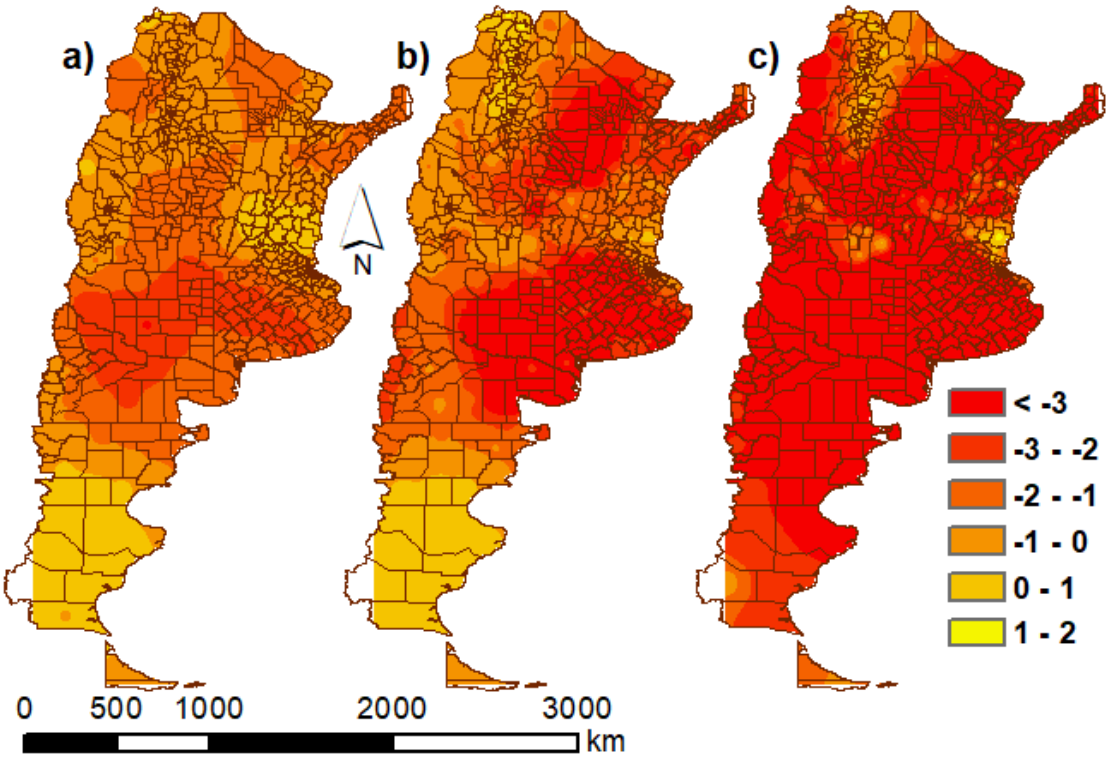


Fig. 5.5 Spatial distribution of droughts in Argentina in March 2009 according to a) SPI, b) SVHI and c) BSPVHI.

3.3 Relationship between drought indices and crop yields

Table 1 shows the results of applying the statistical crop model using the traditional DIs as predictors. Both SVHI and SVCI explained up to 68% of the annual variability of soybean and 67% of maize yields. However, both BSPVCI and BSPVHI improved model performance by explaining approximately 77% of the annual variability for soybean and 73% for maize (Fig. 5.6). The SVHI and BSPVHI introduce temperature in their computation, so it could explain the better performance compared to SVCI or BSPVCI. Nevertheless, both DI could be used to monitor soybean and maize crops at the national level. These indices could be used as predictors of agricultural yields. However, specific studies are needed to analyse this predictive capacity.

Table 5.1 Coefficients of determination (R2) between soybean and maize yields and drought indices 2004-2019.

	SPI	**SVCI**	**SVHI**
Soybean	0.476	0.572	0.686
Maize	0.306	0.407	0.670

4. Conclusions

The bivariate indices BSPVCI and BSPVHI were shown to be more sensitive to the onset and end of droughts across Argentina. Compared to the classical SPI and VCI/VHI, these indices report higher spatial coverage of extreme droughts and better explain the variability of annual maize and soybean yields. The proposed drought indices could be useful for water managers in Argentina and other countries or regions, especially for soybean and maize farmers. This study opens the door to future studies to improve the use of these drought indices, such as assessing the predictive capacity of agricultural yields.

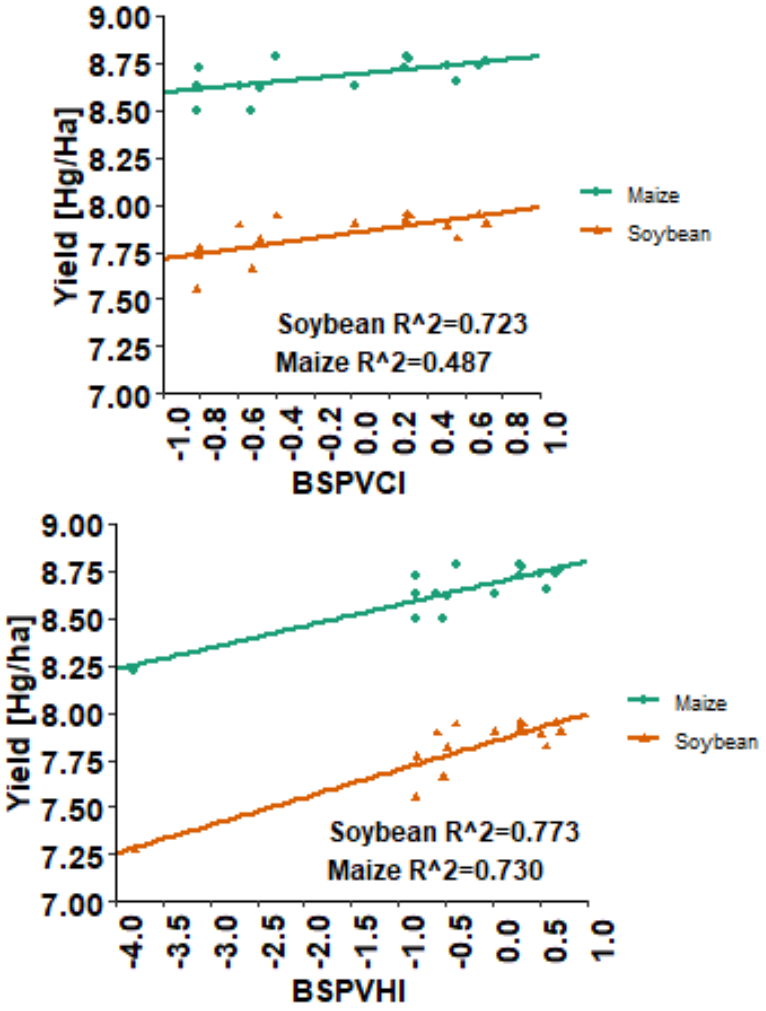


Fig. 5.6 Regression between annual soybean and maize yields with BSPVCI (top) and BSPVHI (bottom).

Acknowledgments

Ronnie Araneda gratefully acknowledges financial support from the Spanish Regional Government of Galicia (Xunta de Galicia) and the European Union through the predoctoral grant reference ED481A-2018/162, and the IACOBUS Program. María Bermúdez was supported by the European Union H2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 754446 and the Research and Transfer Fund of the University of Granada - Athenea3i.

Chapter 6

Practical application: Manual for obtaining and applying hydrometeorological variables from global databases in drought management

Ronnie J. Araneda-Cabrera, Jerónimo Puertas, Manuel Álvarez and Víctor Penas.

Spanish version:

Publication: Universidade da Coruña, Servizo de Publicacións.

Date: February 2022

DOI: https://doi.org/10.17979/spudc.9788497498333

ISBN: 978-84-9749-833-3

Portuguese version:

Publication: Universidade da Coruña, Servizo de Publicacións.

Date: February 2022

DOI: https://doi.org/10.17979/spudc.9788497498340

ISBN: 978-84-9749-834-0

Author contributions:

Ronnie J. Araneda Cabrera: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – original draft, Visualization. Jerónimo Puertas: Conceptualization, Validation, Formal analysis, Writing – review & editing, Supervision. Manuel Álvarez: Conceptualization, Writing – review & editing, Supervision. Víctor Penas: Conceptualization, Resources, Writing – review & editing, Supervision.

Abstract

The main objective of this manual is to provide a basic introductory tool for drought management. This extreme hydrological phenomenon is one of the most important in water resources management, as it causes immense economic and ecological losses worldwide every year, affecting food security and people's quality of life. Due to climate change, drought events will become more frequent and severe in the future. Therefore, understanding, characterising, and monitoring this phenomenon is essential to create contingency plans to reduce risk and become more resilient to droughts. To achieve this purpose, data, tools and procedures are required to understand and model the phenomenon. However, different socio-economic realities, which vary widely in different regions of the world, mean that access to certain data of sufficient quality is limited in some regions and countries. This manual is a tool to overcome this problem through tools developed to obtain the necessary data quickly and efficiently for drought management and monitoring in any region. These data are products of global databases that are free and open access to everyone, so that the application of these data can be used in administrative (countries, cities, etc.) or natural (river basins, nature reserves, etc.) contexts. In addition, the manual presents different uses and applications of these data as concrete drought management tools. As part of the procedure, a brief introduction to the different aspects, concepts, and requirements necessary for the use of the manual is given.

Keywords

Drought, Mozambique, Handbook, Hydrology, Meteorology, Drought management, Water resources, Drought management.

Section I: Introduction

1. General objective

The main objective of this manual is to provide a basic introductory tool for drought management. This extreme hydrological phenomenon is one of the most important in water resources management, as it causes immense economic and ecological losses worldwide every year, affecting food security and people's quality of life.

Due to climate change, drought events will become more frequent and severe in the future. Therefore, understanding, characterising, and monitoring this phenomenon is essential to create contingency plans to reduce risk and become more resilient to droughts. To achieve this purpose, data, tools, and procedures are required to understand and model the phenomenon. However, different socio-economic realities, which vary widely in different regions of the world, mean that access to certain data of sufficient quality is limited in some regions and countries.

This manual is a tool to overcome this problem through tools developed to obtain the necessary data quickly and efficiently for drought management and monitoring in any region. These data are products of global databases that are free and open access to everyone, so that the application of these data can be used in administrative (countries, cities, etc.) or natural (river basins, nature reserves, etc.) contexts. In addition, the manual presents different uses and applications of these data as concrete drought management tools. As part of the procedure, a brief introduction to the different aspects, concepts, and requirements necessary for the use of the manual is given.

2. Scope and limitations of the manual

The conceptual and practical information presented in this manual is intended as an introductory tool for data collection and application in drought monitoring and characterisation. The manual presents information for understanding the phenomenon and its components in a general way, and therefore the practical tools and results are limited to these concepts.

The key concepts synthesised here are generally accepted and recognised by experts and international organisations; however, they are limited to the objectives of the manual and can be expanded and even show differences to other bibliographic sources. Similarly, the tools used, the procedures and interpretations applied for the practical exercises may contain errors and may be improved or applied in other scenarios. All the computational tools used are free for public use and can be replaced by others that serve similar purposes.

3. Document organisation

This document consists of three sections. The first section provides an introduction to the concepts, tools and requirements needed to make use of the manual. First, a compilation of the concepts associated with the study, its classification, characterisation, and management of droughts is presented. Then, a general explanation is given of the technical and computational requirements needed to use the manual. This is followed by an introduction to the various global databases that exist, and the data needed for drought management, with emphasis on those used in the manual. Finally, the specific objectives of this work are explained.

In the second section, the steps necessary to obtain the series of different variables required for drought management are detailed and explained. The third section makes use of the data obtained in Section II and explains the steps to apply them concretely to the characterisation and monitoring of droughts in a specific region. In the annexes, three scripts (programming codes) are provided as a complementary

final product to follow the manual. As case studies we have chosen Mozambique located in the southern cone of Africa and the Licungo river basin in the north-central part of this country.

4. Drought Management

4.1. Water cycle

The hydrological cycle or water cycle is the process of water circulation between the different compartments of the hydrosphere (Fig. 6.1). Water from land and oceans enters the atmosphere by evaporation or sublimation, where it condenses into clouds and precipitates as rain or snow. Water that falls as precipitation runs off into freshwater bodies or infiltrates into the ground. The cycle is completed when surface or groundwater flows back into the ocean. The main processes involved in the water cycle that are important variables in drought analysis are:

- Precipitation (P): The atmosphere loses water through condensation (rain, snow and dew) or reverse sublimation (snow and frost), which, depending on the case, is transferred to the ground, to the sea surface or to the sea ice (ice cover over the oceans).

- Evapotranspiration (ET): On the one hand, water evaporates (E) at the ocean surface, on the ground and also in organisms; in the latter (plants and trees in this case) there is also the phenomenon of transpiration (T), so that ET refers to the combined phenomenon. Potential evapotranspiration (PTE) is the maximum amount of water that can be evapotranspired from a soil completely covered with vegetation, which is growing under optimal conditions, and assuming that there are no limitations in the availability of water. The actual or effective evapotranspiration (ETr) is that which actually occurs under the existing conditions in each case.

- Infiltration: The phenomenon occurs when water reaching the soil penetrates through its pores and becomes groundwater. The proportion of water that infiltrates and that which circulates on the surface (runoff) depends on the permeability of the substrate, the slope and the vegetation cover. The amount of water contained at a given time is the soil moisture and can be measured at different depths.

- Runoff: This term refers to the various means by which liquid water flows downhill over the land surface. When it reaches a body of water (rivers, lakes, reservoirs, etc.) the volume that flows through it for a given time is referred to as the flow rate.

- Underground circulation: It is produced in favour of gravity, like surface runoff. It occurs in aquifers through the pores of permeable rock, involving phenomena such as pressure and capillarity.

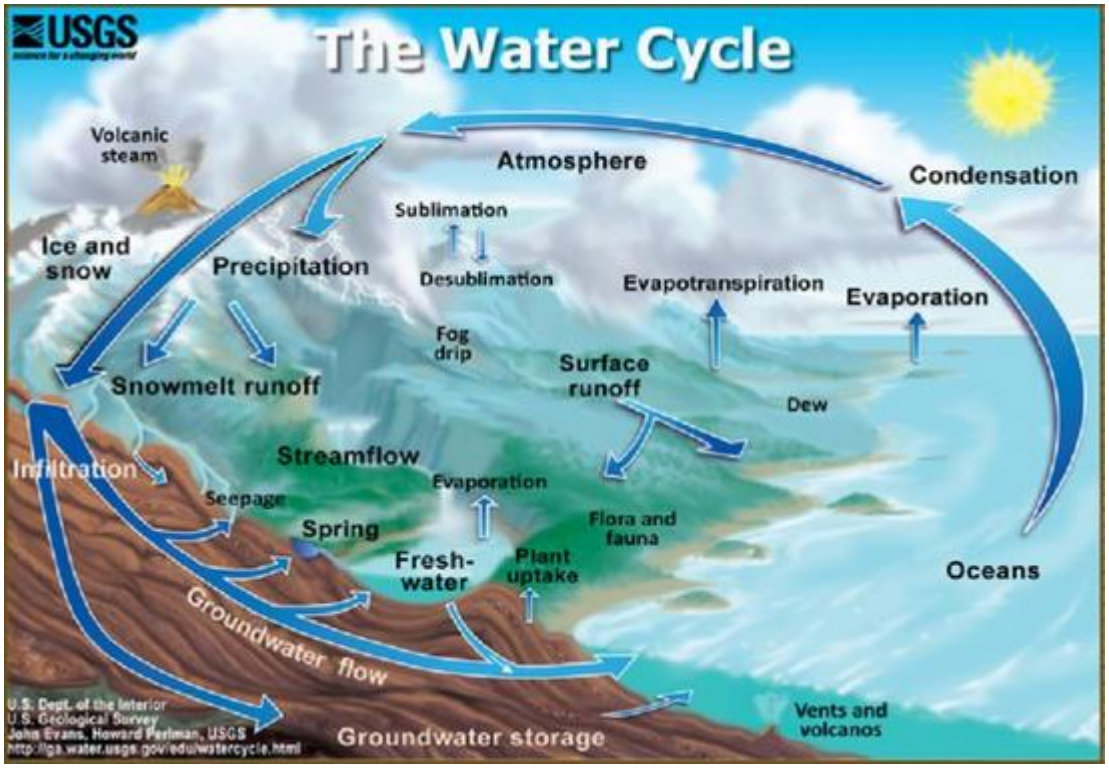


Fig. 6.1 The water cycle. Source: U.S. Geological Survey (USGS).

4.2. Water balance

The water cycle can be formulated mathematically using the water balance, which is derived from the concept of conservation of matter, i.e., it is the balance between all water resources entering and leaving a system (river basin, country, region, etc.) in a given time interval:

$$Status_{t+1} = Status_t + \sum_{i=1}^{n} Inputs_i - \sum_{j=1}^{m} Outflows_j \quad (9)$$

Where the inputs of water are precipitation (P), which can be in the form of rain, hail, snow, or condensation; groundwater input from adjacent groundwater systems; and water transfers from other systems, such as dam discharges or sewage. Outflows can be evapotranspiration; deep seepage feeding aquifers; water diverted to external adjoining systems, human consumption, and industry; and water leaving the system to an external receptor such as a large sea or the sea. The status refers to the volume contained in the system after adding and subtracting inflows and outflows from the volume of water in the previous time step.

In general, for each time step eq. (1) can be formulated as follows:

$$P + I_{in} + U_{in} - I_{out} - A_{out} - ETr - Q_{out} - U_{out} = \Delta V \quad (10)$$

Where P is precipitation, I is infiltration, U the different uses of water (in the same or another system), ETr is real evapotranspiration, Q the outflow of the system and ΔV the variation of water volume in the system. In the long term, the variation tends to zero, however, in each time step there may be times when the system has a deficit or surplus. Each of the processes can be measured and the balance can be completed.

By simplifying the balance, we are able to detect wet and dry months (3); and, to calculate the volume of water available (4):

$$P - ETP = \Delta V \quad (11)$$

$$P - ETr = Q + I + U \quad (12)$$

Where, in eq. 3 the variation ΔV defines a dry month or deficit (ΔV < 0) or a wet month with surplus (ΔV > 0) in the system in a given time period. Meanwhile, in eq. 4 the second part of the equality is called useful precipitation and is the water available to recharge the soil, rivers and other existing water bodies.

The water balance also allows the calculation of the water deficit (D):

$$ETP - ETr = D \quad (13)$$

Where, a deficit (D > 0) is the amount of water lacking to cover the potential water needs for evaporation and transpiration.

4.3. Definition and classification of drought

Drought is an extreme hydrological phenomenon which geographical and temporal limits are difficult to determine, being a particularly dangerous natural disaster when there is a lack of good management of water resources. Drought is understood as a prolonged temporary precipitation anomaly, characterised by a period with precipitation values below normal. It may or may not affect yield crops, soil degradation and/or insufficient water supplies, depending on the level of demand and the characteristics of the water resource exploitation systems.

Many classifications of drought have been developed in recent years. However, the most widely accepted and widespread is that which defines meteorological, agricultural, hydrological, and socio-economic drought. The first three measure drought as a physical phenomenon, while the last deals with drought in terms of supply and demand, based on tracking the effects of water deficit as it churns through socio-economic systems. Each of them is related to a part of the hydrological cycle and develops one after the other initiated by climatic anomalies (precipitation and temperature) that are part of natural climate variability (Fig. 6.2).

Meteorological drought occurs when negative anomalies or absence of precipitation dominate an area. These anomalies may be accompanied by an increase in temperatures, resulting in increased evapotranspiration. Agricultural drought occurs when crops are affected by this lack of precipitation and low soil moisture. Hydrological drought occurs the low available volume of water becomes evident, especially in rivers, streams, reservoirs, and groundwater levels, usually after many months of meteorological drought. Socio-economic drought relates supply and demand for various commodities to drought (water for drinking, irrigation, industries, hydropower, etc.). Meteorological drought can begin and end quickly, while hydrological drought takes much longer to develop and recover from.

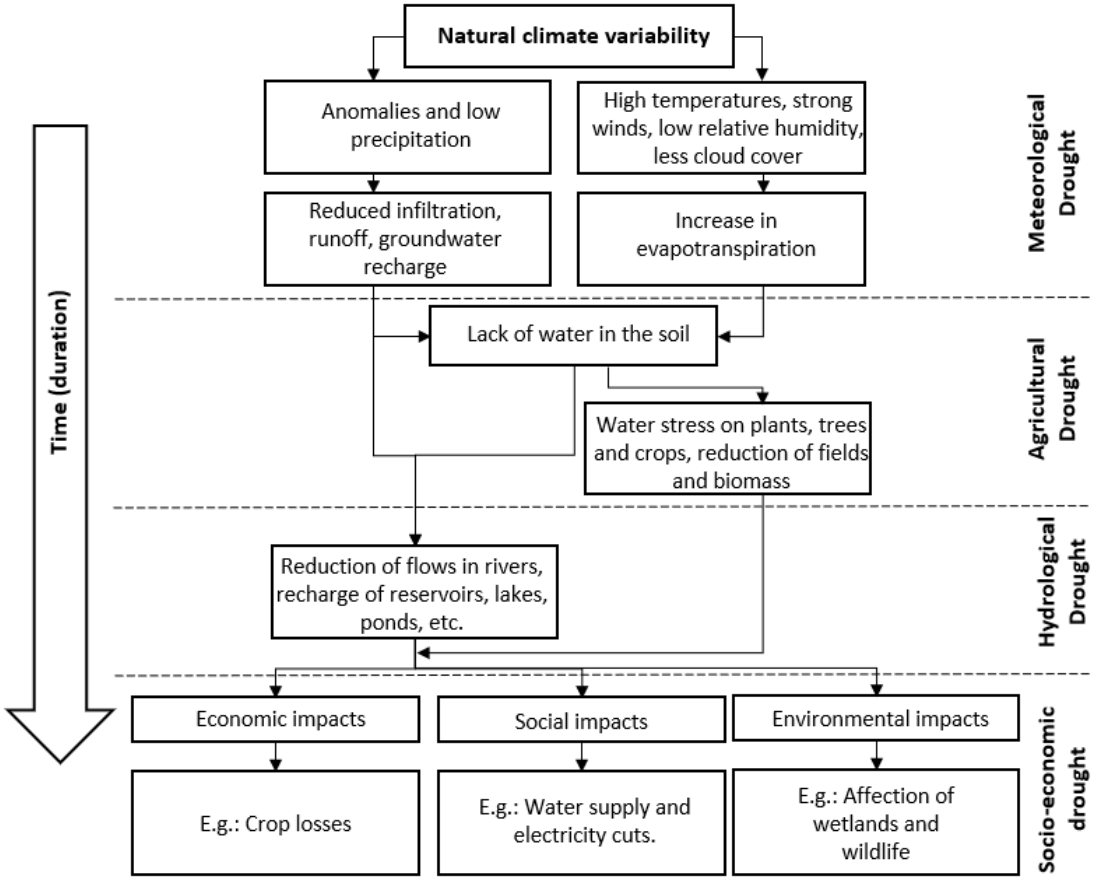


Fig. 6.2 Sequence of occurrence of drought types and their impacts. Source: National Drought Mitigation Center de la Universidad de Nebraska –Lincoln.

Drought must be distinguished from other related phenomena such as scarcity and aridity. Scarcity represents a permanent situation of deficit (calculated with the water balance) in relation to water demand in a regional resource system, characterised either by an arid climate or by a rapid growth in demand. Aridity is a natural structural situation of a region and therefore permanent. Even in these circumstances there should not be a deficit if the exploitation systems are adequately designed and exploited and the demands are kept within reasonable limits, in accordance with the climatic characteristics of the region. This requires planned medium- and long-term actions.

4.4. Drought indices

Indices are numerical representations that display information associated with the drought seriousness, determined by climatic or hydrometeorological data. They are used to quantify the intensity, location, timing, and duration of drought events. Intensity refers to the deviation from normality of an index based on a previously established threshold. In this way we can identify when a drought has started, when it has ended, and the geographical area affected. Location refers to the geographical area in which the drought conditions are recorded. The approximate dates of onset and ending determine the duration of the drought event. Indices can also serve another essential function in that they can provide a historical reference for planners or decision makers. This reference provides users with information on the probability of occurrence or recurrence of droughts of different intensity levels. However, it should be noted that climate change has begun to alter historical trends. It is recommended that the phenomenon be represented at various time scales, as this recognises short-term wet periods within long-term droughts or short-term dry periods within long-term wet periods.

Since the classification of drought responds to analysing anomalies in each stage of the hydrological cycle, there are indices to characterise each type of drought based on the variables in each stage. For example, precipitation for meteorological drought or flows for hydrological drought. In recent years, mixed indices have been developed that attempt to quantify droughts in all or parts of the water cycle. However, using several indices based on several variables is best suited to characterise the state of drought in the whole system.

Due to the complexity of the phenomenon, dozens of indices have been proposed in recent decades. These indices can be calculated with one or several hydrometeorological input variables at different time scales (daily, weekly, or monthly), the most common and accepted being the monthly scale. Only 5 indices, which have been extensively evaluated and validated worldwide, are explained in detail in this manual. For more information on some drought indices, additional literature can be consulted such as the Drought Indicators Handbook (Svodova et al., 2016) where more than 100 drought indices are shown and explained.

4.4.1. Standardised Precipitation Index (SPI)

The SPI was introduced by McKee et al., (1993) and is defined as a value representing the number of standard deviations of precipitation over a defined accumulation period from the mean (standardisation), once the original distribution of precipitation has been transformed to a normal distribution. As an example, a 3-month SPI value of -2.3 would indicate that the amount of precipitation that has been recorded over a period has been 2.3 times the standard deviation below the mean value. The step-by-step for its full calculation can be reviewed in several papers, e.g., Kumar et al., (2009). It is currently recommended by the World Meteorological Organization for drought status monitoring. It is very flexible, easy to interpret, applicable to and comparable across different climate zones and depends on a single variable, precipitation.

Through the use of SPI, it is possible to quantify and compare the intensities of precipitation deficits between areas with very different climates and it has the property that it can be integrated at any time scale, which means that it can be used as an indicator of different types of droughts. Short time scales (2-6 months) capture droughts mostly in the agricultural and forestry sectors and are useful for meteorological and agricultural droughts, while long time scales (12-24 months) are associated with hydrological droughts as they represent drought effects on reservoirs, rivers, and aquifers.

The SPI calculation requires at least 30 years of historical monthly precipitation series. Precipitation series in daily, weekly, or other aggregations are admissible in the computation of the SPI; however, it is recommended to follow the scales in which droughts normally develop, commonly monthly, seasonal or interannual. Once the SPI has been computed over a time series of precipitation, the monthly SPI values can be evaluated according to Table 6.1.

4.4.2. Standardised Precipitation and Evapotranspiration Index (SPEI)

SPEI is a drought index with similar properties to SPI, with the difference that it takes into account temperature through evapotranspiration in its calculation, in addition to precipitation. This difference makes it suitable for analysing trends and changes in drought characteristics due to climate change. Its computation requires monthly historical series of water variation in the system ($\Delta V = P - ETP$), where ETP includes temperature information. The standardisation process followed by SPEI is the same as SPI, with the difference that the ΔV series is previously transformed to a log-normal distribution. Like SPI, SPEI can be calculated for various time scales allowing to assess droughts in various parts of the water cycle and to relate it to different types of droughts. Its evaluation follows Table 6.1.

4.4.3. Vegetation Condition Index (VCI)

The VCI is an index used for the detection and monitoring of agricultural drought. It makes it possible to evaluate the duration, the area covered, the intensity and the impacts of a drought on vegetation. For a given point or region, and for each time step (monthly in our case), the VCI is derived from the Normalized Difference Vegetation Index (NDVI), obtained from the historical series of satellite images obtained by the AVHRR (Advance Very High-Resolution Radiometer) sensor in operation since 1984:

$$VCI = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \quad (14)$$

Where, max and min is the maximum and minimum NDVI value available in the analysed time series and NDVI is the value at each time step. The NDVI is itself a vegetation index used to estimate the quantity, quality and development of vegetation based on the measurement from satellite sensors of the intensity of radiation from certain bands of the electromagnetic spectrum that vegetation emits or reflects. The assessment of the degree of drought intensity according to the VCI is shown in Table 6.1.

4.4.4. Temperature Condition Index (TCI)

Similar to the VCI, the TCI is an index derived from measurements from the AVHRR satellite sensor, in this case of surface temperature. Its use is related to meteorological and agricultural drought monitoring. Its estimation uses the estimated brightness of the thermal infrared band (10.3-11.3 μm) detected by sensors from space. According to the measured brightness, the Brightness Temperature (BT) is estimated, which by itself is an indicator directly correlated with the Earth's surface temperature. Its formulation is:

$$TCI = \frac{BT_{\max} - BT}{BT_{\max} - BT_{\min}} \quad (15)$$

Where, within a time period, BT is the value at each time step (monthly in this case) and min and max refer to the maximum and minimum values of BT within that period. Their evaluation follows Table 6.1.

4.4.5. Vegetation Health Index (VHI)

The VHI is calculated by the weighted combination of two anomalies already described: the VCI and the TCI, both derived from satellite observations. The basic idea of the indicator is as follows: the lower the observed VCI (relatively poor green vegetation) and the lower the observed TCI (relatively warm temperature), the lower the VHI. Low VHI values are indicators of drought, especially when they persist for long periods of time.

The equation used is:

$$VHI = w \cdot VCI + (1 - w) \cdot TCI \quad (16)$$

Where w is the pondered value, the range of which is between 0.0 and 1.0, with 0.5 being the usual value. The assessment of drought intensity according to the VHI is given in Table 6.1.

Table 6.1 Drought categories according to index values.

Categories	SPI/SPEI	VCI/TCI/VHI
Wet	> 2.00	90 - 100
	1.50 - 2.00	75 - 90
	1.00 - 1.50	60 - 75
Normal	0.00 - 1.00	50 - 60
	0.00 - -1.00	40 - 50
Moderate drought	-1.00 - -1.50	25 - 40
Severe drought	- 1.50 - -2.00	10 - 25
Extreme drought	< -2.00	0 - 10

4.5. Drought characteristics

Drought characterisation is essential in drought management. Defining and quantifying the characteristics of each drought event helps to plan continuous monitoring campaigns and apply predictive models. It consists of defining the seriousness of a given event so that a policy maker or technician can make decisions and implement appropriate mitigation plans to avoid severe water shortages, so this methodological component is essential for stakeholders.

Among the different methods proposed to characterise droughts, the "run-theory" method described in Yevjevich, (1969) has become widespread, due to the objectivity in the definition of drought. The method allows analytical derivation of probability distributions of drought characteristics, showing better performance than other approaches. The implementation method can be applied to a time series of interest, which in our case will be the drought indices, assuming a threshold at which an event is framed. The thresholds are defined according to the drought category to be characterised.

A drought event is defined by its duration, intensity, severity and geographic extent. Duration is the length of time in weeks or months that the drought lasts, understood as the time that the index was consecutively below a threshold, defining a start and end of the drought event as it passes through that threshold in time. Maximum or average intensity is the maximum absolute value and the average of the values of an index over the duration of the event. Severity is defined as the sum of the intensities or values taken by the index over the duration of the drought. Fig. 6.3 shows these three characteristics for three drought events. Geographic extent is the area or percentage of area within a region or area under study that is affected by drought at one or more intensity categories during each time step.

In order to quantify these characteristics, it is necessary to calculate drought indices over several points of the territory under study. This spatial distribution should be as dense as possible, with the World Meteorological Organization, (2008) recommending at least one station or geographical point where the index is evaluated every 250km^2.

Suitable drought indices are selected according to the type of drought of interest. Indices can be considered general for the whole system or specific to the part of the hydrological cycle for which they are designed. It is understood that making this distinction is not always easy; therefore, the use of several indices is recommended to obtain a complete characterisation of droughts in a region. Here we will use the indices previously discussed.

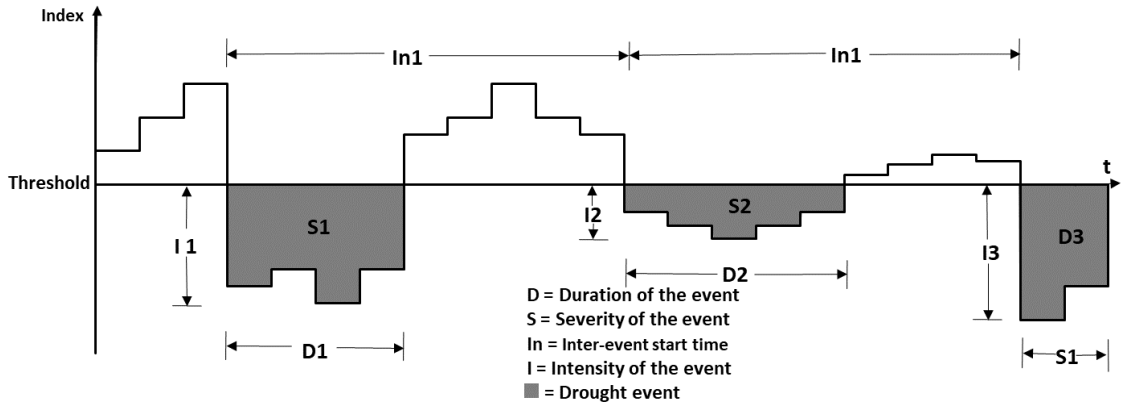


Fig. 6.3 Characteristics of droughts according to the "run-theory".

5. Requirements

For the use and application of this manual, in addition to the theoretical background associated with drought management, it is necessary to be clear about the computational tools that are used. The two groups of tools required are Geographic Information Systems (GIS) and Numerical Computing Systems (NCS). While there are many similar programs to those discussed below, the concepts related to the architecture of these programs are the same as those followed by other tools designed to perform similar functions.

5.1. Geographical Information Systems (GIS)

GIS are computer programmes (software) that provide tools for the collection, processing, management, analysis and representation of data with a cartographic component (geographic information) in digital format. They make it possible to represent the physical space in which we live on maps and tables through thematic layers that can be overlapped. Any record of information with a geographical reference, such as a street address, the name of a city, a land parcel identifier, GPS coordinates, etc., can be located and made available on a map (Fig. 6.4).

GIS can be classified into two main groups exemplified in Fig. 6.4:

Vector GIS: Use vectors for the description of geographical objects. They are commonly presented as shapefiles (common extensions: SHP, SHX, DBF, PRJ, etc.). Geographic features are represented by three basic structures: points, lines, and polygons. Lines are represented by segments that join points and these, in turn, are closed to form polygons. There are therefore three different shapefile types. Each of them has specific qualities so that it is not possible to perform certain processes with them. Thus, it is impossible to calculate areas for points since area is a property of polygons. Or it is not possible to calculate lengths for polygons, but we can fragment polygons into multiple parts.

Raster GIS: Raster Information Systems base their functionality on graphic bitmap formats (common extensions: ECW, JPG, GRID, TIF, GIF, SID, etc.). They (usually) store information by dividing space into a regular grid or matrix of small cells - referred to as pixels - and attribute a numerical value to each cell as a representation of a physical spatial quality or property (altitude, temperature, distance, slope, etc.). If the position of the coordinates of the centre of one of the cells is known, all pixels can be said to be georeferenced.

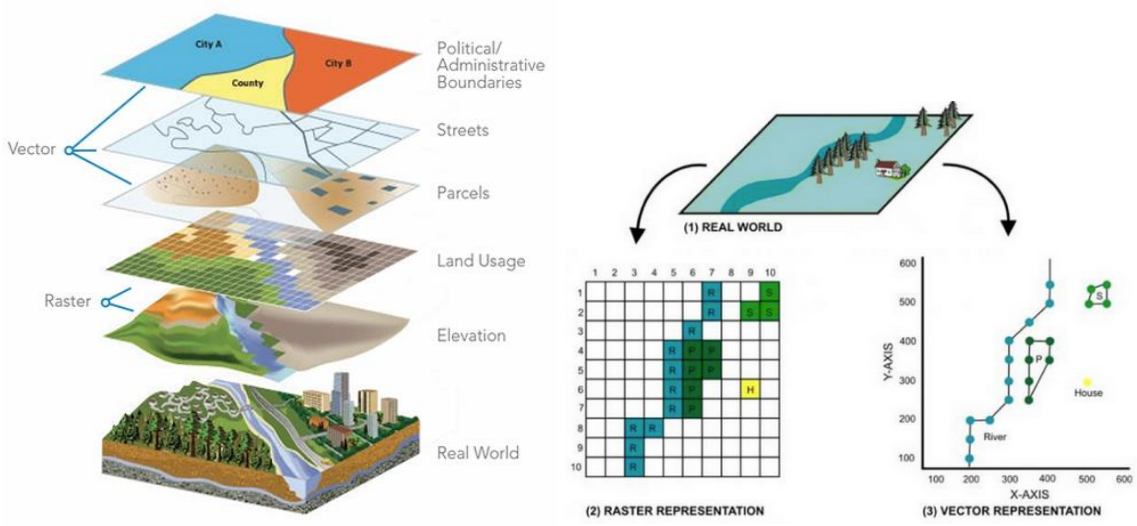


Fig. 6.4 Thematic layers available in a GIS and comparison between raster and vector data. Source: https://geopaisa.blog/2017/03/08/que-es-un-sig/.

Two of the most common and developed GIS software are ArcGIS and QGIS. Both are widely used and cover most users' needs. The main difference is that QGIS is a free open-source software, while ArcGIS is a paid software commercialised by ESRI. However, the geographic information that can be represented, analysed, and displayed in the form of maps in such software can also be manipulated through NCSs that offer these options in their functionalities.

The global data discussed in this manual are in raster formats, while vector files defining an area (country or river basin) are used for extraction in study areas of interest. This information can be downloaded from free databases.

5.2. Numerical Computing Systems (NCS)

NCSs are programmes (software) or tools designed to solve complex numerical problems and calculations through the use of numerical methods. The general procedure consists of introducing data or input variables to an existing function in the software, which will perform pre-programmed calculations and produce a result. NCSs have their own environment and programming language that allow users to use both the tools and functions already existing in the program (simple operations such as addition or subtraction to complex ones such as solving integrals or differential equations) and to create new functions (specific and personalised programming).

There are many NCSs applicable to water engineering and therefore to drought management. By way of example, among the most widely used are MATLAB from the company MathWorks, Phyton, R, C++, among others. Each of them is oriented and specialised in solving problems with specific numerical bases. This manual uses R in its most user-friendly environment RStudio, which is specialised in statistical computing and graphics creation.

5.2.1. R and RStudio

R is a free, open source and free programming environment conceptualised for statistical data processing and analysis (Durán, 2000). It consists of a set of very flexible tools that can be easily extended by means of freely available packages and libraries or by programming its own functions.

R packages are collections of functions and toolkits developed by the user community. They increase the potential of R by improving the existing base functionalities and adding new ones. The thousands of packages that exist are found in repositories (CRAN) from which their user manuals, architecture and

conceptualisation can be downloaded. CRAN (Comprehensive R Archive Network) is the official website of the software, which offers different resources for the use of the programme: user manuals, online courses, general information, package downloads, information on installed packages, etc.

To install R, download the installer according to the operating system to be used from its official website (https://cran.rstudio.com/). In addition, it is advisable to install RStudio (RStudio Team, 2016), a more user-friendly environment that allows computational processes based on the R language. The free version can be downloaded from the programme's website (https://www.rstudio.com/products/rstudio/download/).

RStudio is an integrated development environment (IDE) for the R programming language, dedicated to statistical and graphical computing. It includes a console, syntax editor supporting code execution, as well as tools for plotting, debugging and workspace management. It consists of 4 windows (Fig. 6.5): Top left window: this is the syntax editor; this is the place where we edit the syntax for later execution. Nothing will happen when you write there unless you press a button to execute the commands. This is the window where new scripts are written and where the scripts given in this manual will be opened. Upper right window: this is the working environment of the programme, where the data set and the objects (results, variables, graphs, etc.) that are stored when running different analyses are shown. Lower right window: it has several sub-tabs: (i) the “Files” tab allows viewing the history of files worked with the programme; (ii) the “Plots” tab allows viewing the graphics generated; (iii) the "Packages" tab allows viewing the packages downloaded and saved on the hard disk, as well as managing their installation or update; (iv) the “Help” tab permits accessing the CRAN; (v) the "Viewer" tab shows the results when building reports. Lower left window: this is the console, corresponding to what would be the R software in its basic version. There, the software executes the operations performed from the syntax editor.

