

MANUAL PARA LA OBTENCIÓN Y APLICACIÓN DE VARIABLES HIDROMETEOROLÓGICAS DESDE BASES DE DATOS GLOBALES PARA LA GESTIÓN DE SEQUÍAS

MANUAL PARA LA OBTENCIÓN Y APLICACIÓN DE VARIABLES HIDROMETEOROLÓGICAS DESDE BASES DE DATOS GLOBALES PARA LA GESTIÓN DE SEQUIAS

ARANEDA CABRERA, Ronnie Javier
PUERTAS AGUDO, Jerónimo
ÁLVAREZ ENJO, Manuel Ali
PENAS LÓPEZ, Víctor

A Coruña, 2022
Universidade da Coruña, Servizo de Publicacións

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

N.º de páxinas: 74
Índice: p. 5

ISBN: 978-84-9749-833-3
Depósito legal: C 222-2022

Diseño y maquetación: Tórculo Comunicación Gráfica, S. A.

Esta obra se complementa con los recursos necesarios para el desarrollo del manual, y que se describen en los Capítulos I y II de este documento (p. 27). Pueden encontrarse en la URL <https://ruc.udc.es/dspace/handle/2183/29799> (Repositorio institucional de la UDC, RUC).



PROYECTO REALIZADO CON LA COLABORACIÓN DE



EDICIÓN

Universidade da Coruña, Servizo de Publicacións <<http://www.udc.gal/publicacions>>

© Editora: Servicio de Publicacións. Universidade da Coruña
© Contenido e imáxenes: de los autores



This book is released under a Creative Commons license Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). / Esta obra se edita bajo una licencia Creative Commons CC BY-NC-SA (Atribución-NoComercial-CompartirIgual) 4.0 Internacional.

MANUAL

PARA LA OBTENCIÓN Y APLICACIÓN DE VARIABLES HIDROMETEOROLÓGICAS DESDE BASES DE DATOS GLOBALES PARA LA GESTIÓN DE SEQUÍAS

AUTORES

Ronnie Javier Araneda Cabrera

Jerónimo Puertas Agudo

Manuel Ali Alvarez Enjo

Víctor Penas López

A Coruña, 2022

Servizo de Publicacións

Universidade da Coruña



Índice general

Glosario de siglas y abreviaturas	7	7	Objetivos específicos	26
I Introducción	9	8	Seguimiento del manual	27
1 Objetivo General	9	II Obtención de datos	29	
2 Alcance y limitaciones del manual	9	9 Zona de estudio.....	29	
3 Organización del documento	10	10 Precipitación: CHIRPS.....	32	
4 Gestión de sequías	10	11 Índices de vegetación desde NOAA STAR.....	38	
4.1 Ciclo hidrológico.....	10	12 Varios: TerraClimate	47	
4.2 Balance hídrico.....	11	13 Solución para unificar escalas espaciales:	54	
4.3 Definición y clasificación de la sequía	12	III Aplicaciones a gestión sequías	57	
4.4 Índices de sequía	14	14 Aplicaciones del balance hídrico.....	57	
4.5 Características de las sequías	16	15 Calculo y aplicación de los índices de sequía SPI y SPEI...	61	
5 Requerimientos.....	18	16 Caracterización sequías mediante el SPI	67	
5.1 Sistemas de Información Geográfica (SIG)	18	Conclusiones	71	
5.2 Sistemas de Cómputo Numérico (SCN)	20	Referencias	73	
6 Bases de datos globales	22			
6.1 Bases de datos complementarias	22			
6.2 Bases de datos de variables hidrometeorológicas.....	23			

Índice de figuras

Fig. 1. El ciclo del agua. Fuente: U.S. Geological Survey (USGS)	11
Fig. 2. Secuencia de la ocurrencia de los tipos de sequía y sus impactos	13
Fig. 3. Características de las sequías según el "run-theory"	18
Fig. 4. Capas temáticas disponibles en un SIG y comparación entre datos ráster y vectoriales.	19
Fig. 5. Entorno de RStudio.....	21
Fig. 6. Localización y representación de los límites de Mozambique y la cuenca del río Licungo	31

Índice de tablas

Tabla 1. Categorías de las sequías según los valores de los índices ..	17
--	----



Glosario de siglas y abreviaturas

ETP	Evapotranspiración potencial
ETr	Evapotranspiración real
CHIRPS	Rainfall Estimates from Rain Gauge and Satellite Observations
STAR	The Center for Satellite Applications and Research
NOAA	The National Oceanic and Atmospheric Administration
VCI	Vegetation Condition Index
TCI	Temperature Condition Index
VHI	Vegetation Health Index
NDVI	Normalised Difference Vegetation Index
BT	Brightness Temperature
SPI	Standardized Precipitation Index
SPEI	Standardized Precipitation and Evapotranspiration Index
SIG	Sistema de Información Geográfica
SCN	Sistema de Cómputo Numérico

I. Introducción

1. Objetivo General

Este manual tiene como objetivo principal ser una herramienta básica introductoria para la gestión de sequías. Este fenómeno hidrológico extremo es de los más importantes dentro de la gestión de recursos hídricos, debido a que causan anualmente inmensas pérdidas económicas y ecológicas en todo el mundo condicionando la seguridad alimentaria y calidad de vida de las personas.

Debido al cambio climático, los eventos de sequía serán cada vez más frecuentes y severos en el futuro. Por lo tanto, entender, caracterizar y monitorizar este fenómeno es fundamental para crear planes de contingencia que permitan disminuir el riesgo y ser más resilientes a las sequías. Para lograr este propósito se requieren datos, herramientas y procedimientos que permitan comprender y modelar el fenómeno. Sin embargo, las diferentes realidades socioeconómicas, ampliamente variables en cada región del mundo, hacen que el acceso a ciertos datos con una calidad suficiente sea limitado en algunas regiones y países.

Este manual es una herramienta que permite solventar esta problemática a través de herramientas desarrolladas para obtener rápida y eficazmente los datos necesarios para la gestión y monitorización de sequías en cualquier región. Estos datos son productos de bases de datos globales libres y gratuitos para todo el mundo por lo que la aplicación que se les dé a estos datos puede usarse en contornos administrativos (países, ciudades, etc.) o naturales (cuencas hidrográficas, reservas naturales, etc.). Además, el manual presenta distintos usos y aplicaciones de estos datos como herramientas concretas de gestión de sequías. Como parte del procedimiento, se hace una breve introducción a los distintos aspectos, conceptos y requerimientos necesarios para uso del manual.

2. Alcance y limitaciones del manual

La información conceptual y práctica que se presenta en este manual tiene por objeto ser una herramienta introductoria para la obtención y aplicación de datos a la monitorización y caracterización de sequías. En el manual se presenta la información necesaria para comprender el fenómeno y sus componentes de una manera general, por lo que las herramientas y resultados prácticos están limitadas a estos conceptos.

Los conceptos clave sintetizados son de uso y general generados y reconocidos por expertos y organizaciones internacionales; sin embargo, están limitados a los objetivos del manual, pudiendo ser ampliados e incluso mostrar diferencias a otras fuentes bibliográficas. De igual manera las herramientas usadas, los procedimientos e interpretaciones aplicadas para los ejercicios prácticos, pueden contener errores y ser mejoradas o aplicadas en otros escenarios. Todas las herramientas computacionales usadas son de uso público y gratuito, pudiendo ser reemplazados por otros que presenten fines similares.

3. Organización del documento

Este documento consta de tres capítulos. En el primer capítulo se hace una introducción a los conceptos, herramientas y requerimientos que se necesitan para hacer uso del manual. Primero se hace una recopilación de los conceptos asociados al estudio y gestión de sequías, su clasificación y caracterización. Luego, se explica de manera general los requerimientos técnicos y computacionales que se requieren. A continuación, se presenta una introducción a las varias bases de datos globales que existen y que datos necesarios para la gestión de sequías ofertan, haciendo énfasis en los que se utilizan en el manual. Finalmente, se explican los objetivos específicos que se buscan en este trabajo. En el segundo capítulo se detallan y explican los pasos necesarios para obtener las series de distintas variables que se requieren para la gestión de sequías. En el tercer capítulo se hace uso de los datos obtenidos en el Capítulo II y se exponen los pasos para aplicarlos concretamente en la caracterización y monitorización de sequías en una región específica. En los anexos se entregan tres scripts que son el producto final complementario para seguir el manual. Como casos de estudio hemos elegido Mozambique localizado en el cono sur africano y la cuenca hidrográfica del río Licungo contenida en la parte centro-norte de este país.

4. Gestión de sequías

4.1. Ciclo hidrológico

El ciclo hidrológico o ciclo del agua es el proceso de circulación del agua entre los distintos compartimentos de la hidrosfera (Fig. 1). El agua proveniente de la tierra y los océanos entra a la atmósfera por evaporación o sublimación, donde se condensa en nubes y cae en forma de lluvia o nieve. El agua que cae en forma de precipitación entra a los cuerpos de agua dulce o se infiltra en el suelo. El ciclo se completa cuando el agua superficial o subterránea entra nuevamente al océano. Los principales procesos implicados en el ciclo del agua y que son variables importantes en el análisis de las sequías son:

- Precipitación (P): La atmósfera pierde agua por condensación (lluvia y rocío) o sublimación inversa (nieve y escarcha) que pasan según el caso al terreno, a la superficie del mar o a la banquisa (capa de hielo sobre los océanos).
- Evapotranspiración (ET): El agua se evapora en la superficie oceánica, sobre el terreno y también por los organismos, en el fenómeno de la transpiración, de manera que la ET alude al fenómeno combinado. La evapotranspiración potencial (ETP) es aquella máxima cantidad de agua que puede evaporarse desde un suelo completamente cubierto de vegetación, que se desarrolla en óptimas condiciones, y en el supuesto caso de no existir limitaciones en la disponibilidad de agua. La evapotranspiración real o efectiva (ETr) es la que se produce realmente en las condiciones existentes en cada caso.
- Infiltración: El fenómeno ocurre cuando el agua que alcanza el suelo penetra a través de sus poros y pasa a ser subterránea. La proporción de agua que se infiltra y la que circula en superficie (escorrentía) depende de la permeabilidad del sustrato, de la pendiente y de la cobertura vegetal. La cantidad de agua contenida en un momento determinado es la humedad del suelo y puede medirse a diferentes profundidades.