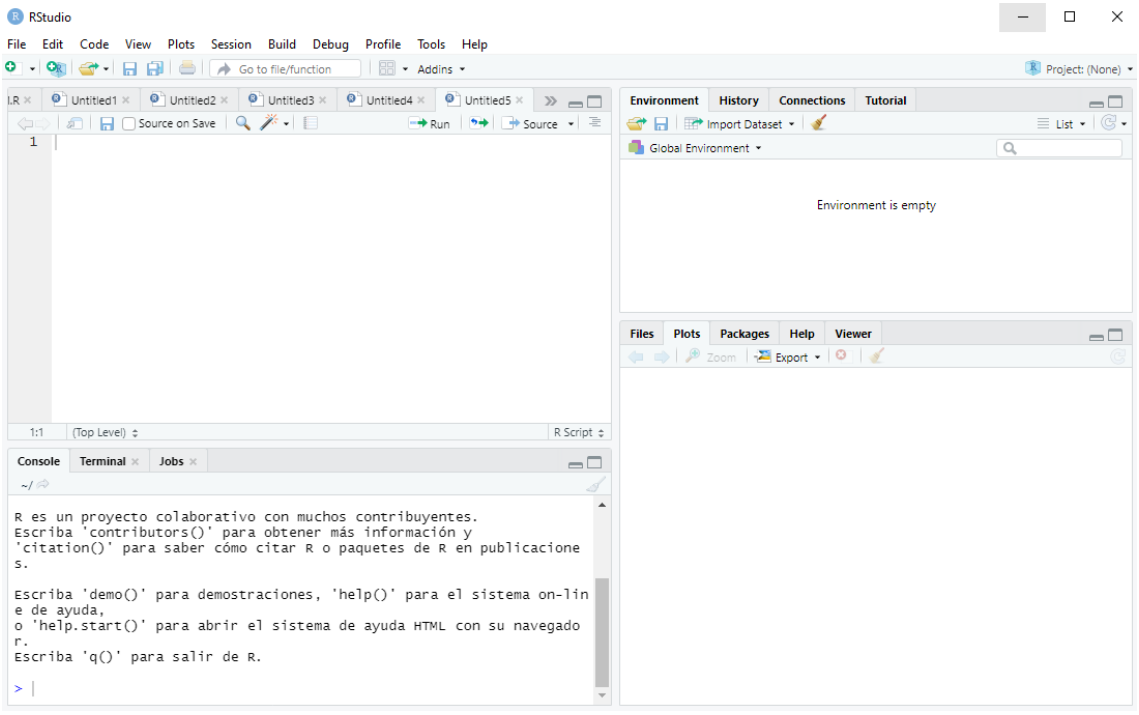


Fig. 6.5 RStudio environment.

Of the thousands of existing packages for R, there are several developed specifically for water resources and drought management (Slater et al., 2019). Some of the packages required in the scripts attached in the annexes are the "SPEI" package developed by Begueria and Vicente-Serrano, (2017), which contains functions that allow direct calculation of SPI and SPEI drought indices; the "ggplot2" package allows

professional graphics, including the graphing of shapefiles and rasters; the "tmap" package allows the R user to manipulate geographic information, typically manipulated through GIS with R; among others.

The scripts presented in this manual (ANNEXES: Scripts in R and RStudio for drought management) are analysed and developed step by step in Sections II and III. The specific information required for each of them is obtained from free databases. The scripts are written in a logical and orderly manner so that they can be used, replicated, manipulated, improved and complemented by the users of this document.

As general information R allows working with all types of data: matrices, vectors, vector and raster geographic information, etc. Each file can be read and assigned a name (using the <- symbol). Different loops and existing algorithms (e.g., sine, cosine functions) can be created. Graphs can be generated and viewed in the viewer and exported to any storage folder on the computer in different formats (.pdf, .jpg, .png, etc.). The following manual is recommended to get started with R programming: R for Beginners (Ahumada, 2003). However, there are hundreds of online manuals and blogs that can be used as a source of guidance.

6. Global databases

As stated above, optimal drought management requires time series of hydrometeorological data for each component of the water cycle. These data must be of long duration (at least 30 years) and have been measured with sufficient distribution over the natural territory (one meteorological station every 250km^2). However, measuring these variables under these conditions is a complex task. The main reason is often their high cost of installation and maintenance, so that in many areas of the world the availability of data is extremely limited.

To obtain data such as precipitation, soil moisture, flow rates, etc., several institutions have developed global databases in which they offer such time series of hydrological and climatic interest openly and at no charge. These products are based on measured point data that have been interpolated, developed, and obtained using statistical techniques, global climate and hydrological models, and remote sensing products from satellite radars.

In addition to specific databases for water and drought management, alternatives are presented to obtain the geographic information necessary for the representation of hydrological variables, maps, and results. These data are vector files corresponding to the administrative boundaries of countries and the boundaries of the main river basins worldwide. There are hundreds of sources of geographic data, products offered by governmental, environmental, and other organisations, which offer this type of information free and openly. You can also find raster files with topographic and climatic information, among others.

6.1. Complementary databases

Vector files (shapefiles) of georeferenced country or river basin boundaries are downloaded from the public geographic databases GADM and FAO (Food and Agricultural Organization). GADM is a high-resolution database of the administrative areas of the entire world, which provides for all countries, all administrative levels and at any time-period. GADM created the spatial data for many countries from spatial databases provided by national governments, NGOs, and/or from maps and lists of names available on the Internet. The database is available in several export formats, including the shapefiles used in most GIS applications. These files can also be used with the R data analysis language. The shapefiles delimit areas for data download and analysis, as well as being necessary for the creation of descriptive data graphics including geographic maps. The web address is https://gadm.org/. It has a very intuitive environment where within the "Data" tab it is possible to choose the country of interest and

download the vector files distributed in levels from 0 to 5 where 0 is the national limit and 5 is the smallest division that each country has (not all of them have 5 levels)

FAO's "GeoNetwork" section offers a wide variety of interactive maps, GIS datasets, satellite imagery and applications related to its work. It provides a wide range of basic global spatial information developed to support decision making, promote multidisciplinary approaches to sustainable development, and improve understanding of the benefits of geographic information. GeoNetwork is an open data source that allows easy sharing of geographically referenced thematic information between different organisations. The web address is http://www.fao.org/geonetwork/srv/en/main.home; where from the "Hydrology and Water Resources" tab it is possible to obtain geo-referenced vector files corresponding to the boundaries of river basins around the world at different levels (basins and sub-basins), riverbeds and others.

6.2. Databases of hydrometeorological variables

6.2.1. Types of databases

This manual explains how to download, extract and apply time series of the following variables: precipitation in mm/month, mean temperature in ˚C per month, ETP in mm/month and ETr in mm/month. In addition to the monthly drought indices VCI, TCI and VHI. Each of these variables is measured or estimated differently, therefore, the databases that provide these products come from different sources. In general, existing data are obtained from 4 techniques: from in situ measurements, data obtained from satellite measurements, data modelled numerically using algorithms and data generated from reanalysis techniques. Some databases integrate these techniques to improve the accuracy of their products.

The traditional way of measuring precipitation and temperature (the most common and basic variables) is using rain gauges and thermometers. Each of these stations is located at a specific point on the territory. In recent years these variables have also been estimated through remote sensing from infrared radars from satellites orbiting the Earth. This last methodology is the mechanism by which the variables from which the VCI, TCI and VHI agricultural drought indices are derived are estimated.

ETr and ETP, soil moisture and flow rates (runoff) are other important variables in the water cycle and thus for drought studies. Measuring these requires lysimeters, drilling in the ground and flow measurement stations at various locations in the watercourse. However, these methods are often costly to install and maintain, making these variables less common to measure, especially in countries and regions with socio-economic limitations. The most common solution to obtain these data is through global (large scale) or local (point scale) hydrological/climate models. The models use measured input variables (precipitation and temperature) to estimate missing variables. Another form of estimation is through mathematical and statistical formulations based on empirical relationships between measured and estimated variables. A clear example is the calculation of ETP using formulations such as Thornthwaite's Thornthwaite, 1948) which estimates ETP based on temperature and precipitation.

The great variety of methodologies for estimating hydrological variables leads to the existence of different databases, generally belonging to large scientific institutions (NASA, Universities, governments, etc.). The products measured in the traditional way (through insitu measurement stations) are usually intermittent (with gaps) and contain errors resulting from the measurement process. In addition, the global coverage is very variable in different parts of the world. To overcome this limitation, institutions have automated the analysis of various point data and calculated a distribution of these values in space with the use of interpolation models, downscaling methods, and reanalysis techniques. The application of these approaches has been supported by data obtained from numerical models and satellite measurements.

Depending on the database, various download formats are available. Although generally, when dealing with time-varying parameters, spatial and temporal information is unified in a single file. This union is performed in layers, as a 3-dimensional matrix (or multi-matrix). For example, a single file may contain a grid of precipitation data for the month of January, another grid of data for the month of February, another for the month of March, etc., with 12 grids (corresponding to the 12 months of a year) being stored one after the other. The grids of the mesh are associated with a value (precipitation, temperature, etc.) comparable to that of a station located at the centroid of each grid, so that several cells can be aggregated within an area or region to form a single average time series.

Due to the large amount of information available in each database, the files containing the data are typically fragmented into several files corresponding to each year, month, or part of the territory. Although there are several types of files, the most common ones we will encounter and the ones we will mostly use in this manual are the "netCDF" (network common data form). This is a file format intended to store multidimensional scientific data (in this case variables of hydrological interest). Each of these variables can be displayed in one dimension (e.g., time). These files (common extension: .NC) are comparable to raster files and are managed with the R package "ncdf4".

For other extensions, however, it is necessary to extract a particular layer and convert it to the desired format (by trimming the information corresponding to the study area), requiring the use of specific algorithms. The processing of the products requires consulting the specific documentation of each database in order to know how to explore the different layers of information. Database managers tend increasingly to facilitate this task and even publish file reading codes.

In this manual we will focus on historical data, however, there are several databases that offer hydrological variables projected up to the year 2100, which can be obtained in a similar way. These projections are supported by global climate models that, based on historical measurements, project estimates according to different human actions in the context of climate change. As an example, the Climate Change Knowledge Portal (https://climateknowledgeportal.worldbank.org/download-data) offers a wide set of variables of hydrological interest that can be downloaded and used in a similar way to what will be explained in chapters 2 and 3 for future periods.

In the following, the databases from which the hydrological variables will be obtained are presented. It is explained where the information comes from, which variables and at which temporal and spatial scales they are offered.

6.2.2. CHIRPS

CHIRPS es el acrónimo de Estimación de las precipitaciones a partir de observaciones de pluviómetros y satélites (Rainfall Estimates from Rain Gauge and Satellite Observations). CHIRPS, de la Universidad de California en Santa Barbara (Funk et al., 2015) ofrece series de precipitación diaria y mensual a una escala de 0.05° (≈ 5.6km en el ecuador) desde 1981. Los productos parten de mediciones de observaciones satelitales de la NOAA (National Oceanic and Atmospheric Administration) y la NASA (National Aeronautics and Space Administration). Posteriormente son calibrados y validados a partir de un gran número de estaciones pluviométricas ubicadas en toda la superficie terrestre.

CHIRPS is the acronym for Rainfall Estimates from Rain Gauge and Satellite Observations. CHIRPS from the University of California, Santa Barbara (Funk et al., 2015) provides daily and monthly precipitation series at a scale of 0.05° (≈ 5.6km at the equator) since 1981. The products are based on measurements from NOAA (National Oceanic and Atmospheric Administration) and NASA (National Aeronautics and Space Administration) satellite observations. They are then calibrated and validated from a large number of rainfall stations located all over the earth's surface.

Version V2.0 (https://www.chc.ucsb.edu/data/chirps) is used here. Its spatial accuracy motivates the use of this database in hydrological studies and drought management as it is particularly useful in small area or catchment studies and can be aggregated to any surface area (e.g., K. Wang, Li, y Wei 2019).

Although the scope is stated to be global, data are only available for the area within 50° north and south of the equatorial plane. Updates are provided on a weekly basis, allowing for near real-time monitoring.

6.2.3. TerraClimate

TerraClimate (http://www.climatologylab.org/terraclimate.html) from the Climatology Lab at the University of Idaho has provided several monthly-scale products since 1958 (Abatzoglou et al., 2018). Using interpolation, downscaling and reanalysis techniques they have created grids at a spatial resolution of 1/24° (≈ 4.6 km at the equator) of precipitation, temperature (maximum and minimum), vapour pressure, wind speed and surface shortwave radiation data over the globe. With this information using the Thornthwaite-Mather global water balance model (WBM) (Willmott et al., 1985), ETP (following the Penman-Montieth formulation), flow or runoff, ETr, water deficit (D), soil moisture and snow water equivalent in each cell have been derived.

The updating of variables is carried out every year, so its use as a source of data for short-term monitoring is not possible. However, due to its high spatial resolution it can be used in the analysis of droughts or water balances in regions of all sizes. Here precipitation, mean temperature (derived from maximum and minimum temperature) and water deficit will be downloaded and processed.

6.2.4. NOAA STAR

El Center for Satellite Applications and Research (STAR) y la National Oceanic and Atmospheric Administration (NOAA) de la NASA proporcionan distintos productos de sensores satelitales globales relacionados con la vegetación (Kogan, 1995). Las principales variables son el NDVI y BT (Kogan, 1997), del que se puede derivar el VCI, el TCI y VHI que también se ofrecen ya computados. Estos productos se han usado en varios estudios relacionados a contenido de agua en el suelo, rendimientos agrícolas y sequías, entre otros (e.g., Agutu et al. 2017; Bento et al. 2018).

Los datos se pueden descargar libremente a una resolución espacial de malla de 4x4 km y una resolución temporal semanal, desde 1983 hasta el presente desde https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. Sus actualizaciones se realizan cada una o dos semanas, permitiendo tener una monitorización en tiempo casi instantánea. Nosotros descargaremos los índices el VCI, el TCI y VHI semanales y los agregaremos a escala mensual.

The Center for Satellite Applications and Research (STAR) and NASA's National Oceanic and Atmospheric Administration (NOAA) provide several global satellite sensor products related to vegetation (Kogan, 1995). The main variables are NDVI and BT (Kogan, 1997), from which VCI, TCI and VHI can be derived and are also available already computed. These products have been used in several studies related to soil water content, agricultural yields and droughts, among others (e.g., Agutu et al. 2017; Bento et al. 2018).

Data can be freely downloaded at 4x4 km grid spatial resolution and weekly temporal resolution, from 1983 to the present from https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. It is updated every one to two weeks, allowing for near real-time monitoring. We will download the weekly VCI, TCI and VHI indices and aggregate them on a monthly scale.

7. Specific objectives:

After a brief introduction to the concepts associated with drought management and a description of the relevant databases, the following specific objectives will be followed in Sections II and III:

a) Download and plot vector files with the boundaries of Mozambique and the Licungo river basin.

b) Download and plot monthly rainfall data from CHIRPS, obtain them within Mozambique and plot their annual averages. Calculate SPI and plot its characteristics.

c) Download weekly VCI, TCI and VHI data from NOAA STAR, obtain them within Mozambique, aggregate them on a monthly scale and plot their development during an event of interest.

d) Download monthly precipitation, ETr and ETP data from TerraClimate, obtain the data within the Licungo river basin, calculate monthly water balance and deficit and plot their annual averages. Calculate the SPEI and tabulate its characteristics.

e) Plot the drought/non-drought status based on the SPI in Mozambique.

f) Plot the area affected by drought according to SPEI in the Licungo river basin.

g) Plot the spatial characteristics of droughts in Mozambique.

To achieve each of these objectives, 3 scripts are presented in the ANNEXES, which are explained step by step in Sections II and III.

8. Following up on the manual

The working environment is designed for the Windows operating system. To follow the manual, it is necessary to install R and RStudio, from where the steps detailed in Sections II and III can be followed according to the scripts provided in the annexes, which are organised as follows:

a) ANNEX I presents the code necessary to extract the information from the different databases. In a destination directory, the data will be saved, sorted and ready for use.

b) ANNEX II contains the necessary tools to read the downloaded and sorted information obtained in Annex I and use it both to characterise droughts through indices and to perform water balances.

c) ANNEX III presents the scripts needed to read the information generated in Annexes I and II, plot it and export it in different formats.

At the beginning of each script, the packages that must be downloaded (only once per CPU used) and activated (each time a new RStudio session is opened) to use these codes are detailed. R has internal functions that allow you to download and activate packages directly from the working environment. To use the codes it is necessary to open a new "R Script" file in RStudio and copy the codes from the annexes. To execute each line of code it is necessary to position yourself at the beginning of each line and press the "run" option at the top right of the work window. Both in the manual and in the code, the step-by-step instructions are specified.

Within each script there are several sections corresponding to each of the subsections of the Manual (packages are required for all the code). If you wanted to follow a single process of downloading data, calculating indices and graphs for any region you could copy each required part of the 3 scripts presented and put them in order in a new file. This manual is intended to be a guide from which you can expand and modify its contents. As you go through the manual you will find similar steps, which will be explained in detail the first time they appear.

The vector files (.shp) for Mozambique and the Licungo river basin need to be downloaded from GADM and FAO and saved in a folder within Windows Explorer. The folder in the scripts is named "Layer_Mozambique", and each of the shapefiles "mozambique.shp" and "basin_Licungo.shp". However, the names of the folders and the files can be changed according to the user's taste..

Section II: Data collection

9. Study area

This section follows the graphics script in which we simply check that the files containing the boundaries of interest are correctly located geographically. This requires a shapefile of all the countries of the world downloaded from the GADM (https://gadm.org/download_world.html).

1) The necessary packages are installed and loaded. Some functions of these packages are used, however, to learn more about them or about all the functions offered by the packages, the manuals provided by the programmers of these packages can be searched in CRAN.

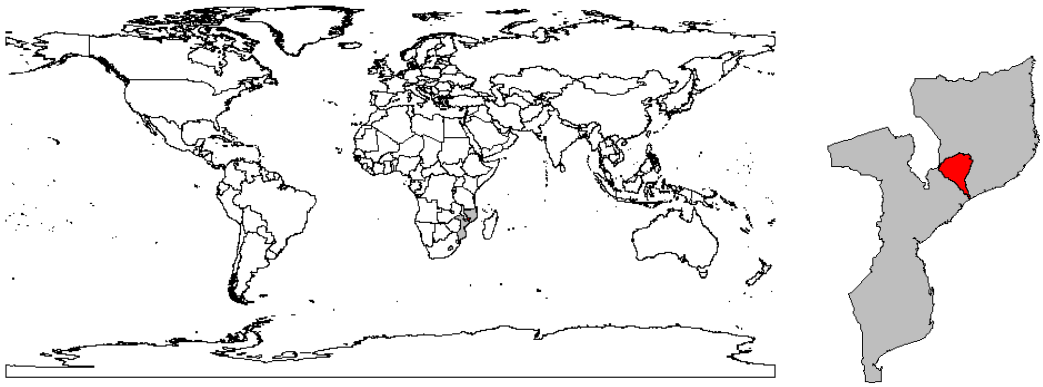
```
#The packages to be used are installed
install.packages("maptools")#Package intended to use shapefiles
install.packages("tmap")#Package for using and graphing geographic information
install.packages("lattice")#Package intended to use spatial graphics
install.packages("RColorBrewer")#Package for colouring graphics
install.packages("ggplot2")#Package for graphs
install.packages("maptools")#Package for using shapefiles
install.packages("ncdf4")#Package intended to use .nc files
install.packages("raster")#Package intended for use with raster files
install.packages("dplyr")#Package for matrix manipulation
install.packages("sf")#Package intended to manipulate geographic information

#The packages to be used are loaded
library(maptools)
library(tmap)
library(lattice)
library(RColorBrewer)
library(ggplot2)
library(maptools)
library(ncdf4)
library(raster)
library(dplyr)
library(sf)
```

2) The shapefiles of Mozambique and the Lacungo river basin are read, the names mz and cl have been assigned respectively. The file with the boundaries of all the countries of the world is also read. Note the address where the file is located and the direction of the slanting line ("/").

```
#The necessary geographical information is read from the folders where they are stored.
mz<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/mozambique.shp")#The Mozambique shapefile is read
cl<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/basin_Licungo.shp")#Read the shapefile of the Licungo Basin
wc<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/World_Countries.shp")#Read shapefile of countries around the world
```

3) The shapefiles are plotted in a simple way to check that the files I have uploaded are correct. The results are displayed in the viewer.

```
#It is plotted and reviewed in the viewer
plot(wc)
plot(mz,add=T,col="grey")
plot(cl,add=T,col="red")#mz and cl are plotted to check that they are the necessary layers
```



4) The files of interest are plotted and imported as .pdf and .png to a folder where they are checked for their appropriate geographical location. Take into account the folder address and the direction of the slanted line. In this case the files are checked to ensure that they correspond to the required boundaries.

```
#The shapefiles are plotted and assigned the name "graph".
graph<-tm_shape(mz) + #Change mz to any other shapefile
  tm_polygons()+
  tm_layout(main.title="Mozambique",main.title.size = 0.8)+ #Change inside "" the title
  tm_grid(lines=T,col="grey",labels.col="black",labels.size=0.8)+#Change within "" the colours
  tm_shape(cl)+tm_polygons(col="red")#A second overlapping shape is painted, delete if not required

#The graphic is exported as png, the destination folder and the size can be modified
png("E:/19_MANUAL_data_global/Graphics/ubicacion_mz_cl.png",height = 270,width = 400)
graph #Assigned name of the graphic to be exported as png
dev.off() #Export closed

#The graphic is exported as pdf, the destination folder and the size can be modified
pdf("E:/19_MANUAL_data_global/Graphics/ubicacion_mz_cl.pdf", width = 3.5, height = 4.5, onefile = FALSE)
graph#Assigned name of the graphic to be exported as pdf
dev.off()#Export closed
```

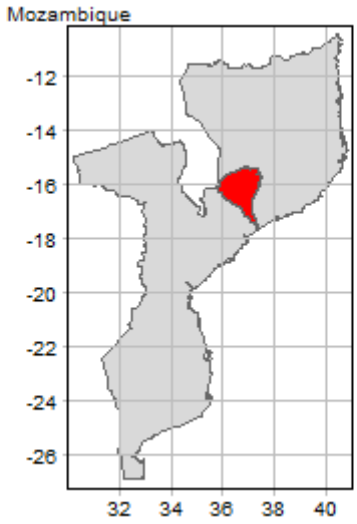


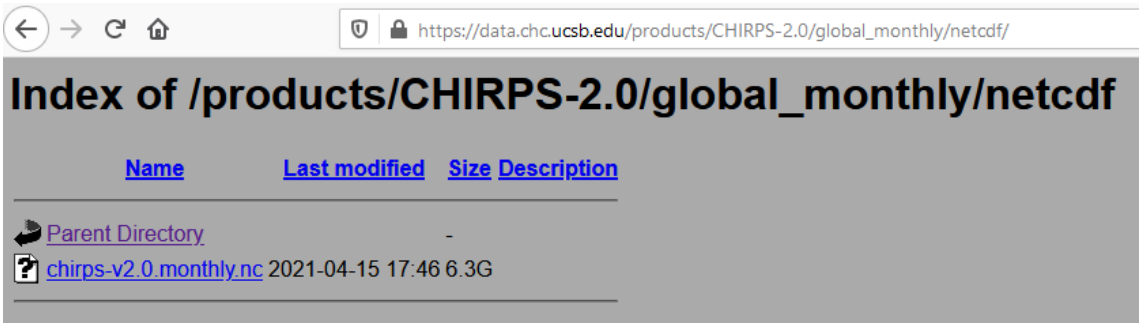
Fig. 6.6 Location and representation of the boundaries of Mozambique and the Licungo river basin.

In case the files require projection changes on the globe or are not correctly located, GIS software should be used to fix the errors.

10. Precipitation: CHIRPS

In this section the files containing the precipitation data are downloaded from the CHIRPS website. The monthly data is then obtained for all pixels within Mozambique. The data are aggregated on an annual scale and plotted in a simple and elegant manner with high quality. The steps to follow are:

1) From the CHIRPS website https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_monthly/netcdf/ download the file: chirps-v2.0.monthly.nc. This file weighs about 6Gb and contains the global monthly precipitation data since 1981. It is stored in a known folder. From this web page, you can download other files such as daily, quarterly, and other precipitation data.



2) The necessary packages are installed and loaded.
3) Mozambique shapefile is read, named mz.
4) The file chirps-v2.0.monthly.nc. is read following the script (variable called "chi" that can be changed to whatever is desired).

```
#The composition of the .nc files is shown
chi <- nc_open("E:/19_MANUAL_data_global/CHIRPS/chirps-v2.0.monthly.nc")#The CHIRPS file is read
print(chi)#The file and its composition are displayed
```

5) The variable "chi" is displayed in the console by means of the print function, which allows to see the characteristics of the file. Here all the necessary information is extracted to understand how the files are stored and thus extract them in the area of interest correctly. This step will be common for all files with .nc extension.

The first section shows the number of variables contained in the file. In this case it contains a single variable called "precip" which stores its data in a three-dimensional array based on longitude, latitude, and time. The unit of each data is mm/month, the monthly time step and for a cell with missing data a value of -9999 is assigned:

```
> print(chi)#Imprimimos el archivo y vemos su composición
File E:/19_MANUAL_datos_globales/CHIRPS/chirps-v2.0.monthly.nc (NC_FORMAT_NETCDF4):

     1 variables (excluding dimension variables):
        float precip[longitude,latitude,time]   (Chunking: [379,106,24])  (Compression: level 5)
            units: mm/month
            standard_name: convective precipitation rate
            long_name: Climate Hazards group InfraRed Precipitation with Stations
            time_step: month
            missing_value: -9999
            _Fillvalue: -9999
            geostatial_lat_min: -50
            geostatial_lat_max: 50
            geostatial_lon_min: -180
            geostatial_lon_max: 180
```

The next section reports the characteristics of the dimensions in which the file stores the information. In this case it divides the world in longitudinal direction into 7000 parts from the east and in

latitudinal direction into 2000 parts from the north. These matrices exist for 468 time steps (months) which are denoted by a number of days since 1 January 1980 (i.e. the number 29 refers to 29 January 1980).

```
3 dimensions:
   longitude  Size:7200
       units: degrees_east
       standard_name: longitude
       long_name: longitude
       axis: X
   latitude  Size:2000
       units: degrees_north
       standard_name: latitude
       long_name: latitude
       axis: Y
   time  Size:468
       units: days since 1980-1-1 0:0:0
       standard_name: time
       calendar: gregorian
       axis: T
```

Finally, the file contains several attributes relating to the source of information. In this case, there are 15 attributes including the version, the title, a contact email to the creator, the date of update, developer institution, etc.

```
15 global attributes:
    Conventions: CF-1.6
    title: CHIRPS Version 2.0
    history: created by Climate Hazards Group
    version: Version 2.0
    date_created: 2020-01-16
    creator_name: Pete Peterson
    creator_email: pete@geog.ucsb.edu
    institution: Climate Hazards Group.  University of California at Santa Barbara
    documentation: http://pubs.usgs.gov/ds/832/
```

6) The variable "time" is extracted from the file "chi" to create a vector with the dates to which each time step belongs. Additional year and month vectors are generated in numerical form.

```
time<-data.frame(as.Date(ncvar_get(b, "time", verbose = F),origin="1980-01-01"))#A vector is extracted de dates
years<-data.frame(as.numeric(unlist(format(time, format = "%Y"))))#Vector of years
months<-data.frame(as.numeric(unlist(format(time, format = "%m"))))#Vector of months
dim(time)#The dimension of the vector is shown: number of rows x number of columns
```

7) At the end of the information analysis and extraction of the dates to which the data belong, the file is closed. It is recommended not to leave it open as these variables normally use a large amount of memory space on the disk, which may slow down the programme or prevent the creation of new variables.

```
nc_close(chi)#The .nc file is closed as it uses a large amount of disk memory
```

8) The precipitation data is extracted. Here the CHIRPS file is opened again with the particularity of reading directly the variable "precip" that was previously analysed. The variable that was assigned this information is examined for class and mode, in which it is checked that it is type "RasterBrick" and "S4" that corresponds to raster information.

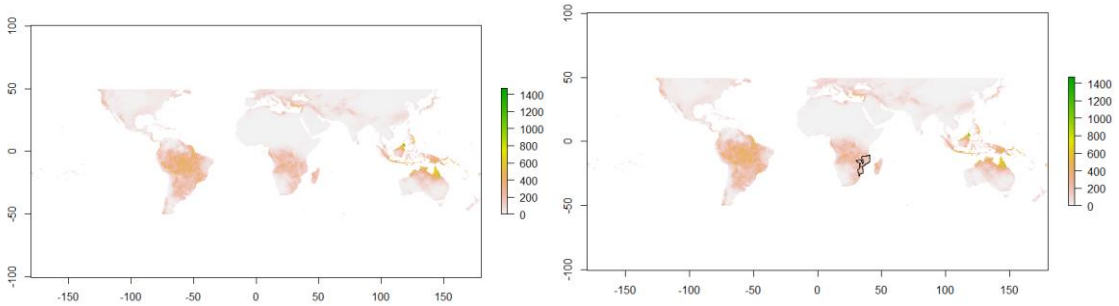
```
#Precipitation information is extracted
precip<-brick("E:/19_MANUAL_data_global/CHIRPS/chirps-v2.0.monthly.nc", varname = "precip")#Precipitation is extracted
class(precip)#It shows what kind of variable it is
mode(precip)#SIt is shown whether the variable information is numeric or text
dim(precip)#The dimension of the raster is displayed number of rows x number of columns x number of months
```

9) As the dimension of each raster has 468-time steps, one of them (the first one in the script) is extracted to analyse its composition. When extracting this information, its dimension is verified.

In addition, it is plotted first on its own and then with the shapefile of the area of interest (Mozambique) to check that it is correctly georeferenced.

```
#A raster is extracted from month 1
precip_1 <-subset(precip,1) #1 is the order of the number of months, comparable with the time step of the variable time
class(precip_1);mode(precip_1)#It is checked that the variable maintains raster characteristics
dim(precip_1)#It is verified that we now have the raster for a single time step (1 in this case)

#Verification of geographical information
plot(precip_1)#The raster is plotted to verify the information
plot(mz,add=T)#Plot the raster and check if the shapefile of interest (mz) is located where it should be
```

In the viewer something like this is obtained:

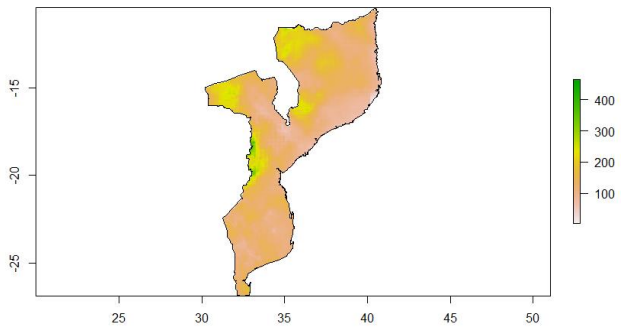


Where the correct spatial coincidence of the raster precipitation information with the "mz" shapefile can be confirmed. It can also be seen how the global precipitation for this particular time step (January 1981) varies from 0 to just over 1400 mm..

10) Once the location of the raster information has been checked against the shapefiles, only the data within the Mozambican area is extracted. The procedure is carried out in three steps:

```
#Precipitation information is extracted within the shapefile of interest (mz)
step1 <- crop(precip_1, mz) #The raster is intersected with the area of interest
step2 <- rasterize(mz, step1) #The shapefile is transformed into a raster
final <- step1*step2 # The final product is created

plot(final)#The raster of interest is plotted
plot(mz, add=TRUE)
```

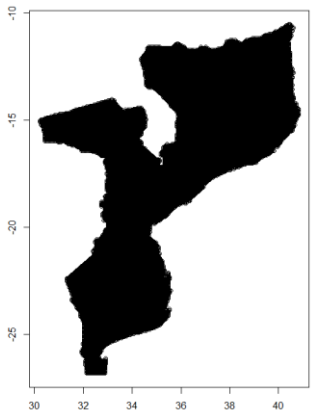
The graph allows to check that the data were correctly extracted:



11) The coordinates of each pixel with precipitation information within the Mozambique shapefile are extracted. Follow the steps in the script and get a matrix with two columns of ordered pairs as longitude and latitude.

```
#The coordinates are extracted according to CHIRPS within the area of interest

aux <-extract(final, mz, cellnumbers=T)#From the final product we extract the order of the coordinates
#within the main raster and the associated precipitation value
coordinates<-data.frame(xyFromCell(final, data.frame(aux)[,1]))#The coordinates of what is inside the
#shape mz are obtained
colnames(coordinates)<-c("long","lat")#Each column of x and y data is named as Longitude and Latitude
#It can be modified
dim(coordinates)#Look at the dimension of the vector which equals the number of cells within the area of interest
plot(coordinates[,1],coordinates[,2])#The coordinates are plotted to check once again that they are correct
head(coordinates)#The first rows of the coordinates are looked at to check that the matrix is correct
```

To find out how many pixels there are, the function dim (26868 pixels) is used, and plotted to see in the viewer if the ordered pairs were extracted correctly:



This matrix (called coordinates) is saved in a folder of interest in .txt format. This file is very important as it contains the location of the cells contained in the shapefile of interest and the order in which it is saved should not be changed later (next step).

```
#The coordinates are saved as a .txt file
write.table(coordinates,"E:/19_MANUAL_data_global/CHIRPS/coordendas_CHIRPS.txt",col.names = T,row.names = F)
#Change the desired location
```

12) During the coordinate extraction process (using the "extract" function), in addition to the coordinates, the precipitation data associated with each coordinate is extracted. In this step a loop is executed where "i" varies from 1 to the number of time iterations contained in the "precip" file. The data extraction process is repeated for each of the time steps and stored in a matrix. This process takes time depending on the number of pixels, i.e., the size of the region under study. At the end of the loop its dimension is checked, which should have one row of data for each time step and a column from left to right corresponding to the coordinate series sorted from top to bottom of the previous step. It is very important to keep the co-ordinate file and the precipitation data file together, as their order is essential for the correct use of the data.

```
#Information is extracted for each time step in a loop (can take 20-30 minutes or more)
for (i in 1:468){#468 are the months of information
  precip_i <-subset(precip,i)
  step1 <- crop(precip_i, mz) #The raster is intersected with the area of interest
  step2 <- rasterize(mz, step1) #The shapefile is converted to raster
  final <- step1*step2 #The final product is created
  if(i==1){
    aux <-data.frame(date=time[1,],year=years[1,],month=months[1,],t(data.frame(extract(final, mz, cellnumbers=T))[,2]))
  }
  if(i!=1){
    aux <-rbind(aux,data.frame(date=time[1,],year=years[1,],month=months[1,],t(data.frame(extract(final, mz, cellnumbers=T))[,2])))
  }
  print(paste(i,"de",468))
  #aux is the result matrix containing the time series of information for each cell
}
dim(aux)#The number of columns of the aux matrix must be equal to the number of cells +3 (3 columns of date, year and month)
#The number of rows equals the number of months
head(aux[,1:10])#The upper left part of the matrix is visible
```

13) The monthly data series are stored in .RDS format, which is an RStudio compatible format that is more efficient than text when the matrices are large. The matrix can be reloaded (function

"readRDS") and with the use of the "head" function the upper left composition of the matrix data is observed, and the correct storage of the data is checked.:

```
#The information is saved as an .RDS file which is used to store large information
#associated with the R environment. Using .txt files may take more time
saveRDS(aux,"E:/19_MANUAL_data_global/CHIRPS/data_CHIRPS.rds")#Change to the desired location

#Monthly data are aggregated to annual data

#Monthly information is read in RDS format

dat_mes<-readRDS("E:/19_MANUAL_data_global/CHIRPS/data_CHIRPS.rds")#File location is used
head(dat_mes[,1:10])#The upper left part of the matrix is visible

class(dat_mes)
mode(dat_mes)
```

14) Monthly data is aggregated to annual data by means of a loop. The code filters the information for each pixel associated with each year (12 months), sums it and accumulates it in a new matrix. This new matrix will have dimensions of number of years per number of pixels. This new variable is saved as an .RDS or .txt file.

```
#Loop aggregating monthly to annual data
for(g in 1981:2019){# g varies for all the years for which information is available
  aux_a<-filter(dat_mes,dat_mes[,2]==g)#The matrix is filtered for data corresponding to each year
  if(g==1981){
    dat_anual<-data.frame(year=g,t(colSums(aux_a[,4:ncol(aux_a)])))#Monthly data are aggregated to annual data
  }
  if(g!=1981){
    dat_anual<-rbind(dat_anual,data.frame(year=g,t(colSums(aux_a[,4:ncol(aux_a)]))))
  }
  print(paste(g,"de",2019))#Loop progress indicator
}
dim(dat_anual)#The dimension of the matrix is checked (number of years x number of cells +1 corresponding to the year)
head(dat_anual[,1:10])#The upper left part of the matrix is checked

#The information is saved as an .RDS file
saveRDS(dat_anual,"E:/19_MANUAL_data_global/CHIRPS/data_anuales_CHIRPS.rds")#Change to the desired location
```

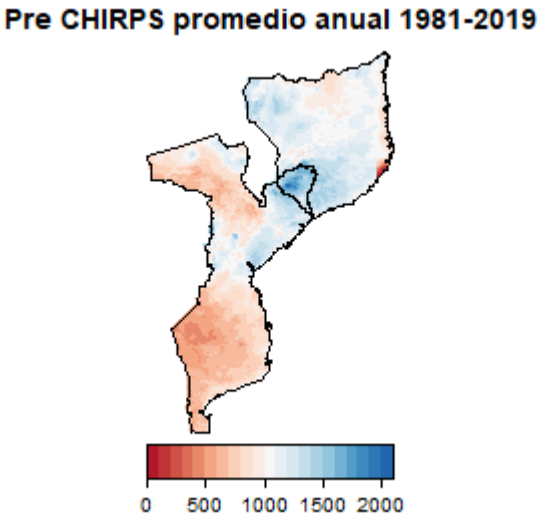
15) From the annual data matrix the mean annual precipitation can be estimated for each pixel. It is computed by generating a 3-column matrix. The first two with the ordered pair of coordinates to which each data belongs and a third one with the vector resulting from averaging the annual precipitation. This file is saved in a folder of preference.

```
#Annual precipitation data are averaged
dat_anual_promedio<-data.frame(coordinates,colMeans(dat_anual[,-1]))#A matrix containing the coordinates of each pixel
# and its corresponding annual average precipitation
#value is formed. Name can be changed
colnames(dat_anual_promedio)<-c("Long","lat","Pre_CHIRPS")#The columns of the matrix are named
head(dat_anual_promedio)#The composition of the matrix is reviewed

#The information is saved as an .RDS file
saveRDS(dat_anual_promedio,"E:/19_MANUAL_data_global/CHIRPS/data_anuales_promedio_CHIRPS.rds")#Change the desired location
```

16) To plot the annual mean precipitation in an elegant and professional way, the file with this information is opened. The matrix is converted into a raster (see code with the functions used), and the "levelplot" function is used to generate a high quality graph. The colour range, the values shown in the legend, whether or not to show geographic quadrants, frames, etc. can be manipulated. This plot can be saved in any desired format.

```
#The matrix containing the data to be plotted is read
dat_grafico<-readRDS("E:/19_MANUAL_data_global/CHIRPS/data_anuales_promedio_CHIRPS.rds")#Change the desired location
dim(dat_grafico)#Their dimensions are revised
head(dat_grafico)#The composition of data is reviewed


#The information is prepared for graphing

raster_1<-rasterFromXYZ(dat_grafico)#The information is rasterised
class(raster_1); mode(raster_1)#The variable is checked to ensure that it is of raster type

#The boundaries within which the information is found are sought

max(dat_grafico[,3],na.rm = T)#Maximum value, 3 is the column where the data of interest is located
min(dat_grafico[,3],na.rm = T)#Minimum value, 3 is the column where the data of interest is located

#The range of colours to be plotted is prepared from low to high rainfall.
cols <- colorRampPalette((brewer.pal(9, "RdBu")))#I define a range of colours, can be changed in ""

#The data is plotted in an elegant way
professional_graphics<-levelplot(raster_1,main="Pre CHIRPS promedio anual 1981-2019",ylab="",xlab="",#Choose the raster to plot
                                                                                                      #The title can be changed
              margin=FALSE,
              colorkey=list(space='bottom',axis.line=list(col='black')),
              par.settings=list(axis.line=list(col='transparent'),#If deleted, the graph is displayed
                                                                  #with grids in its coordinates
                                strip.background=list(col="transparent"),
                                strip.border=list(col="transparent")),#If deleted, the graph is displayed with a border
              scales=list(draw=FALSE),#TRUE shows the geographical coordinates
              col.regions=cols,#Variable defining the colour range
              at=seq(0, 2100,100))+#The limits are changed according to what we want to show
            layer(sp.polygons(mz))+layer(sp.polygons(cl))#Interest limits overlap

#The graphic is exported as png, the destination folder and the size can be modified.
png("E:/19_MANUAL_data_global/Graphics/pre_anual_CHIRPS_profesional.png",height = 270,width = 400)
professional_graphics #Assigned name of the graphic to be exported as png
```

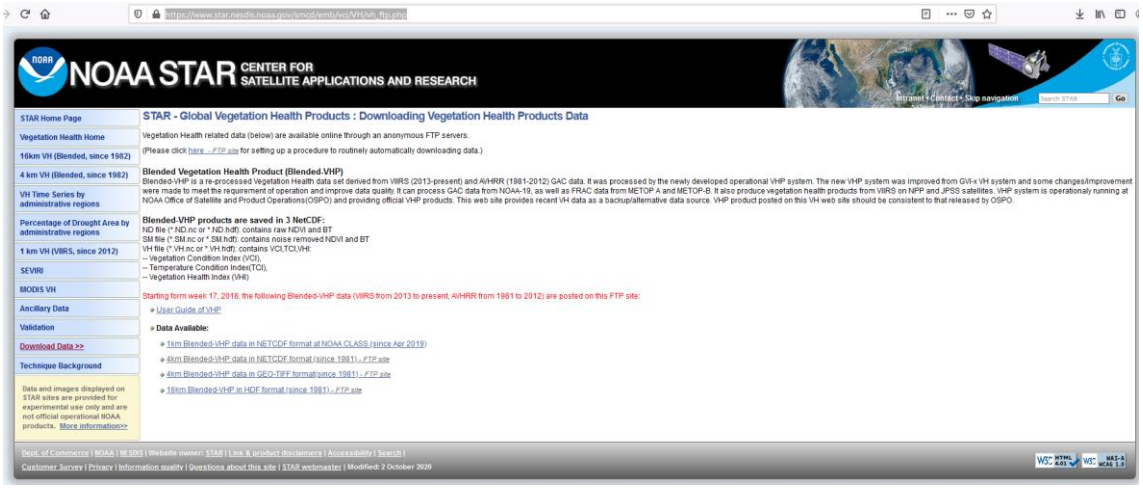
The result of this graphic will look like this, which also includes the shapefile of the Licungo river basin:



11. Vegetation indices from NOAA STAR

In this section the VCI, TCI and VHI drought indices are downloaded for Mozambique from NOAA STAR. The process can be applied to other products such as NDVI or BT. A different data download strategy than the one used from CHIRPS is used. The data series come on a weekly scale, so they are aggregated on a monthly scale to be comparable with precipitation or other drought indices. The monthly state is plotted for 6 consecutive months to analyse its change.

1) The files containing the NOAA STAR products are downloaded from the website: https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. The following information can be found there:



Where, it is specified that the products are stored in 3 NetCDF files (extension .NC). It is indicated which files contain the various data available and there is access to a user manual which is recommended for review. Therefore, the "VH files" containing the VCI, TCI and VHI indexes are required here. Click on the link of interest to download data at a scale of 4km:

Índice de ftp://ftp.star.nesdis.noaa.gov/pub/corp/scsb/wguo/data/Blended_VH_4km/VH/

Subir al directorio superior.

Nombre	Tamaño	Última modificación
Archivo: VHP.G04.C07.NC.P1981035.SM.nc	29574 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981035.VH.nc	54422 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981036.SM.nc	29362 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981036.VH.nc	56737 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981037.SM.nc	29031 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981037.VH.nc	58356 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981038.SM.nc	28745 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981038.VH.nc	59192 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981039.SM.nc	28479 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981039.VH.nc	59522 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981040.SM.nc	28250 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981040.VH.nc	59616 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981041.SM.nc	28070 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981041.VH.nc	59691 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981042.SM.nc	27891 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981042.VH.nc	59750 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981043.SM.nc	27739 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981043.VH.nc	59844 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981044.SM.nc	27599 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981044.VH.nc	59967 KB	07/08/2017 2:00:00
Archivo: VHP.G04.C07.NC.P1981045.SM.nc	27465 KB	07/08/2017 2:00:00

All VH files are downloaded one by one and saved in a preferred folder. Each of these files presents the VCI, VHI and TCI index series worldwide at a 4km scale for each week since 1981. Being a very small grid, NOAA STARR has decided to provide one file for each week. Although there are programming mechanisms to download directly, in this manual we will do it manually.

Additionally, in the folder where we save the files (1946 files from week 35 of 1981 to week 52 of 2019), a text file (.txt) containing the names of each of the downloaded files is required.

2) Once in RStudio, the necessary packages are installed and loaded.
3) The Mozambique shapefile, assigned with the name mz, is read.
4) The file containing the names of the VH files is read. Data on the area of interest is extracted from each of them.

```
#The necessary geographical information is read from the folders where they are stored
mz<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/mozambique.shp")#Mozambique shapefile is read

#Vector with VH filenames is read
vh<-read.table("E:/data_NOAASTAR/VH_extraer.txt")#File with names of downloaded files
dim(vh)#Dimension (number of weeks of data) is revised
```

5) As with the CHIRPS data files, one of the VH files is loaded and printed to view its configuration. Here the list of file names in .txt is used to read the files.

```
#The composition of the files is checked, one is taken as an example

i=1 #i is the row number of the file with names vh
aux <- nc_open(paste0("E:/data_NOAASTAR/doc_pag_web/",vh[i,]))#The first downloaded file is read
print(aux)#The file is displayed and its composition is shown
```

The first thing to note is the number of variables contained in each of these files:

```
> print(aux)#Imprimimos el archivo y vemos su composición
File E:/Datos_NOAASTAR/doc_pag_web/VHP.G04.C07.NC.P1981035.VH.nc (NC_FORMAT_NETCDF4):

     4 variables (excluding dimension variables):
        short VCI[WIDTH,HEIGHT]   (Chunking: [2000,724])  (Compression: level 5)
            long_name: VCI
            coordsys: cartesian
            units: NONE
            range: 0
             range: 100
            _Fillvalue: -999
            scale_factor: 0.00999999977648258
            add_offset: 0
            Remark: Value= scale_factor * (ScaledInteger - add_offset)
        short TCI[WIDTH,HEIGHT]   (Chunking: [2000,724])  (Compression: level 5)
            long_name: TCI
            coordsys: cartesian
            units: NONE
            range: 0
             range: 100
            _Fillvalue: -999
            scale_factor: 0.00999999977648258
            add_offset: 0
            Remark: Value= scale_factor * (ScaledInteger - add_offset)
        short VHI[WIDTH,HEIGHT]   (Chunking: [2000,724])  (Compression: level 5)
            long_name: VHI
            coordsys: cartesian
            units: NONE
            range: 0
             range: 100
            _Fillvalue: -999
            scale_factor: 0.00999999977648258
            add_offset: 0
            Remark: Value= scale_factor * (ScaledInteger - add_offset)
        byte QA[WIDTH,HEIGHT]   (Chunking: [3334,1206])  (Compression: level 5)
            long_name: QA
            coordsys: cartesian
            units: NONE
            range: 0
             range: 255
            _Fillvalue: 1
            Remark: From the least significant bit(LSB):
```

Of the 4 variables or products, the VCI, TCI and VHI are identified by the same name and come with a configuration that depends on the variables "WIDTH" and "HEIGHT". They vary from 0 to 100 and are located on the Cartesian system.

```
    2 dimensions:
        HEIGHT  Size:3616
[1] "vobjtovarid4: **** WARNING **** I was asked to get a varid for dimension named HEIGHT BUT this dimension HAS NO DIMVAR! Code will pro
bably fail at this point"
        WIDTH  Size:10000
[1] "vobjtovarid4: **** WARNING **** I was asked to get a varid for dimension named WIDTH BUT this dimension HAS NO DIMVAR! Code will prob
ably fail at this point"
```

The next thing that can be seen is the size of the dimensions "WIDTH" and "HEIGHT". This means that each matrix containing each variable has 10000 x 3616 cells. Furthermore, you can see that it has no dimensions, this means that it is not possible to directly extract the data using the strategy used in CHIRPS. What is done in these cases is explained below.

```
22 global attributes:
    VERSION: VH (vh.exe,version 1.3, March 21 2012)
    SATELLITE: NC
    INSTRUMENT: AVHRR
    CITATION_TO_DOCUMENTS: User Guide of Vegetation Health(VH) system (version 1.3, March 21 2012)
    CONTACT: NOAA/NESDIS/STAR/EMB
    PRODUCT_NAME: Vegetation Health
    PROJECTION: Plate_Carree
    DATE_BEGIN: 239
    DATE_END: 245
    TIME_BEGIN: 00:00 UTC (use day time data only)
    TIME_END: 23:59 UTC (use day time data only)
    ANCILLARY_FILES: FILE_CONFIGURE:vh.config_NN

            YEAR: 1981
            PERIOD_OF_YEAR: 35
            DAYS_PER_PERIOD: 7
            END_LATITUDE_RANGE: -55.1520004272461
            START_LONGITUDE_RANGE: -180
            START_LATITUDE_RANGE: 75.0240020751953
            END_LONGITUDE_RANGE: 180
            INPUT_FILES: 2
            INPUT_FILENAMES: data/AVHRR_VHP/4km/VH/VHP.G04.C07.NC.P1981035.SM.nc
    data/AVHRR_VHP/4km/climate/VHP.G04.C07.CLIMAT.P35.nc
```

Finally, various other attributes of the file are displayed, such as the measuring instrument, the satellite, the date on which the data collection mission took place, the year and the week to which the data contained therein correspond, among others. Relevant data are the range in which each pixel is geographically located.

6) In this step the VCI values are extracted as an example from this file. This is done to observe its composition and to know how to treat the data later. We store in a variable (vci in this case) the matrix of dimension 10000 x 3616 (we check it with the function "dim"). We also check the class and mode of the variable, which should be an array with numeric values.

```
#A variable is extracted to analyse its composition
vci <- ncvar_get(aux,"VCI",verbose=F)#The VCI is extracted
close(aux)#The aux file is closed as it may use too much memory
class(vci)#The class of the file is shown
mode(vci)#The mode of the file is shown
dim(vci)#The dimension of the file corresponding to the number of cells with data in the world is displayed
```

7) Once the inspection of the data is finished, two vectors are created: one for longitude and latitude corresponding to the dimensions "WIDTH" and "HEIGHT". Since they are not georeferenced, they are created manually using the information extracted in step 5.

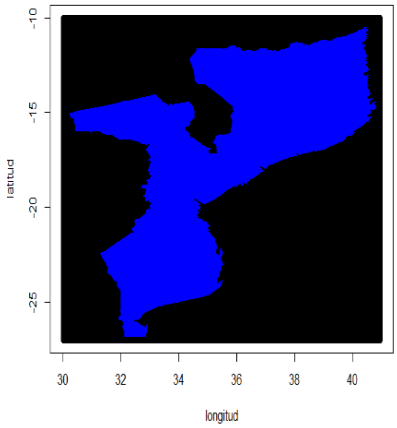
```
#Vectors containing the possible longitudes and latitudes are created
dlong <- 360/10000; dlat <- 360/10000
lon <- matrix(nrow = 10000); lat <- matrix(nrow = 3616)
for (g in 0:9999){ #A matrix with possible length coordinates is created
  lon[g+1,] <- (-180+(g+0.5)*dlong)
}
for (h in 0:3615){  #A matrix is created with possible latitude coordinates
  lat[h+1] <- (75.024-(h+0.5)*dlat)
}
dim(lon)#Number of longitudinal divisions
dim(lat)#Number of latitudinal divisions
head(lon)
head(lat)
```

8) The longitudinal and latitudinal coordinates of all the pixels that conform the database are delimited according to the limits that mark a rectangle containing the shapefile of interest. In this way the matrix containing the coordinates from which the coordinates contained in the shapefile of interest will be extracted will not be excessively large. The longitudinal and latitudinal limits used are those shown in Fig. 6.6.

```
#Limited to a square encompassing Mozambique
long1<-30 #Left longitudinal limit
long2<-41 #Right longitudinal limit
lat1<-(-27) #Lower latitudinal limit
lat2<-(-10) #Upper latitudinal limit

long_a<-lon[lon<=long2]
lon_b<-data.frame(long_a[long_a>=long1])#Lengths are trimmed within limits

lat_a<-lat[lat<=lat2]
lat_b<-data.frame(lat_a[lat_a>=lat1])#Latitudes are cut off within the boundaries
```

9) Using the longitude and latitude vectors already delimited to the shapefile boundaries (lon_b and lat_b), a matrix is created with the ordered pairs corresponding to all the coordinates contained in the proposed boundaries. This matrix is then plotted and superimposed on the shapefile of interest so that it can be visually checked that the coordinates created cover the whole territory of interest.

```
#A matrix is created with ordered pairs of all coordinates within the box of interest
coor <- matrix(ncol=2)
for (ll in 1:nrow(lon_b)){
  for (mm in 1:nrow(lat_b)){
    aux <- matrix(ncol=2)
    aux[1,1] <- lon_b[ll,]
    aux[1,2] <- lat_b[mm,]
    coor<- rbind(coor,aux)
  }
}
coor <- coor[-1,]#coor is the matrix with the ordered pairs of coordinates
colnames(coor)<-c("longitude", "latitude")#Columns are named
dim(coor)#Dimension of the matrix is checked
head(coor)#First rows of the matrix are observed

#Check that the extracted coordinates are well located
plot(coor) #The points are plotted
plot(mz,add=T,col="blue")#Plot the Mozambique shapefile and check that the
                         #points are in the same place and cover the whole polygon
```



10) The next step is to extract the coordinates contained in the shapefile. This can be done using a GIS or directly in R, which is how it is done here. First, the coordinate matrix is transformed to a spatial class file (points with geographic information) where the projection must be defined, in this case the WGS84 geographic projection is used. Then, the coordinates contained in the shapefile of interest, in this case "mz", are extracted. Finally, the file is transformed again to a matrix with numerical data (without geographic information), the dimension of this file is checked, it is plotted to see if the operation has been performed correctly and the matrix is saved in a known folder.

```
#The coordinates within the area of interest are obtained

#The coordinate matrix is transformed into a file containing geographic information
coor_gis <- SpatialPointsDataFrame(coords = as.data.frame(coor),
                                   data = as.data.frame(coor),
                                   proj4string = CRS("+proj=longlat +datum=WGS84 +no_defs"))#The projection is defined
coor_mz<-coor_gis[complete.cases(over(coor_gis, mz)),]#The coordinates intersecting the shapefile mz are extracted
class(mz);class(coor_gis);class(coor_mz)#The type of variable is checked, they must be spatial.
dim(coor_mz)#The dimension is reviewed
#It is drawn to check that they are those points
plot(coor_mz)
coordinates<-as.data.frame(coor_mz)[,1:2]#The spatial file is transformed into a matrix and saved as text

write.table(coordinates,"E:/19_MANUAL_data_global/VCI_TCI_VHI/coordinates_extraer.txt",
            col.names = T,row.names = F)#Change the folder
```

11) Since the variables to be extracted depend on the dimensions "WIDTH" and "HEIGHT", and these vary from 1 to 1000 and 1 to 3616, respectively, the position to which each coordinate contained in the shapefile corresponds is searched. For this search, the decimals of the coordinates of the target matrix (contained in mz) and those of the vectors containing all possible longitudes and latitudes are rounded. The result will be a matrix with a number of rows equal to the number of coordinates contained in mz and 2 columns corresponding to the positions of "WIDTH" and "HEIGHT", they have been named "lo1".

```
#Coordinates are read
coordinates<-read.table("E:/19_MANUAL_data_global/VCI_TCI_VHI/coordinates_extraer.txt",header = T)
dim(coordinates)

#Round the decimal values of the coordinates to match
r1<-round(coordinates[,1:2],3)#Mz coordinates are rounded to 3 decimal places
r2<-round(lon,3)#All longitude values are rounded to 3 decimals
r3<-round(lat,3)#All latitude values are rounded to 3 decimal places

#A matrix is created that locates the positions within the "VCI, TCI and VHI" files
for (i in 1:nrow(r1)){#Loop varies for each coordinate within mz
  if (i==1){
    lo1<-cbind(which(r2[,1]==r1[i,1]),which(r3[,1]==r1[i,2]))
  }
  if (i!=1){
    lo1<-rbind(lo1,cbind(which(r2[,1]==r1[i,1]),which(r3[,1]==r1[i,2])))
  }
}#lo1 is the variable containing the positions within the files to extract for each coordinate
colnames(lo1)<-c("longitude", "latitude")#Name the position columns
dim(lo1)#Check dimension, must be equal to that of coordinates
head(lo1)#Matrix organisation is checked, must be equal to that of coordinates
```

12) Se A loop is executed with a number of iterations equal to the number of weekly files to extract the information. Each iteration follows the order of the vector containing the file names, so care needs to be taken to correctly spell the name of the folders where the original .nc files are stored. The results will be 3 matrices, corresponding to the VCI, TCI and VHI values of all the cells contained in the shapefile. Their dimension will be the number of rows equal to the number of weeks with data and the number of columns equal to the number of pixels. It should be noted that the order of the coordinates from left to right is coincident with the order from top to bottom of the coordinates obtained in step 10.
This step may take several hours, depending on the size of the region of interest.

```
#The weekly series of each index are extracted in a matrix of dimension number of weeks x number of cells
for (i in 1:nrow(sm)){ #The loop opens and extracts the values of each index in the 1946 files
  indvarios <- nc_open(paste0("E:/data_NOAASTAR/doc_pag_web/",vh[i,]))#Each weekly file is read
  vci <- ncvar_get(indvarios,"VCI",verbose=F)#Global VCI information is extracted
  tci <- ncvar_get(indvarios,"TCI",verbose=F)#Global TCI information is extracted
  vhi <-ncvar_get(indvarios,"VHI",verbose=F)#Global VHI information is extracted

  #The dates of each file are extracted from which we derive the year, month and week of the file
  year <- as.data.frame(as.numeric(unlist(ncatt_get(indvarios,0, "YEAR",verbose=F))))[-1,]
  week <- as.matrix(as.numeric(unlist(ncatt_get(indvarios,0, "PERIOD_OF_YEAR"))))[-1,]
  month <- as.numeric(format((as.Date(paste(year,week,"1",sep = "-"),'%Y-%W-%u')),"%m"))

  #The loop that extracts the information from each cell is launched.
  #It can take several hours, depending on the number of cells
  for (j in 1:nrow(lo1)){#The number of steps in the loop is the number of cells we will extract values from
    if (j==1){
      vci1<- data.frame(year,month,week,vci[lo1[j,1],lo1[j,2]])
      tci1<- data.frame(year,month,week,tci[lo1[j,1],lo1[j,2]])
      vhi1<- data.frame(year,month,week,vhi[lo1[j,1],lo1[j,2]])
    }
    if (j!=1){
      vci1<- cbind(vci1,vci[lo1[j,1],lo1[j,2]])#The vector contains the VCI data of the corresponding week for each cell
      tci1<- cbind(tci1,tci[lo1[j,1],lo1[j,2]])#The vector contains the TCI data of the corresponding week for each cell
      vhi1<- cbind(vhi1,vhi[lo1[j,1],lo1[j,2]])#The vector contains the VHI data of the corresponding week for each cell
    }
    print(paste0(j,"de",nrow(lo1)))#Step counter
  }
  colnames(vci1)<-c("year","month","week",(1:(ncol(vci1)-3)))#The columns are named
  colnames(tci1)<-c("year","month","week",(1:(ncol(tci1)-3)))#The columns are named
  colnames(vhi1)<-c("year","month","week",(1:(ncol(vhi1)-3)))#The columns are named
  if (i==1){
    vci2<-vci1
    tci2<-tci1
    vhi2<-vhi1
  }
  if (i!=1){
    vci2<-rbind(vci2,vci1)#The matrix accumulates the corresponding weekly VCI series for each cell
    tci2<-rbind(tci2,tci1)#The matrix accumulates the corresponding weekly TCI series for each cell
    vhi2<-rbind(vhi2,vhi1)#The matrix accumulates the corresponding weekly VHI series for each cell
  }
  nc_close(indvarios)#The file is closed as keeping it open takes up a lot of internal memory
  print(paste(i,"de",nrow(sm)))#Step counter
} # Close loop with list of vh files
dim(vci3)#Check dimension of matrix, must be number of weeks x number of cells +3 (year, month and week)
dim(tci3)#Check dimension of the matrix, must be number of weeks x number of cells +3 (year, month and week)
dim(vh3)#Check dimension of the matrix, must be number of weeks x number of cells +3 (year, month and week)
```

13) The created matrices are saved in .RDS format files. This type of file is recommended over .txt files as it compresses the information resulting in variables that are easier and faster to read in

RStudio. The matrices can be loaded again, the upper left composition of the data can be observed, and it is possible to check that the data storage was correct:

```
#Weekly data matrices are saved as .RDS files
saveRDS(vci3,"E:/19_MANUAL_data_global/VCI_TCI_VHI/vci_weekly_data.rds")
saveRDS(tci3,"E:/19_MANUAL_data_global/VCI_TCI_VHI/tci_weekly_data.rds")
saveRDS(vhi3,"E:/19_MANUAL_data_global/VCI_TCI_VHI/vhi_weekly_data.rds")
```

14) The next step will be to aggregate the weekly data into monthly data by means of a loop. First the weekly data is read and any unusual values that may have been generated are replaced by NA. An NA is an unmeasured but operable value, while others such as NANs represent an error with which no calculations can be performed.

```
#Weekly matrices are added to monthly matrices

#The files containing the VCI, TCI and VHI index matrices are read
vci3<-readRDS("E:/19_MANUAL_data_global/VCI_TCI_VHI/vci_weekly_data.rds")
tci3<-readRDS("E:/19_MANUAL_data_global/VCI_TCI_VHI/tci_weekly_data.rds")
vhi3<-readRDS("E:/19_MANUAL_data_global/VCI_TCI_VHI/vhi_weekly_data.rds")

#The dimensions of the matrices are reviewed
dim(vci2)
dim(tci2)
dim(vhi2)

head(vhi2[,1:10])#The composition of the matrices is reviewed

#Weekly matrices may contain values with errors, we define them as NA

auxil2<-vci2 #New name of the array containing the VCI values
auxil2[is.na(auxil2)]<-NA#Outliers are removed and replaced by NA
dim(auxil2)#The composition of the new matrix is revised

auxil3<-tci2#New name of the array containing the TCI values
auxil3[is.na(auxil3)]<-NA#Outliers are removed and replaced by NA
dim(auxil3)#The composition of the new matrix is revised

auxil4<-vhi2#New name of the array containing the VHI values
auxil4[is.na(auxil4)]<-NA#Outliers are removed and replaced by NA
dim(auxil4)#The composition of the new matrix is revised
```

The code filters the information of each pixel associated to the weeks of each month (12 months), averages it and saves it in a new matrix. This new matrix will have dimensions of number of months by number of pixels. This new variable is saved as an .RDS or .txt file.

```
#Weekly matrices are added to monthly matrices
for (i in 1981:2019){#The loop varies for the years available in the extracted data
  for (j in 1:12){#The weeks of each month in the year in which the loop is found are filtered out
    fir2<-filter(auxil2,auxil2[,1]==i & auxil2[,2]==j)#An auxiliary matrix is created with the weekly VCI
    #values for the year and month according to the loops
    fir3<-filter(auxil3,auxil3[,1]==i & auxil3[,2]==j)#An auxiliary matrix is created with the weekly TCI
    #values for the year and month according to the loops
    fir4<-filter(auxil4,auxil4[,1]==i & auxil4[,2]==j)#An auxiliary matrix is created with the weekly VHI
    #values for the year and month according to the loops
    if (j==1){
      gir2<-data.frame(year=i,month=j,vci=t(data.frame(colMeans(fir2,na.rm = T))))
      gir3<-data.frame(year=i,month=j,tci=t(data.frame(colMeans(fir3,na.rm = T))))
      gir4<-data.frame(year=i,month=j,vhi=t(data.frame(colMeans(fir4,na.rm = T))))
    }
    if (j!=1){
      gir2<-rbind(gir2,data.frame(year=i,month=j,vci=t(data.frame(colMeans(fir2,na.rm = T)))))#CA monthly VCI matrix is created for each year of the loop
      gir3<-rbind(gir3,data.frame(year=i,month=j,tci=t(data.frame(colMeans(fir3,na.rm = T)))))#A monthly TCI matrix is created for each year of the loop
      gir4<-rbind(gir4,data.frame(year=i,month=j,vhi=t(data.frame(colMeans(fir4,na.rm = T)))))#A monthly VHI matrix is created for each year of the loop
    }
  }
  if (i==1981){
    vci_m<-gir2
    tci_m<-gir3
    vhi_m<-gir4
  }
  if (i!=1981){
    vci_m<-rbind(vci_m,gir2)#Monthly VCI matrices are collected at each step of the loop
    tci_m<-rbind(tci_m,gir3)#Monthly TCI matrices are collected at each step of the loop
    vhi_m<-rbind(vhi_m,gir4)#Monthly VHI matrices are collected at each step of the loop
  }
  print(paste(i,"de","2019"))#Loop step counter
}
dim(vhi_m)#dimension of the monthly VCI matrix is revised, should be num. of months x num. of cells +2 (year and month)
dim(vci_m)#Check the diemnsion of the monthly TCI matrix, must be num. of months x num. of cells +2 (year and month)
dim(tci_m)#VHI monthly matrix dimension is checked, must be num. of months x num. of cells +2 (year and month)
head(vhi_m[,1:8])#Matrix composition is checked
head(vci_m[,1:8])#Matrix composition checked
head(tci_m[,1:8])#Matrix composition checked
```

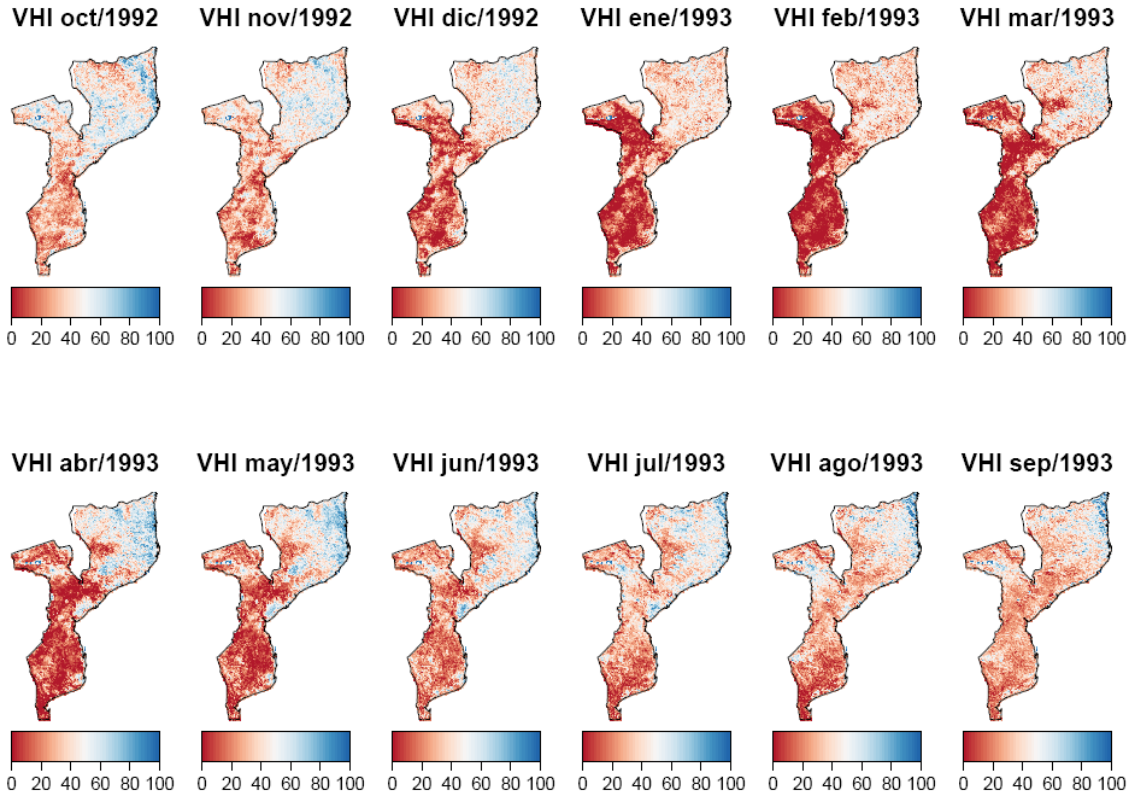
```
#Because some weekly values can be averaged with NA (week with
#undetected value), erroneous values are created and replaced by NA
vci_m[is.na(vci_m)]<-NA
tci_m[is.na(tci_m)]<-NA
vhi_m[is.na(vhi_m)]<-NA

#Monthly data matrices are saved as .RDS files

saveRDS(vci_m,"E:/19_MANUAL_data_global/VCI_TCI_VHI/vci_monthly_data.rds")
saveRDS(tci_m,"E:/19_MANUAL_data_global/VCI_TCI_VHI/tci_monthly_data.rds")
saveRDS(vhi_m,"E:/19_MANUAL_data_global/VCI_TCI_VHI/vhi_monthly_data.rds")
```

15) The last section shows how to generate several graphs corresponding to the state of a variable over the whole territory of interest, how to put them together in a matrix and export them. The monthly VHI is used as an example. The VHI matrix is converted into a raster (see code with the functions used), and the function "levelplot" is applied to generate a high quality graph. The process is repeated for 12 consecutive rows, corresponding to the months from October 1992 to September 1993, a period in which the development of a major drought in Mozambique can be observed.

```
#The coordinates of the centroids of the cells are read
coordinates<-read.table("E:/19_MANUAL_data_global/VCI_TCI_VHI/coordinates_extraer.txt",header = T)#The coordinates are read
dim(coordinates)#Its dimension is reviewed
head(coordinates)#Check that the matrix has the desired composition

#Open the matrix containing the data to be plotted.
vhi<-readRDS("E:/19_MANUAL_data_global/VCI_TCI_VHI/vhi_monthly_data.rds")#The VHI matrix is read
dim(vhi)#The dimension of the matrix is reviewed
head(vhi[,1:5])#Check that the matrix has the desired composition

#The information is prepared for graphing
dat_grafico1<-data.frame(coordinates,t(vhi[98,-c(1:2)]))#A matrix containing the coordinates of each
                                                        #pixel and its corresponding annual vaverage
                                                        #precipitation value is formed. Name can be
                                                        #changed from 98 to 109 are the rows we want to
                                                        #plot corresponding to the dates we want to display.
dat_grafico2<-data.frame(coordinates,t(vhi[99,-c(1:2)]))
dat_grafico3<-data.frame(coordinates,t(vhi[100,-c(1:2)]))
dat_grafico4<-data.frame(coordinates,t(vhi[101,-c(1:2)]))
dat_grafico5<-data.frame(coordinates,t(vhi[102,-c(1:2)]))
dat_grafico6<-data.frame(coordinates,t(vhi[103,-c(1:2)]))
dat_grafico7<-data.frame(coordinates,t(vhi[104,-c(1:2)]))
dat_grafico8<-data.frame(coordinates,t(vhi[105,-c(1:2)]))
dat_grafico9<-data.frame(coordinates,t(vhi[106,-c(1:2)]))
dat_grafico10<-data.frame(coordinates,t(vhi[107,-c(1:2)]))
dat_grafico11<-data.frame(coordinates,t(vhi[108,-c(1:2)]))
dat_grafico12<-data.frame(coordinates,t(vhi[109,-c(1:2)]))

raster_1<-rasterFromXYZ(dat_grafico1)#The information is rasterised
raster_2<-rasterFromXYZ(dat_grafico2)#The information is rasterised
raster_3<-rasterFromXYZ(dat_grafico3)#The information is rasterised
raster_4<-rasterFromXYZ(dat_grafico4)#The information is rasterised
raster_5<-rasterFromXYZ(dat_grafico5)#The information is rasterised
raster_6<-rasterFromXYZ(dat_grafico6)#The information is rasterised
raster_7<-rasterFromXYZ(dat_grafico7)#The information is rasterised
raster_8<-rasterFromXYZ(dat_grafico8)#The information is rasterised
raster_9<-rasterFromXYZ(dat_grafico9)#The information is rasterised
raster_10<-rasterFromXYZ(dat_grafico10)#The information is rasterised
raster_11<-rasterFromXYZ(dat_grafico11)#The information is rasterised
raster_12<-rasterFromXYZ(dat_grafico12)#The information is rasterised

class(raster_1); mode(raster_1)#The variable is checked to ensure that it is of raster type


#The boundaries within which the information is located are sought

max(dat_grafico[,3],na.rm = T)#Maximum value, 3 is the column where the data of interest is located
min(dat_grafico[,3],na.rm = T)#Minimum value, 3 is the column where the data of interest is located

#Se prepara la gama de colores a graficar desde poca a mucha precipitación
cols <- colorRampPalette((brewer.pal(9, "RdBu")))#Defino una gama de colores, se peude cambiar en "".

#Previously obtained raster are plotted with names for each one
graphic_1<-levelplot(raster_1,main="VHI oct/1992",ylab="",xlab="",
                             margin=FALSE,
                             colorkey=list(space='bottom',axis.line=list(col='black')),
                             par.settings=list(axis.line=list(col='transparent'),#If deleted, the graph is displayed
                                                                                 #with grids in its coordinates
                                               strip.background=list(col="transparent"),
                                               strip.border=list(col="transparent")),#If deleted, the graph is displayed with a border
                             scales=list(draw=FALSE),#TRUE shows the geographical coordinates
                             col.regions=cols,#Variable defining the colour range
                             at=seq(0,100,2))+#The limits are changed according to what we want to show, the VHI varies between 0 and 100
  layer(sp.polygons(mz))#Limits of interest are attached
```

Finally, to export the 12 generated graphs, the function "grid.arrange" is used, in which it is necessary to define how to arrange the graphs. Here a 2 x 6 matrix is defined. This array must be defined according to the desired size of the graph. The code needed to perform this step and the final result is:

```
#The graphic is exported as a pdf, the destination folder and size can be changed
pdf("E:/19_MANUAL_data_global/Graphics/VHI_1992_profesional.pdf",height = 7,width = 9)
grid.arrange(graphic_1,graphic_2,graphic_3,graphic_4,graphic_5,graphic_6,
             graphic_7,graphic_8,graphic_9,graphic_10,graphic_11,graphic_12,
             ncol = 6, nrow = 2) #The 12 graphs are arranged in a 6 x 2 matrix, which can be changed
dev.off() #Export closed
```



12. Various: TerraClimate

In this section precipitation, ETr and ETP data are downloaded from the TerraClimate database over the Licungo river basin. A mixed strategy between the two applied above will be used. The data series are on a monthly scale. Although only these three variables are shown, the process will be identical for all the products offered by this database.

1) The TerraClimate files can be downloaded from the website: https://climate.northwestknowledge.net/TERRACLIMATE/index_directDownloads.php. The following information can be found there:

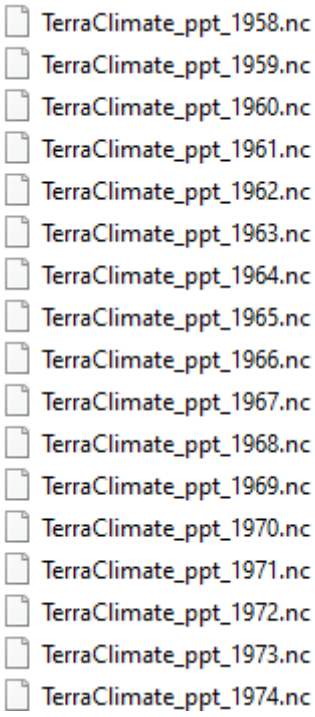
Filter the directory listings:

Select a Year: Show All Years

Select a Variable:

○ Show Variable Description	○ srad (Downward surface shortwave radiation)
◉ Show All Variables	○ swe (Snow water equivalent - at end of month)
○ aet (Actual Evapotranspiration)	○ tmax (Max Temperature)
○ def (Climate Water Deficit)	○ tmin (Min Temperature)
○ pet (Potential evapotranspiration)	○ vap (Vapor pressure)
○ ppt (Precipitation)	○ ws (Wind speed)
○ q (Runoff)	○ vpd (Vapor Pressure Deficit)
○ soil (Soil Moisture)	○ PDSI (Palmer Drought Severity Index)

The variable of interest is chosen, and the required annual files are downloaded one by one. The data extraction from 1958 to 2019 is shown here. The files are saved in a folder of interest. The names of the files vary only in the number of the year to which they correspond.



2) The necessary packages are installed and loaded.
3) The shapefile of the Licungo river basin, assigned with the name cl, is read.
4) As with the other sources of data, a file is read and its composition is analysed.
5) The characteristics of the file are displayed in the console by applying the "print" function to the variable "aux". As for any .nc file, all the necessary information is analysed to understand how the files are stored and to extract them in the area of interest correctly.

The first section shows the number of variables contained in the file. In this case it contains a single variable called "ppt" which stores its data in a three-dimensional array based on longitude, latitude and time. The unit of each data is mm and the time step is monthly:

```
#The composition of the files is analysed, one is taken as an example

aux <- nc_open("E:/data_TerraClimate/ppt_bruto/TerraClimate_ppt_2018.nc")
print(aux)
close(aux)#The aux file is closed as it may use too much memory
```

```
File E:/Datos_TerraClimate/ppt_bruto/TerraClimate_ppt_2018.nc (NC_FORMAT_NETCDF4):

     2 variables (excluding dimension variables):
        short ppt[lon,lat,time]   (Chunking: [1440,720,1])  (Compression: level 9)
            _Fillvalue: -32768
            units: mm
            description: Accumulated Precipitation
            long_name: precipitation_amount
            standard_name: precipitation_amount
            missing_value: -32768
            dimensions: lon lat time
            grid_mapping: crs
            coordinate_system: WGS84,EPSG:4326
            scale_factor: 1
            add_offset: 0
            _Unsigned: true
```

The next section reports the characteristics of the dimensions in which the file stores the information. In this case it divides the world in longitudinal direction into 8640 parts from the east and in latitudinal direction into 4320 parts from the north. These matrices contain 12 time steps (months) which are denoted by a number of days since 1 January 1900 (i.e. the number 29 refers to 29 January 1900). The variable "crs" indicates the projection with which the data is stored.

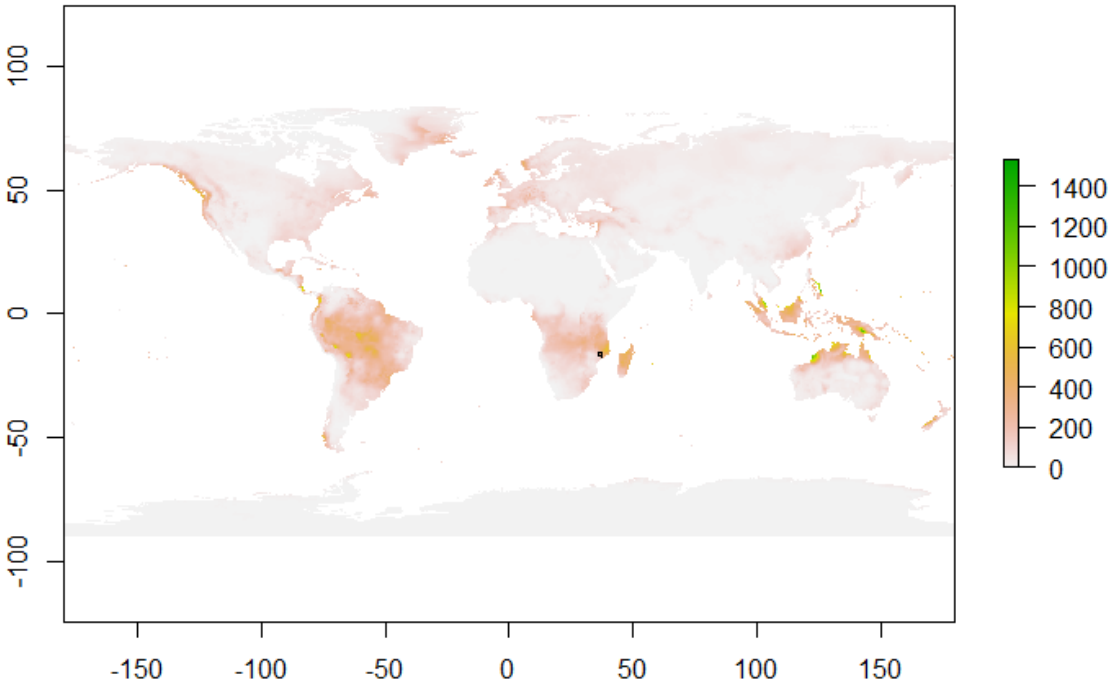
```
4 dimensions:
   lat  Size:4320
        units: degrees_north
        description: Latitude of the center of the grid cell
        long_name: latitude
        standard_name: latitude
        axis: Y
   lon  Size:8640
        units: degrees_east
        description: Longitude of the center of the grid cell
        long_name: longitude
        standard_name: longitude
        axis: X
   time  Size:12
        description: days since 1900-01-01
        units: days since 1900-01-01 00:00:00
        long_name: time
        standard_name: time
        calendar: gregorian
        axis: T
   crs  Size:1
        grid_mapping_name: latitude_longitude
        longitude_of_prime_meridian: 0
        semi_major_axis: 6378137
        inverse_flattening: 298.257223563
        long_name: crs
```

Finally, the file contains several attributes relating to the information source.

6) The next step is to analyse the variables contained in each of the annual files. The precipitation information is extracted as an example. Here the TerraClimate file is read again with the particularity of reading directly the "ppt" data that was previously analysed. This variable is examined for class and mode, where it is checked for "RasterBrick" and "S4" which corresponds to raster information. As it is known that the dimension of each world raster contains 12-time steps, one of them (the first one in the script) is extracted to analyse its composition. When extracting this information, its dimension is verified, which must be of 2 dimensions with values equal to the number of longitudinal and latitudinal divisions in which the database offers its products.