Fig. 1
El ciclo del agua. Fuente: U.S.
Geological Survey (USGS)



- **Escorrentía:** Este término se refiere a los diversos medios por los que el agua líquida se desliza cuesta abajo por la superficie del terreno. Al llegar a un cuerpo de agua (ríos, lagos, embalses, etc.) el volumen que atraviesa por este por un tiempo determinado se lo denomina como caudal.
- **Circulación subterránea:** Se produce a favor de la gravedad, como la escorrentía superficial. Ocurre en los acuíferos a través de los poros de una roca permeable, intervienen fenómenos como la presión y la capilaridad.

4.2. Balance hídrico

El ciclo de agua puede formularse matemáticamente usando el balance hídrico, que se deriva del concepto de conservación de la materia, es decir, que es el equilibrio entre todos los recursos hídricos que entran y salen de un sistema (cuenca hidrográfica, país, región, etc.) en un intervalo de tiempo determinado:

$$Estado_{t+1} = Estado_t + \sum_{i=1}^n Entradas_i - \sum_{j=1}^m Salidas_j \quad (1)$$

Donde, las entradas de agua son la precipitación (P), que puede presentarse en forma de lluvia, granizo, nieve o condensaciones; aporte de aguas subterráneas colindantes al sistema; y, trasvases de agua de otros sistemas, como puede ser descargas de presas o aguas servidas. Las salidas pueden ser la evapotranspiración; las infiltraciones profundas que alimentan acuíferos; el agua se deriva a sistemas externos colindantes, consumo humano e industria; y, las que salen del sistema a un receptor externo como puede ser un gran mar o el mar. El estado se refiere al volumen contenido en el sistema después de sumar y restar las entradas y las salidas al volumen de agua en el paso de tiempo anterior.

De manera general, para cada paso de tiempo la eq. 1 se puede formular de la siguiente manera:

$$P + I_{in} + U_{in} - I_{out} - A_{out} - ETr - Q_{out} - U_{out} = \Delta V \quad (2)$$

En donde P es la precipitación, I la infiltración, U los distintos usos del agua (en el mismo u otro sistema), ETr la evapotranspiración real, Q el caudal de salida del sistema y ΔV la variación de volumen de agua en el sistema. Al largo plazo la variación tiende a cero, sin embargo, en cada paso de tiempo puede haber épocas en que el sistema tenga déficit o superávit. Cada uno de los procesos se puede medir y llegar a completar el balance.

Simplificando el balance somos capaces de detectar meses húmedos y secos (3); y, de calcular el volumen de agua disponible (4):

$$P - ETP = \Delta V \quad (3)$$

$$P - ETr = Q + I + U \quad (4)$$

En donde, en la eq. 3 la variación ΔV que define un mes seco o déficit ($\Delta V < 0$) o un mes húmedo con superávit ($\Delta V > 0$) en el sistema en un periodo de tiempo determinado. Mientras, en la eq. 4 la segunda parte de la igualdad se denomina precipitación útil y es el agua disponible recargar el suelo, los ríos y otros cuerpos de agua existentes.

El balance hídrico nos permite además calcular el déficit hídrico (D):

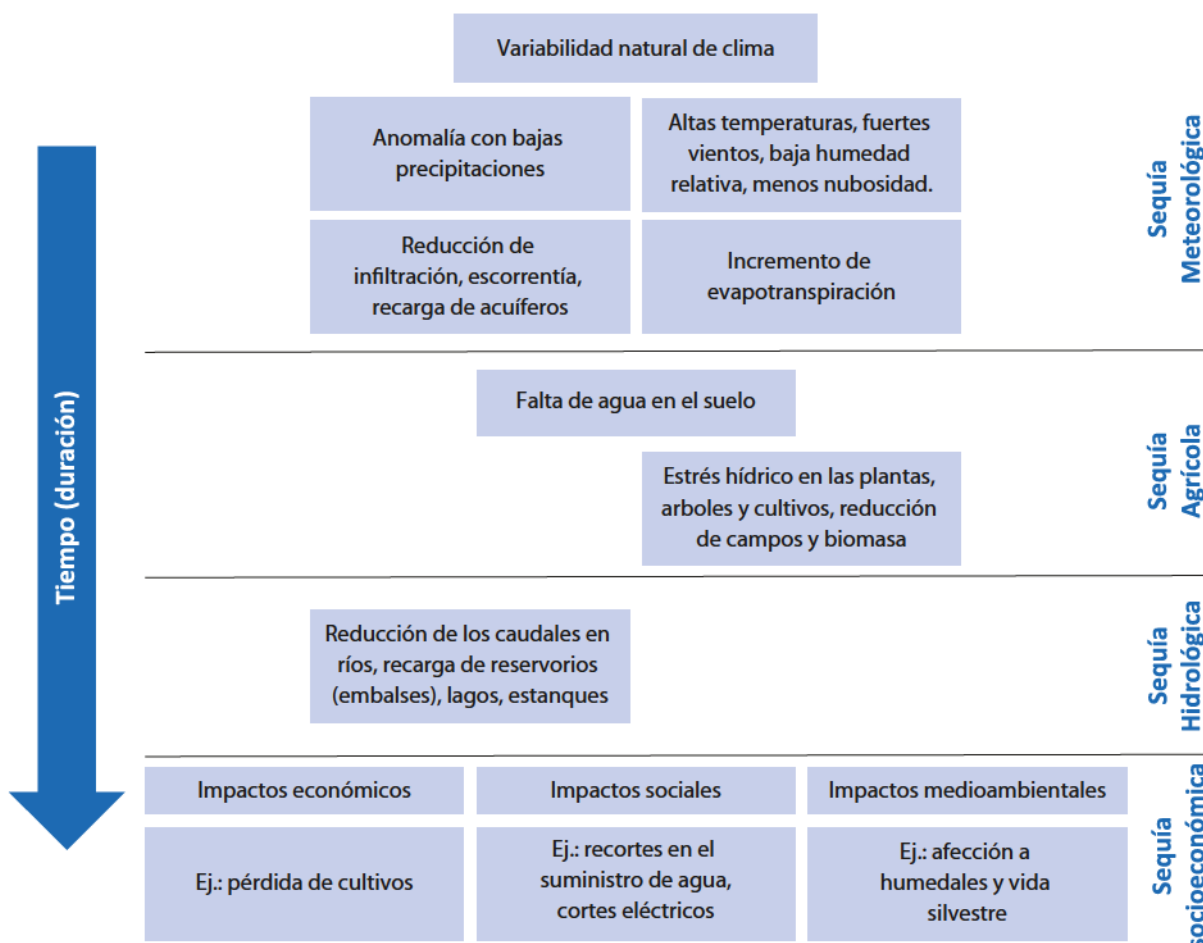
$$ETP - ETr = D \quad (5)$$

En donde, un déficit ($D > 0$) es la cantidad de agua que falta para cubrir las necesidades potenciales del agua para evaporar y transpirar.

4.3. Definición y clasificación de la sequía

La sequía es un fenómeno hidrológico extremo cuyos límites geográficos y temporales son difíciles de determinar, siendo un desastre natural especialmente peligroso cuando no existe una buena gestión de los recursos hídricos. La sequía se entiende como una anomalía temporal más o menos prolongada de precipitación, caracterizada por un período de tiempo con valores de precipitaciones inferiores a las normales. Puede producir, o no, afecciones a cultivos, degradaciones en los suelos y/o situación de insuficiencia en los suministros de agua, en función del nivel de demanda y de las características de los sistemas de explotación del recurso hídrico.

Fig. 2
Secuencia de la ocurrencia de los tipos de sequía y sus impactos
 Fuente: National Drought Mitigation Center de la Universidad de Nebraska-Lincoln



En los últimos años se han desarrollado muchas clasificaciones de sequía. Sin embargo, la más aceptada y difundida es la de que define la sequía meteorológica, agrícola, hidrológica y socioeconómica. Las tres primeras miden la sequía como un fenómeno físico, mientras la última se ocupa de la sequía en términos de oferta y demanda, basada en el seguimiento de los efectos del déficit de agua a medida que se agita a través de los sistemas socioeconómicos. Cada una de ellas está relacionada a una parte del ciclo hidrológico y se desarrolla una de tras de la otra iniciando por anomalías climáticas que son parte de la variabilidad natural del clima (Fig. 2).

La sequía meteorológica se produce cuando anomalías de ausencia de precipitaciones dominan una zona. Estas anomalías pueden venir acompañadas de un aumento en las temperaturas, lo que se traduce en aumento de evapotranspiración. La sequía agrícola se produce cuando los cultivos se ven afectados por esta falta de precipitaciones y bajas humedades en el suelo. La sequía hidrológica se produce cuando el bajo suministro de agua se hace evidente, especialmente en los arroyos, embalses y niveles de agua subterránea, normalmente después de muchos meses de sequía

meteorológica. La sequía socioeconómica relaciona la oferta y la demanda de diversos productos básicos con la sequía (agua para consumo humano, riego, industrias, hidroeléctricas, etc.). La sequía meteorológica puede empezar y terminar rápidamente, mientras que la sequía hidrológica tarda mucho más en desarrollarse y recuperarse.

La sequía debe diferenciarse de otros fenómenos relacionados como son la escasez y la aridez. La escasez representa una situación permanente de déficit (calculada con el balance hídrico) en relación con la demanda de agua en un sistema de recursos de ámbito regional, caracterizado, bien por un clima árido o bien por un rápido crecimiento de las demandas. La aridez, es una situación estructural natural de una región y por tanto permanente. Aún en estas circunstancias no debería haber déficit si los sistemas de explotación estuvieran adecuadamente diseñados y explotados y las demandas se mantuvieran en límites razonables, acordes con las características climáticas de la región. Ello precisa actuaciones planificadas a medio y largo plazo.

4.4. Índices de sequía

Los índices son representaciones numéricas que muestran información asociada a la gravedad de las sequías, determinadas mediante datos climáticos o hidrometeorológicos. Se utilizan para cuantificar la gravedad, la localización, el desarrollo cronológico y la duración de los episodios de sequía. La gravedad se refiere a la desviación de la normalidad de un índice tomando como base un umbral previamente establecido. De esta manera podemos identificar cuando una sequía ha comenzado, el momento en que ha terminado y la zona geográfica afectada. La localización se refiere a la zona geográfica en la que se registran las condiciones de sequía. Las fechas aproximadas de inicio y cese fijan la duración del evento de sequía. Hay índices que también pueden cumplir otra función esencial, en el sentido de que pueden facilitar una referencia histórica a los planificadores o las instancias decisorias. Esta referencia brinda a los usuarios información sobre la probabilidad de que se produzcan o se repitan sequías de distintos niveles de gravedad. No obstante, cabe destacar que el cambio climático comenzará a alterar las tendencias históricas. Se recomienda representar el fenómeno en diversas escalas temporales, ya que así se reconocen los períodos húmedos de corta duración dentro de las sequías de larga duración o los períodos secos de corta duración dentro de los períodos húmedos de larga duración.

Dado que la clasificación de sequía responde a analizar anomalías en cada parte del ciclo hidrológico, existen índices destinados a caracterizar cada tipo de sequía basados en variables propias de cada una. Por ejemplo, precipitación para la sequía meteorológica o caudales para la sequía hidrológica. En los últimos años se han desarrollado índices mixtos que intentan cuantificar las sequías en todo el ciclo del agua o partes de él. Sin embargo, utilizar varios índices basados en varias variables es lo más recomendable para caracterizar el estado de la sequía en todo el sistema.

Debido a la complejidad del fenómeno, en las últimas décadas se han propuesto decenas de índices. Estos índices pueden ser calculados con una o varias variables hidrometeorológicas de entrada a distintas escalas temporales (diario, semanal o mensual), siendo la más común y aceptada la escala mensual. En este manual explicaremos a detalle solamente 5 índices, que han sido ampliamente evaluados y validados en todo el mundo. Para mayor información de algunos índices de sequía, se puede acudir a bibliografía adicional como por ejemplo, el manual de Indicadores de sequía (Svodova et al., 2016) donde se muestran y explican más de 100 índices de sequía.

4.4.1. Índice Estandarizado de Precipitación (SPI)

El SPI fue introducido por McKee et al., (1993) y se define como un valor que representa el número de desviaciones estándar de precipitaciones a lo largo de un período de acumulación definido, respecto a la media (estandarización), una vez que la distribución original de precipitación ha sido transformada a una distribución normal. A modo de ejemplo se puede indicar que un valor de SPI para 3 meses de -2.3 indicaría que la cantidad de precipitación que se ha registrado a lo largo de un período se ha situado a 2.3 veces la desviación estándar por debajo del valor medio. El paso a paso para su completo cálculo puede ser revisado en varios documentos, por ejemplo, en Kumar et al., (2009). Está actualmente recomendado por la Organización Mundial de Meteorología para la monitorización del estado de sequías. Es muy flexible, de fácil interpretación, aplicable a y comparable en distintas zonas climáticas y depende de una sola variable, la precipitación.

A través del uso del índice SPI es posible cuantificar y comparar las intensidades de los déficits de precipitación entre zonas con climas muy diferentes y tiene la propiedad de que puede integrarse a cualquier escala temporal, lo que hace que pueda ser utilizado como indicador de diferentes tipos de sequía. Escalas temporales de corta duración (2 - 6 meses) reproducen sequías principalmente sobre los sectores agrícola y forestal siendo útil para sequías meteorológicas y agrícolas, mientras los de larga duración (12-24 meses) se asocian a sequías hidrológicas ya que representan efectos de sequía en embalses, ríos y acuíferos.

El cálculo del SPI requiere partir de series históricas de precipitaciones mensuales de al menos 30 años. Series de precipitación en agregaciones diarias, semanales u otras son admisibles en el cómputo del SPI; sin embargo, se recomienda seguir las escalas en las que normalmente se desarrollan las sequías, siendo esta comúnmente mensual, estacional o interanual. Una vez realizado el cómputo del SPI sobre una serie temporal de precipitaciones, los valores mensuales del SPI pueden evaluarse siguiendo la Tabla 1.

4.4.2. Índice Estandarizado de Precipitación y Evapotranspiración (SPEI)

El SPEI es un índice de sequía con propiedades similares al SPI, con la diferencia que toma en cuenta la temperatura a través de la evapotranspiración en su cálculo, además de la precipitación. Esta diferencia lo hace adecuado para analizar tendencias y cambios en las características de las sequías debido al cambio climático. Su cómputo requiere de series históricas mensuales de la variación de agua en el sistema ($\Delta V = P - ETP$), donde la ETP incluye la información de la temperatura. El proceso de estandarización que sigue el SPEI es el mismo al SPI, con la diferencia que la serie de ΔV es previamente transformada a una distribución log-normal. De igual manera que el SPI, el SPEI puede calcularse para varias escalas temporales permitiendo evaluar las sequías en varias partes del ciclo del agua y relacionarla con los diferentes tipos de sequía. Su evaluación sigue la Tabla 1.

4.4.3. Índice de Condición de Vegetación (VCI)

El VCI es un índice utilizado para la detección y monitoreo de la sequía agrícola. Permite evaluar la duración, el área cubierta, la intensidad y los impactos que genera una sequía sobre la vegetación. Para un determinado punto o región, y para cada paso de tiempo (mensual en nuestro caso), el VCI se deriva del Índice de vegetación de diferencia normalizada (NDVI – Normalized Difference Vegetation Index), obtenido de la serie histórica de imágenes satelitales obtenidas por el sensor AVHRR (Advance Very High Resolution Radiometer) en funcionamiento desde 1984:

$$VCI = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}}$$

Donde, el max y min es el valor de NDVI máximo y mínimo presente en la serie temporal analizada y el NDVI es el valor que se tenga en cada paso de tiempo. El NDVI es por sí mismo un índice de vegetación que se utiliza para estimar la cantidad, calidad y desarrollo de la vegetación con base a la medición desde sensores satelitales de la intensidad de la radiación de ciertas bandas del espectro electromagnético que la vegetación emite o refleja. La evaluación del grado de intensidad de la sequía según el VCI se muestra en la Tabla 1.

4.4.4. Índice de Condición de la Temperatura (TCI)

Similar al VCI, el TCI es un índice derivado de las mediciones desde el sensor satelital AVHRR, en este caso de la temperatura superficial. Su uso está relacionado con el monitoreo de sequías meteorológicas y agrícolas. Su estimación utiliza el brillo estimado por la banda infrarroja termal (10.3-11.3 μm) detectado por los sensores desde el espacio. De acuerdo con el brillo medido, se estima el Brightness Temperature (BT), que por sí solo es indicador correlacionado directamente con la temperatura de la superficie terrestre. Su formulación es:

$$TCI = \frac{BT_{\max} - BT}{BT_{\max} - BT_{\min}}$$

Donde, dentro de un período de tiempo, BT es el valor en cada paso de tiempo (mensual en nuestro caso) y el min y max se refieren a los valores máximos y mínimos de BT dentro de dicho período. Su evaluación sigue la Tabla 1.

4.4.5. Índice de Salud de la Vegetación (VHI)

El VHI es calculado por la combinación ponderada de dos anomalías ya descritas: el VCI y el TCI, ambos derivados desde observaciones obtenidas por sensores satelitales. La idea básica del indicador es la siguiente: cuanto menor sea el VCI observado (vegetación verde relativamente pobre) y menor sea el TCI observado (temperatura relativamente cálida), menor será el VHI. Los valores de VHI bajos son indicadores de sequía, especialmente cuando persisten por largos períodos de tiempo.

La ecuación utilizada es:

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

Donde w es el peso ponderado, cuyo valor se encuentra entre 0.0 y 1.0, siendo habitualmente 0.5. La evaluación de la intensidad de sequía según el VHI se expone en la Tabla 1.

4.5. Características de las sequías

La caracterización de las sequías es fundamental en la gestión de las sequías. Definir y cuantificar las características a cada evento de sequía, permite planificar campañas de monitoreo continuo y aplicar modelos de predicción. Consiste en definir la gravedad de un evento determinado de manera un gestor o un técnico pueda tomar decisiones y aplicar planes de mitigación adecuadas para evitar una grave escasez de agua, por lo que este componente metodológico es esencial para las partes interesadas.