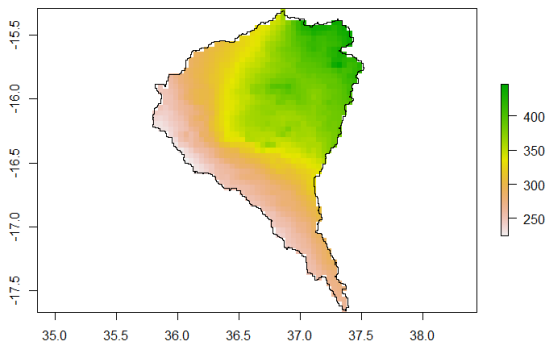
```
#A variable is extracted to analyse its composition

# Precipitation is extracted from the example file
pre <- brick(paste0("E:/data_TerraClimate/ppt_bruto/TerraClimate_ppt_2018.nc",sep = ""), varname = "ppt")
class(pre)#Shows what kind of file it is
mode(pre)#Shows the mode of the file
dim(pre)#Displays the dimension of the file corresponding to the number of cells with data in the world

#A raster is extracted from month 1
precip_1 <-subset(pre,1) # 1 is the order of the number of months, comparable with the time step of the variable time
class(precip_1);mode(precip_1)#Revised to maintain raster characteristics
dim(precip_1)#It is confirmed that you now have the raster for only one time step (1 in this case)

#Geographic information check
plot(precip_1) #Plot the raster to verify the information
plot(cl,add=T)#Plot the raster and check if the shapefile of interest (cl) is located where it is supposed to be
```

To check the location of the information and the shapefile of interest, the raster "precip_1" is plotted with the shapefile of the area of interest (Licungo river basin). In the viewer something like this is obtained:



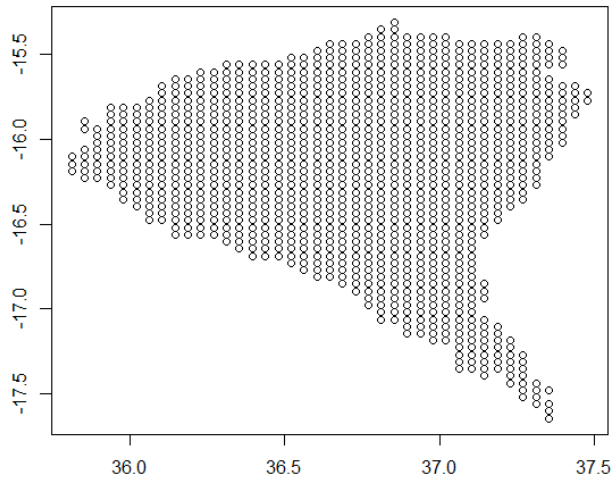
7) Once it has been verified that the raster information is correctly georeferenced, only the data within the area of interest is extracted, following similar steps to those used in the extraction of data from CHIRPS. The procedure is performed in three steps:

```
#Precipitation information is extracted within the shapefile of interest (cl)
step1 <- crop(precip_1, cl) #Intersect the raster with the area of interest
step2 <- rasterize(cl, step1) #Rasterize the shapefile
final <- step1*step2 #Create final product

plot(final)#Plot the raster of interest
plot(cl, add=TRUE)
```



8) The coordinates of each pixel with precipitation information within the shapefile of the Licungo river basin are extracted. Following the steps of the script, a matrix with two columns of ordered longitude and latitude pairs is obtained. Using the dim function, the number of cells contained

in the area (1119 pixels) is obtained, and plotted to check in the viewer if the ordered pairs are correct:



This matrix (called coordinates) is saved in a folder of interest in .txt format. This file is very important as it contains the location of the cells inside the shapefile of interest and the order in which it is saved should not be changed later (step 10).

```
#TerraClimate coordinates within the area of interest are extracted

aux2 <-extract(final, cl, cellnumbers=T)#From the final product we extract the order of the coordinates
#within the main raster and the associated precipitation value
coordinates<-data.frame(xyFromCell(final, data.frame(aux2)[,1]))#The coordinates of what is inside the
#shape cl are obtained
colnames(coordinates)<-c("long","lat")#Each column of x and y data is named as Longitude and Latitude
#and can be modified
dim(coordinates)#Look at the dimension of the vector that equals the number of cells within the area of interest
plot(coordinates[,1],coordinates[,2])#Plot the coordinates to check once again that they are correct
head(coordinates)#Look at the first few rows of the coordinates to check that the matrix is correct

#The coordinates are saved as a .txt file
write.table(coordinates,"E:/19_MANUAL_data_global/TerraClimate/coordendas_TC_Licungo.txt",col.names = T,row.names = F)
#Change the desired location
```

9) Since each .nc file has 12 raster files containing the variable information for each month, a vector of dates, years and months corresponding to the period of interest is created. All files from 1958 to 2019 are downloaded here.

```
#Vectors are created with dates and their corresponding years and months

#Vectors are created with the dates defining the start, end and time step
dates<-data.frame(seq(as.Date("1958/1/15"), as.Date("2019/12/15"), "months"))
year<-data.frame(format(seq(as.Date("1958/1/15"), as.Date("2019/12/15"), "months"),"%Y"))#Dates to years
month<-data.frame(format(seq(as.Date("1958/1/15"), as.Date("2019/12/15"), "months"),"%m"))#Dates to months
```

10) A loop is executed with a number of iterations equal to the number of annual files we want to extract the information. The variable "k" varies according to the year corresponding to each file, so each iteration will follow the annual order of the files according to their names. Care shall be taken to correctly spell the name of the folders where the original .nc files are stored. In each iteration of the loop, the composition of the data in each file is checked and the variable information within the shapefile of interest is stored in a matrix. The result will be 3 matrices, corresponding to the precipitation, ETP and ETr values of all the pixels contained in the shapefile. The dimension of each matrix will be a number of rows equal to the number of months with data and a number of columns equal to the number of pixels. Note that the order of the coordinates from left to right is coincident with the order of the coordinates from top to bottom obtained in step 8.
This step may take several hours, depending on the size of the region of interest.

```
#A loop is launched for the extraction of the variables of interest

for (k in 1958:2019){#The loop depends on the period of years the data needs to be extracted
  #The files containing each variable are opened from the folder containing the files
  aa1 <- brick(paste0("E:/data_TerraClimate/ppt_bruto/TerraClimate_ppt_",k,".nc",sep = ""), varname = "ppt")
  aa2 <- brick(paste0("E:/data_TerraClimate/pet_bruto/TerraClimate_pet_",k,".nc",sep = ""), varname = "pet")
  aa3 <- brick(paste0("E:/data_TerraClimate/aet_bruto/TerraClimate_aet_",k,".nc",sep = ""), varname = "aet")

  for(j in 1:12){#This loop extracts the raster of the 12 months within each annual file
    pal1 <-subset(aa1,j)#The raster of the month varying from 1 to 12 is extracted when "j" varies
    pal2 <-subset(aa2,j)
    pal3 <-subset(aa3,j)

    step1_1 <- crop(pal1, cl) #Intersect the raster with the shapefile of interest
    step1_2 <- crop(pal2, cl)
    step1_3 <- crop(pal3, cl)

    step2_1 <- rasterize(cl, step1_1) #The shapefile is converted to raster
    step2_2 <- rasterize(cl, step1_2)
    step2_3 <- rasterize(cl, step1_3)

    final1 <- step1_1*step2_1 #The final product is created
    final2 <- step1_2*step2_2
    final3 <- step1_3*step2_3

    ext1 <- extract(final1, cl, cellnumbers=T)#From the final product, the order of the coordinates within the main raster and the
    #raster and the value of the variable associated with these coordinates are extracted
    ext2 <- extract(final2, cl, cellnumbers=T)
    ext3 <- extract(final3, cl, cellnumbers=T)

    if(j==1){
      cor_data1<-data.frame(t(data.frame(ext1)[,2]))
      cor_data2<-data.frame(t(data.frame(ext2)[,2]))
      cor_data3<-data.frame(t(data.frame(ext3)[,2]))
    }
    if(j!=1){
      cor_data1<-rbind(cor_data1,data.frame(t(data.frame(ext1)[,2])))
      cor_data2<-rbind(cor_data2,data.frame(t(data.frame(ext2)[,2])))
      cor_data3<-rbind(cor_data3,data.frame(t(data.frame(ext3)[,2])))
    }
    print(paste(j, "de", 12, "en" ,k))#Contador de pasos
  }
  aa1<-0#The auxiliary variable is replaced to reduce memory usage
  aa2<-0#The auxiliary variable is replaced to reduce memory usage
  aa3<-0#The auxiliary variable is replaced to reduce memory usage

  if(k==1958){
    pre_data<-cor_data1
    etp_data<-cor_data2
    etr_data<-cor_data3
  }
  if(k!=1958){
    pre_data<-rbind(pre_data,cor_data1)#The matrices of each domain containing the
    #variable of all pixels within cl are accumulated
    etp_data<-rbind(etp_data,cor_data2)
    etr_data<-rbind(etr_data,cor_data3)
  }
  print(paste(k,"de 2019"))#Step counter
}
```

11) To the three matrices that contain data, the 3 vectors of dates, years and months are added and saved in .RDS format.

```
#A vector of dates, years and months is attached to the beginning of each matrix
pre_data<-cbind(dates,year,month,pre_data);colnames(pre_data)<-c("dates","year","month",c(1:(ncol(pre_data)-3)))
etp_data<-cbind(dates,year,month,etp_data);colnames(etp_data)<-c("dates","year","month",c(1:(ncol(etp_data)-3)))
etr_data<-cbind(dates,year,month,etr_data);colnames(etr_data)<-c("dates","year","month",c(1:(ncol(etr_data)-3)))

#The dimension of the extracted matrices is checked
dim(pre_data)
dim(etp_data)
dim(etr_data)

#The composition of each of the matrices is reviewed
head(pre_data[,1:10])
head(etp_data[,1:10])
head(etr_data[,1:10])

#The matrices are stored in a known folder
saveRDS(pre_data,"E:/19_MANUAL_data_global/TerraClimate/pre_data_cl.rds")
saveRDS(etp_data,"E:/19_MANUAL_data_global/TerraClimate/etp_data_cl.rds")
saveRDS(etr_data,"E:/19_MANUAL_data_global/TerraClimate/etr_data_cl.rds")
```

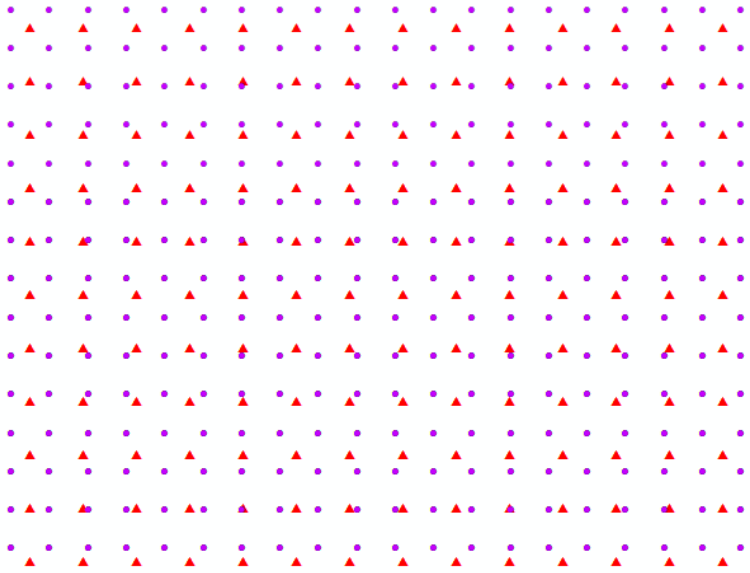
In the console the results of applying the head function to the three matrices can be observed, where the final composition of the matrices can be seen and checked.

```
> head(pre_data[,1:10])
      fechas  ano mes   1   2   3   4   5   6   7
1 1958-01-15 1958  01 311 298 317 306 306 312 323
2 1958-02-15 1958  02 309 314 319 320 323 327 331
3 1958-03-15 1958  03 105 106 111 109 113 118 124
4 1958-04-15 1958  04  74  84  86  89  96 100 102
5 1958-05-15 1958  05  33  48  46  60  64  60  58
6 1958-06-15 1958  06  42  54  55  60  66  70  72
> head(etp_data[,1:10])
      fechas  ano mes   1   2   3   4   5   6   7
1 1958-01-15 1958  01 108 112 108 110 112 113 112
2 1958-02-15 1958  02  86  89  86  88  89  90  89
3 1958-03-15 1958  03 104 108 104 106 108 108 108
4 1958-04-15 1958  04  94  98  95  96  97  97  96
5 1958-05-15 1958  05  91  93  90  89  90  92  92
6 1958-06-15 1958  06  65  65  62  63  64  63  64
> head(etr_data[,1:10])
      fechas  ano mes   1   2   3   4   5   6   7
1 1958-01-15 1958  01 108 112 108 110 112 113 112
2 1958-02-15 1958  02  86  89  86  88  89  90  89
3 1958-03-15 1958  03 103 105 104 105 108 108 108
4 1958-04-15 1958  04  85  92  90  92  95  96  96
5 1958-05-15 1958  05  67  75  72  77  79  79  79
6 1958-06-15 1958  06  53  59  58  61  63  63  64
```

13. Solution to unify spatial scales:

This section will explain simple solutions in the case of requiring two variables to have exactly the same coordinates. Two options are analysed, the first one will be the case of requiring a single time series representative of a specific area; and the other one when all the variables, even if they come from different databases, need to have the same number of pixels.

To exemplify these cases, the precipitation series and the monthly VHI obtained for Mozambique from CHIRPS and NOAA STAR are used. The following graph shows a 60 x 60 km random area zoom within Mozambique, where the blue dots indicate NOAA STAR coordinates and the red triangles indicate CHIRPS coordinates.



1) Coordinates are read from the CHIRPS and NOAA STAR databases. As example variables, precipitation and VHI are also read. The dimensions of the coordinates are different between the two databases since the pixel size is different.

```
#The coordinates of the NOAA STAR cell centroids are read
coor_vhi<-read.table("E:/19_MANUAL_data_global/VCI_TCI_VHI/coordinates_extraer.txt",header = T)#The coordinates are read
dim(coor_vhi)#Dimension checked
head(coor_vhi)#Check that the matrix has the desired composition

#The coordinates of the centroids of the CHIRPS cells are read
coor_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordendas_CHIRPS.txt",header = T)
dim(coor_chi)#Dimension checked
head(coor_chi)#Check that the matrix has the desired composition

#Se lee la matriz que contenga los data del VHI
vhi<-readRDS("E:/19_MANUAL_data_global/VCI_TCI_VHI/vhi_monthly_data.rds")#The VHI matrix is read
dim(vhi)#Dimension checked
head(vhi[,1:5])#Check that the matrix has the desired composition

#Se lee la matriz que contenga los data del VHI
chi_data<-readRDS("E:/19_MANUAL_data_global/CHIRPS/data_CHIRPS.rds")#The precipitation matrix is read
dim(chi_data)#Dimension checked
head(chi_data[,1:5])#Check that the matrix has the desired composition
```

2) In this step the precipitation values for the whole of Mozambique are averaged. Following the code, a matrix is obtained with the same number of rows (corresponding to the number of months), and only columns corresponding to the dates and the average rainfall of the whole country. In order to average several sub-sections of the country, it would be necessary to filter the coordinates of interest and use the same function. This work is not shown here but can be done using various tools shown in the manual.

```
#Averaging within the CHIRPS focus area
chi_mz<-data.frame(chi_data[,1:3],pre=rowMeans(chi_data[,4:ncol(chi_data)],na.rm=T))
#Matrix containing dates and average rainfall across the country
dim(chi_mz)
head(chi_mz)

#The date matrix and representative rainfall vector for Mozambique are stored
write.table(chi_mz,"E:/19_MANUAL_data_global/change_coordinates/CHIRPS_mz.txt",row.names = F)
```

```
> head(chi_mz)
       fecha  ano mes       pre
1 1981-01-01 1981   1 138.37133
2 1981-01-01 1981   2 246.64878
3 1981-01-01 1981   3 129.98534
4 1981-01-01 1981   4  54.99255
5 1981-01-01 1981   5  38.58981
6 1981-01-01 1981   6  12.78778
```

3) From this section onwards, an artifice to obtain data of one variable in the coordinates of another database will be explained. There are many techniques that can be used such as linear or bilinear interpolation. However, when the grid cell size is relatively small compared to the area of interest, this process does not generate significant accuracy, but it does demand a lot of time and computational effort. Here, the information (VHI) of the coordinate closest to the coordinate of interest (CHIRPS) will be searched and that vector of information will be assigned to the CHIRPS coordinate. As an example, first a matrix of target coordinates is defined, in this case the CHIRPS coordinates, and the matrix with initial coordinates, which will be the NOAA STAR coordinates. The minimum distance "d" from the first initial coordinate to all the target coordinates is calculated, then the position of the smallest distance will be the position of the initial coordinate from which the information will be extracted.

```
#Se muestra como buscaremos la posición de la coordenada más cercana con un ejemplo
objetivo<-data.frame(coor_chi[,1:2])#Se define las coordenadas objetivo
iniciales<-data.frame(coor_vhi)#Se definen las coordenadas a buscar
d<-data.frame(which(data.frame(pointDistance(objetivo[1,],iniciales,lonlat = T))
                    ==min(data.frame(pointDistance(objetivo[1,],iniciales,lonlat = T)))))
#d es la posición de la coordenada más cercana
```

The target and initial coordinates are displayed on the console. It can be seen that they are very close in latitude and longitude.

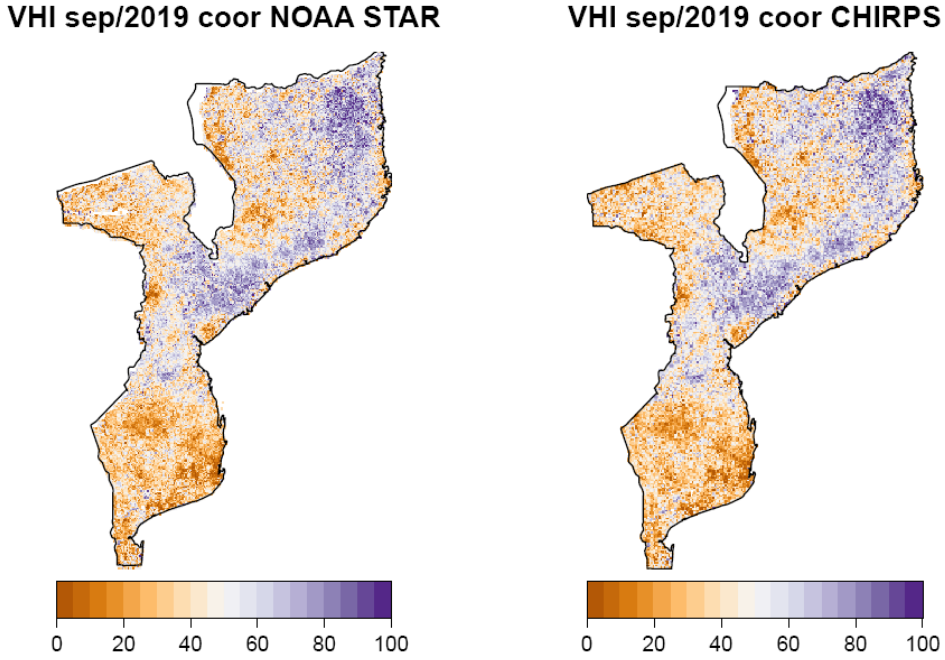
```
#Nearby coordinates are displayed on the console
objetive[1,]
initials[d[,1,1],]
```

```
> objetivo[1,]
    long     lat
1 40.425 -10.525
> iniciales[d[,1,1],]
     longitud latitud
7949   32.886 -26.838
```

4) This process is repeated for all initial coordinates (VHI) with a loop varying from the first to the last coordinate. At each step, the VHI vector belonging to the closest coordinate is extracted and stored in a new matrix. This matrix (called "di1") shall have the same initial number of rows as the VHI data, but shall have the same number of columns as the number of CHIRPS coordinates. The order of the CHIRPS coordinates from top to bottom indicates the order of VHI vectors from left to right. Finally, the matrix is saved in a known folder in .RDS format.

```
#The nearest coordinates of the whole CHIRPS coordinate matrix are searched
objetive<-data.frame(coor_chi[,1:2])#Target coordinates are defined
initials<-coor_vhi#The coordinates to be searched for are defined
for(i in 1:nrow(coor_chi)){
  d<-data.frame(which(data.frame(pointDistance(objetive[i,],initials,lonlat = T))
                      ==min(data.frame(pointDistance(objetive[i,],initials,lonlat = T)))))
  if(i==1){
    di1<-data.frame(vhi[,1:2],data.frame(vhi[,-c(1:2)])[,d[1,1]])
  }
  if(i!=1){
    di1<-cbind(di1,data.frame(data.frame(vhi[,-c(1:2)])[,d[1,1]]))#A new VHI matrix is created where the left to
    #right order of the columns corresponds to the
    #top to bottom coordinate of the CHIRPS
  }
  print(paste(i,"de", nrow(coor_chi)))#Step counter
}
colnames(di1)<-c("year","month",c(1:(ncol(di1)-2)))#The columns of the new matrix are renamed
dim(di1)#Dimension must be equal to CHIRPS data
head(di1[,1:15])

#The new VHI data matrix is saved but corresponding to the CHIRPS coordinates
saveRDS(di1,"E:/19_MANUAL_data_global/change_coordinates/vhi_coordinates_chirps.txt")
```

5) With the aim of comparing the new version of VHI data with the original, both databases are plotted for a random date (sep/2019 in this case). The process is followed to plot the raster in space following the code. The result will be:



The original resolution of the VHI data is finer than that of CHIRPS. However, the spatial distribution of the information has not decreased in significant detail.

Section III: Drought management applications

14. Water balance applications

This section is aimed at the calculation and representation in maps of the dry months (eq. 3), the water deficit (eq. 4) and the volume of available water (eq. 5) over the Licungo river basin. Precipitation, ETr and ETP variables extracted from TerraClimate will be used. Three different applications are shown that could be used in other ways according to the needs of the users.

1) The necessary packages are installed and loaded.
2) The shapefile of the Licungo river basin is read, it has been assigned the name cl.
3) The file with the TerraClimate coordinates that are within the Licungo river basin is read.

```
#TerraClimate coordinates from the Licungo basin are read
coor<-read.table("E:/19_MANUAL_data_global/TerraClimate/coordinate_TC_Licungo.txt",header=T)
head(coor)
dim(coor)
```

4) The files needed for the calculations are read, these are the precipitation, ETP and ETr matrices. Their dimensions and composition are checked.

```
#The necessary variables are read
pre<-readRDS("E:/19_MANUAL_data_global/TerraClimate/pre_data_cl.rds")#Precipitation
etr<-readRDS("E:/19_MANUAL_data_global/TerraClimate/etr_data_cl.rds")#ETr
etp<-readRDS("E:/19_MANUAL_data_global/TerraClimate/etp_data_cl.rds")#EtP
#Dimensions are revised
dim(pre)
dim(etr)
dim(etp)
#The compositions of the matrices are reviewed
head(pre[,1:5])
head(etr[,1:5])
head(etp[,1:5])
```

5) The months with deficit or super-habit in each of the cells are calculated by subtracting the ETP matrix from the precipitation matrix. This variable is called "dry".

```
#The deficit or surplus is calculated (Pre-ETP)
dry<-data.frame(pre[,1:3],pre[,4:ncol(pre)]-etp[,4:ncol(etp)])#Dry months with values below zero
```

6) The volume of available water per cell is calculated by subtracting the ETr matrix from the precipitation matrix. This variable is called "available".

```
#The volume of available water is calculated (Pre-ETr)
available<-data.frame(pre[,1:3],pre[,4:ncol(pre)]-etr[,4:ncol(etr)])
```

7) The water deficit per cell is calculated by subtracting the ETr matrix from the ETP matrix. This variable is called "deficit".

```
#The water deficit is calculated(ETr-ETP)
deficit<-data.frame(pre[,1:3],etr[,4:ncol(etr)]-etp[,4:ncol(etp)])#Deficit with values less than zero
```

8) The three resulting matrices are checked for dimensions that must match the original precipitation, ETP and ETr matrices, and saved as .RDS files.

```
#Dimensions are revised
dim(dry)
dim(available)
dim(deficit)
#Matrix compositions are checked
head(dry[,1:5])
head(available[,1:5])
head(deficit[,1:5])

#Arrays are saved as .RDS files.
saveRDS(dry,"E:/19_MANUAL_data_global/Balance_hidrico/dry.rds")
saveRDS(available,"E:/19_MANUAL_data_global/Balance_hidrico/available.rds")
saveRDS(deficit,"E:/19_MANUAL_data_global/Balance_hidrico/deficit.rds")
```

9) With the matrix of dry months, the percentage of time (number of months within the period of analysis) with dry months is calculated. The corresponding coordinates are added to this vector and the matrix is saved in a known folder. As always, the dimension and composition of the variable is checked before saving.

```
#Calculation of percentage of dry months in the whole period

dry_percentage<-data.frame(colSums(dry[,-c(1:3)]<0))#Numbers of months with values less than zero are counted
dry_percentage<-(dry_percentage/(nrow(dry)-3)*100)#The percentage of months with values less than zero is calculated

#A matrix is created with TerraClimate coordinates and if corresponding percentage of months less than zero
dry_percentage2<-data.frame(coor,dry_percentage);colnames(dry_percentage2)<-c("long","lat","percentage_dry")
dim(dry_percentage2)
head(dry_percentage2)

#The matrix is saved
write.table(dry_percentage2,"E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_dry.txt",row.names = F)
```

```
       long       lat porcentaje_secos
X1 36.85417 -15.31250         60.99865
X2 36.81250 -15.35417         60.05398
X3 36.85417 -15.35417         59.51417
X4 36.77083 -15.39583         58.56950
X5 36.81250 -15.39583         57.75978
X6 36.85417 -15.39583         57.48988
```

10) To the available water matrix, filter the data corresponding to each month (January, February, etc.) and average all the values (of 62 years in this case). The 12 vectors are added, one for each month, with their respective coordinates and the matrices are stored in a known folder. The dimension and composition of the variable is checked before saving.

```
#A matrix is created with TerraClimate coordinates and if corresponding percentage of months less than zero
for(i in 1:12){#Varies the loop for each month of the year
  aux<-filter(available,available[,3]==i)#A matrix is filtered with the data for the 62 years of each month
  if(i==1){
    #Values are averaged over 62 years
    vol_average<-data.frame(colMeans(aux[,-c(1:3)]))#Numbers of months with values less than zero are counted
  }
  if(i!=1){
    #The averages for each month of the 62 years are accumulated
    vol_average<-cbind(vol_average,data.frame(colMeans(aux[,-c(1:3)])))
  }
}
vol_average1<-data.frame(coor,vol_average)#The coordinates are added to the monthly averages
#Se nombran las columnas
colnames(vol_average1)<-c("long","lat","ene","feb","mar","abr","may","jun","jul","ago","sep","oct","nov","dic")
head(vol_average1)
dim(vol_average1)

#The matrix containing the averages of available water for each month is stored
write.table(vol_average2,"E:/19_MANUAL_data_global/Balance_hidrico/vol_average_months.txt",row.names = F)
```

```
       long       lat      ene      feb      mar      abr       may       jun       jul       ago sep oct      nov      dic
X1 36.85417 -15.31250 190.4387 185.7774 167.6532 22.90806 0.0983871 0.6306452 0.5419355 0.06451613   0   0 24.75968 134.7500
X2 36.81250 -15.35417 175.4919 186.3355 168.9226 30.36290 0.9467742 1.5209677 1.5483871 0.54838710   0   0 25.17903 126.9323
X3 36.85417 -15.35417 195.1048 194.2581 183.2371 32.70000 0.8048387 1.6322581 1.9435484 0.51612903   0   0 30.09032 141.5823
X4 36.77083 -15.39583 186.3113 191.6839 181.1032 36.21613 3.0032258 2.2129032 2.4032258 0.85483871   0   0 27.70323 133.5048
X5 36.81250 -15.39583 181.7935 192.6516 186.8871 41.50968 3.6370968 2.7483871 2.5161290 1.14516129   0   0 28.23226 130.1839
X6 36.85417 -15.39583 184.8161 196.0065 196.5194 44.15968 2.6532258 3.2370968 2.4612903 1.50000000   0   0 30.17581 131.2984
> |
```

11) The deficit matrix is transformed from mm/month to km3. For this procedure, an area per cell is calculated, using the cell size to which each pixel corresponds, in the case of TerraClimate it is ≈ 21.43 km^2. Subsequently, the average of all cells is calculated and a vector in km^3 of the average deficit in each of the monthly time steps is obtained. The corresponding coordinates are added to this vector and the final matrix is stored. The dimension and composition of the variable is checked before saving.

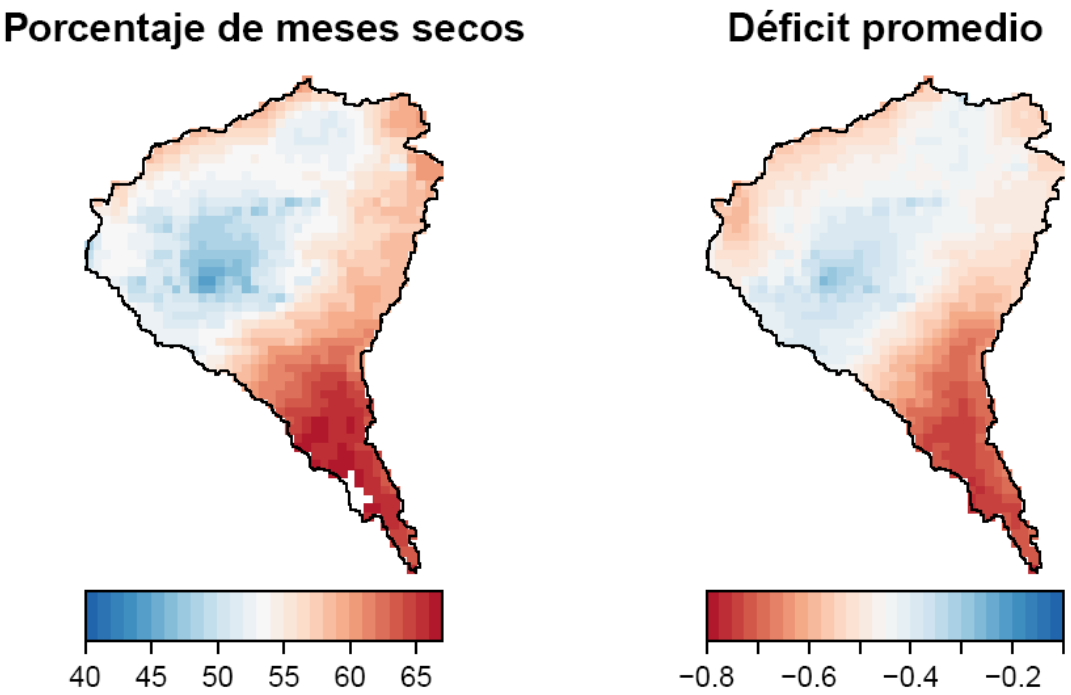
```
#A constant is defined which transforms the mm of cell to km3
area_celda_km3<-((c((111.1/24)^2)*1000000)/1000000000)

#The deficit matrix is transformed to km3
deficit_km3<-data.frame(deficit[,1:3],data.frame(deficit[,4:ncol(deficit)])*area_celda_hm3)
dim(deficit_km3)
head(deficit_km3[,1:5])

#Calculation of average deficit in km3

stress_percentage<-data.frame(colMeans(deficit_km3[,-c(1:3)]))

#A matrix is created with TerraClimate coordinates and if corresponding percentage of months less than zero
stress_percentage2<-data.frame(coor,stress_percentage);colnames(stress_percentage2)<-c("long","lat","stress_percentage")
dim(stress_percentage2)
head(stress_percentage2)

#The matrix is saved
write.table(stress_percentage2,"E:/19_MANUAL_data_global/Balance_hidrico/stress_monthly_percentage2.txt",row.names = F)
```

	long	lat	porcentaje_estres
X1	36.85417	-15.31250	-0.6322185
X2	36.81250	-15.35417	-0.6007084
X3	36.85417	-15.35417	-0.5488060
X4	36.77083	-15.39583	-0.5469050
X5	36.81250	-15.39583	-0.5418358
X6	36.85417	-15.39583	-0.5346351

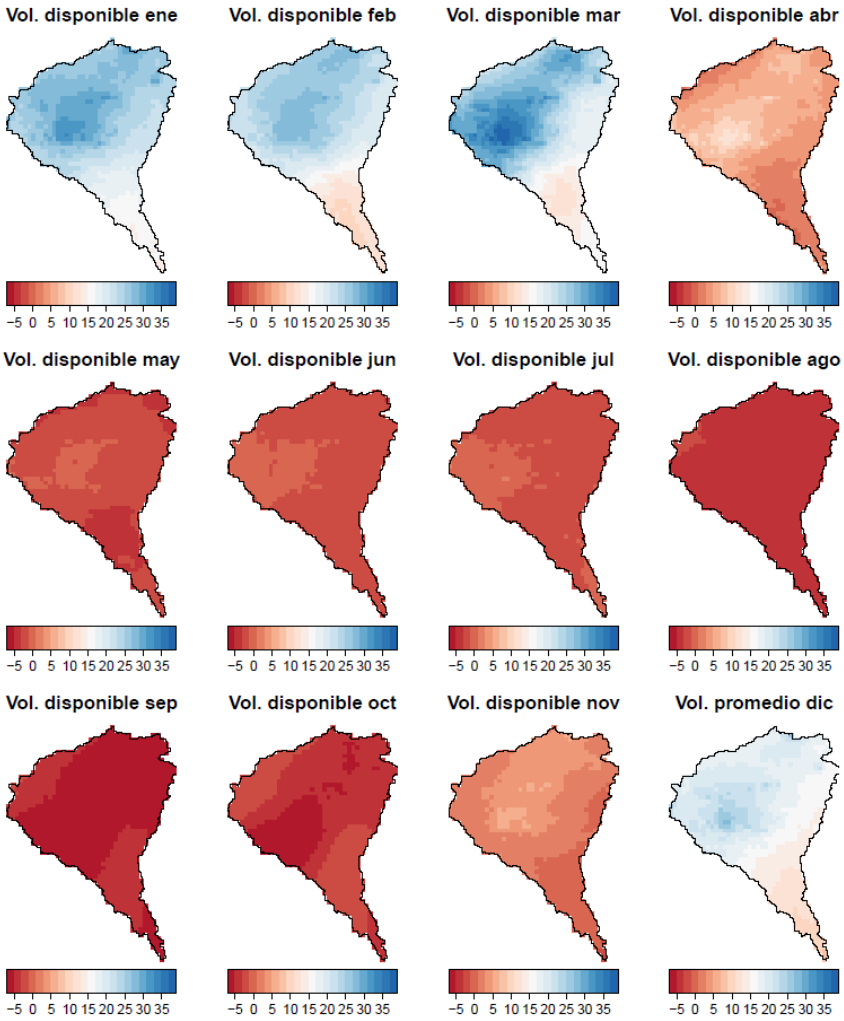
12) For graphing, the shapefile of the Licungo river basin and the matrices of percentage of dry months, available water, and deficit in km^3 are read.

```
#The necessary geographical information is read from the folders where they are stored
cl<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/basin_Licungo.shp")#The Mozambique shapefile is read
plot(cl)

#The percentage of dry months in the Licungo river basin is read as follows
dry_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_dry.txt",header=T)
dim(dry_percentage2)
head(dry_percentage2)

#The average volume per month in the Licungo river basin is read as follows
vol_promedio2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/volumen_promedio_months.txt",header=T)
dim(vol_promedio2)
head(vol_promedio2)

#The percentage of deficit in km3 in the Licungo river basin is read as follows
stress_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_stress.txt",header=T)
dim(stress_percentage2)
head(stress_percentage2)
```

13) The plotting steps used previously (in the example of step 16 in section 10) are followed to plot the dry months percentage and deficit km3 matrices. The result should look something like this:



14) Following the steps of the graphs defined for the VHI (in the example of step 15 of section 11), the average monthly volumes in km3 for each month in the Licungo river basin are plotted. The result should look something like this:



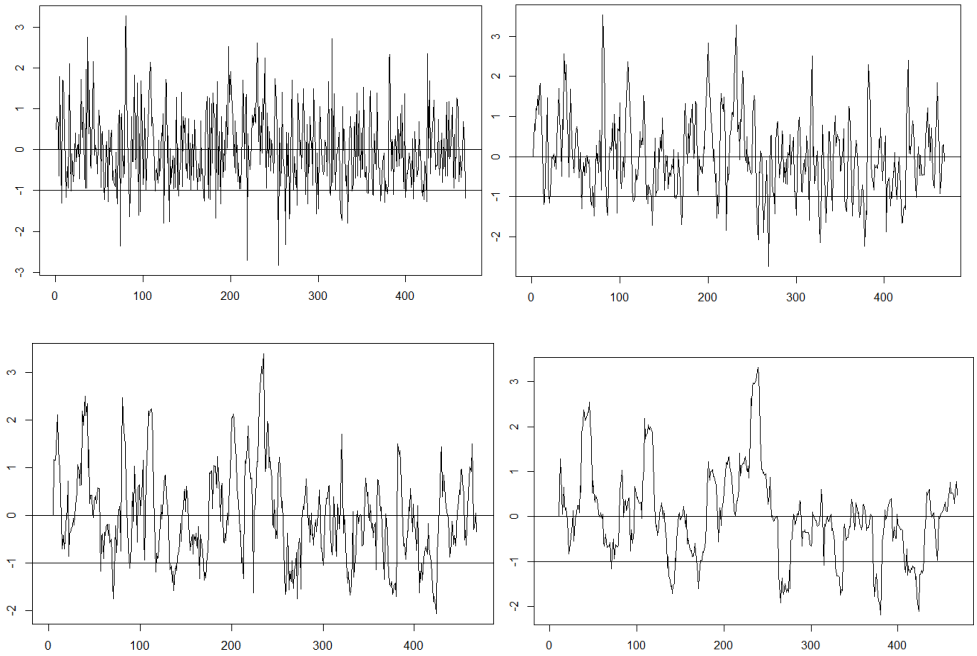
15. Calculation and application of SPI and SPEI drought indices

In this part, the calculation of the SPI and SPEI drought indices is explained step by step. The CHIRPS precipitation data obtained over Mozambique will be used for SPI, and the precipitation minus ETP (dry and wet months) data from TerraClimate over the Licungo river basin will be used for SPEI. The process for computing a single time series in each case is shown first, and then for the entire matrix of pixels contained in the areas of interest. This section ends by showing two common and very useful applications of the indices. The first is the calculation of percentage of area affected by drought at each time step and the second, spatial plots where drought/non-drought states can be shown for a specific period.

1) The necessary packages are installed and loaded. The main package is "SPEI", in which the necessary functions for the direct computation of SPI and SPEI have been programmed.
2) The shapefiles of the Licungo and Mozambique river basin are read, they have been assigned with the name cl and mz respectively.
3) The coordinates of the CHIRPS data located within Mozambique and of the TerraClimate data corresponding to the Licungo river basin are read.
4) For a random vector of the CHIRPS precipitation matrix, understood as a monthly time series at a point within Mozambique, the SPI is calculated. First, a vector containing the time scales to be calculated is defined; here, as an example, the scales of 1, 3, 6 and 12 months have been considered. The result will be a vector of SPI-1, -3, -6 and -12 months corresponding to the

original precipitation vector. The 4 vectors are plotted to compare the difference, and the threshold of 0 and -1 (moderate drought) is delimited.

```
#SPI is calculated for a random cell (35)
vectorSPISPEI<-matrix(c(1,3,6,12)) #Vector with the required time scales to be calculated
SPI <- matrix(nrow=nrow(pre_chi),ncol = nrow(vectorSPISPEI))#Empty matrix to be filled with SPIs
for (j in vectorSPISPEI){   #4 SPIs will be calculated depending on the predefined vector
  spii<- spi(pre_chi[,35],j)#The function "spi" is applied and the input variables are a precipitation
  #vector and the time scale at which it is required to be calculated
  tr <- as.numeric(fitted(spii))#The spi data is extracted from the class spi
  tr[is.na(tr)] <- 0#NAs that can be generated are changed to zeros
  SPI[,which(vectorSPISPEI==j)] <- tr #The spi vector is stored in the SPI array
  colnames(SPI)[vectorSPISPEI==j] <- j #Each column is named with the number of the time scale
}
#SPI is reviewed results
head(SPI)
dim(SPI)

#It is graphed to show the difference
plot(c(1:nrow(SPI)),SPI[,1],"l")#SPI1
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,2],"l")#SPI3
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,3],"l")#SPI6
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,4],"l")#SPI12
abline(h=c(0,-1))#Crosses the threshold 0 y -1
```



5) For a random vector of the matrix of dry and wet months, understood as a monthly time series at a point within the Licungo river basin, the SPEI is calculated. As for the SPI, a vector containing the time scales to be calculated is first defined, here the scales of 1, 3, 6 and 12 months have been proposed as an example. The result will be a vector of SPEIs at scales of -1, -3, -6 and -12 months corresponding to the vector of the original D deficit. The 4 vectors are plotted (not shown) to compare the difference, and the threshold of 0 and -1 (moderate drought) is delimited.

```
#SPEI is calculated for a random cell (35)
vectorSPISPEI<-matrix(c(1,3,6,12))#Vector with the required time scales to be calculated
SPEI <- matrix(nrow=nrow(dry),ncol = nrow(vectorSPISPEI))#Empty matrix to be filled with SPEIs
for (j in vectorSPISPEI){    #4 SPEIs are to be calculated depending on the predefined vector
  spii<- spei(dry[,35],j)#The function "spei" is applied and the input variables are a deficit
  #vector (D) and the time scale at which the calculation is required
  tr <- as.numeric(fitted(spii))#spei data is extracted from class spi
  tr[is.na(tr)] <- 0            #NAs that can be generated are changed to zeros
  SPEI[,which(vectorSPISPEI==j)] <- tr#The spi vector is stored in the SPEI matrix
  colnames(SPEI)[vectorSPISPEI==j] <- j #Each column is named with the number of the time scale
}
head(SPEI)
dim(SPEI)

#SPEIs are graphed to show the difference
plot(c(1:nrow(SPEI)),SPEI[,1],"l")#SPEI1
abline(h=c(0,-1))#crosses the threshold 0 y -1
plot(c(1:nrow(SPEI)),SPEI[,2],"l")#SPEI3
abline(h=c(0,-1))#crosses the threshold 0 y -1
plot(c(1:nrow(SPEI)),SPEI[,3],"l")#SPEI6
abline(h=c(0,-1))#crosses the threshold 0 y -1
plot(c(1:nrow(SPEI)),SPEI[,4],"l")#SPEI12
abline(h=c(0,-1))#crosses the threshold 0 y -1
```

6) Using a loop that varies for all columns of the precipitation matrix (all time series of all coordinates), the same SPI calculation procedure is performed for each of them. The resulting matrices when following the code will be 4 (in this case), one for each time scale that the index has been calculated. Following the code will generate in a relatively short time (variable to the number of coordinates) matrices that should be of the same dimension as the original CHIRPS precipitation matrix. The matrices are saved as .RDS files in destination folders.

```
#SPI is calculated for the whole matrix containing CHIRPS
vectorSPISPEI<-matrix(c(1,3,6,12))#Vector with the required time scales to be calculated
for(i in 4:ncol(pre_chi)){
  SPI <- matrix(nrow=nrow(pre_chi),ncol = nrow(vectorSPISPEI))
  for (j in vectorSPISPEI){
    spii<- spi(pre_chi[,i],j)
    tr <- as.numeric(fitted(spii))
    tr[is.na(tr)] <- 0
    SPI[,which(vectorSPISPEI==j)] <- tr
    colnames(SPI)[vectorSPISPEI==j] <- j
  }
  if (i==4){
    SPI1 <- data.frame(pre_chi[,1:3],SPI[,1])
    SPI3 <- data.frame(pre_chi[,1:3],SPI[,2])
    SPI6 <- data.frame(pre_chi[,1:3],SPI[,3])
    SPI12 <- data.frame(pre_chi[,1:3],SPI[,4])
  }
  if (i!=4){
    SPI1 <- data.frame(SPI1,data.frame(SPI[,1]))
    SPI3 <- data.frame(SPI3,data.frame(SPI[,2]))
    SPI6 <- data.frame(SPI6,data.frame(SPI[,3]))
    SPI12 <- data.frame(SPI12,data.frame(SPI[,4]))
  }
  print(paste(i,"de",ncol(pre_chi)))
}
colnames(SPI1)<-c("dates","year","month",c(1:(ncol(pre_chi)-3)))#The columns are named
colnames(SPI3)<-c("dates","year","month",c(1:(ncol(pre_chi)-3)))
colnames(SPI6)<-c("dates","year","month",c(1:(ncol(pre_chi)-3)))
colnamonth(SPI12)<-c("dates","year","month",c(1:(ncol(pre_chi)-3)))
dim(SPI1)#Dimensions are checked, they must be equal to the original precipitation matrices
dim(SPI3)
dim(SPI6)
dim(SPI12)
head(SPI1[,1:50])#The composition of the matrices is reviewed
head(SPI3[,1:50])
head(SPI6[,1:5])
head(SPI12[,1:5])

#The matrices corresponding to the SPIs are stored in a folder known as .RDS
saveRDS(SPI1,"E:/19_MANUAL_data_global/SPI_SPEI/SPI1_mz.rds")
saveRDS(SPI3,"E:/19_MANUAL_data_global/SPI_SPEI/SPI3_mz.rds")
saveRDS(SPI6,"E:/19_MANUAL_data_global/SPI_SPEI/SPI6_mz.rds")
saveRDS(SPI12,"E:/19_MANUAL_data_global/SPI_SPEI/SPI12_mz.rds")
```

7) By means of a loop that varies for all the columns of the matrix of dry and wet months deficits, (D), the same SPEI calculation procedure is performed for each of them. The resulting matrices when following the code will be 4 (in this case), one for each time scale that the index has been calculated. Following the code will generate in a relatively short time (variable to the number of coordinates) matrices that should be of equal dimension to the deficit matrix. The matrices are saved as .RDS files in destination folders.

```
#SPEI is calculated for the whole matrix containing
#Precipitation minus ETP data in the Licungo river basin
vectorSPISPEI<-matrix(c(1,3,6,12))#Vector with the required time scales to be calculated
for(i in 4:ncol(dry)){
  SPEI <- matrix(nrow=nrow(dry),ncol = nrow(vectorSPISPEI))
  for (j in vectorSPISPEI){
    spii<- spei(dry[,i],j)
    tr <- as.numeric(fitted(spii))
    tr[is.na(tr)] <- 0
    SPEI[,which(vectorSPISPEI==j)] <- tr
    colnames(SPEI)[vectorSPISPEI==j] <- j
  }
  if (i==4){
    SPEI1 <- data.frame(dry[,1:3],SPEI[,1])
    SPEI3 <- data.frame(dry[,1:3],SPEI[,2])
    SPEI6 <- data.frame(dry[,1:3],SPEI[,3])
    SPEI12 <- data.frame(dry[,1:3],SPEI[,4])
  }
  if (i!=4){
    SPEI1 <- data.frame(SPEI1,data.frame(SPEI[,1]))
    SPEI3 <- data.frame(SPEI3,data.frame(SPEI[,2]))
    SPEI6 <- data.frame(SPEI6,data.frame(SPEI[,3]))
    SPEI12 <- data.frame(SPEI12,data.frame(SPEI[,4]))
  }
  print(paste(i,"de",ncol(dry)))
}
colnames(SPEI1)<-c("dates","year","month",c(1:(ncol(dry)-3)))#The columns are named
colnames(SPEI3)<-c("dates","year","month",c(1:(ncol(dry)-3)))
colnames(SPEI6)<-c("dates","year","month",c(1:(ncol(dry)-3)))
colnames(SPEI12)<-c("dates","year","month",c(1:(ncol(dry)-3)))
dim(SPEI1)#Dimensions are checked, they must be equal to the original D matrices
dim(SPEI3)
dim(SPEI6)
dim(SPEI12)
head(SPEI1[,1:5])#The composition of the matrices is reviewed
head(SPEI3[,1:5])
head(SPEI6[,1:5])
head(SPEI12[,1:5])

#The matrices corresponding to the SPIs are stored in a folder known as .RDS
saveRDS(SPEI1,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI1_cl.rds")
saveRDS(SPEI3,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI3_cl.rds")
saveRDS(SPEI6,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI6_cl.rds")
saveRDS(SPEI12,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI12_cl.rds")
```

8) In this step the percentage area affected by drought is calculated. This is to quantify for each time step (each month) the number of pixels that are in drought and calculate their percentage against the total. A threshold is required according to the intensity of drought to be studied. In the example, moderate drought (values less than -1), severe drought (values less than -1.5) and severe drought (values less than -2) are used as threshold values. The final composition of the matrix that is saved is shown here. The SPI at a time scale of 6 months over Mozambique is used as an example.

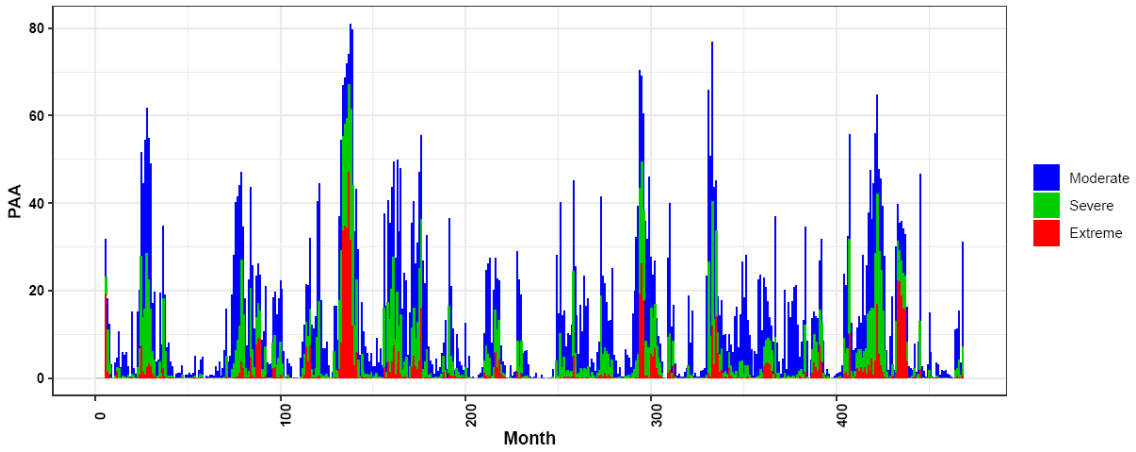
```
#Mozambique SPI 6 is read based on CHIRPS data
spi6<-readRDS("E:/19_MANUAL_data_global/SPI_SPEI/SPI6_mz.rds")
dim(spi6)
head(spi6[,1:5])

#The percentage of area affected under moderate drought is calculated
percentage_1<-data.frame(rowSums(spi6[,-c(1:3)]<(-1)))#The number of cells with values less than
#-1 is quantified, where -1 is the threshold
percentage_1<-(percentage_1/(ncol(spi6)-3)*100)#The percentage of cells under the threshold is defined
dim(percentage_1)
head(percentage_1)

#The percentage of affected area under severe drought is calculated
percentage_2<-data.frame(rowSums(spi6[,-c(1:3)]<(-1.5)))#The number of cells with values less than
#-1.5 is quantified, where -1.5 is the threshold
percentage_2<-(percentage_2/(ncol(spi6)-3)*100)#The percentage of cells under the threshold is defined
dim(percentage_2)
head(percentage_2)

#The percentage of affected area under extreme drought is calculated
percentage_3<-data.frame(rowSums(spi6[,-c(1:3)]<(-2)))#The number of cells with values less than
#-2 is quantified, where -2 is the threshold
percentage_3<-(percentage_3/(ncol(spi6)-3)*100)#The percentage of cells under the threshold is defined
dim(percentage_3)
head(percentage_3)

#The different percentages for each drought intensity are joined with their respective dates.
percentages<-data.frame(spi6[,1:3],percentage_1,percentage_2,percentage_3)
colnames(percentages)<-c("dates","year","month","seq_moderate","seq_severe","seq_extreme")
dim(percentages)
head(percentages)

#The calculated percentages are saved in a known folder
write.table(percentages,"E:/19_MANUAL_data_global/SPI_SPEI/percentages.txt",row.names = F)
```

```
> dim(porcentajes)
[1] 468   6
> head(porcentajes)
      Fechas  ano mes seq_moderada seq_severa seq_extrema
1 1981-01-01 1981   1      0.00000    0.00000     0.00000
2 1981-01-01 1981   2      0.00000    0.00000     0.00000
3 1981-01-01 1981   3      0.00000    0.00000     0.00000
4 1981-01-01 1981   4      0.00000    0.00000     0.00000
5 1981-01-01 1981   5      0.00000    0.00000     0.00000
6 1981-01-01 1981   6     31.73664   23.03112    18.97052
```

9) The percentages are plotted as a time series following the code in ANNEX III. In this case the function "ggplot" is used. It is necessary to read the percentages of area affected by droughts, which are defined here as PAA. In case it is desired to display the time series without shading, step 4 can be applied..

```
#The necessary geographical information is read from the folders where they are stored
cl<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/basin_Licungo.shp")#Read the shapefile of the Licunco basin
mz<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/mozambique.shp")#The Mozambique shapefile is read
plot(mz)
plot(cl,add=T)

#Drought percentages are read for the different intensities
percentages<-read.table("E:/19_MANUAL_data_global/SPI_SPEI/percentages.txt",header=T)
dim(percentages)
head(percentages)

#The limits to be plotted are calculated
spi1<-data.frame(c(1:nrow(percentages)),(percentages[,4]-percentages[,5]),
                 (percentages[,5]-percentages[,6]),percentages[,c(6)])#Calculation
colnames(spi1)<-c("date","Moderate","Severe","Extreme")

#The created time series are plotted
aux1<-melt(data = spi1, id.vars = "date",value.name = "por")#"date" must be equal to the column name of spi1
                                                            #"por" can be changed
head(aux1)
spi<-ggplot(aux1, aes(fill=variable, y=por, x=date)) +#"por" must be equal to that defined above
  geom_bar(position="stack", stat="identity")+
  scale_fill_manual(values=c(Moderate = "blue", Severe = "green3",Extreme="red"))+#Colours are defined for each column name to be plotted
  labs(x = "Mes", y = "PAA",title="SPI-6")+
  theme_bw()+
  theme(plot.title = element_blank(),legend.title = element_blank(),axis.text = element_text(face="bold",colour = "black"),
        axis.title = element_text(face="bold",colour = "black"),axis.text.x = element_text(angle = 90))


#The graphic is exported as pdf, the destination folder and the size can be modified
pdf("E:/19_MANUAL_data_global/Graphics/percentage_area_afectada.pdf",height = 4,width = 10)
spi
dev.off() #Export closed
```

The resulting graphic will look something like this:



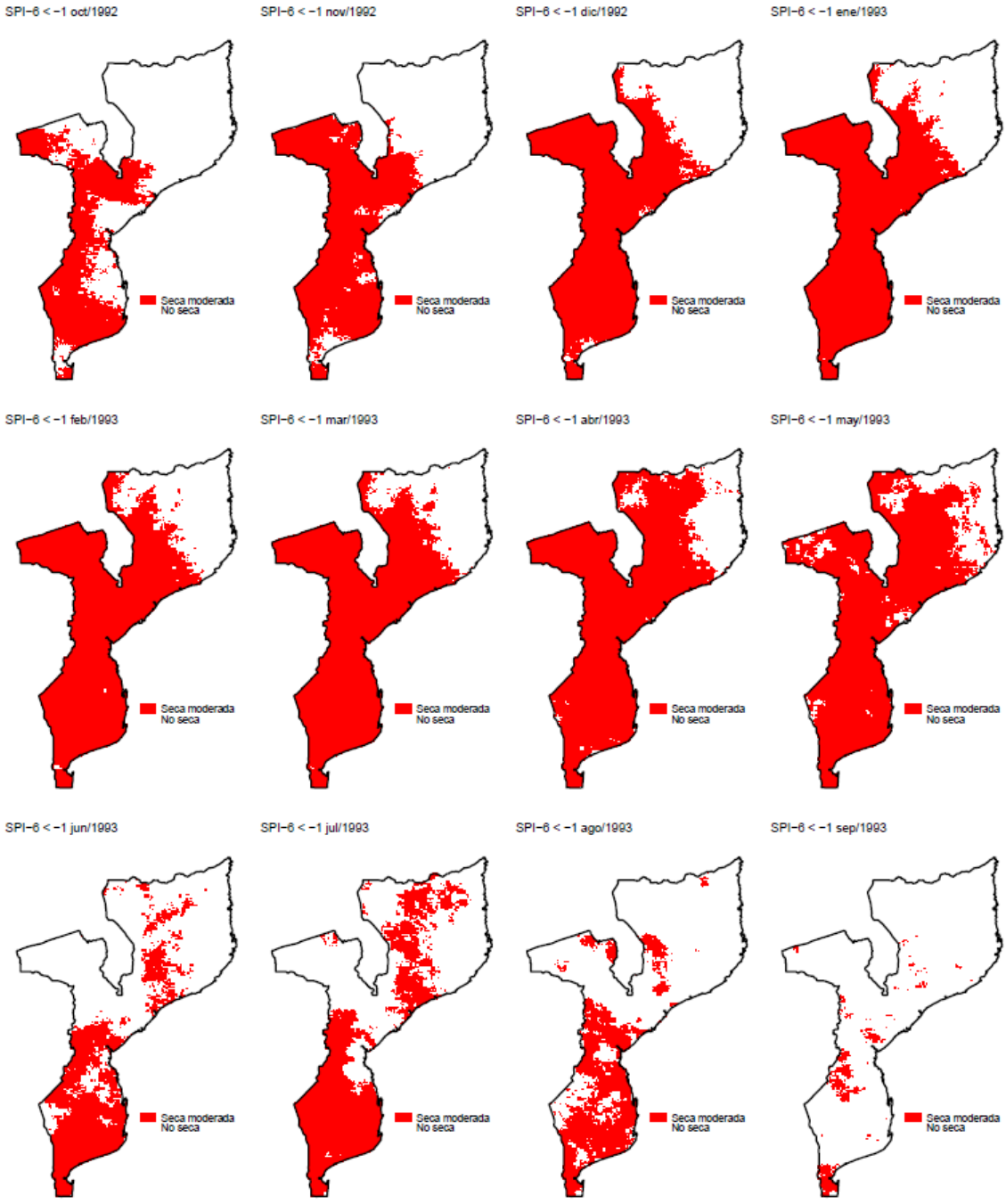
10) In this step, the drought or non-drought status over Mozambique is calculated. This is to transform all values below a drought threshold as 1 and those above the threshold as 0. The SPI is used as an example for a time scale of 6 months over Mozambique and for a moderate drought intensity, i.e. a threshold SPI value of less than -1.

```
#Mozambique SPI 6 is read using CHIRPS data
spi6<-readRDS("E:/19_MANUAL_data_global/SPI_SPEI/SPI6_mz.rds")
dim(spi6)
head(spi6[,1:5])

#CHIRPS coordinates are read inside Mozambique
coor_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordinate_CHIRPS.txt",header=T)
head(coor_chi)
dim(coor_chi)

#Transforms to 0 where there are values with moderate drought
#(values less than -1) and values greater than -1 as 1
spi6_a<-spi6[,-c(1:3)]#An auxiliary matrix is defined
spi6_a[spi6_a>(-1)]<-1#SAll values greater than -1 (drought threshold) are converted to 1.
spi6_a[spi6_a<=(-1)]<-0#All values less than or equal to -1 (drought threshold) are transformed to 1
dim(spi6_a)
head(spi6_a[,1:8])

#Dates are added to the matrix with zeros and ones
spi6_b<-data.frame(spi6[,c(1:3)],spi6_a)
dim(spi6_b)
head(spi6_b[,1:8])

#The matrices are stored with this information of zeros (drought) and 1 (no drought)

saveRDS(spi6_b,"E:/19_MANUAL_data_global/SPI_SPEI/seq_no_seq.rds")
```

11) Previously used graphing steps (the example of step 16 in section 10) are followed and adapted to plot drought or non-drought states. The objective is to show for a series of months the areas where droughts occur. As an example, the period from October 1992 to September 1993 is taken, when there was a major drought in Mozambique. The "ggplot" function is used to create graphs following the code, where titles, colours and other characteristics can be modified. Unlike the functions that have been used for other graphs (such as the use of the "levelplot" function), here it is necessary to transform the geographical information (mz shapefile) into a data matrix compatible with the "ggplot" method; this is done using the "fortify" function.

```
#The necessary geographical information is read from the folders where they are stored
mz<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/mozambique.shp")#The Mozambique shapefile is read
plot(mz)
plot(cl,add=T)
cd<-fortify(mz)

#CHIRPS coordinates are read inside Mozambique
coor_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordinates_CHIRPS.txt",header=T)
head(coor_chi)
dim(coor_chi)

#The matrix is read with information of zeros (drought) and 1 (no drought)
spi6_b<-readRDS("E:/19_MANUAL_data_global/SPI_SPEI/seq_no_seq.rds")
dim(spi6_b)
head(spi6_b[,1:5])

#The information is prepared for graphing
dat_graphic<-data.frame(coor_chi,t(spi6_b[132:143,-c(1:3)]))#A matrix is created with the months corresponding to rows
                                                            #98 to 109 corresponding to the dates to be displayed
colnames(dat_graphic)<-c("long","lat","a","b","c","d","e","f","g","h","i","j","k","l")
dim(dat_graphic)
head(dat_graphic)

#The states are plotted for each date by varying the name of the column
graphic_1<-ggplot(dat_graphic)+
  geom_tile(aes(x = long, y = lat, fill=factor(dat_graphic$a)))+#The name of the column is changed here
  theme_classic()+
  theme(legend.position = c(0.75, 0.25),legend.box="horizontal",legend.title = element_blank(),
        legend.text=element_text(size = rel(1)),axis.title.x=element_blank(),axis.title.y=element_blank(),axis.line=element_blank(),
        axis.ticks=element_blank(),axis.text=element_blank(),legend.key.width = unit(0.6, "cm"),
        legend.key.size = unit(0.2, "cm"))+
  labs(title="SPI-6 < -1 oct/1992")+#Se cambian los titulos
  geom_polygon(data = cd, aes(x=long, y = lat,z=c(rep(1,nrow(cd)))), fill = NA, color = "black",linetype = "solid",size=1)+
  scale_fill_manual(values=c("1"="white","0"="red"),labels=c("1"="No seca","0"="Seca Moderate"))+#Legends and colours are changed
  coord_fixed()
```

A graph is created for each month and saved as a PDF (or other file type). The result will look something like this:



16. Drought characterisation using SPI

This section establishes a procedure for calculating the average characteristics of droughts. These characteristics are duration, maximum intensity and severity; the number of events occurring and the probability of having a drought month are also calculated. Drought events are defined when at least two consecutive months were below the defined intensity threshold. The procedure is performed for an SPI time series (SPI-6 in Mozambique as an example) from which the averages of each characteristic within the time series are extracted. The thresholds can be changed, and the code presented is suitable for any other drought index that is standardised (mean 0 and standard deviation 1).

1) The necessary packages are installed and loaded.

2) The matrix containing the time series of the indicator to be characterised is read. In this case the Mozambique SPI-6 computed with CHIRPS precipitation.

```
#Mozambique SPI 6 is read based on CHIRPS data
spi6<-readRDS("E:/19_MANUAL_data_global/SPI_SPEI/SPI6_mz.rds")
dim(spi6)
head(spi6[,1:5])
```

3) The coordinates of the CHIRPS data located within Mozambique are read.

```
#CHIRPS coordinates are read inside Mozambique
coor_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordinate_CHIRPS.txt",header=T)
head(coor_chi)
dim(coor_chi)
```

4) The loop that calculates the characteristics of all the SPIs in the target matrix has very few variables to modify if required. The threshold is defined prior to executing the loop, it should be noted that the loop only identifies dry not wet season characteristics, so the values must be negative. The start of the loop follows the order from left to right, where the first column is the number from which the auxiliary variable "l" varies, this value must be modified if the target matrix starts in another column. Also the starting number must be changed at the end of the code following the instructions shown in the code. If it is necessary to extract the characteristics of a single time series or a limited number, the "ncol(spi6)" indicating the number of columns of the matrix called "spi6" must also be changed to the position of the last column from which the characteristics are to be obtained.

```
#The threshold is defined
umbral<-(-1) #Threshold
for (l in 4:ncol(spi6)){#l is the sequence of the SPI6 of each cell, each column where 4 is the first SPI column
  y<-data.matrix(spi6[,l])#Se lee el vector de acuerdo al orden del que se extraen las características
  if (nrow(data.frame(which(y<=umbral)))==0){#The vector is read according to the order from which the characteristics are extracted
    ini<-data.frame()
    fin<-data.frame()
  }
  if (nrow(data.frame(which(y<=umbral)))!=0){#The onset and end of drought events are calculated
```

… (refer to the entire loop in the code provided in the annexes)

```
      int<-rbind(int,data.frame(min(aux)))
      sev<-rbind(sev,data.frame(sum(aux)))
    }
  }#End if there are drought events

  #Averaging of events with duration >=2 months
  casi1<-data.frame(du,int,sev,freq,prob)
  if(nrow(data.frame(which(casi1[,1]==1))) != 0){
    casi1<-casi1[-which(casi1[,1]==1),]
  }

  if(nrow(casi1)==0){
    casi1<-data.frame(0,0,0,0,0)
  }
  if (l==4){#Change the 4 when you change the start of the initial loop
    dat<-t(data.frame(colMeans(casi1)))
  }
  if (l!=4){#Change the 4 when you change the start of the initial loop
    dat<-rbind(dat,t(data.frame(colMeans(casi1))))
  }
  print(paste(l,"de",ncol(spi6)))#Step counter, one for each column
} #fin de l
colnames(dat)<-c("Duration","Intensity","Severity","number","probability")
dat #Dat is the matrix containing the average characteristics of all SPI series
head(dat[,1:5])
dim(dat)#Dimension equal to the number of coordinates
```

The matrix "dat" is the result of the loop and contains the average characteristics of all drought events of each SPI time series. Its dimension has a number of rows equal to the number of CHIRPS coordinates and 5 columns corresponding to each characteristic:

```
> head(dat[,1:5])
                Duracion Intensidad Severidad numero probabilidad
colMeans.casi1. 7.529412  -1.705905 -7.935980     19     1.608849
colMeans.casi1. 7.470588  -1.754820 -8.184736     18     1.596280
colMeans.casi1. 9.384615  -1.743732 -9.567796     14     2.005260
colMeans.casi1. 7.750000  -1.734323 -8.028418     18     1.655983
colMeans.casi1. 7.687500  -1.735632 -8.060110     18     1.642628
colMeans.casi1. 7.176471  -1.737271 -8.029506     18     1.533434
>
```

The corresponding coordinates are added to this matrix and stored in a known destination folder.

```
#Se adhieren las coordenadas respectivas a sus características
carac<-data.frame(coor_chi,dat)
head(carac)
dim(carac)

#Se guarda la matriz en una carpeta conocida
write.table(carac,"E:/19_MANUAL_datos_globales/caracterizacion/carac_spi.txt",col.names = T,row.names = F)
```

5) The procedure for plotting these average characteristics is similar to the one used for plotting drought and non-drought states in section 1. The geographical information for Mozambique and the matrix containing the drought characteristics according to SPI6 (in this case) and their coordinates are read. Only the durations and probabilities are plotted as an example.

```
#The respective coordinates are attached to their characteristics
carac<-data.frame(coor_chi,dat)
head(carac)
dim(carac)

#The matrix is saved in a known folder
write.table(carac,"E:/19_MANUAL_data_global/characterisation/carac_spi.txt",col.names = T,row.names = F)
```

6) The vector of the feature to be plotted is extracted, where the number inside the square brackets is the position of the column containing the information to plot.

```
#The duration are extracted
duration<-data.frame(carac[,3])
```

7) This vector is transformed into sections that are within its maximum and minimum limits. The number assigned to it must be within the existing limits.

```
#The durations are extracted
duration<-data.frame(carac[,3])

#Maximums and minimums are visible
max(carac[,3])
min(carac[,3])

#They are transformed into a number of options
duration[duration>=13]<-14
duration[duration<13 & duration>=11]<-12
duration[duration<11 & duration>=9]<-10
duration[duration<9 & duration>=7]<-8
duration[duration<7 & duration>=5]<-6
duration[duration<=5]<-4
```

8) The results are plotted using the "ggplot" function. The colours, the name of the section (previously defined), among others, can be modified. It is important to define the values of "x" and "y" as the longitude and latitude coordinates contained in the coordinate matrix. The graph is saved with a name for later export.

```
#The durations are plotted by varying the limits and colours
graphic_1<-ggplot(duration)+
  geom_tile(aes(x = carac[,1], y = carac[,2], fill=factor(duration[,1])))+#Here you change the name of the column
  theme_classic()+
  theme(legend.position = c(0.75, 0.25),legend.box="horizontal",legend.title = element_blank(),
        legend.text=element_text(size = rel(1)),axis.title.x=element_blank(),axis.title.y=element_blank(),axis.line=element_blank(),
        axis.ticks=element_blank(),axis.text=element_blank(),legend.key.width = unit(0.6, "cm"),
        legend.key.size = unit(0.2, "cm"))+
  labs(title="SPI-6 < -1 Duration")+#Titles are changed
  geom_polygon(data = cd, aes(x=long, y = lat,z=c(rep(1,nrow(cd)))), fill = NA, color = "black",linetype = "solid",size=1)+
  scale_fill_brewer(palette = "YlOrRd",labels =c("<5", "5-7", "7-9","9-11","11-13",">13"))+#Limits and colours are changed
                                                                                          #and the number of names must be
                                                                                          #equal to the limits
  coord_fixed()
```

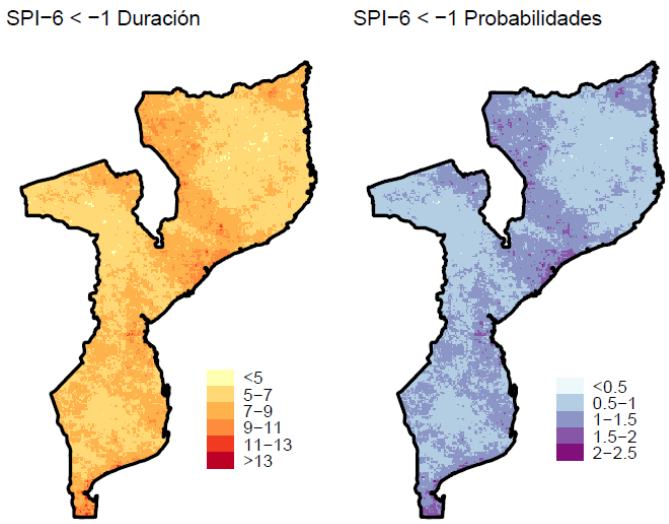
9) The above process is repeated with the probability matrix.

```
#Probabilities are extracted
probabilities<-data.frame(carac[,7])

#The maximum and minimum values are analysed
max(probabilities)
min(probabilities)

#The probabilities are transformed into a given number of options
probabilities[probabilities>=2.5]<-3
probabilities[probabilities<2.5 & probabilities>=2]<-2.25
probabilities[probabilities<2 & probabilities>=1.5]<-1.75
probabilities[probabilities<1.5 & probabilities>=1]<-1.25
probabilities[probabilities<1 & probabilities>=0.5]<-0.75
probabilities[probabilities<=0.5]<-0.4

graphic_2<-ggplot(probabilities)+
  geom_tile(aes(x = carac[,1], y = carac[,2], fill=factor(probabilities[,1])))+#Here you change the name of the column
  theme_classic()+
  theme(legend.position = c(0.75, 0.25),legend.box="horizontal",legend.title = element_blank(),
        legend.text=element_text(size = rel(1)),axis.title.x=element_blank(),axis.title.y=element_blank(),axis.line=element_blank(),
        axis.ticks=element_blank(),axis.text=element_blank(),legend.key.width = unit(0.6, "cm"),
        legend.key.size = unit(0.2, "cm"))+
  labs(title="SPI-6 < -1 Probabilities")+#Titles are changed
  geom_polygon(data = cd, aes(x=long, y = lat,z=c(rep(1,nrow(cd)))), fill = NA, color = "black",linetype = "solid",size=1)+
  scale_fill_brewer(palette = "BuPu",labels =c("<0.5", "0.5-1", "1-1.5","1.5-2","2-2.5",">2.5"))+#Boundaries and colours are changed
  coord_fixed()
```

10) Graphics are exported as PDF.

```
#The graphic is exported as pdf, the destination folder and the size can be modified
pdf("E:/19_MANUAL_data_global/Graphics/characteristics.pdf",height = 8,width = 6)
grid.arrange(graphico_1,graphic_2,
             nrow = 1) #The 2 graphs are arranged in a 1 x 2 matrix, can be changed
dev.off() #Export closed
```

The result will look like this:



Conclusions

This document is an introductory manual for obtaining hydrometeorological variables in any region and applying them to drought management. Droughts are an extreme hydrological phenomenon, naturally occurring in any climatic regime. Their impacts on the environment cause immense economic and ecological losses every year, so their characterisation and monitoring has become increasingly important. However, the characterisation of droughts is a complex process, because in addition to affecting different parts of the water cycle, there is no clear indication of the onset or end of droughts. Added to this is the problem that the tools required to model the phenomenon, such as hydrological and climatic data (precipitation, temperature, etc.), do not always meet the required quality.

This work collects a series of drought management concepts widely accepted among the scientific and technical community and applies them through several scripts (programming codes) programmed in R software. These scripts divided into 3 sections show the step-by-step for i) downloading, extracting, manipulating and saving hydrometeorological variables from global databases; ii) using and applying these variables for drought management; and iii) presenting results through graphs in different formats.

By following the step-by-step tools provided in this manual, much useful information can be generated for any region of the world. The results that can be obtained may be of interest to managers and technicians who do not always have easy access to this type of information. A great advantage of this manual is the flexibility of its tools, which can be used in different places and allow to create new scripts and generate new ways of using the information downloaded by the users. The methodologies discussed in this document can be extended in future work, such as obtaining data from climate change projection models, trend analysis, representation of results for regions within a country (cities or provinces) or river basins (sub-basins), among others.

ANNEXES: Scripts in R and RStudio for drought management

ANNEX I: Downloading and extracting variables from global databases

ANNEX II: Calculation of drought indices and water balances

ANNEX III: Graphs

Please refer to for download the original scripts:

https://ruc.udc.es/dspace/handle/2183/29799

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2017.191

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. Rev. Geophys. 53, 452–480. https://doi.org/10.1002/2014RG000456

Agutu, N.O., Awange, J.L., Zerihun, A., Ndehedehe, C.E., Kuhn, M., Fukuda, Y., 2017. Assessing multi-satellite remote sensing, reanalysis, and land surface models' products in characterizing agricultural drought in East Africa. Remote Sens. Environ. 194, 287–302. https://doi.org/10.1016/j.rse.2017.03.041

Ahumada, J.A., 2003. R para Principiantes.

Alley, W.M., 1984. The Palmer Drought Severity Index: Limitations and Assumptions. J. Clim. Appl. Meteorol. https://doi.org/10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2

Anderson, W., Seager, R., Baethgen, W., Cane, M., 2018. Trans-Pacific ENSO teleconnections pose a correlated risk to agriculture. Agric. For. Meteorol. 262, 298–309. https://doi.org/10.1016/j.agrformet.2018.07.023

Anderson, W., Seager, R., Baethgen, W., Cane, M., 2017. Life cycles of agriculturally relevant ENSO teleconnections in North and South America. Int. J. Climatol. 37, 3297–3318. https://doi.org/10.1002/joc.4916

Aragón, R., Jobbágy, E.G., Viglizzo, E.F., 2011. Surface and groundwater dynamics in the sedimentary plains of the Western Pampas (Argentina). Ecohydrology 4, 433–447. https://doi.org/10.1002/eco.149

Araneda-Cabrera, R.J., Bermudez, M., Puertas, J., 2021a. Benchmarking of drought and climate indices for agricultural drought monitoring in Argentina. Sci. Total Environ. 790, 148090. https://doi.org/10.1016/j.scitotenv.2021.148090

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2021b. Assessment of the performance of drought indices for explaining crop yield variability at the national scale : Methodological framework and application to Mozambique. Agric. Water Manag. 246. https://doi.org/10.1016/j.agwat.2020.106692

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2020. Unified framework for drought monitoring and assessment in a transboundary river basin, in: River Flow 2020; Uijttewaal, W, Franca M, Valero D, Chavarrias V, Arbós C, Schielen R and Crosato A, Eds. Taylor & Francis Group, London, pp. 1081–1086. https://doi.org/10.1201/b22619

Ayantobo, O.O., Li, Y., Song, S., Yao, N., 2017. Spatial comparability of drought characteristics and related return periods in mainland China over 1961–2013. J. Hydrol. 550, 549–567. https://doi.org/10.1016/j.jhydrol.2017.05.019

Backlund, P., Janetos, A., Schimel, D., 2008. The effects of climate change on agriculture, land resources, water resources, and biodiversity in the United States. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. U.S. Department of Agriculture, Washington, DC., USA, 362 pp.

Barros, V.R., Silvestri, G.E., 2002. The relation between sea surface temperature at the subtropical South-Central Pacific and precipitation in Southeastern South America. J. Clim. 15, 251–267. https://doi.org/10.1175/1520-0442(2002)015<0251:trbsst>2.0.co;2

Beck, H.E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., De Roo, A., 2017a. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21, 589–615. https://doi.org/10.5194/hess-21-589-2017

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., Dijk, A.I.J.M. van, Weedon4, G.P., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci. 21, 6201–6217. https://doi.org/https://doi.org/10.5194/hess-21-6201-2017

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-geiger climate classification maps at 1-km resolution. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2018.214

Begueria, S., Vicente-Serrano, S.M., 2017. Package 'SPEI'. version 1.7. ftp://tucows.icm.edu.pl/packages/cran/web/packages/SPEI/SPEI.pdf. A case study Birkoor Kortigiri Mandals. https://doi.org/10.1175/2009JCLI2909.1.http

Ben-Ari, T., Makowski, D., 2014. Decomposing global crop yield variability. Environ. Res. Lett. 9. https://doi.org/10.1088/1748-9326/9/11/114011

Bento, V.A., Gouveia, C.M., DaCamara, C.C., Trigo, I.F., 2018. A climatological assessment of drought impact on vegetation health index. Agric. For. Meteorol. 259, 286–295. https://doi.org/10.1016/j.agrformet.2018.05.014

Bolar, K., 2019. Package 'STAT'. version 0.1.0. https://cran.r-project.org/web/packages/STAT/STAT.pdf.

Borchers, H.W., 2019. Package 'pracma'. version 2.2.9. http://mirrors.ucr.ac.cr/CRAN/web/packages/pracma/pracma.pdf.

Brida, A.B., Owiyo, T., 2013. Loss and damage from the double blow of flood and drought in Mozambique. Int. J. Glob. Warm. 5, 514. https://doi.org/10.1504/IJGW.2013.057291

Bryant, E.A., Head, L.M., Morrison, R., 2005. Planning for Natural Hazards — How Can We Mitigate the Impacts ?, in: R.J. Morrison, S. Quin and E.A. Bryant (Eds.), Planning for Natural Hazards — How Can We Mitigate the Impacts?, Proceedings of a Symposium, 2-5 February 2005, University of Wollongong, GeoQuEST Research Centre, 2005, 1-11.

Cangelosi, R., Goriely, A., 2007. Component retention in principal component analysis with application to cDNA microarray data. Biol. Direct 2, 1–21. https://doi.org/10.1186/1745-6150-2-2

Carcedo, A.J.P., Gambin, B.L., 2019. Sorghum drought and heat stress patterns across the Argentinean temperate central region. F. Crop. Res. 241, 107552. https://doi.org/10.1016/j.fcr.2019.06.009

Champagne, C., White, J., Berg, A., Belair, S., Carrera, M., 2019. Impact of Soil Moisture Data Characteristics on the Sensitivity to Crop Yields Under Drought and Excess Moisture Conditions. Remote Sens. 11, 372. https://doi.org/10.3390/rs11040372

Chen, J.L., Wilson, C.R., Tapley, B.D., Longuevergne, L., Yang, Z.L., Scanlon, B.R., 2010. Recent la Plata basin drought conditions observed by satellite gravimetry. J. Geophys. Res. Atmos. 115, 1–12. https://doi.org/10.1029/2010JD014689

Chen, T., Xia, G., Liu, T., Chen, W., Chi, D., 2016. Assessment of drought impact on main cereal crops using a standardized precipitation evapotranspiration index in Liaoning Province, China. Sustain. 8, 1–16. https://doi.org/10.3390/su8101069

Cherlet, M., Hutchinson, Charles Reynolds, J., Hill, J., Sommer, S., Von Maltitz, G., 2018. World atlas of desertification. Rethinking land degradation and sustainable land management., Third Edit. ed, Publications Office of the European Union. https://doi.org/10.2760/06292

Conselho de Ministros, 2020. BR Nº 160 de 20.08.20, Boletim da República - I Serie. Publicação oficial da República de Moçambique. Maputo, Mozambique. https://www.inm.gov.mz/pt-br/content/br-n%C2%BA-160-de-200820-boletim-da-rep%C3%BAblica-i-serie.

Cumani, M. (FAO), Rojas, O. (FAO, 2016. Characterization of the agricultural drought prone areas on global a global scale at scale. Using the FAO Agricultural Stress Index System (ASIS) to enhance the understanding of, and boost resilience to, water stress conditions in drought-prone areas.

D'Ambrosio, G.T., Bohn, V.Y., Piccolo, M.C., 2013. Evaluation of the 2008-2009 drought in the west of the Pampean region (Argentina). Cuad. Geográficos 52, 29–45.

D'Arrigo, R., Smerdon, J.E., 2008. Tropical climate influences on drought variability over Java, Indonesia. Geophys. Res. Lett. 35, 1–5. https://doi.org/10.1029/2007GL032589

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

Daryanto, S., Wang, L., Jacinthe, P.A., 2016. Drought effects on root and tuber production: A meta-analysis. Agric. Water Manag. 176, 122–131. https://doi.org/10.1016/j.agwat.2016.05.019

De La Casa, A.C., Ovando, G.G., 2006. The influence of El Niño Southern Oscillation (ENSO) episodes on rainfall and corn yields in Córdoba Province, Argentina. Agric. Técnica 66, 80–89. https://doi.org/10.4067/S0365-28072006000100009

Deering, D.W., 1978. Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. Ph. D Dissertation. Texas A&M Univ.

Díaz, E., García, M., Rodríguez, A., Dölling, O., Ochoa, S., Bertoni, J., 2018. Temporal evolution of hydrological drought in Argentina and its relationship with macroclimatic indicators. Tecnol. y Ciencias del Agua 9, 1–32. https://doi.org/10.24850/j-tyca-2018-05-01

Dikshit, A., Pradhan, B., Alamri, A.M., 2021. Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model. Sci. Total Environ. 755, 142638. https://doi.org/10.1016/j.scitotenv.2020.142638

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sens. Environ. 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Du, T.L.T., Du Bui, D., Nguyen, M.D., Lee, H., 2018. Satellite-based, multi-indices for evaluation of agricultural droughts in a highly dynamic tropical catchment, Central Vietnam. Water (Switzerland) 10. https://doi.org/10.3390/w10050659

Durán, A., 2000. Agua en el suelo.