Tabla 1
Categorías de las
sequías según los
valores de los índices

Categorías	SPI/SPEI	VCI/TCI/VHI
Húmedo	> 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
Sequía moderada	-1.00 - -1.50	25 - 40
Sequía severa	-1.50 - -2.00	10 - 25
Sequía extrema	< -2.00	0 - 10

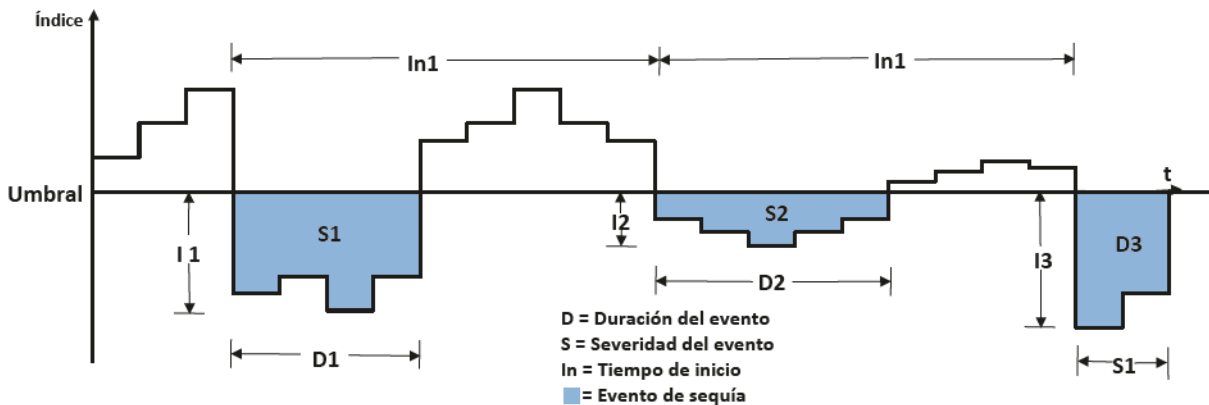
Entre los diferentes métodos propuestos para caracterizar las sequías, el método de la “run-theory” descrito en Yevjevich, (1969) se ha generalizado, debido a la objetividad en la definición de la sequía. El método permite realizar una derivación analítica de las distribuciones de probabilidad de las características de la sequía, mostrando un mejor desempeño que otros enfoques. El método de ejecución puede aplicarse a una serie temporal de interés, que en nuestro caso serán los índices de sequía, asumiendo un umbral representativo de la gravedad del evento. Los umbrales se definen según la categoría de sequía que se requiera caracterizar.

Un evento de sequías se define por su duración, intensidad, severidad y extensión geográfica. La duración es el período de tiempo en semanas o meses que dura la sequía, entendida como el tiempo en que el índice estuvo de manera consecutiva bajo un umbral, definiéndose un inicio y fin del evento de sequías al pasar por dicho umbral en el tiempo. La intensidad máxima o promedio es el valor absoluto máximo y el promedio entre los valores que tomó un índice durante la duración del evento. La severidad se define como la suma de las intensidades o valores tomados por el índice durante el período que duró la sequía. En la Fig. 3 se indican estas tres características para tres eventos de sequía. La extensión geográfica es el área o porcentaje de área dentro de una región o zona bajo estudio que se encuentra afectada por sequías a una o varias categorías de intensidad durante cada paso de tiempo.

Para cuantificar estas características se requieren calcular índices de sequías sobre varios puntos del territorio bajo estudio. Esta distribución espacial debe ser lo más densa posible, recomendándose en World Meteorological Organization, (2008) al menos una estación o punto geográfico donde se evalúe el índice cada 250 km².

Los índices de sequía adecuados se seleccionan en función del tipo de sequía de interés. Los índices pueden considerarse generales para todo el sistema o específicos en función de la parte del ciclo hidrológico para los que fueron diseñados. Se entiende que hacer esta distinción no siempre es fácil; por lo que, el uso de varios índices es recomendable para obtener una caracterización completa de las sequías en una región. Aquí usaremos los índices previamente expuestos.

Fig. 3
Características de las sequías
según el "run-theory"



5. Requerimientos

Para el uso y aplicación de este manual, además de las bases teóricas asociadas a la gestión de sequías, es necesario tener claro cuáles son las herramientas computacionales que se utilizan. Los dos grupos de herramientas que se requieren son los Sistemas de Información Geográfica (SIG) y Sistemas de Cómputo Numérico (SCN). Si bien existen muchos softwares similares a los que se explican a continuación, los conceptos alrededor de la arquitectura de estos programas son los mismos que siguen otras herramientas destinadas a cumplir funciones similares.

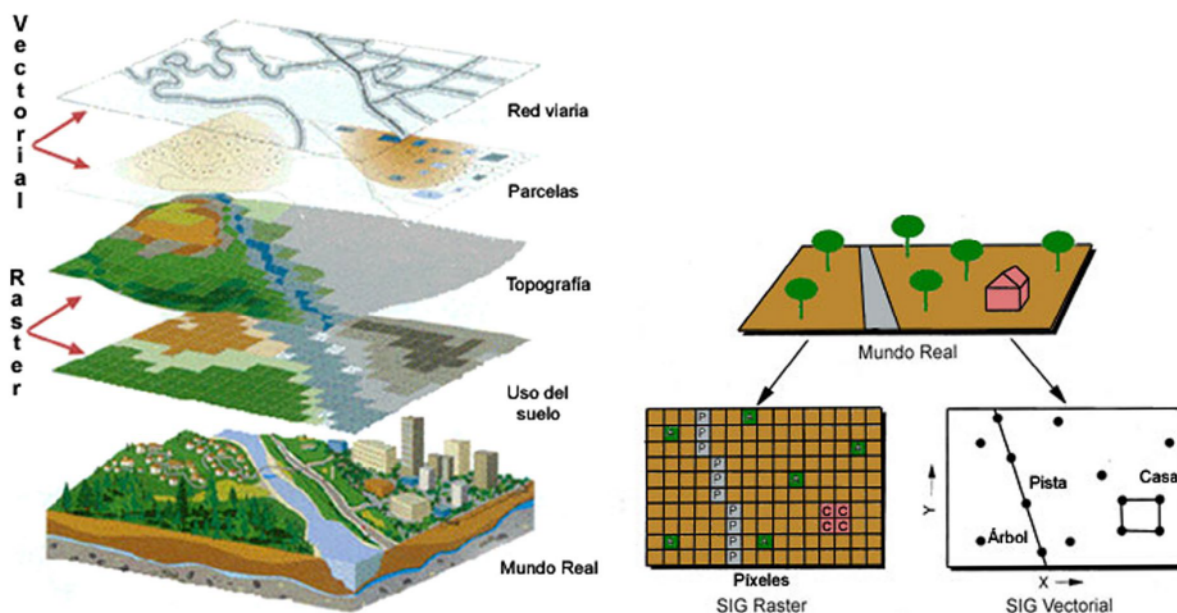
5.1. Sistemas de Información Geográfica (SIG)

Los SIG son programas informáticos (softwares) que proporcionan herramientas para la recopilación, procesamiento, gestión, análisis y representación de datos con una componente cartográfica (información geográfica) en formato digital. Nos permiten representar el espacio físico en el que vivimos sobre mapas y tablas a través de capas temáticas que pueden superponerse. Cualquier registro de información con una referencia geográfica, como una dirección de calle, el nombre de una ciudad, un identificador de parcela de tierra, coordenadas GPS, etc., puede localizarse y estar disponible en un mapa (Fig. 4).

Los SIG se pueden clasificar en dos grupos principales ejemplificadas en la Fig. 4:

SIG Vectoriales: Utilizan vectores para la descripción de los objetos geográficos. Son comúnmente presentados como archivos shapefile (extensiones comunes: .SHP, .SHX, .DBF, .PRJ, etc.). Los elementos geográficos se representan a partir de tres estructuras básicas: puntos, líneas y polígonos. Las líneas quedan representadas por segmentos que unen puntos y éstas, a su vez, se cierran para formar polígonos. Como resultado obtenemos tres naturalezas de shapefile diferentes. Cada una de ellas presenta unas cualidades específicas de manera que no es posible realizar ciertos procesos con ellos. De esta forma, no es imposible calcular superficies para puntos, pero sí para polígonos.

Fig. 4
Capas temáticas disponibles en un SIG y
comparación entre datos ráster y vectoriales
 Fuente: <https://geopaisa.blog/2017/03/08/que-es-un-sig/>



O no es posible calcular longitudes para polígonos, pero podemos fragmentar polígonos en múltiples partes.

SIG Raster: Los Sistemas de Información Raster basan su funcionalidad en los formatos gráficos de mapas de bits (extensiones comunes: .ECW, .JPG, .GRID, .TIF, .GIF, .SID). Su forma de almacenar información es dividiendo el espacio en una retícula o matriz (generalmente) regular de pequeñas celdas - a las que se denomina píxeles - y atribuir un valor numérico a cada celda como representación de una cualidad o propiedad espacial física (altitud, temperatura, distancia, pendiente, etc.). Si cocemos la posición de las coordenadas del centro de una de las celdas, se puede decir que todos los píxeles están georreferenciados.

Dos de los softwares GIS más comunes y desarrollados son ArcGIS y QGIS. Ambos son ampliamente utilizados y cubren la mayoría de las necesidades de los usuarios. La principal diferencia es que QGIS es un software libre gratuito, mientras ArcGIS es un programa de pago comercializado por ESRI. No obstante, la información geográfica que puede representarse, analizarse y mostrarse en forma de mapas en dichos softwares, puede ser también manipulada a través de SCN que ofrezcan estas opciones en sus características.

Los datos globales que se analizan en este manual están en formatos ráster, mientras que para su extracción en zonas de estudio de interés se usan archivos vectoriales que definen un área (país o cuenca hidrográfica). Esta información puede ser descargada desde bases de datos gratuitas.

5.2. Sistemas de Cómputo Numérico (SCN)

Los SCN son softwares o herramientas destinadas a resolver problemas y cálculos numéricos complejos a través del uso de métodos numéricos. El procedimiento generalmente consiste en introducir datos o variables de entrada a una función existente en el software, que desarrollará cálculos previamente programados y que arrojará un resultado. Los SCN cuentan con su propio entorno y lenguaje de programación que permiten a los usuarios utilizar tanto las herramientas y funciones ya existentes en el programa (operaciones sencillas como sumar o restar hasta complejas como resolución de integrales o ecuaciones diferenciales) como crear unas nuevas funciones (programación específica y personalizada).

Existen muchos SCN aplicables a la ingeniería del agua y por tanto a la gestión de sequías. A forma de ejemplo, entre los más extendidos en uso se encuentran MATLAB de la empresa Mathworks, Phyton, R, C++, entre otros. Cada uno de ellos está orientado y especializado en la resolución de problemas con bases numéricas específicas. En este manual se usa R en su entorno más amigable RStudio, que está especializado en la computación estadística y creación de gráficos.

5.2.1. R y RStudio

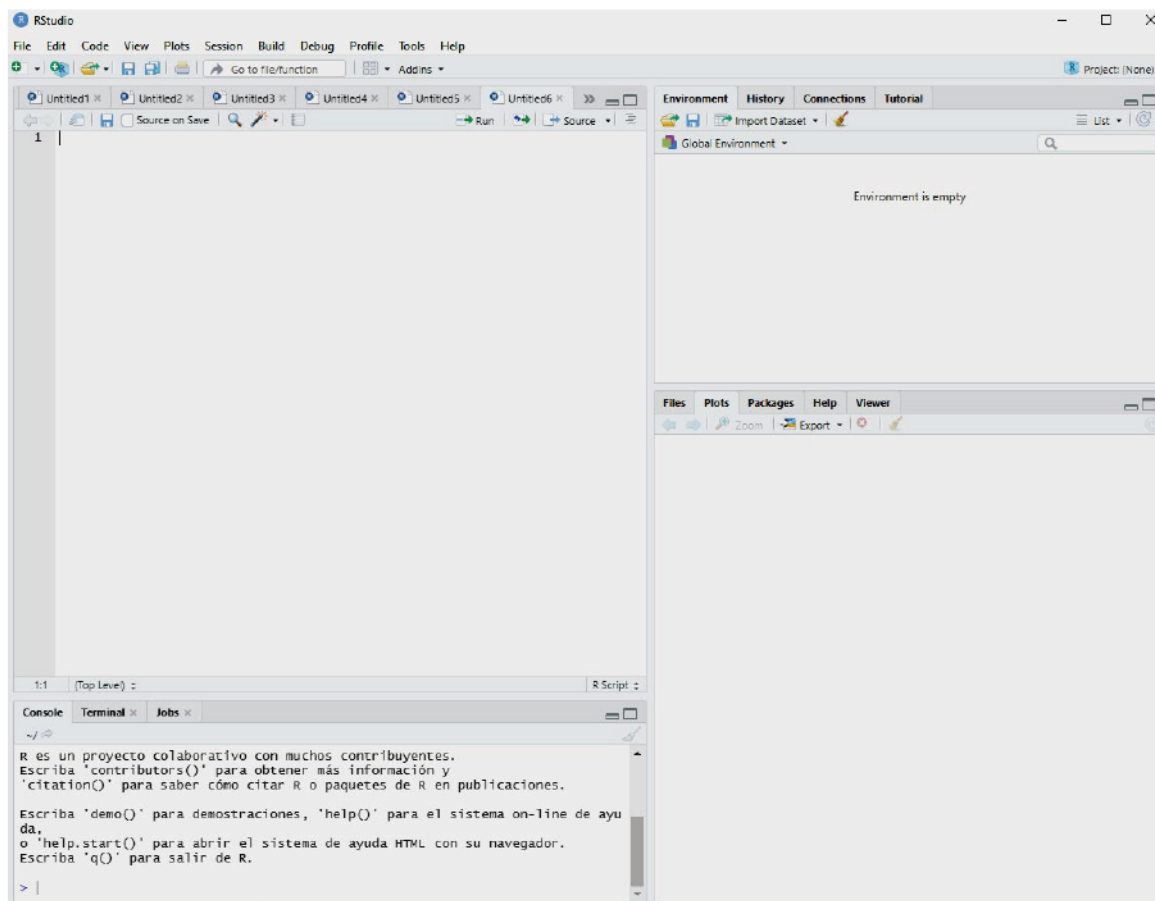
R es un entorno de programación libre, de código abierto y gratuito conceptualizado para el procesamiento y análisis estadístico de datos (Durán, 2000). Está formado por un conjunto de herramientas muy flexibles que pueden ampliarse fácilmente mediante paquetes y librerías de libre acceso o programando nuestras propias funciones.

Los paquetes en R son colecciones de funciones y conjunto de herramientas desarrollados por la comunidad de usuarios. Estos incrementan la potencialidad de R mejorando las funcionalidades base ya existentes, añadiendo unas nuevas. Los miles de paquetes que existen se encuentran en repositorios (CRAN) desde los cuales se pueden descargar sus manuales de uso, su arquitectura y conceptualización. CRAN (Comprehensive R Archive Network) es la página oficial del software donde se ofrece diferentes recursos para el uso del programa: manuales para el usuario, cursos online, información general, descarga de paquetes, información de los paquetes instalados, etc.

Para instalar R hay que descargar el instalador de acuerdo al sistema operativo que se vaya a usar desde su página web oficial (<https://cran.rstudio.com/>). Además, es recomendable instalar RStudio (RStudio Team, 20016), un entorno más amigable que permite hacer procesos computacionales basándose en el lenguaje R. La versión gratuita se puede descargar desde la página web del programa (<https://www.rstudio.com/products/rstudio/download/>).

RStudio es un entorno de desarrollo integrado (EDI) para el lenguaje de programación R, dedicado a la computación estadística y gráfica. Incluye una consola, editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo. Consta de 4 ventanas (Fig. 5): Ventana superior izquierda: es el editor de sintaxis, se trata del lugar donde editamos la sintaxis para posteriormente ejecutarla. Al escribir allí no sucederá nada, a no ser que se apriete algún botón para ejecutar los comandos. En esta ventana se escriben nuevos scripts y donde se abrirán los entregados en este manual. Ventana superior derecha: es el entorno de trabajo del programa, en este lugar se muestra el conjunto de datos y los objetos (resultados, variables, gráficos, etc.) que se almacenan al ejecutar diferentes análisis. Ventana inferior derecha: tiene varias subpestañas: (i) la pestaña files permite ver el historial de archivos trabajados con el programa; (ii) la pestaña plots permite visualizar los gráficos

Fig. 5
Entorno de RStudio



que se generen; (iii) la pestaña packages permite ver los paquetes descargados y guardados en el disco duro, así como gestionar su instalación o actualización; (iv) la ventana help permite acceder al CRAN; (v) la ventana viewer muestra los resultados al construir reportes. Ventana inferior izquierda: es la consola, corresponde a lo que sería el software R en su versión básica. Allí el software ejecuta las operaciones realizadas desde el editor de sintaxis.

De los miles de paquetes existentes para R, existen varios desarrollados específicamente a la gestión de recursos hídricos y de sequías (Slater et al., 2019). Algunos de los paquetes requeridos en los scripts adjuntos en los anexos son el “SPEI” desarrollado por Begueria and Vicente-Serrano, (2017), el cual contiene funciones que permiten calcular de manera directa los índices de sequía SPI y SPEI; el paquete “ggplot2” permite realizar gráficos de manera profesional, permitiendo incluso graficar shapefiles y rasters; el paquete “tmap” permite al usuario de R manipular información geográfica, típicamente manipulable a través de SIG con R; entre otros.

Los scripts presentados en este manual (ANEXOS: Scripts en R y RStudio para gestión de sequías) que se analizan y desarrollan paso a paso en los capítulos II y III. La información específica que se requiere para cada uno de ellos se obtiene de bases de datos gratuitas. Los

scripts están escritos de forma lógica y ordenada para poder ser usados, replicados, manipulados, mejorados y complementados por los usuarios de este documento.

Como información general R permite trabajar con todo tipo de datos: matrices, vectores, información geográfica vectorial y ráster, etc. Cada archivo se puede leer y asignar a un nombre (utilizando el símbolo <-). Se pueden crear diferentes bucles y algoritmos existentes (e.g. funciones seno, coseno). Los gráficos se pueden generar y observar en el viewer y exportar a cualquier carpeta del explorador de Windows en distintos formatos (.pdf, .jpg, png, etc.). Se recomienda el siguiente manual para iniciar con la programación en R: R para Principiantes (Ahumada, 2003). Sin embargo, existen cientos de manuales y blogs online que pueden usarse como fuente guía.

6. Bases de datos globales

Como se ha expuesto previamente, para una óptima gestión de sequías se requieren series temporales de datos hidrometeorológicos correspondientes a cada componente del ciclo del agua. Estos datos deben ser de larga duración (al menos 30 años) y haberse medido con suficiente distribución sobre el territorio natural (una estación meteorológica cada 250km²). Sin embargo, la medición de estas variables con estas condiciones es una tarea compleja. La principal razón suele ser su alto costo de instalación y mantenimiento, por lo que en muchas zonas del mundo la disponibilidad de datos es extremadamente limitada.