Dutra, E., Di Giuseppe, F., Wetterhall, F., Pappenberger, F., 2013. Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index. Hydrol. Earth Syst. Sci. 17, 2359–2373. https://doi.org/10.5194/hess-17-2359-2013

Easterling, D.R., 2013. Global Data Sets for Analysis of Climate Extremes., in: Water Science and Technology Library (Ed.), Global Data Sets for Analysis of Climate Extremes. Springer, Dordrecht, pp. 347–361. https://doi.org/doi:10.1007/978-94-007-4479-0_12

El Kenawy, A.M., McCabe, M.F., Vicente-Serrano, S.M., López-Moreno, J.I., Robaa, S.M., 2016. Cambios en la frecuencia y severidad en las sequías hidrológicas de Etiopía entre 1960 y 2013. Cuad. Investig. Geogr. 42, 145–166. https://doi.org/10.18172/cig.2931

EM-DAT, 2021. The Emergency Events Database, Universite catholique de Louvain, Brussels, Belgium. [Available at http://www. emdat.be/, Accessed November 2020.].

Eriksen, S., Silva, J.A., 2009. The vulnerability context of a savanna area in Mozambique: household drought coping strategies and responses to economic change. Environ. Sci. Policy 12, 33–52. https://doi.org/10.1016/j.envsci.2008.10.007

Espinosa, L.A., Portela, M.M., Rodrigues, R., 2019. Spatio-temporal variability of droughts over past 80 years in Madeira Island. J. Hydrol. Reg. Stud. 25, 100623. https://doi.org/10.1016/j.ejrh.2019.100623

FAO, 2019. World Food and Agriculture - Statistical Pocketbook 2019, Food and Agriculture Organization of the United Nations. Rome. https://doi.org/10.4060/ca1796en

FAO, 2017. Country fact sheet on food and agriculture policy trends. Argentina, Food and Agriculture Organization of the United Nations. I7752EN/1/08.17.

FAO, 2016a. AQUASTAT Country Profile – Mozambique. FAO., Food and Agriculture Organization of the United Nations (FAO). Rome, Italy.

FAO, 2016b. Regional Roundtable on the World Programme for the Census of Agriculture 2020 (WCA 2020). 12-16 December 2016, Montevideo, Uruguay. In: World Programme for the Census of Agriculture [online]. Rome, Italy. http://www.fao.org/world-census-agriculture/events/wca2020montevideo/en/.

FAO, 2006. The State of Food and Agricultural Food aid for food security? FAO. Rome, Italy.

FAO, PROSAP, Ministerio de Agricultura Ganadería y Pesca de Argentina, 2015. Estudio del potencial de ampliación del riego en Argentina. Buenos Aires, http://www.fao.org/3/a-i5183s.pdf.

FEWS NET Moçambique, 2014. MOÇAMBIQUE Descrição das zonas de Formas de Vida, USAID https://fews.net/sites/default/files/documents/reports/MZ%20LHdescriptions%202013%20pt.pdf.

Fleming, S.W., Marsh Lavenue, A., Aly, A.H., Adams, A., 2002. Practical applications of spectral analysis of hydrologic time series. Hydrol. Process. 16, 565–574. https://doi.org/10.1002/hyp.523

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., G., H., 2015. A global satellite-assisted precipitation climatology. Earth Syst. Sci. Data 7, 275–287. https://doi.org/10.5676/DWD

Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J.D., Romero, B.E., Husak, G.J., Michaelsen, J.C., Verdin, A.P., 2014. A Quasi-Global Precipitation Time Series for Drought Monitoring. U.S. Geol. Surv. Data Ser. 832, 4.

Gao, X., Dong, S., Li, S., Xu, Y., Liu, S., Zhao, H., Yeomans, J., Li, Y., Shen, H., Wu, S., Zhi, Y., 2020.

Using the random forest model and validated MODIS with the field spectrometer measurement promote the accuracy of estimating aboveground biomass and coverage of alpine grasslands on the Qinghai-Tibetan Plateau. Ecol. Indic. 112. https://doi.org/10.1016/j.ecolind.2020.106114

García-Garizábal, I., 2017. Rainfall variability and trend analysis in coastal arid Ecuador. Int. J. Climatol. 37, 4620–4630. https://doi.org/10.1002/joc.5110

García-León, D., Contreras, S., Hunink, J., 2019. Comparison of meteorological and satellite-based drought indices as yield predictors of Spanish cereals. Agric. Water Manag. 213, 388–396. https://doi.org/10.1016/j.agwat.2018.10.030

George, H., 2006. Agro-MAPS: A global spatial database of sub-national agricultural land-use statistics, in: Robinson, T., Thornton, P. (Eds.), Global Mapping of Agricultural Production Systems. Pro-Poor Livestock Policy Initiative, pp. 22–23.

Golnaraghi, M., Etienne, C., Sapir, D.G., Below, R., 2014. Atlas of Mortality and Economic Losses From Weather, Climate and Water Extremes (1970-2012), WMO-No. 1123, World Meteorological Organization, Geneva, Switzerland. https://www.preventionweb.net/files/38413_wmo1123atlas120614.pdf.

Gouhier, T.C., Grinsted, A., Simko, V., 2016. Package 'biwavelet'. version 0.20.11. https://mran.revolutionanalytics.com/snapshot/2018-01-06/web/packages/biwavelet/biwavelet.pdf.

Gringorten, I.I., 1963. A Plotting Rule for Extreme Probability Paper. J. Geophys. Res. 68, 813–814. https://doi.org/doi:10.1029/JZ068i003p00813

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Process. Geophys. 11, 561–566. https://doi.org/10.5194/npg-11-561-2004

Gu, Y., Liu, H., Traoré, D.D., Huang, C., 2020. ENSO-related droughts and ISM variations during the last millennium in tropical southwest China. Clim. Dyn. 54, 649–659. https://doi.org/10.1007/s00382-019-05019-1

Guha-Sapir, D., D'Aoust, O., Vos, F., Hoyois, P., 2015. "The frequency and impact of natural disasters" in The Economic Impacts of Natural Disasters, D. Guha-Sapir, I. Santos, Eds. (Oxford University Press, Oxford, United Kingdom, 2013). pp. 7–27. https://doi.org/10.1080/00213624.2015.1013902

Gupta, V., Jain, M.K., 2021. Unravelling the teleconnections between ENSO and dry/wet conditions over India using nonlinear Granger causality. Atmos. Res. 247, 105168. https://doi.org/10.1016/j.atmosres.2020.105168

Guttman, N.B., 1999. Accepting the Standardized Precipitation Index: a Calculation Algorithm. JAWRA J. Am. Water Resour. Assoc. 35, 311–322. https://doi.org/10.1111/j.1752-1688.1999.tb03592.x

Guttman, N.B., 1998. Comparing the palmer drought index and the standardized precipitation index. J. Am. Water Resour. Assoc. 34, 113–121. https://doi.org/10.1111/j.1752-1688.1998.tb05964.x

Hagenlocher, M., Meza, Isabel, Carl Anderson, Annika Min, Fabrice G. Renaud, Y., Walz, S.S., Sebesvari, Z., 2019a. Drought vulnerability and risk assessments: state of the art, persistent gaps, and research agenda. Environ. Res. Lett. 4. https://doi.org/https://doi.org/10.1088/1748-9326/ab225d Manuscript

Hagenlocher, M., Meza, I., Anderson, C., Min, A., Renaud, F.G., Walz, Y., Siebert, S., Sebesvari, Z., 2019b. Drought vulnerability and risk assessments: state of the art, persistent gaps, and research agenda. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/ab225d

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 1998. Multivariate Data Analysis, 7th ed, Pearson. Prentice Hall, Englewood Cliffs, New Jersey, USA.

Hao, Z., AghaKouchak, A., 2014. A Nonparametric Multivariate Multi-Index Drought Monitoring Framework. J. Hydrometeorol. 15, 89–101. https://doi.org/10.1175/JHM-D-12-0160.1

Hao, Z., AghaKouchak, A., 2013. Multivariate Standardized Drought Index: A parametric multi-index model. Adv. Water Resour. 57, 12–18. https://doi.org/10.1016/j.advwatres.2013.03.009

Hao, Z., Singh, V.P., 2015. Drought characterization from a multivariate perspective: A review. J. Hydrol. 527, 668–678. https://doi.org/10.1016/j.jhydrol.2015.05.031

Hao, Z., Singh, V.P., Xia, Y., 2018. Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. Rev. Geophys. 56, 108–141. https://doi.org/10.1002/2016RG000549

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. data 7, 109. https://doi.org/10.1038/s41597-020-0453-3

Hassan, W.U., Nayak, M.A., 2020. Global teleconnections in droughts caused by oceanic and atmospheric circulation patterns. Environ. Res. Lett. 16. https://doi.org/10.1088/1748-9326/abc9e2

Hayes, M., Svoboda, M., Wall, N., Widhalm, M., 2011. The lincoln declaration on drought indices: universal meteorological drought index recommended. Bull. Am. Meteorol. Soc. 92, 485–488. https://doi.org/10.1175/2010BAMS3103.1

Hellmuth, M.E., Moorhead, A., Thomson, M.C., Williams, J., 2007. Climate Risk Management in Africa: Learning from Practice, International Research Institute for Climate and Society (IRI). Columbia University, New York, USA.

Hipel, K.W., McLeod, A.I., 1994. Time Series Modelling of Water Resources and Environmental Systems. Elsevier Science, New York.

Hosking, J.R.M., Wallis, J.R., 1993. Some Statistics Useful in Regional Frequency Analys. Water Resour. Res. 29, 271–281.

Huang, S., Huang, Q., Leng, G., Liu, S., 2016. A nonparametric multivariate standardized drought index for characterizing socioeconomic drought: A case study in the Heihe River Basin. J. Hydrol. 542, 875–883. https://doi.org/10.1016/j.jhydrol.2016.09.059

Huang, S., Li, P., Huang, Q., Leng, G., Hou, B., Ma, L., 2017. The propagation from meteorological to hydrological drought and its potential influence factors. J. Hydrol. 547, 184–195. https://doi.org/10.1016/j.jhydrol.2017.01.041

Hudgins, L., Friehe, C.A., Mayer, M.E., 1993. Wavelet transforms and atmopsheric turbulence. Phys. Rev. Lett. 71, 3279. https://doi.org/https://doi.org/10.1103/PhysRevLett.71.3279

Hurst, H.E., 1956. Methods of using long-term storage in reservoirs., in: Proceedings of the Institution

of Civil Engineers 5 (5). pp. 519–543. https://doi.org/https://doi.org/10.1680/iicep.1956.11503

Hurtado, S.I., Agosta, E.A., 2020. El Niño Southern Oscillation-related precipitation anomaly variability over eastern subtropical South America: Atypical precipitation seasons. Int. J. Climatol. 1–20. https://doi.org/10.1002/joc.6559

Iizumi, T., Luo, J.J., Challinor, A.J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M.E., Yamagata, T., 2014. Impacts of El Niño Southern Oscillation on the global yields of major crops. Nat. Commun. 5, 1–7. https://doi.org/10.1038/ncomms4712

INGC, 2009. Synthesis report. INGC Climate Change Report: Study on the impact of climate change on disaster risk in Mozambique. [van Logchem B and Brito R (ed.)]. INGC, Mozambique. Maputo, Mozambique. http://pure.iiasa.ac.at/id/eprint/9007/.

Inguane, R., Gallego-Ayala, J., Juízo, D., 2014. Decentralized water resources management in Mozambique: Challenges of implementation at the river basin level. Phys. Chem. Earth 67–69, 214–225. https://doi.org/10.1016/j.pce.2013.08.004

IPCC, 2014. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, Cambridge University Press. Cambridge, United Kingdom and New York, NY, USA.

IPCC, 2014: Climate Change, 2014. Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp. https://doi.org/10.1017/CBO9781107415324

Jayanthi, H., Husak, G.J., Funk, C., Magadzire, T., Chavula, A., Verdin, J.P., 2013. Modeling rain-fed maize vulnerability to droughts using the standardized precipitation index from satellite estimated rainfall-Southern Malawi case study. Int. J. Disaster Risk Reduct. 4, 71–81. https://doi.org/10.1016/j.ijdrr.2013.02.001

Jehanzaib, M., Kim, T.W., 2020. Exploring the influence of climate change-induced drought propagation on wetlands. Ecol. Eng. 149, 105799. https://doi.org/10.1016/j.ecoleng.2020.105799

Jehanzaib, M., Shah, S.A., Yoo, J., Kim, T.W., 2020. Investigating the impacts of climate change and human activities on hydrological drought using non-stationary approaches. J. Hydrol. 588, 125052. https://doi.org/10.1016/j.jhydrol.2020.125052

Jury, M.R., 2013. Climate trends in southern Africa. S. Afr. J. Sci. 109, 1–11. https://doi.org/10.1590/sajs.2013/980

Kasnakoglu, H., Mayo, R., 2004. FAO Statistical Data Quality Framework: A multi-layered approach to monitoring and assessment., in: Conference on Data Quality for Interna-Tional Organizations, Wiesbaden, Germany, 27 and 28 May 2004.

Keyantash, J., Dracup, J.A., 2002. Quantification of Drought An Evaluation of Drought Indices. Am. Meteorol. Soc. 1167–1180. https://doi.org/10.1175/1520-0477-83.8.1167

Kim, T.W., Jehanzaib, M., 2020. Drought risk analysis, forecasting and assessment under climate change. Water (Switzerland) 12, 1–7. https://doi.org/10.3390/W12071862

Kogan, F.N., 1997. Global Drought Watch from Space. Bull. Am. Meteorol. Soc. 78, 621–636. https://doi.org/10.1175/1520-0477(1997)078<0621:GDWFS>2.0.CO;2

Kogan, F.N., 1995. Application of vegetation index and brightness temperature for drought detection. Adv. Sp. Res. 15, 91–100. https://doi.org/10.1016/0273-1177(95)00079-T

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Zeitschrift 15, 259–263. https://doi.org/10.1127/0941-2948/2006/0130

Koutsoyiannis, D., 2005. Hydrologic Persistence and The Hurst Phenomenon, in: Water Encyclopedia. pp. 210–221. https://doi.org/10.1002/047147844x.sw434

Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon, and hydrological statistics. Hydrol. Sci. J. 48, 3–24. https://doi.org/10.1623/hysj.48.1.3.43481

Kumar, N.M., Murthy, C.S., Sesha Sai, M.V.R., Roy, P.S., 2009. On the use of Standardized Precipitation Index (SPI) for drought intensity assessment. Meteorol. Appl. 16, 381–389. https://doi.org/10.1002/met.136

Kuwayama, Y., Thompson, A., Bernknopf, R., Zaitchik, B., Vail, P., 2018. Estimating the Impact of Drought on Agriculture Using the U.S. Drought Monitor. Am. J. Agric. Econ. 101, 193–210. https://doi.org/10.1093/ajae/aay037

Landerer, F.W., Flechtner, F.M., Save, H., Webb, F.H., Bandikova, T., Bertiger, W.I., Bettadpur, S. V., Byun, S.H., Dahle, C., Dobslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G.L.H., Loomis, B.D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strekalov, D., Tamisiea, M.E., Wang, F., Watkins, M.M., Wen, H.Y., Wiese, D.N., Yuan, D.N., 2020. Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance. Geophys. Res. Lett. 47, 1–10. https://doi.org/10.1029/2020GL088306

Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. Water Resour. Res. 48, 1–11. https://doi.org/10.1029/2011WR011453

Lawal, S., Lennard, C., Jack, C., Wolski, P., Hewitson, B., Abiodun, B., 2019. The observed and model-simulated response of southern African vegetation to drought. Agric. For. Meteorol. 279, 107698. https://doi.org/10.1016/j.agrformet.2019.107698

Leng, G., Hall, J., 2019. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. Sci. Total Environ. 654, 811–821. https://doi.org/10.1016/j.scitotenv.2018.10.434

Lima, C.H.R., AghaKouchak, A., 2017. Droughts in Amazonia: Spatiotemporal Variability, Teleconnections, and Seasonal Predictions. Water Resour. Res. 53, 10824–10840. https://doi.org/10.1002/2016WR020086

Lisboa, I.P., Damian, M., Cherubin, M.R., Barros, P.P.S., Fiorio, P.R., Cerri, C.C., Cerri, C.E.P., 2018. Prediction of sugarcane yield based on NDVI and concentration of leaf-tissue nutrients in fields managed with straw removal. Agronomy 8. https://doi.org/10.3390/agronomy8090196

Liu, S., Shi, H., Sivakumar, B., 2020. Socioeconomic Drought Under Growing Population and Changing Climate: A New Index Considering the Resilience of a Regional Water Resources System. J. Geophys. Res. Atmos. 125. https://doi.org/10.1029/2020JD033005

Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011a. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat. Clim. Chang. 1, 42–45. https://doi.org/10.1038/nclimate1043

Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to

climate change. Agric. For. Meteorol. 150, 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008

Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011b. Climate Trends and Global Crop Production Since 1980. Science (80-.). 333, 616–621. https://doi.org/10.7551/mitpress/8876.003.0036

Lovino, M., García, N.O., Baethgen, W., 2014. Spatiotemporal analysis of extreme precipitation events in the Northeast region of Argentina (NEA). J. Hydrol. Reg. Stud. 2, 140–158. https://doi.org/10.1016/j.ejrh.2014.09.001

Lovino, M.A., Müller, O. V., Müller, G. V., Sgroi, L.C., Baethgen, W.E., 2018. Interannual-to-multidecadal hydroclimate variability and its sectoral impacts in northeastern Argentina. Hydrol. Earth Syst. Sci. 22, 3155–3174. https://doi.org/10.5194/hess-22-3155-2018

Lyon, B., Barnston, A.G., 2005. ENSO and the spatial extent of interannual precipitation extremes in tropical land areas. J. Clim. 18, 5095–5109. https://doi.org/10.1175/JCLI3598.1

Ma, M., Ren, L., Yuan, F., Jiang, S., Liu, Y., Kong, H., Gong, L., 2014. A new standardized Palmer drought index for hydro-meteorological use. Hydrol. Process. 28, 5645–5661. https://doi.org/10.1002/hyp.10063

Macarringue, L.S., Sano, E.E., Chaves, J.M., Bolfe, E.L., 2017. Considerações Sobre Precipitação, Relevo E Solos E Análise Do Potencial De Expansão Agrícola Da Região Norte De Moçambique. Soc. Nat. 29, 109–122. https://doi.org/10.14393/sn-v29n1-2017-7

Magrin, G.O., Travasso, M.I., Rodríguez, G.R., 2005. Changes in climate and crop production during the 20th century in Argentina. Clim. Change 72, 229–249. https://doi.org/10.1007/s10584-005-5374-9

Manatsa, D., Chingombe, W., Matarira, C.H., 2008a. The impact of the positive Indian Ocean dipole on Zimbabwe droughts. Int. J. Climatol. 28. https://doi.org/10.1002/joc

Manatsa, D., Chingombe, W., Matsikwa, H., Matarira, C.H., 2008b. The superior influence of Darwin Sea level pressure anomalies over ENSO as a simple drought predictor for Southern Africa. Theor. Appl. Climatol. 92, 1–14. https://doi.org/10.1007/s00704-007-0315-3

Manhique, A.J., Reason, C.J.C., Rydberg, L., Fauchereau, N., 2011. ENSO and Indian Ocean sea surface temperatures and their relationships with tropical temperate troughs over Mozambique and the Southwest Indian Ocean. Int. J. Climatol. 31, 1–13. https://doi.org/10.1002/joc.2050

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometria 13, 245–259.

Masih, I., Maskey, S., Mussá, F.E.F., Trambauer, P., 2014. A review of droughts on the African continent: A geospatial and long-term perspective. Hydrol. Earth Syst. Sci. 18, 3635–3649. https://doi.org/10.5194/hess-18-3635-2014

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The Relationship of Drought Frequency and Duration to Time Scales, Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA. https://doi.org/10.1088/1755-1315/5

McPhaden, M.J., Busalacchi, A.J., Cheney, R., Donguy, J.R., Gage, K.S., Halpern, D., Ji, M., Julian, P., Meyers, G., Mitchum, G.T., Niiler, P.P., Picaut, J., Reynolds, R.W., Smith, N., Takeuchi, K., 1998. The Tropical Ocean-Global Atmosphere observing system: a decade of progress. J. Geophys. Res. Ocean. 103, 14169–14240. https://doi.org/10.1029/97jc02906

Mehran, A., Mazdiyasni, O., Aghakouchak, A., 2015. A hybrid framework for assessing socioeconomic

drought: Linking. J. Geophys. Res. Atmos. 1–14. https://doi.org/10.1002/2015JD023147.Received

Meresa, H.K., Osuch, M., Romanowicz, R., 2016. Hydro-meteorological drought projections into the 21-st century for selected polish catchments. Water (Switzerland) 8. https://doi.org/10.3390/w8050206

Midgley, S., Dejene, A., Mattick, A., 2012. Adaptation to Climate Change in Semi-arid Environments: Experience and Lessons from Mozambique, Environment and Natural Management Series, Food and Agricultural Organization (FAO), Rome, Italy.

Ministério da Agricultura e Segurança Alimentar, 2015. Anuário de Estatísticas Agrárias 2015.

Mishra, A.K., Singh, V.P., 2011. Drought modeling - A review. J. Hydrol. 403, 157–175. https://doi.org/10.1016/j.jhydrol.2011.03.049

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391, 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012

Mo, K.C., 2008. Model-based drought indices over the United States. J. Hydrometeorol. 9, 1212–1230. https://doi.org/10.1175/2008JHM1002.1

Monteleone, B., Bonaccorso, B., Martina, M., 2020. A joint probabilistic index for objective drought identification: The case study of Haiti. Nat. Hazards Earth Syst. Sci. 20, 471–487. https://doi.org/10.5194/nhess-20-471-2020

MunichRE, 2018. NatCatSERVICE.

Nam, W.H., Hayes, M.J., Svoboda, M.D., Tadesse, T., Wilhite, D.A., 2015. Drought hazard assessment in the context of climate change for South Korea. Agric. Water Manag. 160, 106–117. https://doi.org/10.1016/j.agwat.2015.06.029

Nashwan, M.S., Shahid, S., Dewan, A., Ismail, T., Alias, N., 2020. Performance of five high resolution satellite-based precipitation products in arid region of Egypt: An evaluation. Atmos. Res. 236, 104809. https://doi.org/10.1016/j.atmosres.2019.104809

Noor, N.M., Al Bakri Abdullah, M.M., Yahaya, A.S., Ramli, N.A., 2015. Comparison of linear interpolation method and mean method to replace the missing values in environmental data set, in: Materials Science Forum. Trans Tech Publications Ltd., pp. 278–281. https://doi.org/10.4028/www.scientific.net/MSF.803.278

Oguntunde, P.G., Abiodun, B.J., Lischeid, G., 2017. Impacts of climate change on hydro-meteorological drought over the Volta Basin, West Africa. Glob. Planet. Change 155, 121–132. https://doi.org/10.1016/j.gloplacha.2017.07.003

Oguntunde, P.G., Lischeid, G., Abiodun, B.J., 2018. Impacts of climate variability and change on drought characteristics in the Niger River Basin, West Africa. Stoch. Environ. Res. Risk Assess. 32, 1017–1034. https://doi.org/10.1007/s00477-017-1484-y

Oñate-Valdivieso, F., Uchuari, V., Oñate-Paladines, A., 2020. Large-Scale Climate Variability Patterns and Drought: A Case of Study in South – America. Water Resour. Manag. 34, 2061–2079. https://doi.org/10.1007/s11269-020-02549-w

Osbahr, H., Twyman, C., Neil Adger, W., Thomas, D.S.G., 2008. Effective livelihood adaptation to climate change disturbance: Scale dimensions of practice in Mozambique. Geoforum 39, 1951–1964. https://doi.org/10.1016/j.geoforum.2008.07.010

Palmer, W.C., 1965. Meteorological Drought, Research paper no. 45. US Weather Bur. Washington, DC. p-and-precip/drought/docs/palmer.pdf 58.

Patt, A.G., Schröter, D., 2008. Perceptions of climate risk in Mozambique: Implications for the success of adaptation strategies. Glob. Environ. Chang. 18, 458–467. https://doi.org/10.1016/j.gloenvcha.2008.04.002

Peña-Gallardo, M., Martín Vicente-Serrano, S., Domínguez-Castro, F., Beguería, S., 2019a. The impact of drought on the productivity of two rainfed crops in Spain. Nat. Hazards Earth Syst. Sci. 19, 1215–1234. https://doi.org/10.5194/nhess-19-1215-2019

Peña-Gallardo, M., Vicente-Serrano, S.M., Quiring, S., Svoboda, M., Hannaford, J., Tomas-Burguera, M., Martín-Hernández, N., Domínguez-Castro, F., El Kenawy, A., 2019b. Response of crop yield to different time-scales of drought in the United States: Spatio-temporal patterns and climatic and environmental drivers. Agric. For. Meteorol. 264, 40–55. https://doi.org/10.1016/j.agrformet.2018.09.019

Peña-Gallardo, M., Vicente-Serrano, S.M., Quiring, S., Svodova, M., Hannaford, J., 2018. Effectiveness of drought indices in identifying impacts on major crops across the USA. Clim. Res. 75, 221–240. https://doi.org/10.3354/cr01519

Podestá, G.P., Messina, C.D., Grondona, M.O., Magrin, G.O., 1999. Associations between grain crop yields in Central-Eastern Argentina and El Niño-Southern oscillation. J. Appl. Meteorol. 38, 1488–1498. https://doi.org/10.1175/1520-0450(1999)038<1488:ABGCYI>2.0.CO;2

Pohlert, T., 2020. Package 'trend'. version 1.1.4. https://cran.r-project.org/web/packages/trend/trend.pdf.

Potopová, V., Trnka, M., Hamouz, P., Soukup, J., Castraveț, T., 2020. Statistical modelling of drought-related yield losses using soil moisture-vegetation remote sensing and multiscalar indices in the south-eastern Europe. Agric. Water Manag. 236, 106168. https://doi.org/10.1016/j.agwat.2020.106168

Quiring, S.M., Papakryiakou, T.N., 2003. An evaluation of agricultural drought indices for the Canadian prairies. Agric. For. Meteorol. 118, 49–62. https://doi.org/10.1016/S0168-1923(03)00072-8

Ragab, R., Prudhomme, C., 2002. Climate change and water resources management in arid and semi-arid regions: Prospective and challenges for the 21st century. Biosyst. Eng. 81, 3–34. https://doi.org/10.1006/bioe.2001.0013

Räsänen, T.A., Lindgren, V., Guillaume, J.H.A., Buckley, B.M., Kummu, M., 2016. On the spatial and temporal variability of ENSO precipitation and drought teleconnection in mainland Southeast Asia. Clim. Past 12, 1889–1905. https://doi.org/10.5194/cp-12-1889-2016

Rencher, A.C., 2002. Methods of Multivariate Analysis, 2nd ed, John Wiley & Sons. https://www.ipen.br/biblioteca/slr/cel/0241.

Rivera, J.A., Araneo, D.C., Penalba, O.C., Villalba, R., 2018. Regional aspects of streamflow droughts in the Andean rivers of Patagonia, Argentina. Links with large-scale climatic oscillations. Hydrol. Res. 49, 134–149. https://doi.org/10.2166/nh.2017.207

Rivera, J.A., Otta, S., Lauro, C., Zazulie, N., 2021. A Decade of Hydrological Drought in Central-Western Argentina. Front. Water 3, 1–20. https://doi.org/10.3389/frwa.2021.640544

Robledo, F.A., Penalba, O.C., Bettolli, M.L., 2013. Teleconnections between tropical-extratropical oceans and the daily intensity of extreme rainfall over Argentina. Int. J. Climatol. 33, 735–745.

https://doi.org/10.1002/joc.3467

Rojas, O., Vrieling, A., Rembold, F., 2011. Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. Remote Sens. Environ. 115, 343–352. https://doi.org/10.1016/j.rse.2010.09.006

Romm, J., 2011. The next dust bowl. Nature 450–451, 1–2. https://doi.org/https://doi.org/10.1038/478450a

Rouault, M., Richard, Y., 2005. Intensity and spatial extent of droughts in southern Africa. Geophys. Res. Lett. 32, 2–5. https://doi.org/10.1029/2005GL022436

RStudio Team, 2016. Integrated Development Environment for R. RStudio, Inc. , Boston, MA URL http://www.rstudio.com/.

Rudolf, B., Becker, A., Schneider, U., 2011. New GPCC full data reanalysis version 5 provides high-quality gridded monthly precipitation data. GEWEX News 2010, 2009–2010.

Santos, J.F., Pulido-Calvo, I., Portela, M.M., 2010. Spatial and temporal variability of droughts in Portugal. Water Resour. Res. 46, 1–13. https://doi.org/10.1029/2009WR008071

Santos, M.S., Costa, V.A.F., Fernandes, W.D.S., de Paes, R.P., 2019. Time-space characterization of droughts in the São Francisco river catchment using the Standard Precipitation Index and continuous wavelet transform. Rev. Bras. Recur. Hidricos 24, 1–12. https://doi.org/10.1590/2318-0331.241920180092

Schauberger, B., Ben-Ari, T., Makowski, D., Kato, T., Kato, H., Ciais, P., 2018. Yield trends, variability and stagnation analysis of major crops in France over more than a century. Sci. Rep. 8, 1–12. https://doi.org/10.1038/s41598-018-35351-1

Seibert, M., Merz, B., Apel, H., 2017. Seasonal forecasting of hydrological drought in the Limpopo Basin: A comparison of statistical methods. Hydrol. Earth Syst. Sci. 21, 1611–1629. https://doi.org/10.5194/hess-21-1611-2017

Seiler, R.A., Kogan, F., Wei, G., Vinocur, M., 2007. Seasonal and interannual responses of the vegetation and production of crops in Cordoba - Argentina assessed by AVHRR derived vegetation indices. Adv. Sp. Res. 39, 88–94. https://doi.org/10.1016/j.asr.2006.05.024

Sheffield, J., Wood, E.F., 2012. Drought: Past problems and future scenarios. Drought Past Probl. Futur. Scenar. 9781849775, 1–234. https://doi.org/10.4324/9781849775250

Shi, W., Tao, F., Zhang, Z., 2013. A review on statistical models for identifying climate contributions to crop yields. J. Geogr. Sci. 23, 567–576. https://doi.org/10.1007/s11442-013-1029-3

Singh, R.M., Shukla, P., 2020. Drought Characterization Using Drought Indices and El Niño Effects. Natl. Acad. Sci. Lett. 43, 339–342. https://doi.org/10.1007/s40009-019-00870-6

Singh, V.P.E., 2012. Hydrology of Disasters, Springer Science & Business Media. https://doi.org/10.1007/978-94-015-8680-1

Slater, L.J., Thirel, G., Harrigan, S., Delaigue, O., Hurley, A., Khouakhi, A., Prodoscimi, I., Vitolo, C., Smith, K., 2019. Using R in hydrology: a review of recent developments and future directions. Hydrol. Earth Syst. Sci. Discuss. 1–33. https://doi.org/10.5194/hess-2019-50

Sorensen, E.J., 2005. Vegetable Crops, EM4830E.

Spinoni, J., Barbosa, P., De Jager, A., McCormick, N., Naumann, G., Vogt, J. V., Magni, D., Masante, D., Mazzeschi, M., 2019. A new global database of meteorological drought events from 1951 to 2016. J. Hydrol. Reg. Stud. 22, 100593. https://doi.org/10.1016/j.ejrh.2019.100593

Stagge, J.H., Kingston, D.G., Tallaksen, L.M., Hannah, D.M., 2017. Observed drought indices show increasing divergence across Europe. Sci. Rep. 7, 1–10. https://doi.org/10.1038/s41598-017-14283-2

Stenseth, N.C., Ottersen, G., Hurrell, J.W., Mysterud, A., Lima, M., Chan, K.S., Yoccoz, N.G., Ådlandsvik, B., 2003. Studying climate effects on ecology through the use of climate indices: The North Atlantic Oscillation, El Niño Southern Oscillation and beyond. Proc. R. Soc. B-Biological Sci. 270, 2087–2096. https://doi.org/10.1098/rspb.2003.2415

Sun, L., Mitchell, S.W., Davidson, A., 2012. Multiple drought indices for agricultural drought risk assessment on the Canadian prairies. Int. J. Climatol. 32, 1628–1639. https://doi.org/10.1002/joc.2385

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. Rev. Geophys. 56, 79–107. https://doi.org/10.1002/2017RG000574

Svodova, M., Funchs, B.A., Integrated Drought Management Programme (IDMP), 2016. Handbook of drought indicators and indices, Drought Mitigation Center Faculty Publications. 117. https://doi.org/10.1007/s00704-016-1984-6

Tan, R., Perkowski, M., 2015. Wavelet-Coupled Machine Learning Methods for Drought Forecast Utilizing Hybrid Meteorological and Remotely-Sensed Data, in: Proceedings Conference on Data Mining, DMIN15. Http://Archives.Pdx.Edu/Ds/Psu/19370.

Tapley, B.D., Bettadpur, S., Ries, J., Thompson, P.F., Watkins, M.M., 2004. Grace measurements of Mass variability in the Earth system: supporting online material. Science (80-.). 503–505, 503–506. https://doi.org/10.1126/science.1099192

Tester, M., Langridge, P., 2010. Breeding Technologies to Increase. Science (80-.). 818, 818–822. https://doi.org/10.1126/science.1183700

Thomas, A.C., Reager, J.T., Famiglietti, J.S., Rodell, M., 2014. A GRACE-based water storage deficit approach for hydrological drought characterization. Geophys. Res. Lett. Res. 41, 3307–3314. https://doi.org/10.1002/2014GL061184.Received

Tian, L., Yuan, S., Quiring, S.M., 2018. Evaluation of six indices for monitoring agricultural drought in the south-central United States. Agric. For. Meteorol. 249, 107–119. https://doi.org/10.1016/j.agrformet.2017.11.024

Tigkas, D., Vangelis, H., Tsakiris, G., 2019. Drought characterisation based on an agriculture-oriented standardised precipitation index. Theor. Appl. Climatol. 135, 1435–1447. https://doi.org/10.1007/s00704-018-2451-3

Torrence, C., Compo, G.P., 1997. A Practical Guide to Wavelet Analysis, American Meteorological Society. https://doi.org/10.1016/j.biopha.2017.10.142

Trambauer, P., Maskey, S., Werner, M., Pappenberger, F., Van Beek, L.P.H., Uhlenbrook, S., 2014. Identification and simulation of space-time variability of past hydrological drought events in the Limpopo River basin, southern Africa. Hydrol. Earth Syst. Sci. 18, 2925–2942. https://doi.org/10.5194/hess-18-2925-2014

Trambauer, P., Werner, M., Winsemius, H.C., Maskey, S., Dutra, E., Uhlenbrook, S., 2015. Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa. Hydrol. Earth Syst. Sci. 19, 1695–1711. https://doi.org/10.5194/hess-19-1695-2015

Trapletti, A., Hornik, K., LeBaron, B., 2020. Package ‘ tseries ’. version 0.10-48. https://cran. r-project. org/web/packages/tseries/tseries. pdf.

Uele, D.I., Lyra, G.B., de Oliveira Júnior, J.F., 2017. Variabilidade Espacial e Intranual das Chuvas na Região Sul de Moçambique, África Austral. Rev. Bras. Meteorol. 32, 473–484. https://doi.org/10.1590/0102-77863230013

United Nations, 2009. Global Assessment Report on Disaster Risk Reduction:Risk and poverty in a changing climate 1–197. https://doi.org/10.1037/e522342010-005

Van Loon, A.F., 2015. Hydrological drought explained. Wiley Interdiscip. Rev. Water 2, 359–392. https://doi.org/10.1002/wat2.1085

Vicario, L., García, C.M., Teich, I., Bertoni, J.C., Ravelo, A., Rodríguez, A., 2015. Characterization of meteorological droughts in the central Argentina. Tecnol. y Ciencias del Agua 6, 153–165.

Vicente-Serrano, S.M., 2006a. Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula. Water Resour. Manag. 20, 37–60. https://doi.org/10.1007/s11269-006-2974-8

Vicente-Serrano, S.M., 2006b. Spatial and temporal analysis of droughts in the Iberian Peninsula (1910-2000). Hydrol. Sci. J. 51, 83–97. https://doi.org/10.1623/hysj.51.1.83

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

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., Angulo, M., El Kenawy, A., 2010b. A New Global 0.5° Gridded Dataset (1901–2006) of a Multiscalar Drought Index: Comparison with Current Drought Index Datasets Based on the Palmer Drought Severity Index. J. Hydrometeorol. 11, 1033–1043. https://doi.org/10.1175/2010JHM1224.1

Vicente-Serrano, S.M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J.J., López-Moreno, J.I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., Sanchez-Lorenzo, A., 2012. Performance of drought indices for ecological, agricultural, and hydrological applications. Earth Interact. 16. https://doi.org/10.1175/2012EI000434.1

Vicente-serrano, S.M., Beguerı, S., Camarero, J., Lorenzo-lacruz, J., Azorin-molina, C., Lo, J.I., 2012. Performance of Drought Indices for Ecological , Agricultural , and Hydrological Applications. Earth Interact. 16. https://doi.org/10.1175/2012EI000434.1

Vicente-Serrano, S.M., Martínez, E.A.R., Aguilar, E., Martínez, R., Martín-Hernández, N., Azorin-Molina, C., Sanchez-Lorenzo, A., Kenawy, A. El, Tomás-Burguera, M., Moran-Tejeda, E., López-Moreno, J.I., Revuelto, J., Beguería, S., Nieto, J.J., Drumond, A., Gimeno, L., Nieto, R., 2017. The complex influence of ENSO on droughts in Ecuador. Clim. Dyn. 48, 405–427. https://doi.org/10.1007/s00382-016-3082-y

Wang, B., Feng, P., Waters, C., Cleverly, J., Liu, D.L., Yu, Q., 2020. Quantifying the impacts of pre-occurred ENSO signals on wheat yield variation using machine learning in Australia. Agric. For. Meteorol. 291, 108043. https://doi.org/10.1016/j.agrformet.2020.108043

Wang, K., Li, T., Wei, J., 2019. Exploring drought conditions in the three river headwaters region from

2002 to 2011 using multiple drought indices. Water (Switzerland) 11. https://doi.org/10.3390/w11020190

Wang, Y., Liu, G., Guo, E., 2019. Spatial distribution and temporal variation of drought in Inner Mongolia during 1901–2014 using Standardized Precipitation Evapotranspiration Index. Sci. Total Environ. 654, 850–862. https://doi.org/10.1016/j.scitotenv.2018.10.425

WFP, 2007. World Food Programme. Be part of the solution. Rome, Italy.

Wilhite, D.A., Glantz, M.H., 1985. Understanding: The drought phenomenon: The role of definitions. Water Int. 10, 111–120. https://doi.org/10.1080/02508068508686328

Wilhite, D.A., Sivakumar, M.V.K., Wood, D.A., 2000. Early Warning Systems for Drought Preparedness and Drought Management, in: Proceedings of an Expert Group Meeting Held 5-7 September, 2000, in Lisbon, Portugal. Http://Www.Wamis.Org/Agm/Pubs/Agm2/Agm02.Pdf. pp. 182–199.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, Second. ed. Academic Press, San Diego, CA.

Willmott, C.J., Rowe, C.M., Mintz, Y., 1985. Climatology of the terrestrial seasonal water cycle. J. Climatol. 5, 589–606. https://doi.org/10.1002/joc.3370050602

WMO, 2006. Drought monitoring and early warning : concepts , progress and future challenges. World Meteorol. Organ. 24.