Para obtener datos como precipitación, humedad del suelo, caudales, etc. varias instituciones han desarrollado bases de datos globales en las que ofrecen de manera abierta y gratuita este tipo de series temporales que son de interés hidrológico y climático. Estos productos están basados en datos puntuales medidos que han sido interpolados mediante técnicas estadísticas, en modelos climáticos e hidrológicos globales, y, en información basada en la teledetección provenientes de radares satelitales.

Además de bases de datos específicos para la gestión del agua y sequías, se presentan alternativas para obtener con información geográfica necesaria también para la representación de variables hidrológicas, mapas y resultados. Estos datos son archivos vectoriales correspondientes a los límites administrativos de los países y los límites de las principales cuencas hidrográficas del mundo. Existen cientos de fuentes de datos geográficos, productos ofertados por organizaciones gubernamentales, ambientales, entre otras, que ofrecen de manera gratuita y abierta este tipo de información. También se pueden encontrar archivos ráster con información topográfica y climática, entre otras.

6.1. Bases de datos complementarias

Los archivos vectoriales (shapefiles) de los límites de países o cuencas hidrográficas georreferenciados se descargan desde las bases de datos geográficos públicos GADM y FAO (Food and Agricultural Organization).

GADM es una base de datos de alta resolución de las áreas administrativas del mundo entero, que ofrece para todos los países, todos sus niveles administrativos y en cualquier período de tiempo. GADM creó los datos espaciales de muchos países a partir de bases de datos espaciales proporcionadas por los gobiernos nacionales, ONG, y/o de mapas y listas de nombres disponibles en Internet. La base de datos está disponible en varios formatos de exportación, incluidos los

archivos shapefiles que se utilizan en la mayoría de las aplicaciones SIG. Estos archivos pueden usarse también con el lenguaje de análisis de datos R. Los shapefiles delimitan las áreas para la descarga y análisis de datos, además de ser necesarios para creación de gráficos de datos descriptivos que incluyan mapas geográficos. La dirección web es <https://gadm.org/>. Cuenta con un entorno muy intuitivo en el que dentro de la pestaña Datos se elige el país de interés y se descargan los archivos vectoriales distribuidos en niveles de 0 a 5 donde 0 es el límite nacional y 5 es la división menor son la que cada país cuente (no todos tienen 5 niveles).

La FAO en su sección “geonetwork” ofrece una gran variedad de mapas interactivos, conjuntos de datos SIG, imágenes de satélite y aplicaciones relacionadas a su labor. Cuenta con una amplia variedad de información espacial global básica desarrollada para apoyar la toma de decisiones, promover enfoques multidisciplinarios para el desarrollo sostenible, y mejorar la comprensión de los beneficios de la información geográfica. GeoNetwork es una fuente de datos abierta que permite compartir fácilmente información temática referenciada geográficamente entre diferentes organizaciones. La dirección web es <http://www.fao.org/geonetwork/srv/en/main.home>; en la que se puede acceder a la pestaña “Hydrology and Water Resources” y obtener archivos vectoriales georreferenciados correspondientes a los límites de cuencas hidrográficas de todo el mundo a distintos niveles (cuencas y subcuencas), cauces de ríos y otros.

6.2. Bases de datos de variables hidrometeorológicas

6.2.1. Tipos de bases de datos

En este manual se muestra como descargar, extraer y aplicar series temporales de las siguientes variables: precipitación en mm/mes, temperatura media en °C al mes, ETP en mm/mes y ETr en mm/mes. Además de los índices de sequía VCI, TCI y VHI mensuales. Cada una de estas variables se mide o estima de manera diferente, por lo tanto, las bases de datos que ofrecen estos productos provienen de distintas fuentes. De manera general, los datos existentes se obtienen a partir 4 técnicas: a partir de mediciones in situ, datos obtenidos desde mediciones de satélites, datos modelados numéricamente mediante algoritmos y datos generados a partir de técnicas de reanálisis. Algunas bases de datos integran estas técnicas para mejorar la precisión de sus productos.

La forma tradicional de medir la precipitación y la temperatura (variables más comunes y básicas) es mediante el uso de pluviómetros y termómetros. Cada una de estas estaciones se localiza en un punto específico sobre el territorio. En los últimos años estas variables también se han estimados través de teledetección desde radares infrarrojos desde satélites que orbitan la Tierra. Esta última metodología es el mecanismo por el que se estiman las variables de las cuales se derivan los índices de sequía VCI, TCI y VHI.

La ETr y ETP, humedad del suelo y caudales (escorrentía) son otras de las variables importantes en el ciclo del agua y por ende para los estudios de sequías. Para medirlos se requieren lisímetros, perforaciones en el terreno y estaciones de medición de caudales en distintas ubicaciones de los cauces. Sin embargo, la aplicación de estos métodos suele ser costosas de instalar y mantener, por lo que estas variables no son menos comunes de medir, en especial en países y regiones con limitaciones socioeconómicas. La solución más común para obtener estos datos es a través de modelos hidrológicos/climáticos globales (grande escala) o locales (escala puntual). Los modelos utilizan variables de entrada medidas (precipitación y temperatura) para estimar las variables faltantes. Otra forma de estimación es a través de formulaciones matemáticas y estadísticas

basadas en relaciones empíricas entre las variables medidas y estimadas. Un ejemplo claro es el cálculo de la ETP mediante formulaciones como la de Thornthwaite (Thornthwaite, 1948) que estima la ETP en base a la temperatura y precipitación.

La gran variedad de metodologías para estimar variables hidrológicas hace que existan distintas bases de datos generalmente pertenecientes a grandes instituciones científicas (NASA, Universidades, gobiernos, etc.). Los productos medidos de manera tradicional (a través de estaciones de medición insitu) suelen ser intermitentes (con vacíos) y contener errores resultado de su proceso de medición. Además, que la cobertura a nivel global es muy distinta en diferentes partes del globo. Para solucionar este limitante, las instituciones han automatizado el análisis de varios datos puntuales y calculado una distribución de estos valores en el espacio con el uso de modelos de interpolación, downscaling y de técnicas de reanálisis. La aplicación de estas técnicas ha sido apoyada por datos obtenidos a partir de modelos numéricos y a partir de mediciones satelitales.

Dependiendo de la base de datos, se puede encontrar varios formatos de descarga. Aunque generalmente, al tratarse de parámetros variables en el tiempo, se unifica, en un único archivo, información espacial y temporal. Esta unión se realiza por capas, como una matriz de 3 dimensiones (o multimatriz). Por ejemplo, en un solo archivo se puede encontrar una malla de datos de precipitación para el mes de enero, otra malla de datos para el mes de febrero, otra para el mes de marzo, etc., guardándose 12 mallas (correspondientes a los 12 meses de un año), una detrás de otra. Las cuadrículas de la malla están asociadas a un valor (de precipitación, temperatura, etc.) comparable con el de una estación ubicada en el centroide de cada una, por lo que varias celdas pueden ser agregadas dentro de un área o región para formar una sola serie temporal promedio.

Debido a la gran cantidad de información que presenta cada base de datos, los archivos que contienen los datos están típicamente fragmentados en varios archivos correspondientes a los datos de cada año, mes u otros sobre el territorio. Aunque existen varios tipos de archivos, los más comunes que encontraremos y los que utilizaremos mayormente en este manual son los netCDF (formulario de datos comunes en red). Este es un formato de archivo destinado a almacenar datos científicos multidimensionales (en este caso variables de interés hidrológico). Cada una de estas variables se puede mostrar en una dimensión (por ejemplo, tiempo). Estos archivos (extensión común: .NC) son comparables a los archivos ráster y se gestionan con el paquete de R “ncdf4”.

Para otras extensiones, sin embargo, es necesario extraer una capa en concreto y convertirla al formato deseado (recortando la información correspondiente a nuestra zona de estudio), requiriéndose del uso de algoritmos específicos. El procesamiento de los productos exige consultar la documentación concreta de cada base de datos para conocer cómo explorar las diferentes capas de información. Los gestores de las bases de datos tienden cada vez más a facilitar esta tarea e incluso publicando códigos de lectura de archivos.

En este manual nos centraremos en datos históricos, sin embargo, existen varias bases de datos que ofrecen variables hidrológicas proyectadas hasta el año 2100, que pueden obtenerse de manera similar. Estas proyecciones están apoyadas en modelos climáticos globales que, basados en las mediciones históricas, proyectan las estimaciones acordes a distintas actuaciones del ser humano frente al cambio climático. Como ejemplo citaremos el portal “Climate Change Knowledge Portal” (<https://climateknowledgeportal.worldbank.org/download-data>) que ofrece una amplia

gama de variables de interés hidrológico que pueden descargarse y usarse de manera similar a la que explicaremos en los capítulos II y III.

A continuación, se presentan las bases de datos de las cuales se obtendrán las variables hidrológicas. Se explica de donde proviene la información, que variables y a cuáles escalas temporales y espaciales las ofrecen.

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° (≈ 5.6 km 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.

Aquí se usa la versión V2.0 (<https://www.chc.ucsb.edu/data/chirps>). Su precisión espacial motiva el uso de esta base de datos en estudios hidrológicos y gestión de sequías ya que es particularmente útil en estudios en zonas o cuencas de pequeño tamaño, y puede agregarse a cualquier superficie (e.g., K. Wang, Li, y Wei 2019). A pesar de que se indica que el alcance es global, únicamente se disponen datos de la superficie situada a menos de 50° al norte y al sur del plano ecuatorial. Las actualizaciones se realizan cada semana, permitiendo tener una monitorización en tiempo casi instantánea.

6.2.3. TerraClimate

El TerraClimate (<http://www.climatologylab.org/terraclimate.html>) procedente del Climatology Lab de la Universidad de Idaho ofrece varios productos a escala mensuales desde 1958 (Abatzoglou et al., 2018). Mediante técnicas de interpolación, downscaling y reanálisis han creado mallas a una resolución espacial de $1/24^\circ$ (≈ 4.6 km en el ecuador) de datos de precipitación, temperatura (máxima y mínima), presión de vapor, velocidad de viento y radiación de onda corta superficial en todo el mundo. Con esta información usando el modelo de balance hídrico global (WBM) de Thornthwaite-Mather (Willmott et al., 1985), se ha derivado la ETP (sigueiendo la formulación de Penman-Montieth), el caudal o escorrentía, ETr, déficit hídrico (D), humedad del suelo y agua de nieve equivalente en cada una de celdas.

La actualización de variables se hace cada año, por lo que su uso como fuente de datos para monitorizar a corto plazo no es posible. No obstante, debido a su alta resolución espacial puede ser usado en el análisis de sequías o balances hídricos en regiones de todo tamaño. Aquí se descargarán y usarán la precipitación, la temperatura media (derivada de la temperatura máxima y mínima) y el déficit hídrico.

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.

6.2.5. MODIS

El Moderate Resolution Imaging Spectroradiometer (MODIS) es un sensor a bordo de los satélites Terra y Aqua de la NASA (Vermote y Wolfe 2015). Ofrece, entre varios otros, productos satelitales de temperatura media y evapotranspiración real diaria y semanal desde el año 2001 hasta la actualidad. La resolución espacial es de 500 m y de 1 km, lo que la hace una base de datos de aplicación muy recomendable en cuencas hidrográficas o regiones muy pequeñas. La actualización de los datos es diaria, aunque en ocasiones puede sufrir ligeros retrasos que son subsanados en los días consecutivos. Los datos pueden encontrarse en <https://modis.gsfc.nasa.gov/data/dataproduct/mod09.php>. Nosotros descargaremos la temperatura y la ETr y la agregaremos a una escala mensual.

7. Objetivos específicos

Después de hacer una breve introducción a los conceptos asociados a la gestión de sequías y, de exponer las bases de datos de interés, en los capítulos II y III seguiremos los siguientes objetivos específicos:

- a) Descargar y graficar archivos vectoriales con los límites de Mozambique y de la cuenca hidrográfica del río Licungo.
- b) Descargar y datos mensuales de precipitación desde CHIRPS, obtenerlos dentro de Mozambique y graficar sus promedios anuales. Calcular el SPI y graficar sus características.
- c) Descargar datos semanales de VCI, TCI y VHI desde la NOAA STAR, obtenerlos dentro de Mozambique, agregarlos a escala mensual y graficar su desarrollo durante una época de interés.
- d) Descargar datos mensuales de precipitación, ETr y ETP desde el TerraClimate, obtener los datos dentro de la cuenca río Licungo, calcular balance y déficit hídrico mensual y graficar sus promedios anuales. Calcular el SPEI y tabular sus características.
- e) Graficar el estado de sequía/no sequía en base al SPI en Mozambique.
- f) Graficar el área afectada por sequías según el SPEI en la cuenca del río Licungo.
- g) Graficar las características espaciales de las sequías en Mozambique.

Para conseguir cada uno de estos objetivos se presentan 3 scripts disponibles en los ANEXOS, los mismos que se explican paso a paso en los capítulos II y III.

8. Seguimiento del manual

El ambiente de trabajo está pensado para el sistema operativo Windows. Para seguir el manual se requiere instalar R y RStudio, desde donde se pueden seguir los pasos detallados en los capítulos II y III según los scripts proporcionados en los anexos y que se organizan de la siguiente manera:

- a) El ANEXO I presenta el código necesario para extraer la información desde las distintas bases de datos. En una carpeta de destino, se guardarán los datos y quedarán ordenados y listos para su uso.
- b) El ANEXO II contiene las herramientas necesarias para leer la información descargada y ordenada obtenida en el anexo I y utilizarlo tanto para caracterizar sequías a través de índices como para realizar balances hídricos.
- c) El ANEXO III presenta los scripts necesarios para leer la información generada en los anexos I y II, graficarla y exportarla en distintos formatos.

Al inicio de cada script se detalla los paquetes que deberán descargarse (una sola vez por CPU utilizado) y activarse (cada vez que se abre una nueva sesión de RStudio) para usar esos códigos. R tiene funciones internas que permiten realizar la descarga y activación de paquetes de manera directa desde el entorno de trabajo. Para utilizar los códigos es necesario abrir un nuevo archivo “R Script” en RStudio y copiar los códigos de los anexos. Para ejecutar cada línea de código es necesario posicionarse al principio de cada línea y presionar la opción “run” en la parte superior derecha de la ventana de trabajo. Tanto en el manual como en el código se especifica el paso a paso.

Dentro de cada script se encuentran varias secciones que corresponden a cada uno de los subcapítulos del Manual (los paquetes se requieren para todo el código). Si se deseara seguir un solo proceso de descarga de datos, cálculo de índices y gráficos para cualquier región se podría copiar cada parte que se requiera de los 3 scripts presentados y ponerlos en orden en un nuevo archivo. Este manual pretende ser una guía desde la cual se pueda ampliar y modificar su contenido. Al avanzar con el manual se encontrarán pasos similares, los mismos que serán explicados al detalle la primera vez que aparezcan.

Se requiere tener descargado desde el GADM y la FAO y guardado en una carpeta dentro del explorador de Windows los archivos vectoriales (.shp) de Mozambique y la cuenca del río Licungo. La carpeta en los scripts se denomina “Layer_Mozambique”, y cada uno de los archivos shapefile “mozambique.shp” y “basin_Licungo.shp”. Sin embargo, los nombres de las carpetas y de los archivos pueden modificarse según gusto del usuario.



II. Obtención de datos

9. Zona de estudio

En esta sección se sigue el script de gráficos en el que simplemente se comprueba que los archivos que contienen los contornos de interés estén correctamente situados geográficamente. Para esto se requiere de un shapefile de todos los países del mundo descargados desde el GADM (https://gadm.org/download_world.html).

- 1) Se instalan y cargan los paquetes necesarios. De estos paquetes se utilizan algunas funciones, no obstante para profundizar en ellas o saber sobre todas funciones que ofrecen los paquetes se puede buscar en el CRAN los manuales ofrecidos por los programadores de dichos paquetes.

```
#The packages to be used are installed
install.packages("mapproj")#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("mapproj")#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(mapproj)
library(tmap)
library(lattice)
library(RColorBrewer)
library(ggplot2)
library(mapproj)
library(ncdf4)
library(raster)
library(dplyr)
library(sf)
```

- 2) Se leen los shapefiles de Mozambique y de la cuenca del río Lacungo, se han asignado los nombres mz y cl respectivamente. Se lee también el archivo con los límites de todos los países del mundo. Tener en cuenta la dirección donde se encuentra la carpeta y la dirección de la línea inclinada ("/").

```
#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) Se grafican los shapefiles de manera básica para comprobar que los archivos que he cargado sean correctos. Los resultados los vemos en el 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) Se grafican e importan los archivos de interés como .pdf y .png a una carpeta donde se comprueba que su ubicación geográfica sea la adecuada. Tener en cuenta la dirección donde



se encuentra la carpeta y la dirección de la línea inclinada. En este caso se comprueba que los archivos corresponden a los límites requeridos.

```
#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(c1)+tm_polygons(col="red")#A second overlapping shape is painted, delete if not required