World Meteorological Organization, 2008. Guide to Hydrological Practices. Volume I: Hydrology– From Measurement to Hydrological Information, WMO-No. 168. Geneva. https://doi.org/10.1080/02626667.2011.546602

World Meteorological Organization and Global Water Partnership, 2016. Handbook of Drought Indicators and Indices (M. Svoboda and B.A. Fuchs). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2, Geneva. https://doi.org/10.1201/9781315265551-12

Yao, N., Li, Y., Lei, T., Peng, L., 2018. Drought evolution, severity and trends in mainland China over 1961–2013. Sci. Total Environ. 616–617, 73–89. https://doi.org/10.1016/j.scitotenv.2017.10.327

Yevjevich, V., 1969. An objective approach to definitions and investigations of continental hydrologic droughts. J. Hydrol. https://doi.org/10.1016/0022-1694(69)90110-3

Yihdego, Y., Vaheddoost, B., Al-Weshah, R.A., 2019. Drought indices and indicators revisited. Arab. J. Geosci. 12. https://doi.org/10.1007/s12517-019-4237-z

You, L., Wood, S., Wood-Sichra, U., 2009. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. Agric. Syst. 99, 126–140. https://doi.org/10.1016/j.agsy.2008.11.003

Yue, S., Ouarda, T.B.M.J., Bobée, B., Legendre, P., Bruneau, P., 1999. The Gumbel mixed model for flood frequency analysis. J. Hydrol. 226, 88–100. https://doi.org/10.1016/S0022-1694(99)00168-7

Zambreski, Z.T., Lin, X., Aiken, R.M., Kluitenberg, G.J., Pielke, R.A., 2018. Identification of hydroclimate subregions for seasonal drought monitoring in the U.S. Great Plains. J. Hydrol. 567, 370–381. https://doi.org/10.1016/j.jhydrol.2018.10.013

Zampieri, M., Ceglar, A., Dentener, F., Toreti, A., 2017. Wheat yield loss attributable to heat waves , drought and water excess at the global , national and subnational scales. Environ. Res. Lett. 12, 1293. https://doi.org/10.1088/1748-9326/aa723b

Zargar, A., Sadiq, R., Naser, B., Khan, F.I., 2011. A review of drought indices. Environ. Rev. 19, 333–349. https://doi.org/10.1139/a11-013

Zeleke, T.T., Giorgi, F., Diro, G.T., Zaitchik, B.F., 2017. Trend and periodicity of drought over Ethiopia. Int. J. Climatol. 37, 4733–4748. https://doi.org/10.1002/joc.5122

Zelenhasić, E., Salvai, A., 1987. A method of streamflow drought analysis. Water Resour. Res. 23, 156–168. https://doi.org/10.1029/WR023i001p00156

Zhou, W., Guan, K., Peng, B., Shi, J., Jiang, C., Wardlow, B., Pan, M., Kimball, J.S., Franz, T.E., Gentine, P., He, M., Zhang, J., 2020. Connections between the hydrological cycle and crop yield in the rainfed U.S. Corn Belt. J. Hydrol. 590, 125398. https://doi.org/10.1016/j.jhydrol.2020.125398

Appendix A:

Extended summary of the thesis in Spanish

Resumen extendido de la tesis en español

Resumen extendido de la tesis

Las sequías son uno de los fenómenos naturales más comunes en todo el mundo, y pueden ocurrir en cualquier lugar bajo cualquier régimen (Bryant et al., 2005; Sheffield and Wood, 2012). Sin embargo, es uno de los peligros menos comprendidos debido a su complejidad y difícil cuantificación (Hagenlocher et al., 2019a). Las sequías pueden ser percibidas por los impactos en las cadenas de producción de energía, alimentos, suministro de agua, etc., siendo generalmente detectadas cuando las consecuencias del fenómeno son difíciles de mitigar, por lo que es el desastre que más pérdidas socioeconómicas provoca a nivel mundial (WMO, 2006). Los impactos en la agricultura están directamente relacionados con la seguridad alimentaria y la calidad de vida de gran parte de la población mundial. Según Cumani and Rojas, (2016), un tercio de la población se dedica directamente a la agricultura. Por lo tanto, se necesitan estudios específicos sobre la sequía y sus impactos a nivel regional o local para apoyar a los responsables de la gestión de la sequía y a todas las partes interesadas.

Las sequías suelen evaluarse por eventos, que pueden cuantificarse y compararse. Las características y extensión territorial de un evento de sequía pueden definirse según la "run theory" (Yevjevich, 1969). Estas características pueden entonces relacionarse con los impactos correspondientes, por ejemplo en la agricultura (Yihdego et al., 2019). Tal y como afirma el IPCC, (2014) existen tendencias de aumento de la intensidad y la frecuencia de las sequías en todo el mundo, siendo posiblemente las zonas áridas y semiáridas las más afectadas. Países como Mozambique y Argentina son propensos a posibles impactos del cambio climático que pueden traer consigo fenómenos meteorológicos extremos más graves que afecten a los sectores naturales y socioeconómicos. Por lo tanto, en las regiones particularmente expuestas y vulnerables a la sequía, se necesitan estudios específicos sobre la sequía para ayudar a mitigar sus impactos. Esta tesis es una contribución a la gestión de las sequías y sus impactos, específicamente en la agricultura.

El estudio de la sequía consta de diferentes componentes, empezando por la caracterización de la sequía en una zona determinada, es posible considerar el seguimiento, el análisis, la predicción, la visualización y la evaluación de los impactos de uno o varios eventos de sequía. Según Wilhite et al., (2000), el proceso de estudio de las sequías en una zona determinada (país o cuenca hidrográfica) debe empezar por tener una división espacial en regiones según las características de la sequía, luego debe desarrollarse un sistema de vigilancia de la sequía para crear alertas tempranas de las condiciones de sequía emergentes y, por último, la predicción de la sequía. En este último paso, es necesario comprender los factores climáticos que desencadenan los fenómenos de sequía en la región y utilizar esta información como herramienta de previsión.

Para analizar los diferentes componentes de un estudio de sequía, se han desarrollado varios índices de sequía (World Meteorological Organization and Global Water Partnership, 2016). Los índices se calculan utilizando como insumos variables climáticas e hidrológicas. Por ejemplo, la precipitación (McKee et al., 1993), la humedad del suelo (Hao and AghaKouchak, 2013), el estado de la vegetación (Kogan, 1995) y la escorrentía (Mo, 2008). También se han creado algunos índices multivariantes que han dado buenos resultados. Por ejemplo, utilizando la precipitación y la evapotranspiración (Vicente-Serrano et al., 2010a), el estado de la vegetación y la temperatura del suelo (Kogan, 1995), y el almacenamiento en los embalses y la demanda de agua (Mehran et al., 2015). Los índices de sequía son herramientas muy poderosas que identifican los déficits de agua a lo largo del tiempo y controlan las características de los eventos pasados y presentes. Observar la variabilidad climática, de la vegetación y de los cultivos es otro hecho que los índices detectan con una precisión sobresaliente, aun cuando la sequía es un fenómeno muy complejo.

Por tanto, para la correcta aplicación de los índices de sequía, la gestión óptima de la misma requiere series temporales de datos hidrometeorológicos para cada componente del ciclo hidrológico. Esta tarea implica que los insumos utilizados en el cálculo de los índices de sequía requieren una calidad de alto nivel. Según la World Meteorological Organization, (2008) se necesitan al menos 30 años de variables

capaces con una distribución espacial de mediciones in situ de al menos una cada 250km^2. Estas condiciones son especialmente difíciles de cumplir en los países en desarrollo o en determinadas regiones del mundo debido a las condiciones socioeconómicas, medioambientales y de otro tipo (Easterling, 2013). La razón principal suele ser el elevado coste de la puesta en marcha y el mantenimiento de la instrumentación necesaria para medir las variables necesarias para calcular los índices de sequía en grandes territorios, por lo que en muchas zonas del mundo la disponibilidad de datos es extremadamente limitada.

Para obtener datos como la precipitación, la humedad del suelo, el caudal de los arroyos, etc., varias instituciones han desarrollado bases de datos mundiales en las que ofrecen estas series temporales de interés hidrológico y climático de forma abierta y gratuita. Estos productos se basan en datos puntuales aforados (medidos) que han sido interpolados mediante técnicas estadísticas; en resultados de modelos climáticos e hidrológicos globales; y, en información de teledetección procedente de radares de satélite. Estos conjuntos de datos (casi) globales, además de ofrecer productos que cumplen con la calidad requerida para los estudios de sequía, han demostrado ser una valiosa fuente de datos oportunos y espacialmente continuos con información mejorada sobre el seguimiento climático e hidrológico (Araneda-Cabrera et al., 2020; Funk et al., 2014; Harris et al., 2020).

El objetivo principal de la tesis fue desarrollar metodologías, a nivel técnico y procedimental, para la mejora de la gestión de los recursos hídricos y de la sequía que puedan ser replicadas en cualquier región del mundo. Las investigaciones se realizaron en Mozambique y Argentina como casos de estudio. En un entorno de cooperación y desarrollo a través de proyectos en los que la Universidad da Coruña participa ("Proyecto de Implementación de índices de sequías como herramientas de vigilancia y alerta temprana para el desarrollo de planes de preparación y gestión de sequías en el Norte de Mozambique (Proyecto SECARA)" y el "Fortalecimiento de capacidades técnicas y operativas para la mejora de la gestión de los recursos hídricos en Mozambique, Aqua-Moz": Fase 1 y 2.), algunas acciones se centraron en el fortalecimiento de capacidades a través de la implementación y creación de herramientas técnicas y de gestión. Además, a lo largo del desarrollo de la tesis, se consideró el uso de fuentes de datos hidrometeorológicos alternativos de libre acceso, y que las metodologías puedan ser aplicadas a otros países o regiones a cualquier escala espacial.

Este objetivo principal se desglosa en objetivos más específicos acorde a la necesidad de comprender las características espaciotemporales de las sequías y sus desencadenantes, a la validación de las herramientas para el seguimiento de las sequías y sus impactos con un enfoque en la agricultura, y que los beneficiarios locales se conviertan progresivamente en los protagonistas de la transferencia de conocimientos a todos los implicados en la gestión de la sequía. Las investigaciones se trataron entonces en la caracterización de las condiciones espaciotemporales y las tendencias de la sequía, en identificar la influencia de los factores climáticos a gran escala que desencadenan eventos de sequía, en comparar y validar con los registros históricos de sequía, en analizar la variabilidad temporal entre varios índices de sequía para identificar aquellos que están fuertemente correlacionados y que potencialmente proporcionan información redundante, en desarrollar modelos estadísticos de cultivos basados en series temporales para explicar y posteriormente pronosticar los rendimientos nacionales utilizando los índices de sequía como predictores, en proponer un nuevo índice para mejorar la evaluación y el seguimiento de las sequías agrícolas, en satisfacer la necesidad de obtener rápida y eficazmente los datos necesarios para la gestión y seguimiento de la sequía en cualquier región, y en crear una herramienta probada para la gestión de la sequía.

Mozambique (Fig. A.1) es uno de los países más pobres del mundo. Su localización es el sur de África en una de las zonas más propensas a la sequía del mundo (Eriksen and Silva, 2009; 2014: Climate Change IPCC, 2014; Osbahr et al., 2008; Patt and Schröter, 2008). Según la Base de Datos Internacional de Desastres (EM-DAT, 2021), Mozambique ha sufrido varios eventos de sequía que han afectado a más de 20 millones de personas. Los eventos más importantes se registraron en 1982-1983, 1991-1993 y 2016. La población total del país era de más de 30 millones de habitantes en diciembre de 2019 según

el Instituto Nacional de Estadística ("Instituto Nacional de Estatística", www.ine.gov.mz), de los cuales más del 60% viven en zonas rurales y tienen su principal actividad económica relacionada con la agricultura (Ministério da Agricultura e Segurança Alimentar, 2015). De hecho, históricamente los agricultores mozambiqueños han experimentado pérdidas recurrentes debido a la escasez de agua y a la variabilidad climática, y las personas más vulnerables de las zonas rurales están especialmente expuestas a este fenómeno climático (Eriksen and Silva, 2009).

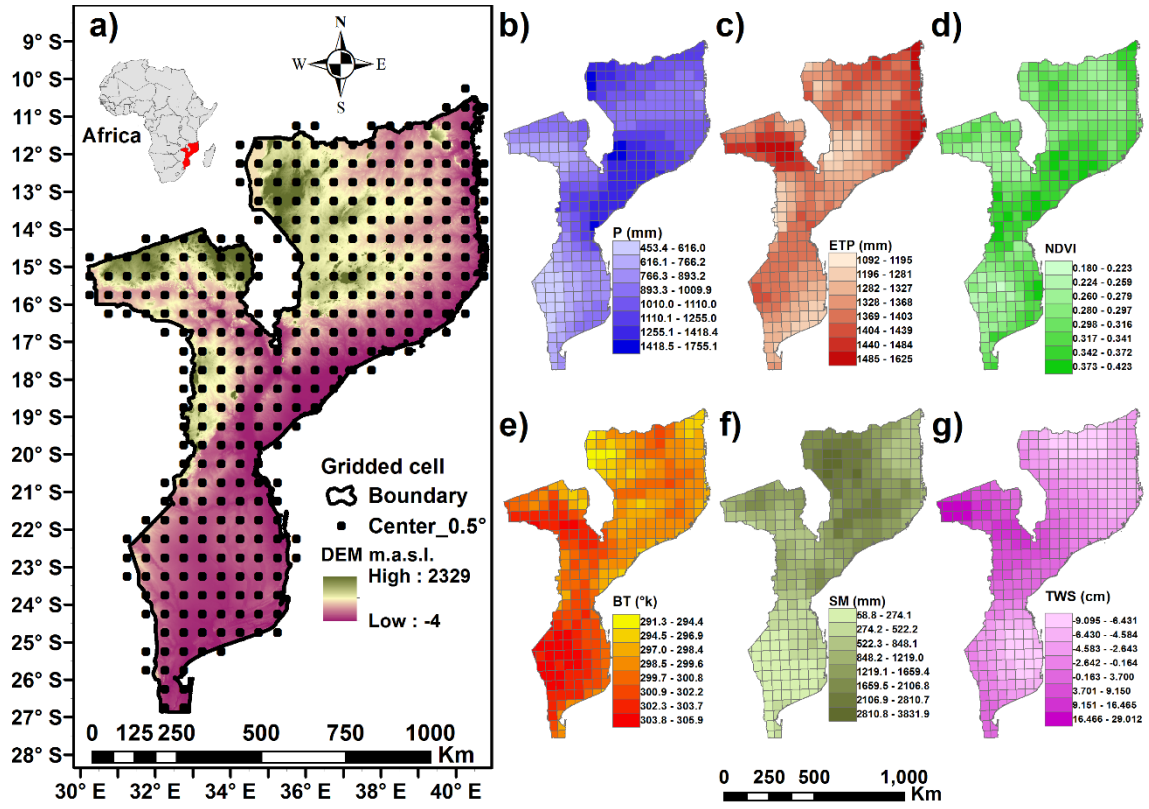


Fig. A.1 (a) Ubicación de Mozambique en África y su topografía. Los puntos negros ilustran los puntos de la cuadrícula del CRU (0,5◦ × 0,5◦). Distribución espacial de los valores medios anuales de: (b) precipitación (1973-2017), (c) evapotranspiración potencial (ETP, 1973-2017), (d) índice de vegetación de diferencia normalizada (NDVI, 1983-2017), (e) temperatura de brillo (BT, 1983-2017), (f) humedad del suelo (1973-2017), y (g) almacenamiento de agua terrestre (TWS, 2002-2017), en todo el país.

Mozambique, a nivel científico, es un escenario muy complejo para realizar estudios específicos sobre la sequía. El país no cuenta con una red de medición operativa que cumpla con los criterios óptimos establecidos por Easterling, (2013). Además, a pesar de ser un país altamente dependiente de la mano de obra agrícola de secano, los datos sobre los rendimientos agrícolas son limitados. En general, las investigaciones previas sobre los impactos de la sequía son escasas, por lo que las metodologías propuestas durante el desarrollo de esta tesis debían ser probadas y validadas en otros escenarios donde se dispone de datos de mejor calidad, pero donde se requieren estudios específicos de sequías similares para contribuir a la gestión del agua. Se eligió Argentina como caso de estudio complementario adecuado para esta tesis.

Argentina tiene una superficie de 2791810 km^2 y está dividida en cinco regiones administrativas principales según el Instituto Nacional de Estadística y Censos de la República Argentina (INDEC: www.indec.gob.ar), 24 provincias (incluyendo la Ciudad Autónoma de Buenos Aires como provincia) y 525 departamentos (Fig. A.2). Argentina tiene la mayor producción de cultivos per cápita del mundo (FAO, 2019) y al mismo tiempo presenta una importante variabilidad climática anual e interanual (Barros and Silvestri, 2002). Estos hechos hacen que la probabilidad de que los rendimientos de los

cultivos en Argentina se reduzcan debido a la sequía oscile hasta el 80% en algunos casos (Leng and Hall, 2019). Argentina es uno de los principales productores mundiales de cereales (FAO, 2017), que se cultivan en gran medida en la Pampa argentina y más de la mitad de ellos son de secano (Cherlet et al., 2018). El valor de la producción de cereales de Argentina fue de 10 200 millones de USD en 2013, lo que representa el 8,3 % de su PIB (FAO, 2017). El país es vulnerable a varios fenómenos naturales (terremotos, inundaciones, etc.); sin embargo, las sequías representan el mayor riesgo de pérdidas agrícolas (Cherlet et al., 2018). Además, los factores desencadenantes de los eventos de sequía en Argentina sólo se han estudiado en áreas específicas (provincias o cuencas hidrográficas), pero no a nivel nacional o regional de la producción agrícola.

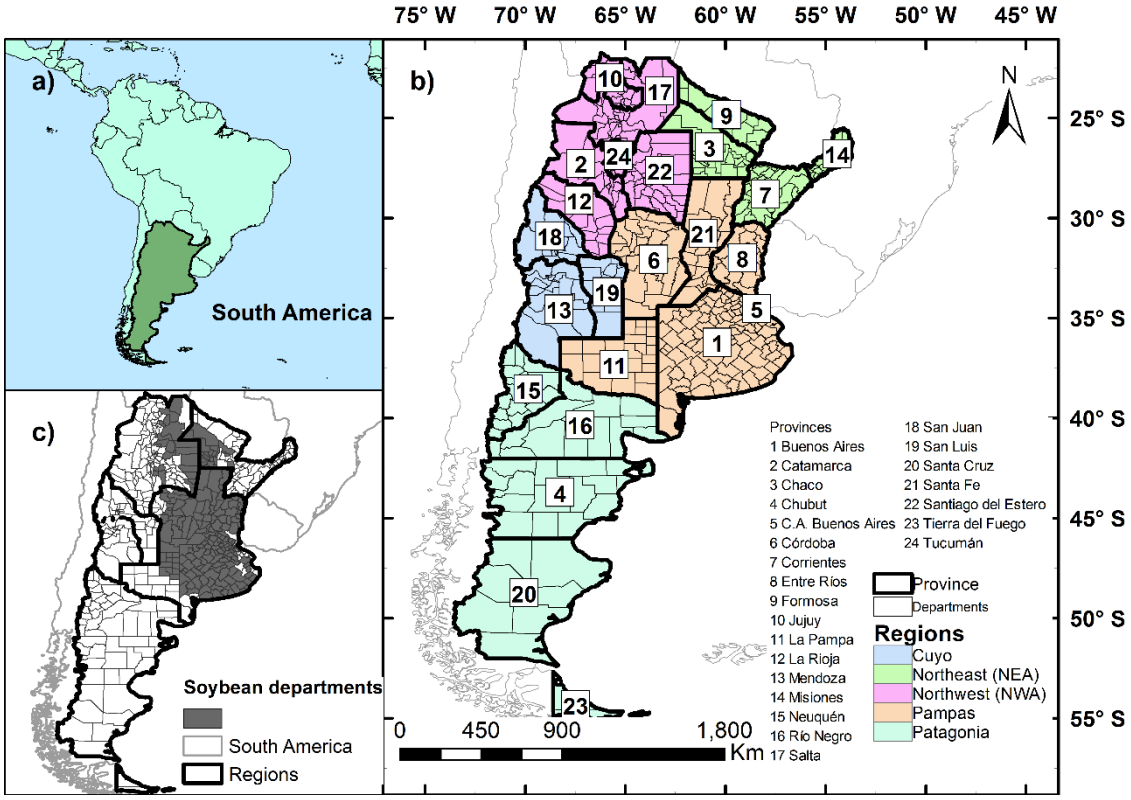


Fig. A.2 Ubicación de a) Argentina; b) divisiones por regiones, provincias y departamentos; y c) departamentos con producción de soja.

A diferencia de Mozambique, Argentina dispone de una base de datos agrícolas mejor y más fiable. El Ministerio de Agricultura, Ganadería y Pesca de Argentina documenta los rendimientos anuales de los cultivos a nivel departamental (subdivisión administrativa de segundo orden de las provincias). Esta base de datos incluye la superficie sembrada, la superficie cosechada y la producción total de 30 cultivos diferentes desde 1961 hasta la fecha. Sin embargo, Argentina, al igual que Mozambique, tampoco cumple los requisitos óptimos de las bases de datos climáticos e hidrológicos establecidos. Por estas razones, se ha elegido este país como segundo caso de estudio para el desarrollo de esta tesis, con el fin de comparar y validar las metodologías aplicadas en Mozambique, a la vez que aportar y crear conocimiento adicional para los gestores del agua y la sequía en Argentina.

La disertación de esta tesis se presenta en la modalidad de compendio de trabajos científicos. Esta subsección resume los principales resultados de la tesis. Los resultados de este trabajo han sido publicados en cuatro revistas internacionales, tres de las cuales están indexadas por el Journal Citation Report (Journal of Hydrology: Regional Studies, Agriculture Water Management y Science of the Total Environment) y la cuarta por el Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal, Latindex (Revista Hidrolatinoamericana). Al mismo

tiempo, parte de los resultados aquí presentados forman parte de un manuscrito para una comunicación en un congreso internacional y de un manual de acceso abierto (libro). Cada capítulo de la tesis se organiza de forma que cada uno de ellos pueda considerarse como un estudio individual, incluyendo su propio estado del arte, metodología, resultados y conclusiones.

El estudio presentado en el Capítulo 2 tenía como objetivo caracterizar las condiciones y tendencias de la sequía en Mozambique y evaluar la influencia de los principales factores climáticos como desencadenantes de eventos de sequía. Las condiciones de sequía se estudiaron utilizando el Índice de Precipitación y Evapotranspiración Estandarizado (SPEI) y la "run theory". Se aplicó la técnica de análisis de componentes principales y el método de agrupación de k-means para definir regiones de sequía homogéneas. Para definir las características temporales de la sequía se utilizaron la prueba de tendencia de Mann-Kendall y el análisis estadístico Rescaled Range. El método de correlación cruzada, un análisis espectral basado en la transformada rápida de Fourier y un análisis Cross-Wavelet se utilizaron para identificar posibles indicadores climáticos como desencadenantes de la sequía. Se utilizaron varios índices climáticos para analizar su relación con las sequías.

Se descargaron los datos mensuales de precipitación y evapotranspiración potencial de la Unidad de Investigación Climática (CRU) (http://www.cru.uea.ac.uk/data) para calcular el SPEI-12. El CRU proporciona series temporales climáticas mensuales con una resolución de 0,5˚ (≈ 55 km en el ecuador) en todo el mundo (Harris et al., 2014). Con esta resolución, se calcularon un total de 343 series temporales de SPEI en todo Se definieron 3 regiones homogéneas de sequía situadas en el norte, el sur y el centro del país (Fig. A.3).

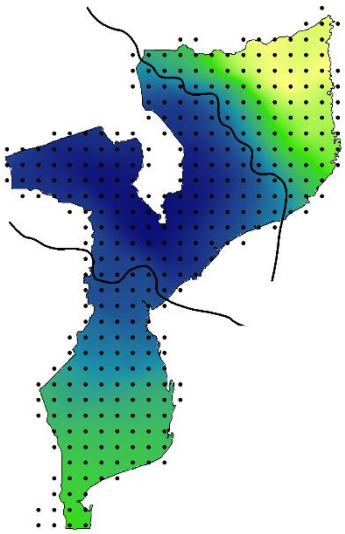


Fig. A.3 Regiones homogéneas de sequía definidas en Mozambique. Los puntos son los centroides de las celdas según la cuadrícula del CRU.

En cada una de las regiones se han identificado y enumerado los eventos de sequía más importantes, que han coincidido con los principales a nivel nacional, siendo las regiones Sur y Centro las que han tenido los eventos de sequía más intensos y severos en el pasado. En las tres regiones, las tendencias del SPEI fueron negativas, aunque en la región Norte fueron estadísticamente no significativas al nivel del 5% para el índice anual. En cuanto a la persistencia, el índice de Hurst (H) sugirió que las tendencias negativas a largo plazo persistirán en un futuro próximo en todo el país, especialmente en las regiones Sur y Central.

Los índices climáticos se basan en las fluctuaciones de la presión atmosférica a nivel del mar (SLP) de diferentes lugares del mundo (índices Darwin SLP, Tahití SLP, SOI y NAO), y la temperatura de la

superficie del mar (SST) del Atlántico (índices TNA, TSA, NAT, SAT y TASI), del Pacífico (índices ENSO Niño 1 +2, Niño 3, Niño 4, Niño 3 4 y PDO) y el Océano Índico (índices SWIO, WTIO, SETIO y DMI). Según las correlaciones cruzadas, las anomalías de las series temporales de los índices climáticos fueron negativas con Darwin, Niño 3.4, Niño 3, Niño 4, WTIO y SETIO, y positivas con el SOI, mostrando picos alrededor de 2-3 meses antes del pico negativo del SPEI. La región Norte mostró correlaciones persistentemente bajas con los índices climáticos analizados, mientras que las correlaciones más fuertes se obtuvieron en las regiones Sur y Central.

De acuerdo con los análisis espectrales (Fig. A.4), los índices climáticos con mejor correlación (Darwin SLP y El Niño 4) y la serie SPEI tienen una periodicidad asociada a energías altas entre 40 y 120 meses (3,5 y 10 años). En la región norte, se encontraron periodos entre 35 y 60 meses (3 y 5 años). Estos periodos son consistentes y similares a los reportados en otros estudios en África (Oguntunde et al., 2018). Tanto el SLP de Darwin como los eventos de El Niño 4 mostraron fuertes impactos en las series mensuales de SPEI, especialmente en las regiones Sur y Central, indicando que juegan un papel relevante en las características de la evolución de la sequía en estas regiones (algo que ya se había visto con las correlaciones).

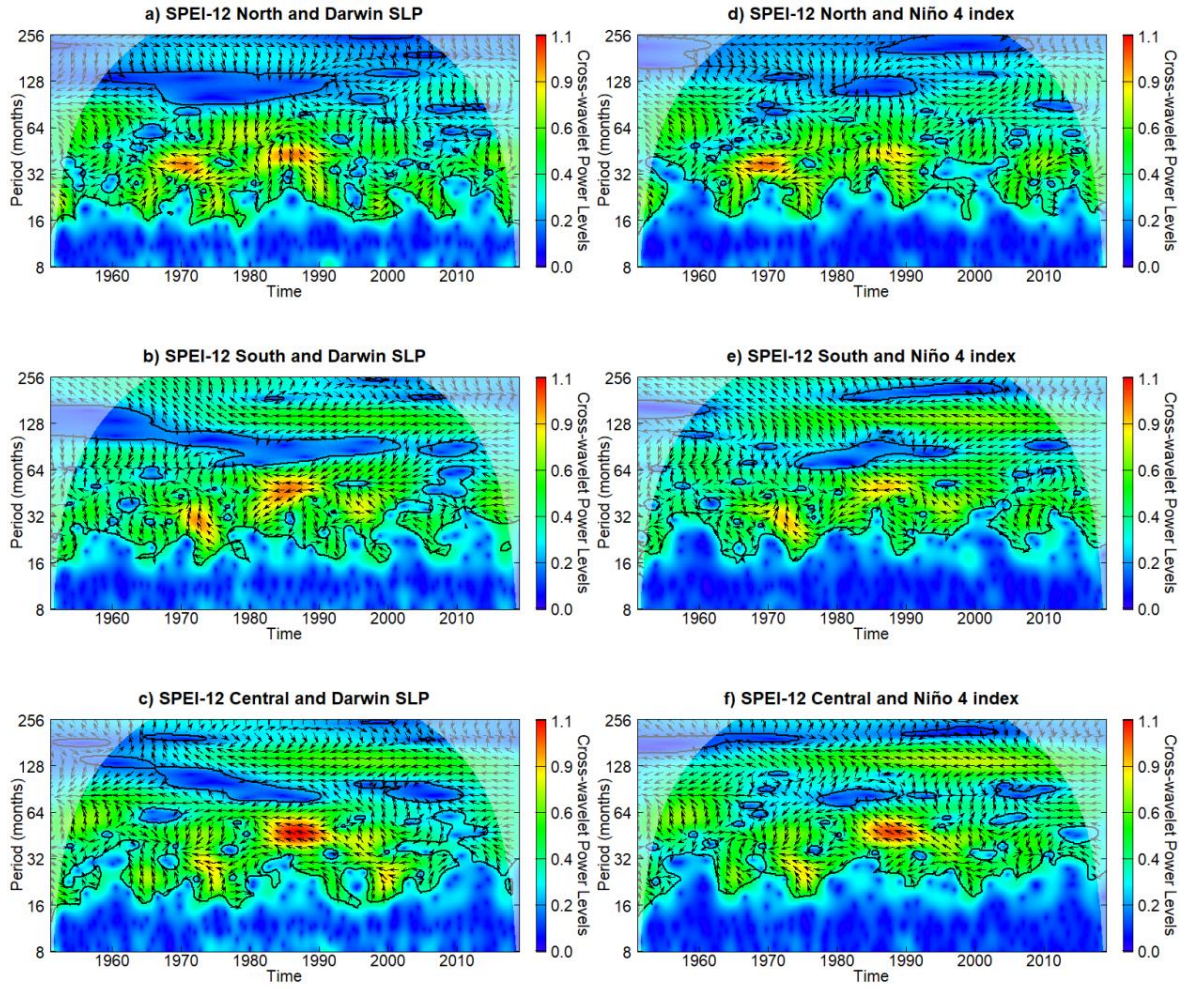


Fig. A.4 Transformada cross wavelet entre el SPEI de las regiones a) Norte, b) Sur, c) Centro y el índice climático Darwin PSL; y de las regiones d) Norte, e) Sur, f) Centro y el índice climático Niño 4. El eje y equivale a los periodos definidos con la Transformada rápida de Fourier (Periodo = 1/Frecuencia); la barra de color denota la densidad de energía (rojo más alta densidad de energía); el nivel de confianza del 5% contra el ruido rojo se muestra en un contorno con la línea negra gruesa; y la relación de fase relativa se representa con flechas (con la anti-fase señalando a la izquierda, la en-fase señalando a la derecha).

En el Capítulo 3 se propone una metodología para identificar los índices de sequía y las fuentes de datos más adecuados para el seguimiento de las sequías y su impacto en los cultivos. Se utiliza Mozambique como caso de estudio, ya que representa un ejemplo desafiante debido a sus características y limitaciones. Se calcularon un total de siete indicadores de sequía estandarizados: Índice de Precipitación Estandarizado (SPI), Índice de Precipitación y Evapotranspiración Estandarizado (SPEI), Índice de Humedad del Suelo Estandarizado (SSI), Índice de Condición de la Vegetación Estandarizado (SVCI), Índice de Condición de la Temperatura Estandarizado (STCI), Índice de Salud de la Vegetación Estandarizado (SVHI) y Almacenamiento de Agua Terrestre Estandarizado (STWS) a diferentes escalas (1, 3, 6 y 12 meses) a partir de bases de datos globales basadas en mediciones in situ y satelitales: el CRU, el Centro de Aplicaciones e Investigación por Satélite y los satélites ambientales de la Administración Oceánica y Atmosférica de Estados Unidos (NOAA STAR), el conjunto de datos TerraClimate (Abatzoglou et al., 2018) y los satélites GRACE (Tapley et al., 2004).

Estos índices se compararon y evaluaron como herramientas de gestión de la sequía y como potenciales predictores de la variabilidad anual de los rendimientos agrícolas a nivel nacional. Se utilizó un modelo estadístico de los rendimientos de los cultivos basado en series temporales para medir el poder explicativo de cada índice. Utilizando métricas de rendimiento basadas en la probabilidad de detección (POD) o tasa de aciertos, y la probabilidad de falsa detección (POFD) o tasa de falsas alarmas (Wilks, 2006) se compararon los registros históricos de sequía del EMDAT (EM-DAT, 2021) y del IRI (Hellmuth et al., 2007) con las series temporales de los valores medios nacionales de los índices de sequía y el porcentaje de superficie afectada por las sequías. Ambas series mostraron una capacidad de detección similar a la de los registros históricos. En general, los índices meteorológicos (SPEI y SPI) fueron los que mejor se ajustaron a los registros históricos, seguidos del índice hidrológico SSI (Fig. A.5).

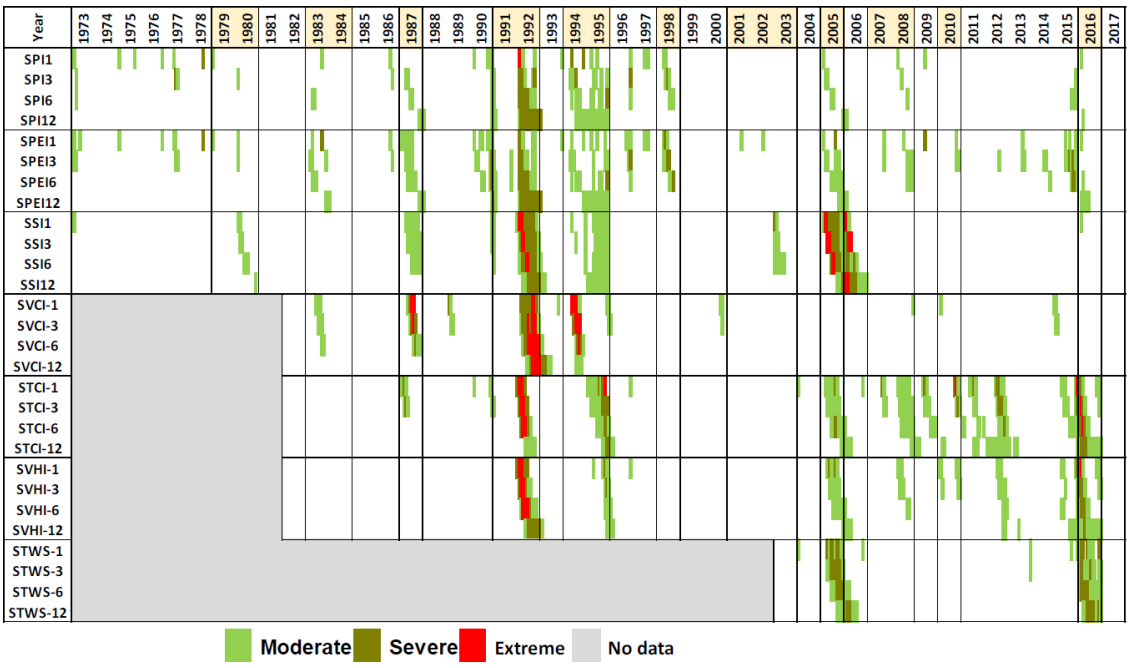


Fig. A.5 Evolución temporal mensual de SPI, SPEI, SSI, SVCI, STCI, SVHI y STWS a escala nacional en Mozambique (agregaciones de -1, -3, -6 y -12 meses). Los niveles de intensidad pueden interpretarse según McKee et al., (1993). Los años de sequía históricos según los registros se destacan en amarillo.

Posteriormente, se realizó un análisis de correlación entre los índices de sequía, que mostró la coherencia con el significado físico de los diferentes índices. Aunque existe un desfase temporal entre el forzamiento meteorológico y las respuestas hidrológicas, los índices basados en variables meteorológicas con acumulaciones de 6 y 12 meses mostraron fuertes correlaciones con los índices

basados en datos de vegetación/hidrológicos. Los resultados sugieren que sólo uno de los índices meteorológicos SPI y SPEI con acumulaciones de 6 y 12 meses podría utilizarse solo en cualquier análisis posterior. El SCVI puede utilizarse, pero teniendo en cuenta que no tiene una correlación significativa con los indicadores hidrológicos y climáticos, debería utilizarse junto con otro índice. Los resultados observados en esta sección validaron en primer lugar el uso del SPEI en el Capítulo 1 como índice de sequía representativo de las condiciones hidrometeorológicas en Mozambique.

Con un modelo estadístico de rendimiento de los cultivos nacionales, se supuso que el rendimiento anual de 12 cultivos (maíz, mijo, sorgo, trigo, anacardos, yuca, patatas, caña de azúcar, té, tabaco y hortalizas) era la respuesta de una función de variables independientes, que en este caso eran los índices de sequía y los porcentajes de superficie afectada por la sequía según las diferentes categorías definidas en McKee et al., (1993). Los resultados indicaron que los mejores predictores candidatos eran diferentes para cada cultivo (Fig. A.6). Esto se debe a que no todos los cultivos son igualmente sensibles a la sequía, ni tienen la misma capacidad de recolección o almacenamiento de agua. Para la mayoría de los cultivos, la variabilidad fue explicada por dos indicadores genéricos: SPEI-3 para los cereales (maíz, mijo y sorgo) y SSI-12 para otros cultivos (anacardos, yuca, patatas, té, tabaco y hortalizas). El SPEI-6 es el que mejor explica la variabilidad del trigo y el SVCI-6 la de la caña de azúcar. Los mejores predictores de la variabilidad del rendimiento agrícola fueron los que incorporaban información espacial, ya que los índices nacionales perdían información espacial en su cálculo.

Los resultados presentados en los capítulos 2 y 3 motivaron la investigación presentada en el Capítulo 4. Por un lado, se probaron las correlaciones entre los índices de sequía y los índices climáticos en otros escenarios a escalas espaciales más pequeñas y se compararon como posibles predictores de la variabilidad de los cultivos. Por otro lado, dado que en general se observó que la introducción de la desagregación espacial mejoraba los resultados, se consideró necesario validar el modelo estadístico con una desagregación de la evolución de los cultivos por áreas, asumiendo a priori que los resultados serían mucho más precisos. El estudio se aplicó en Argentina y se evaluaron ocho índices de sequía, siete ya conocidos (SPI, SPEI, SSI, SVCI, SVCI, STCI, SVHI, STWS) y adicionalmente un octavo, el Índice Estandarizado de Severidad de Sequía de Palmer (SPDSI) (Ma et al., 2014), todos ellos computados a escalas de 3, 6 y 12 meses. Se relacionaron con 19 índices climáticos (similares a los usados en el Capítulo 2 en Mozambique) y se comparó su rendimiento para explicar la variabilidad de la producción de soja. Se eligió la soja como cultivo de secano representativo de los impactos de la sequía en la agricultura.

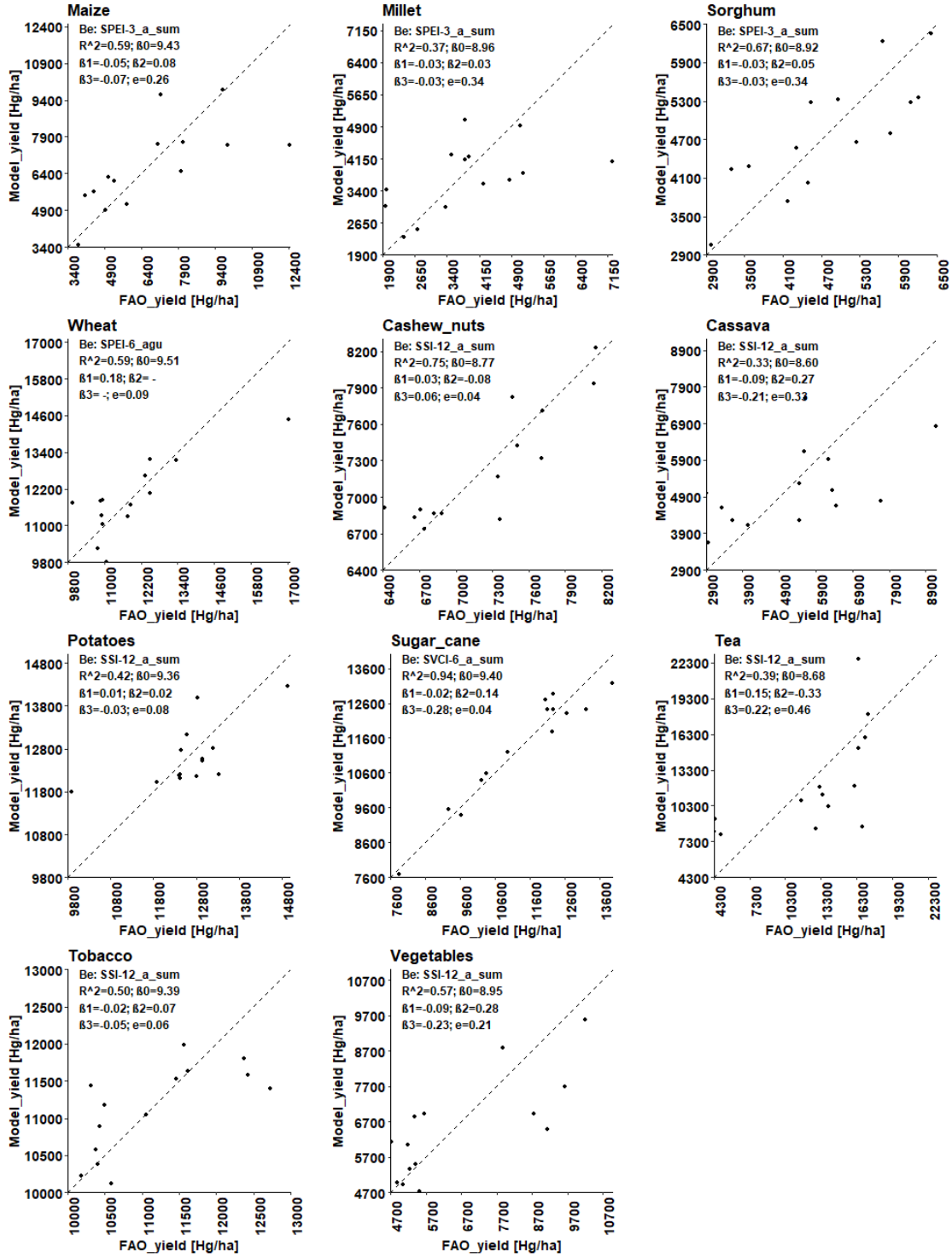


Fig. A.6 Rendimiento de los cultivos medido por la FAO frente al rendimiento de los cultivos calculado con la mejor variable explicativa candidata (indicada como Be). También se muestran los parámetros ajustados. La línea discontinua corresponde a la línea 1:1.

Se analizaron las correlaciones de Pearson (r) entre los índices climáticos y los índices de sequía agregados 12 meses a nivel departamental en todo el país. Luego, se realizó un análisis estacional (verano, otoño, invierno y primavera) con los índices de sequía agregados de 6 meses y un número determinado de índices climáticos seleccionados en el paso anterior. El análisis estacional se realizó para todo el país y para todos los departamentos productores de soja, priorizando el análisis de verano por ser el periodo de siembra y crecimiento de la soja. Los índices climáticos ubicados en el Océano Pacífico ecuatorial como Tahití, SOI, El Niño 3.4 (Fig. A.7) y El Niño 4 mostraron mayores correlaciones con los índices de sequía en Argentina. Se ha encontrado que estos índices indican el desencadenamiento de sequías en otras partes del mundo (Gupta and Jain, 2021; Vicente-Serrano et al., 2017), demostrando la importancia global de su variabilidad en los eventos hidrológicos extremos.

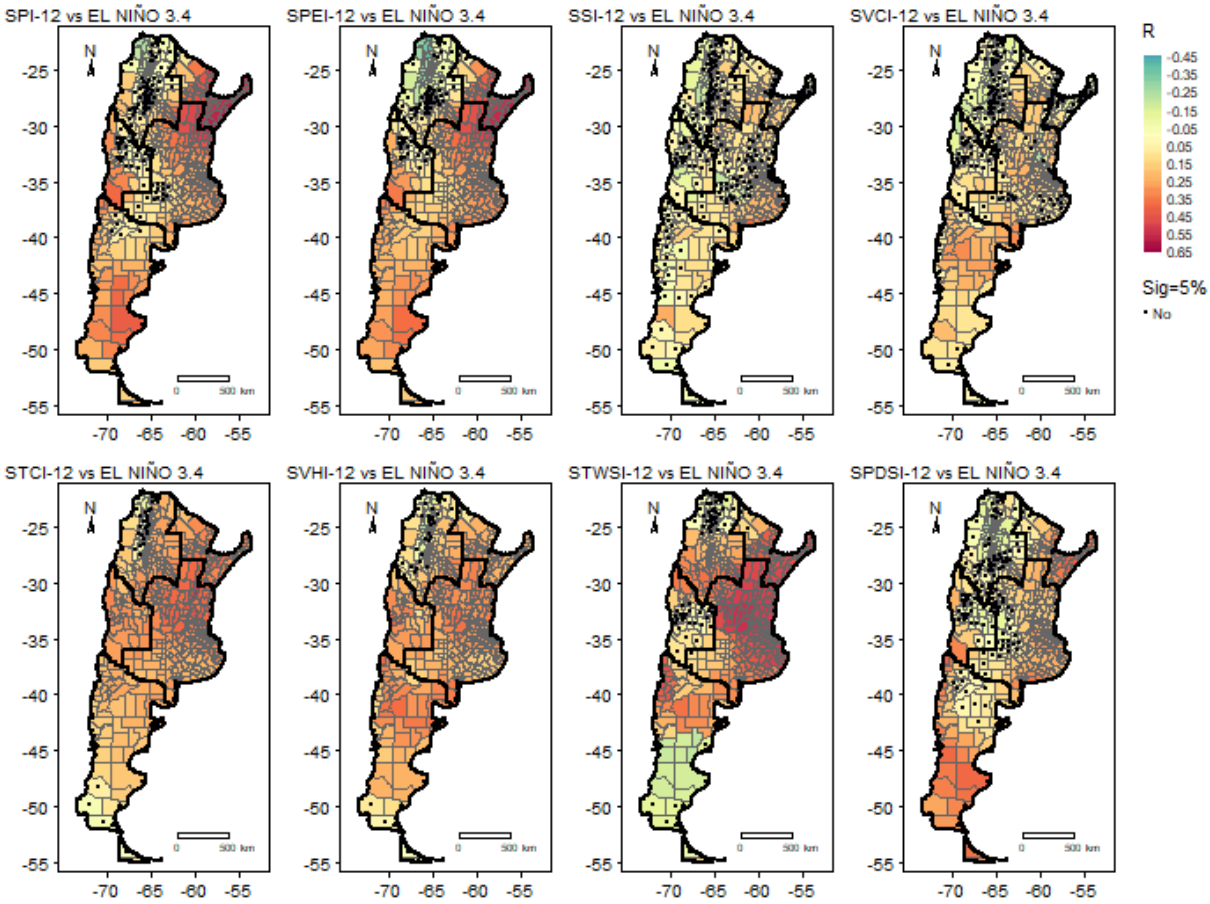


Fig. A.7 Patrón espacial de las correlaciones entre los 8 índices de sequía con un periodo de agregación de 12 meses y el índice climático El Niño 3.4 basado en los datos de 1982-2019 (STWSI 2003-2019). Los puntos negros indican los casos en que las correlaciones no fueron estadísticamente significativas.

Se observaron correlaciones negativas persistentes (con muy pocas excepciones) durante la primavera y el verano y correlaciones positivas durante el otoño y el invierno entre El Niño 3.4 y los índices de sequía. Se observaron patrones similares para los otros tres índices seleccionados (Tahití, SOI, El Niño 3.4 y El Niño 4). Espacialmente, las regiones con correlaciones débiles tuvieron el mayor número de departamentos con correlaciones no significativas, y este número aumentó en las estaciones de invierno y primavera. Por ejemplo, la Fig. A.8 ilustra los patrones espaciales de las correlaciones entre SPEI-6 y El Niño 3.4. Considerando sólo los departamentos productores de soya y la temporada de suma de verano, cuando ocurre la siembra y el crecimiento de la soya, SPEI y STWSI tuvieron las correlaciones más fuertes.

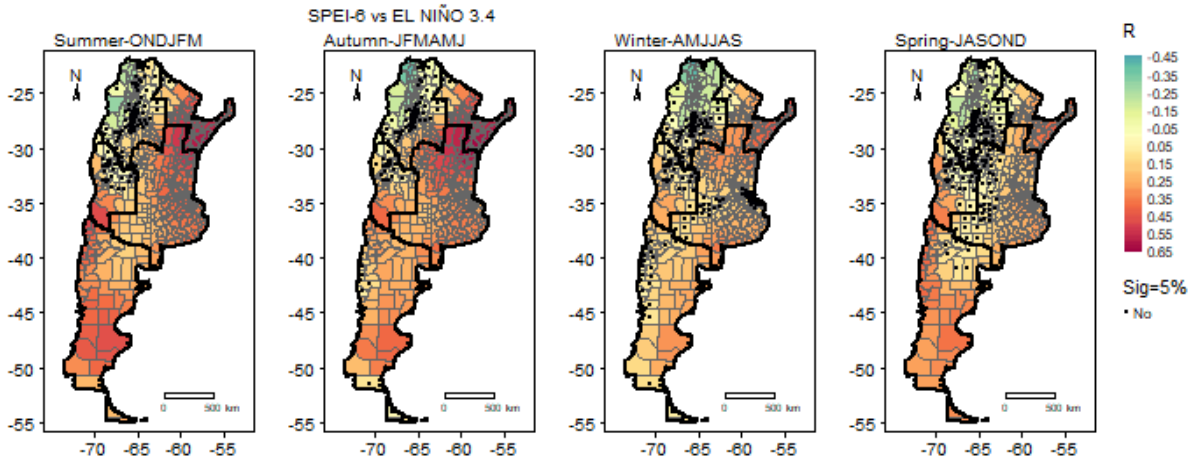


Fig. A.8 Patrón espacial de las correlaciones entre las series temporales estacionales SPEI-6 y El Niño 3.4 basado en datos de 1982-2019.

A continuación, se evaluó la eficacia de los índices de sequía y de los índices climáticos para explicar la variabilidad anual de los rendimientos de la soja, entendidos como impactos de las sequías agrícolas, mediante modelos de regresión estadística. El SVHI, el SVCI, el STCI, el SPEI y el SPI, que se basan tanto en variables meteorológicas como de vegetación, fueron los que mejor explicaron la variabilidad de los rendimientos de la soja en los tres modelos aplicados (series temporales, modelo de panel y modelos de sección transversal). Las distribuciones espaciales de los coeficientes de determinación (R^2) utilizando el modelo de series temporales con predictores agregados de 6 meses se muestran en la Fig. A.9. La variabilidad del rendimiento de la soja (impactos asociados a las sequías agrícolas) respondió mejor a los índices de sequía que a los índices meteorológicos. El SVHI y el SPEI agregados para 6 meses y correspondientes al mes de marzo (temporada de cultivo de la soja) resultaron ser los que mejor explican el estado de la producción de soja en las regiones seleccionadas.

De acuerdo con los resultados obtenidos en los primeros capítulos en las secciones relacionadas con el análisis de los índices de sequía como predictores de la variabilidad anual de los cultivos en Mozambique y Argentina, y siguiendo los objetivos planteados, en el Capítulo 5 se proponen nuevos índices para mejorar la evaluación y seguimiento de las sequías agrícolas a través de la explicación de la variabilidad anual de los cultivos. Se trata de dos índices de sequía compuestos multivariantes (bivariantes) que tienen en cuenta las condiciones meteorológicas y de sequía agrícola, combinando en un marco probabilístico el SPI (a través de la precipitación) con el VCI y el VHI. La metodología para validar y comparar los nuevos índices consistió en utilizar los modelos estadísticos de cultivos empleados en los capítulos 3 y 4.

El método desarrollado para caracterizar las sequías agrícolas fue un enfoque multivariante que depende de dos variables individuales (Mehran et al., 2015): la precipitación acumulada 3 meses y el VCI o VHI. Con la formulación de Gringorten, (1963) se calculan las probabilidades de no superación de ambas variables y se combinan utilizando el marco multivariante explicado en (Yue et al., 1999). Las probabilidades empíricas se transforman en un índice estandarizado y se obtienen los índices multivariantes estandarizados (o bivariantes) de precipitación y estado de la vegetación y salud (MSPVI y MSPHI). Estos índices se interpretan de forma similar al SPI original (McKee et al., 1993).

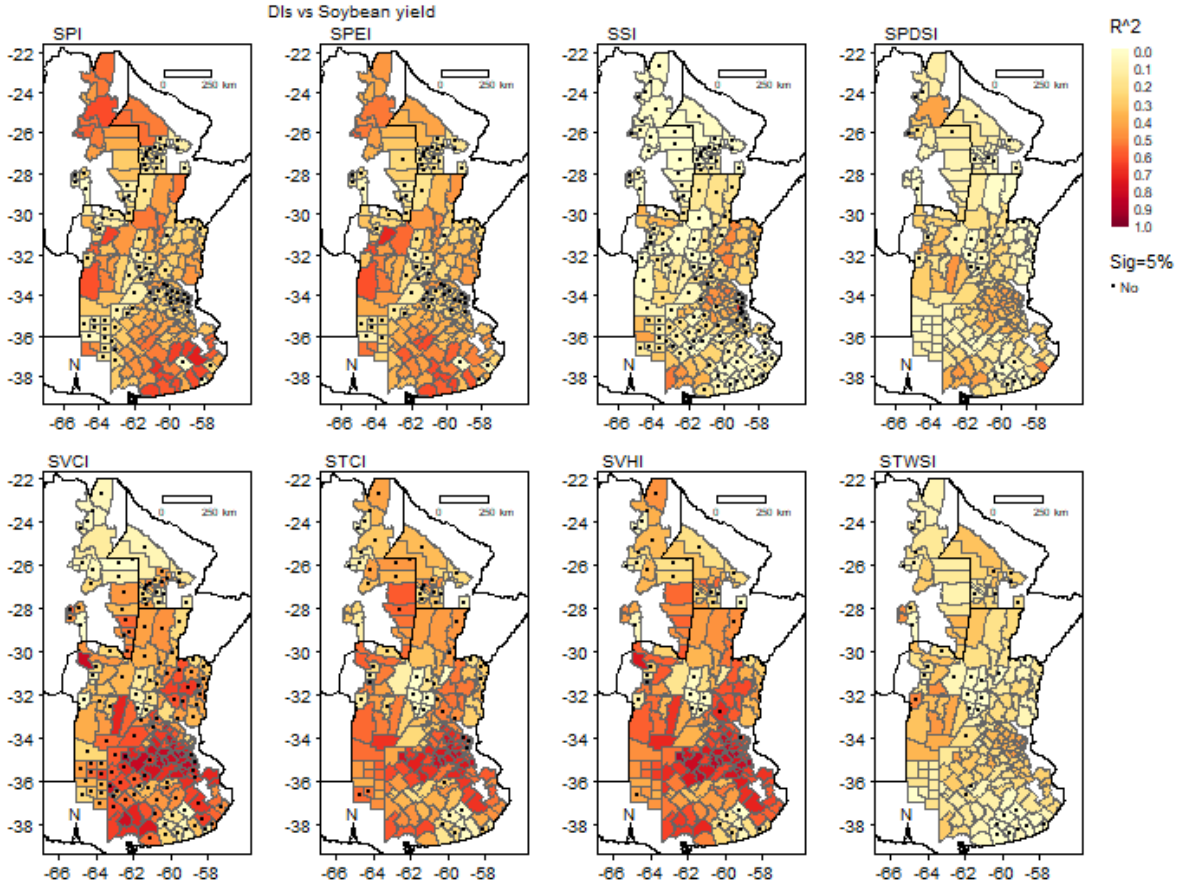


Fig. A.9 Patrones espaciales de los resultados de los coeficientes de determinación (R^2) de los modelos de series temporales entre los índices de sequía agregados de 6 meses (ONDJFM) y el rendimiento de la soja en base a los departamentos productores de soja y los datos de 2004-2019.

La investigación se llevó a cabo en Mozambique y Argentina. En el primer caso, se utilizaron datos agrícolas de maíz, soja y trigo de la FAO y de precipitaciones del CRU; en el segundo, se emplearon datos de cultivos de soja y maíz del Ministerio de Agricultura, Ganadería y Pesca de Argentina y precipitaciones del TerraClimate. Para ambos casos, el VCI y el VHI se obtuvieron de la NOAA STAR.

Los patrones temporales de los índices propuestos (caso de Mozambique para el MSPVCI se muestra en la Fig. A.10) indican que los índices muestran información crítica sobre el inicio y la recuperación de los eventos de sequía. Los nuevos índices detectan eventos de sequía importantes según los registros históricos de la base de datos internacional de Desastres EM-DAT, donde, por ejemplo, el periodo 1991-1992 fue uno de los eventos de sequía más dañinos (en Mozambique). Dentro de este periodo, el MSPVCI informó de casi el 50% del territorio mozambiqueño bajo sequía extrema [índice < -2,0 según McKee et al., (1993)] mientras que el SVCI informó del 22% y el SPI del 11% (la distribución espacial del evento de sequía de 1992 en mayo se representa en la Fig. A.11). El MSPVHI mostró patrones espaciotemporales muy similares a los del MSPVCI, lo que indica que los índices mostraron una mayor sensibilidad a estos eventos y que utilizarlos para vigilar las sequías significaría estar en el lado seguro (Monteleone et al., 2020).

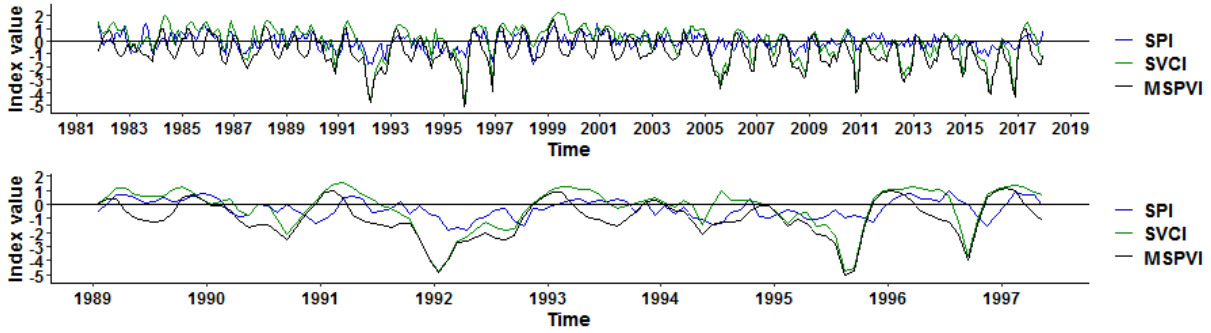


Fig. A.10 Las series SPI-3, SVCI y MSPVCI en 1982-2017 (arriba) y 1989-1997 (abajo) en Mozambique.

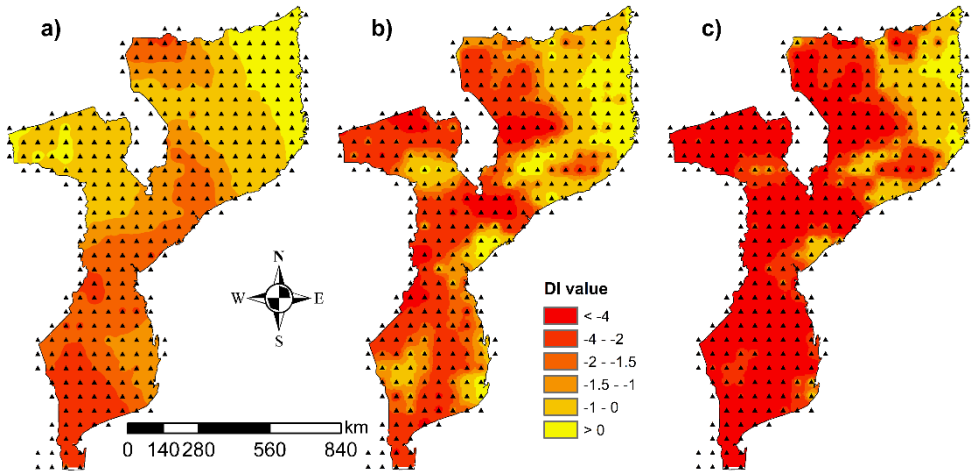


Fig. A.11 Distribución espacial del a) SPI-3, b) SVCI y c) MSPVCI durante el evento de sequía de mayo de 1992.

En la comparación de los índices propuestos con los convencionales para explicar la variabilidad anual de los cultivos en ambos territorios, los resultados fueron similares (Argentina y Mozambique). En el caso argentino, los resultados de la aplicación del modelo estadístico se presentan en la Tabla A.1 y se representan en la Fig. A.12. En comparación con el SPI clásico y el SVCI/SVHI, los índices multivariados reportaron explicar mejor la variabilidad de los rendimientos anuales de maíz y soja (en Mozambique sorgo, maíz y trigo).

Tabla A.1 Coeficientes de determinación (R^2) entre los rendimientos de la soja y el maíz y los índices de sequía clásicos 2004-2019.

	SPI	SVCI	SVHI
Soja	0.476	0.572	0.686
Maíz	0.306	0.407	0.670

Finalmente, para concluir esta tesis, como material final de la investigación realizada, en el Capítulo 6 se presenta un manual titulado "Manual para la obtención y aplicación de variables hidrometeorológicas de bases de datos globales en la gestión de la sequía". Este trabajo recoge una serie de conceptos y técnicas relacionadas con la gestión de la sequía que se han desarrollado a lo largo de la tesis que se aplican a través de varios scripts (códigos de programación) codificados en el Software R (RStudio Team, 2016). Estos scripts divididos en 3 secciones muestran el paso a paso para i) descargar, extraer, manipular y guardar las variables hidrometeorológicas de las bases de datos globales; ii) utilizar y aplicar estas variables para la gestión de la sequía; y iii) presentar los resultados a través de gráficos en diferentes formatos. Se utilizó Mozambique como ejemplo de estudio de caso.

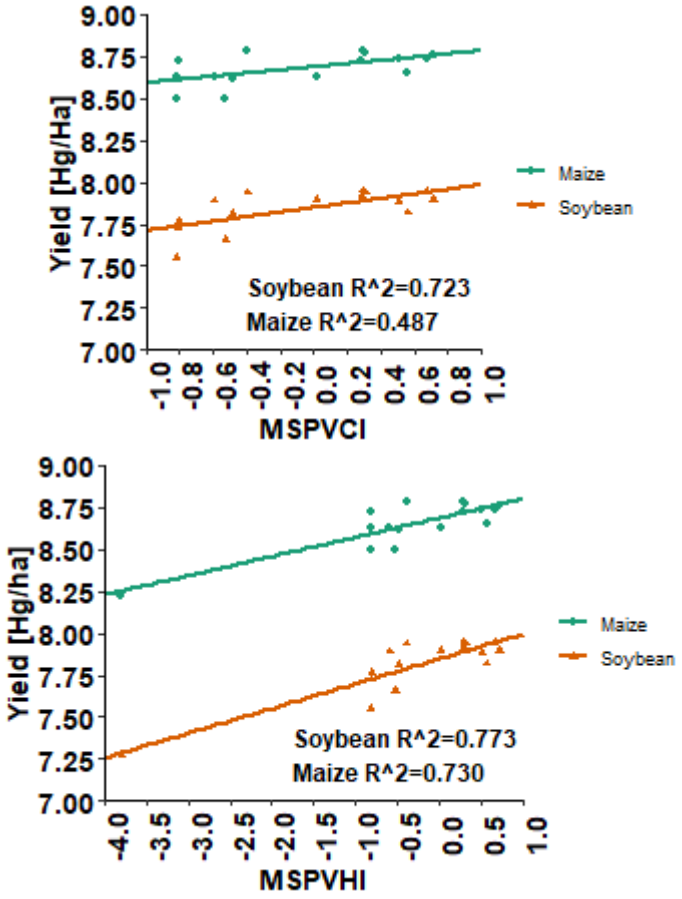


Fig. A.12 Regresión lineal y coeficiente de determinación (R2) entre los rendimientos anuales de soja y maíz con MSPVCI (arriba) y MSPVHI (abajo) en Argentina.

Como resultado adicional de este trabajo, el doctorando impartió un curso para estudiantes de Ingeniería Civil, Geografía e Ingeniería Ambiental de la Universidade Eduardo Mondlane - ESUDER en Maputo (capital de Mozambique) y a los funcionarios de las Administraciones Regionales del Agua del Sur (Administração Regional da Água do Sur) - ARA Sur de Mozambique. Los principales objetivos del curso eran enseñar conceptos y herramientas relacionados con la vigilancia de la sequía; obtener y manipular datos hidrometeorológicos en Mozambique; y aplicar balances hídricos e índices de sequía en Mozambique. El curso se realizó en línea durante 12 horas en noviembre de 2021 (Fig. A.13).

Como conclusión, en esta tesis se desarrollaron varias metodologías con el objetivo de aumentar el conocimiento de los fenómenos de la sequía y aportar soluciones para la gestión de los recursos hídricos y de la sequía. Se consideró el uso de fuentes de datos hidrometeorológicos alternativos de libre acceso, de manera que las metodologías pueden aplicarse a cualquier país o región del mundo y a (casi) cualquier escala espacial. Los casos de estudio fueron Mozambique y Argentina, ya que son territorios muy propensos a la sequía, sufren continuamente sus efectos, especialmente en la agricultura, y no disponen de las herramientas de gestión necesarias para controlar y gestionar el fenómeno. Las metodologías respondieron a la necesidad de entender las características espaciotemporales de las sequías y sus desencadenantes; a la validación de herramientas para el monitoreo de las sequías y sus impactos con un enfoque especial en la actividad agrícola; y, de acuerdo con la necesidad de que los beneficiarios locales se conviertan progresivamente en los protagonistas de la transferencia de conocimiento a todos los involucrados en la gestión de la sequía.

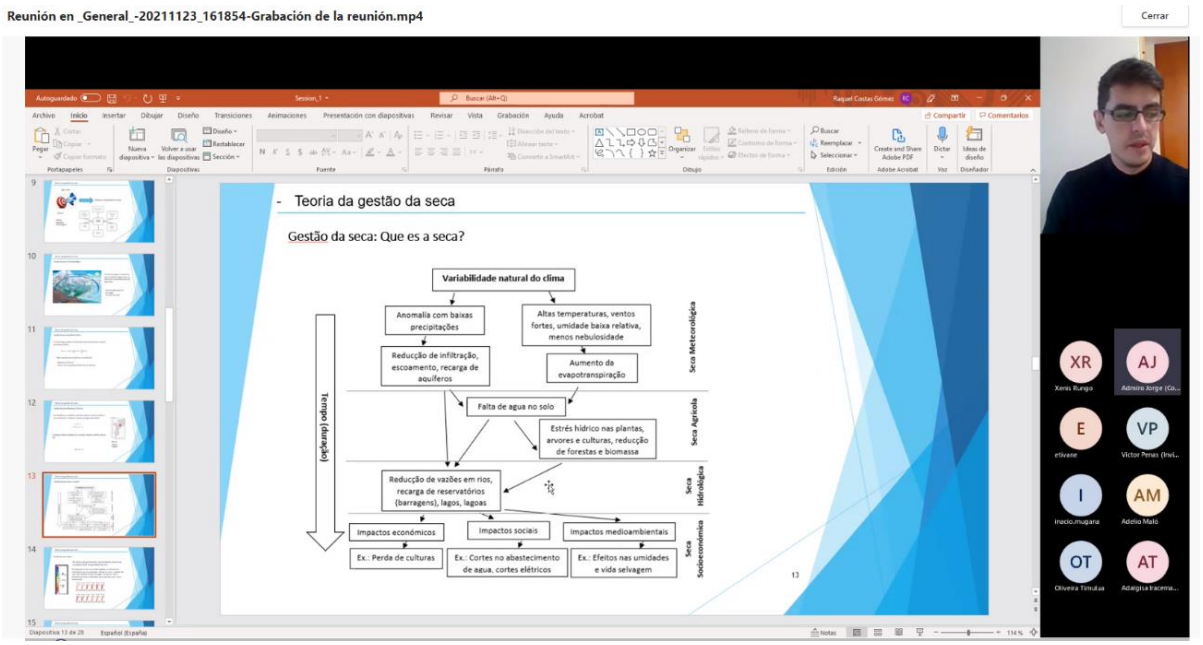


Fig. A.13 Captura de pantalla de la segunda sesión del curso online "Obtenção e aplicação de variáveis hidrometeorológicas de bancos de dados globais para a monitororia de secas" ("Obtención y aplicación de variables hidrometeorológicas de bancos de datos globales para la monitorización de sequías") a estudiantes de la Universidade Eduardo Mondlane - ESUDER y empleados públicos de la ARA Sur de Mozambique, impartida por el doctorando.

Los estudios futuros que se proponen a partir de los resultados y conclusiones generados en esta tesis están asociados principalmente a la predicción de la sequía y sus impactos. Estos estudios podrían complementarse considerando varios aspectos que constituyen un reto actual para los científicos, técnicos y decisores en la gestión de la sequía. Por ejemplo, el cambio climático, la gran variedad de nuevas bases de datos y las diferentes fuerzas institucionales y gubernamentales. Este amplio abanico de posibles investigaciones futuras podría seguir generando cursos y documentos como el manual que se presenta en esta tesis, para que el conocimiento y las herramientas sean accesibles, utilizados y compartidos por los diferentes actores implicados no sólo a nivel nacional sino también local.

Referencias

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2017.191

Araneda-Cabrera, R.J., Bermúdez, M., Puertas, J., 2020. Unified framework for drought monitoring and assessment in a transboundary river basin, in: River Flow 2020; Uijttewaal, W, Franca M, Valero D, Chavarrias V, Arbós C, Schielen R and Crosato A, Eds. Taylor & Francis Group, London, pp. 1081–1086. https://doi.org/10.1201/b22619

Barros, V.R., Silvestri, G.E., 2002. The relation between sea surface temperature at the subtropical South-Central Pacific and precipitation in Southeastern South America. J. Clim. 15, 251–267. https://doi.org/10.1175/1520-0442(2002)015<0251:trbsst>2.0.co;2

Bryant, E.A., Head, L.M., Morrison, R., 2005. Planning for Natural Hazards — How Can We Mitigate the Impacts ?, in: R.J. Morrison, S. Quin and E.A. Bryant (Eds.), Planning for Natural Hazards — How Can We Mitigate the Impacts?, Proceedings of a Symposium, 2-5 February 2005, University of Wollongong, GeoQuEST Research Centre, 2005, 1-11.

Cherlet, M., Hutchinson, Charles Reynolds, J., Hill, J., Sommer, S., Von Maltitz, G., 2018. World atlas of desertification. Rethinking land degradation and sustainable land management., Third Edit. ed, Publications Office of the European Union. https://doi.org/10.2760/06292

Cumani, M. (FAO), Rojas, O. (FAO, 2016. Characterization of the agricultural drought prone areas on global a global scale at scale. Using the FAO Agricultural Stress Index System (ASIS) to enhance the understanding of, and boost resilience to, water stress conditions in drought-prone areas.

Easterling, D.R., 2013. Global Data Sets for Analysis of Climate Extremes., in: Water Science and Technology Library (Ed.), Global Data Sets for Analysis of Climate Extremes. Springer, Dordrecht, pp. 347–361. https://doi.org/doi:10.1007/978-94-007-4479-0_12

EM-DAT, 2021. The Emergency Events Database, Universite catholique de Louvain, Brussels, Belgium. [Available at http://www. emdat.be/, Accessed November 2020.].

Eriksen, S., Silva, J.A., 2009. The vulnerability context of a savanna area in Mozambique: household drought coping strategies and responses to economic change. Environ. Sci. Policy 12, 33–52. https://doi.org/10.1016/j.envsci.2008.10.007

FAO, 2019. World Food and Agriculture - Statistical Pocketbook 2019, Food and Agriculture Organization of the United Nations. Rome. https://doi.org/10.4060/ca1796en

FAO, 2017. Country fact sheet on food and agriculture policy trends. Argentina, Food and Agriculture Organization of the United Nations. I7752EN/1/08.17.

Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J.D., Romero, B.E., Husak, G.J., Michaelsen, J.C., Verdin, A.P., 2014. A Quasi-Global Precipitation Time Series for Drought Monitoring. U.S. Geol. Surv. Data Ser. 832, 4.

Gringorten, I.I., 1963. A Plotting Rule for Extreme Probability Paper. J. Geophys. Res. 68, 813–814. https://doi.org/doi:10.1029/JZ068i003p00813

Gupta, V., Jain, M.K., 2021. Unravelling the teleconnections between ENSO and dry/wet conditions over India using nonlinear Granger causality. Atmos. Res. 247, 105168. https://doi.org/10.1016/j.atmosres.2020.105168

Hagenlocher, M., Meza, Isabel, Carl Anderson, Annika Min, Fabrice G. Renaud, Y., Walz, S.S., Sebesvari, Z., 2019. Drought vulnerability and risk assessments: state of the art, persistent gaps, and research agenda. Environ. Res. Lett. 4. https://doi.org/https://doi.org/10.1088/1748-9326/ab225d Manuscript

Hao, Z., AghaKouchak, A., 2013. Multivariate Standardized Drought Index: A parametric multi-index model. Adv. Water Resour. 57, 12–18. https://doi.org/10.1016/j.advwatres.2013.03.009

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. data 7, 109. https://doi.org/10.1038/s41597-020-0453-3

Hellmuth, M.E., Moorhead, A., Thomson, M.C., Williams, J., 2007. Climate Risk Management in Africa: Learning from Practice, International Research Institute for Climate and Society (IRI). Columbia University, New York, USA.

IPCC, 2014. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III

to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, Cambridge University Press. Cambridge, United Kingdom and New York, NY, USA.

IPCC, 2014: Climate Change, 2014. Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp. https://doi.org/10.1017/CBO9781107415324

Kogan, F.N., 1995. Application of vegetation index and brightness temperature for drought detection. Adv. Sp. Res. 15, 91–100. https://doi.org/10.1016/0273-1177(95)00079-T

Leng, G., Hall, J., 2019. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. Sci. Total Environ. 654, 811–821. https://doi.org/10.1016/j.scitotenv.2018.10.434

Ma, M., Ren, L., Yuan, F., Jiang, S., Liu, Y., Kong, H., Gong, L., 2014. A new standardized Palmer drought index for hydro-meteorological use. Hydrol. Process. 28, 5645–5661. https://doi.org/10.1002/hyp.10063

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The Relationship of Drought Frequency and Duration to Time Scales, Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA. https://doi.org/10.1088/1755-1315/5

Mehran, A., Mazdiyasni, O., Aghakouchak, A., 2015. A hybrid framework for assessing socioeconomic drought: Linking. J. Geophys. Res. Atmos. 1–14. https://doi.org/10.1002/2015JD023147.Received

Ministério da Agricultura e Segurança Alimentar, 2015. Anuário de Estatísticas Agrárias 2015.

Mo, K.C., 2008. Model-based drought indices over the United States. J. Hydrometeorol. 9, 1212–1230. https://doi.org/10.1175/2008JHM1002.1

Monteleone, B., Bonaccorso, B., Martina, M., 2020. A joint probabilistic index for objective drought identification: The case study of Haiti. Nat. Hazards Earth Syst. Sci. 20, 471–487. https://doi.org/10.5194/nhess-20-471-2020

Oguntunde, P.G., Lischeid, G., Abiodun, B.J., 2018. Impacts of climate variability and change on drought characteristics in the Niger River Basin, West Africa. Stoch. Environ. Res. Risk Assess. 32, 1017–1034. https://doi.org/10.1007/s00477-017-1484-y

Osbahr, H., Twyman, C., Neil Adger, W., Thomas, D.S.G., 2008. Effective livelihood adaptation to climate change disturbance: Scale dimensions of practice in Mozambique. Geoforum 39, 1951–1964. https://doi.org/10.1016/j.geoforum.2008.07.010

Patt, A.G., Schröter, D., 2008. Perceptions of climate risk in Mozambique: Implications for the success of adaptation strategies. Glob. Environ. Chang. 18, 458–467. https://doi.org/10.1016/j.gloenvcha.2008.04.002

RStudio Team, 2016. Integrated Development Environment for R. RStudio, Inc. , Boston, MA URL http://www.rstudio.com/.

Sheffield, J., Wood, E.F., 2012. Drought: Past problems and future scenarios. Drought Past Probl. Futur. Scenar. 9781849775, 1–234. https://doi.org/10.4324/9781849775250

Tapley, B.D., Bettadpur, S., Ries, J., Thompson, P.F., Watkins, M.M., 2004. Grace measurements of Mass variability in the Earth system: supporting online material. Science (80-.). 503–505, 503–

506. https://doi.org/10.1126/science.1099192

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

Vicente-Serrano, S.M., Martínez, E.A.R., Aguilar, E., Martínez, R., Martín-Hernández, N., Azorin-Molina, C., Sanchez-Lorenzo, A., Kenawy, A. El, Tomás-Burguera, M., Moran-Tejeda, E., López-Moreno, J.I., Revuelto, J., Beguería, S., Nieto, J.J., Drumond, A., Gimeno, L., Nieto, R., 2017. The complex influence of ENSO on droughts in Ecuador. Clim. Dyn. 48, 405–427. https://doi.org/10.1007/s00382-016-3082-y

Wilhite, D.A., Sivakumar, M.V.K., Wood, D.A., 2000. Early Warning Systems for Drought Preparedness and Drought Management, in: Proceedings of an Expert Group Meeting Held 5-7 September, 2000, in Lisbon, Portugal. Http://Www.Wamis.Org/Agm/Pubs/Agm2/Agm02.Pdf. pp. 182–199.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, Second. ed. Academic Press, San Diego, CA.

WMO, 2006. Drought monitoring and early warning : concepts , progress and future challenges. World Meteorol. Organ. 24.

World Meteorological Organization, 2008. Guide to Hydrological Practices. Volume I: Hydrology– From Measurement to Hydrological Information, WMO-No. 168. Geneva. https://doi.org/10.1080/02626667.2011.546602

World Meteorological Organization and Global Water Partnership, 2016. Handbook of Drought Indicators and Indices (M. Svoboda and B.A. Fuchs). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2, Geneva. https://doi.org/10.1201/9781315265551-12

Yevjevich, V., 1969. An objective approach to definitions and investigations of continental hydrologic droughts. J. Hydrol. https://doi.org/10.1016/0022-1694(69)90110-3

Yihdego, Y., Vaheddoost, B., Al-Weshah, R.A., 2019. Drought indices and indicators revisited. Arab. J. Geosci. 12. https://doi.org/10.1007/s12517-019-4237-z

Yue, S., Ouarda, T.B.M.J., Bobée, B., Legendre, P., Bruneau, P., 1999. The Gumbel mixed model for flood frequency analysis. J. Hydrol. 226, 88–100. https://doi.org/10.1016/S0022-1694(99)00168-7

Appendix B:

Poster 6th IAHR Europe Congress, 2020, Warsaw, Poland: "A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation index".

Póster del 6to Congreso Europeo de la IAHR, 2020, Varsovia, Polonia: "A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation index".

A hybrid framework for assessing agricultural drought: a multivariate standardized precipitation and vegetation index

Ronnie Araneda-Cabrera[1*], María Bermúdez[2], Jerónimo Puertas[3]
[1,3]University A Coruña, A Coruña, Spain; [2]University of Granada, Granada, Spain

6th IAHR Europe Congress
Warsaw Poland 2020
Hydro-Environment Research and Engineering
NO FRAMES NO BORDERS
February 15 - 18, 2021

Introduction

Agriculture, which is related to food security and socio-economic activities, is specially affected by droughts. In order to characterize and monitor agricultural droughts and reduce the associated risk various drought indices (DI) have been developed. The Standardized Precipitation Index, SPI (McKee et al., 1993) and the Vegetation Condition Index, VCI (Kogan, 1995) has been popularized. Here, we propose integrate both meteorological and vegetation variables into a new DI, the Multivariate Standardized Precipitation and Vegetation Index (MSPVI). We used open global datasets over Mozambique as a case of study.

Methodology

- Precipitation (P) and the VCI is averaging in a desirable scale and the non-exceedance probabilities of both variables are calculated (Gringorten, 1963):

$$P_t = \frac{I - 0.44}{N + 0.12} \qquad (1)$$

Where, Pt is the cumulative frequency at month t (1…N). I is the position from the smallest to largest and N denotes the sample size.

- We transform the empirical probabilities into a standardized index (SI):

$$SI_t = \varphi^{-1}(P_t) \qquad (2)$$

Where φ denotes the standard normal distribution function.

- The two univariate indicators are combined using the multivariate framework (Yue et al., 1999).
- Changing the position, I for the number of occurrences of the pair (SI(P),SI(VCI)) for SI(P) ≤ SI(P_t) and SI(VCI) ≤ SI(VCI_t), we obtained the joint bivariate empirical probability at month t using (1).
- Calculate the multivariate standardized precipitation and vegetation index (MSPVI) through (2). The lower the values, the greater the intensity of the drought.

Results

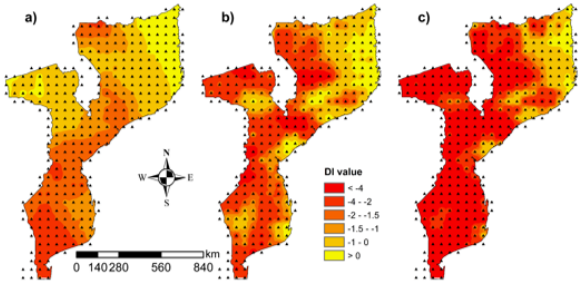


Fig. 1. Spatial distribution of the a) SPI-3, b) SVCI and c) MSPVI during May/1992 drought event.

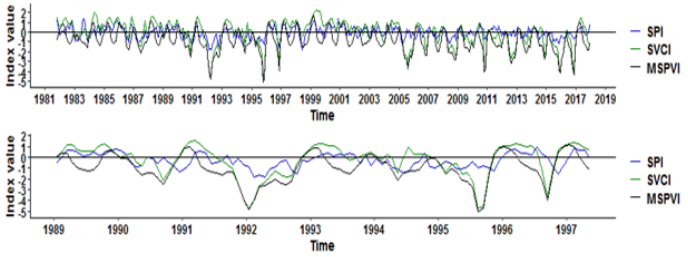


Fig. 2 The SPI-3, SVCI and the MSPVI series in 1982-2017 (above) and 1989-1997 (below).

Crop	MSPVCI	SPI-3	SVCI
Maize	0.593	0.515	0.172
Sorghum	0.833	0.3735	0.247
Wheat	0.538	0.363	0.003

Table 1. Pearson correlation coefficients between Mozambican annual crop yields and drought indices from July.

- VCI from the NOAA STAR Center for and precipitation data from the Climate Research Unit (CRU) were downloaded over CRU grid of 0.5˚ (343 cells) from 1981 to 2019.
- VCI was standardized in order to evaluate all DIs in a a common scale.
- Spatial distribution of the 1992 (May) drought event are plotted in Fig. 1. It can be seen the MSPVCI identified a major area affected by droughts.
- Fig. 2 shows the temporal distribution of the national aggregated DIs. Negative trends can be appreciated.
- At national scale correlation between annual benchmark crops yields and DI of July are sowed in Table 1. Improvements using MSPVCI instead SPI or SVCI are noted.

Conclusions

The proposed agricultural drought index is more sensitive to the onset and recovery of droughts, while it detects a more conservative spatial distribution of droughts than the SPI and SVCI. The MSPVI series has significantly negative trends higher than SPI and SVCI. Further it explains better the annual variability of maize and sorghum yield. The MSPVCI may be used for evaluating and monitoring agricultural drought being a useful tool for decision-makers.

References

- Gringorten, I. I. (1963). A Plotting Rule for Extreme Probability Paper. Journal of Geophysical Research, 68(3), 813–814.
- Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. Advances in Space Research, 15(11), 91–100.
- McKee, T. B., Doesken, N. J., Kleist, J. (1993). The Relationship of Drought Frequency and Duration to Time Scales, Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA.
- Yue, S., Ouarda, T. B. M. J., Bobée, B., Legendre, P., Bruneau, P. (1999). The Gumbel mixed model for flood frequency analysis. Journal of Hydrology, 226(1–2), 88–100.

*Correspondence: ronnie.aranedac@udc.es

Water and Environmental Engineering Group
XUNTA DE GALICIA