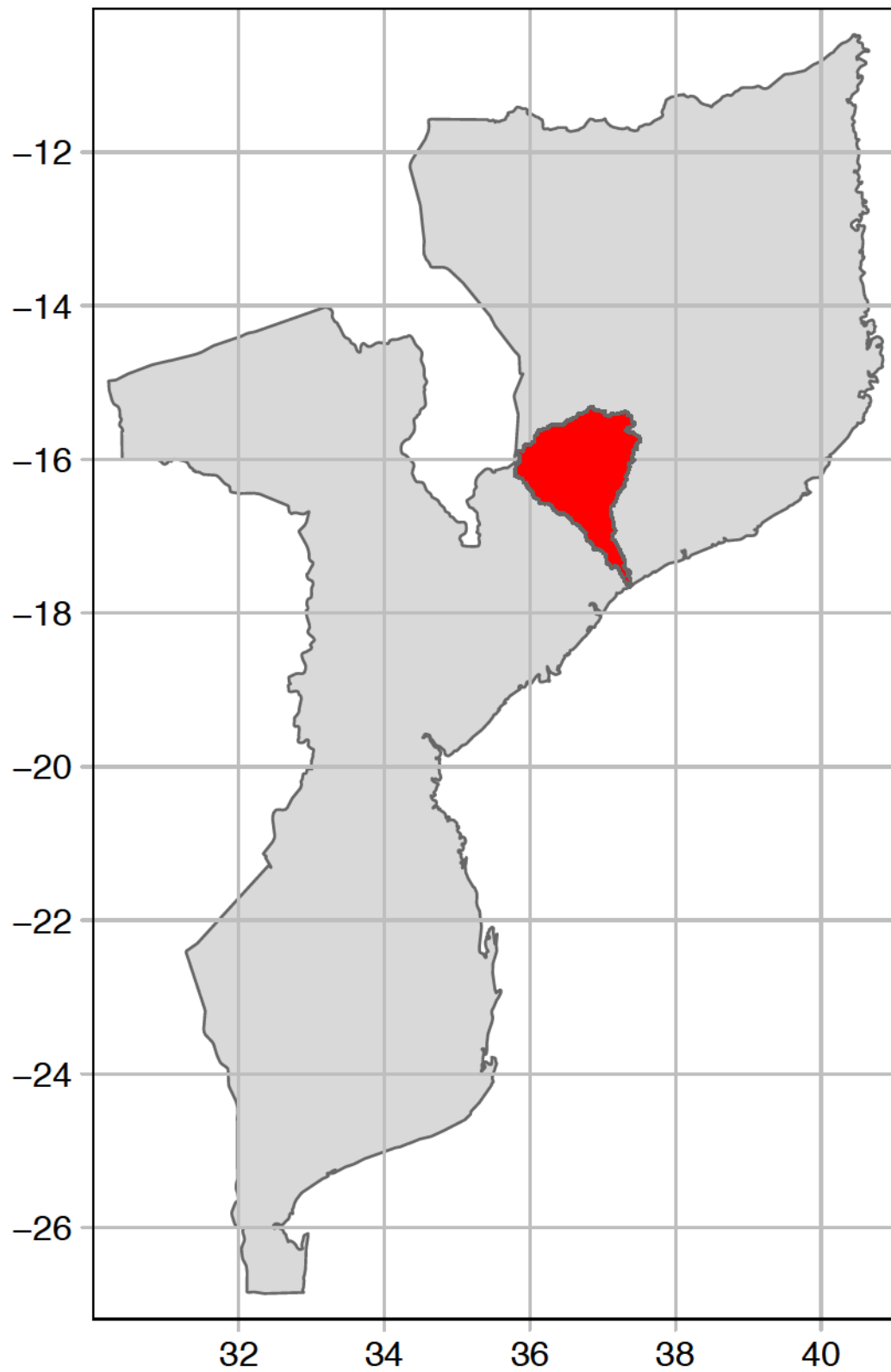
#The graphic is exported as png, the destination folder and the size can be modified
png("E:/19_MANUAL_data_global/Graphics/ubicacion_mz_cl.png",height = 270,width = 400)
graph #Assigned name of the graphic to be exported as png
dev.off() #Export closed

#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
```

En caso de que los archivos requieran cambios de proyección en el 1globo terrestre o no estén correctamente ubicados se deberá usar un software GIS para emendar los errores.



Fig. 6
Localización y representación
de los límites de Mozambique
y la cuenca del río Licungo



10. Precipitación: CHIRPS

En esta sección se descargan los archivos que contienen la información de precipitación desde la página web de CHIRPS. Luego se obtiene la información mensual de todos los píxeles que se encuentran dentro de Mozambique. Se agregan los datos a escala anual y se grafican de una manera sencilla y elegante con alta calidad. Los pasos para seguir son:

- 1) Desde la página de CHIRPS https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_monthly/netcdf/ se descarga el archivo: `chirps-v2.0.monthly.nc`. Este archivo pesa alrededor de 6Gb y contiene la información global de precipitación mensual desde 1981. Se guarda en una carpeta conocida. Desde esta página web, se pueden descargar otros archivos como los datos de precipitación diaria, trimestral, entre otras.



- 2) Se instalan y cargan los paquetes necesarios.
- 3) Se lee el shapefile de Mozambique, lo hemos asignado con el nombre `mz`.
- 4) Se lee el archivo `chirps-v2.0.monthly.nc`, siguiendo el script (variable denominada “`chi`” se puede cambiar por el que se desee).

```
#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) Se muestra en la consola la variable “`chi`” mediante la función `print`, que permite ver las características del archivo. Aquí se extrae toda la información necesaria para comprender como están almacenados los archivos y de esta manera extraerlos en la zona de interés correctamente. Este paso será común para todos los archivos con extensión `.nc`.

En la primera sección se muestra el número de variables que contiene el archivo. En este caso contiene una sola variable denominada “`precip`” y que almacena sus datos en una matriz tridimensional basada en longitud, latitud y tiempo. La unidad de cada dato es mm/mes, el paso de tiempo mensual y para una celda con dato perdido se asigna el valor de -9999:

```
> print(chi)
File E:/19_MANUAL_data_global/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
```


La siguiente sección informa las características de las dimensiones en la que el archivo guarda la información. En este caso divide al mundo en dirección longitudinal en 7000 partes desde el este y en dirección latitudinal en 2000 partes desde el norte. Estas matrices existen para 468 pasos de tiempo (meses) que se denominan con un número de días desde el 1 de enero de 1980 (es decir por ejemplo que el número 29 se refiere al 29 de enero de 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
```

Finalmente, el archivo contiene varios atributos relativos a la fuente de información. En este caso, existen 15 atributos entre ellos la versión, el título, un email de contacto al creador, la fecha de actualización, institución desarrolladora, 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) Se extrae la variable “time” del archivo “chi” (tiempo) para crear un vector con las fechas a la que pertenece cada paso de tiempo. Se generan vectores del año y mes adicionales en forma numérica.

```
time<-data.frame(as.Date(ncvar_get(chi, "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) Al finalizar el análisis de información y extracción de las fechas a la cual los datos pertenecen, se cierra el archivo. Se recomienda no dejarlo abierto ya que normalmente estas variables usan gran espacio de memoria en disco, pudiendo relentecer el programa o evitar la creación de nuevas variables.

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

- 8) Se extrae la información de precipitación. Aquí se abre nuevamente el archivo CHIRPS con la particularidad de leer directamente la variable “precip” que se analizó previamente. A la variable que se le asigno esta información se le examina la clase y modo, en la que se comprueba que sea tipo “RasterBrick” y “S4” que corresponde a información ráster.

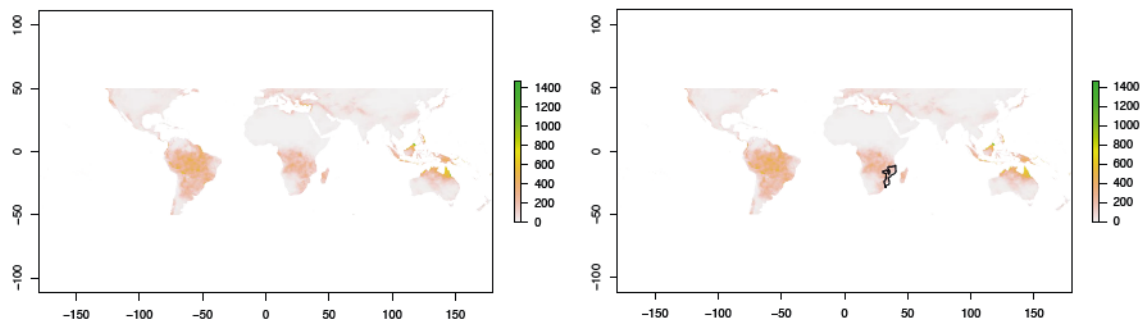
```
#Precipitation information is extracted
precip<-brick("E:/19_MANUAL_data_global/CHIRPS/chirps-v2.0.monthly.nc", varname = "precip")#Precipitation is extracted
class(precip)#It shows what kind of variable it is
mode(precip)#It is shown whether the variable information is numeric or text
dim(precip)#The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

- 9) Como la dimensión de cada ráster tiene 468 pasos de tiempo, se extrae uno de ellos (el primero en el script) para analizar su composición. Al extraer esta información se verifica su dimensión. Además, se grafica primero por si sola y luego con el shapefile del área de interés (Mozambique) para comprobar que este correctamente georreferenciado.

```
#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
```

En el viewer obtendremos algo así:



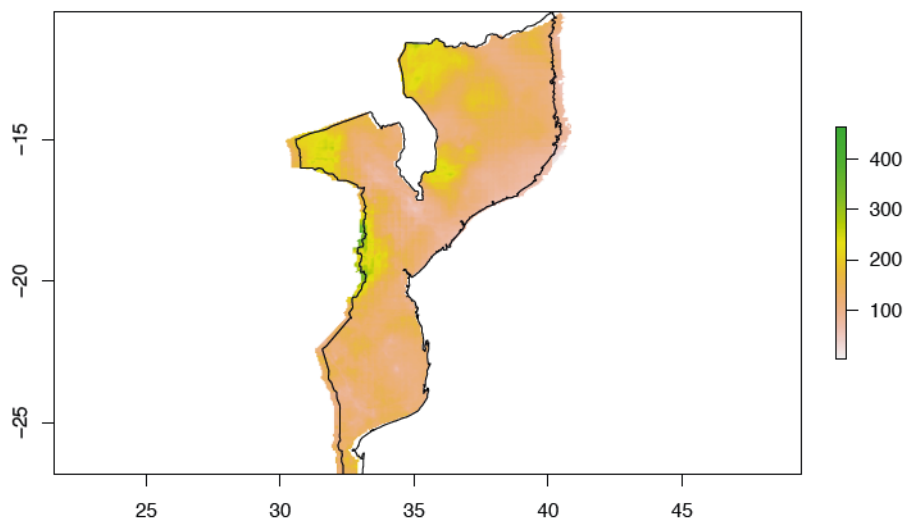
Donde se puede confirmar la correcta coincidencia espacial de la información de precipitación ráster con el shapefile “mz”. Se puede ver también como la precipitación mundial para este paso de tiempo en particular (enero de 1981) varía desde 0 hasta un poco más 1400 mm.

- 10) Una vez comprobado que coincide la localización de la información ráster con los shapefiles, se extraen solamente los datos dentro del área mozambiqueña. El procedimiento se realiza en tres pasos:

```
#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)
```

El gráfico permite comprobar que los datos fueron correctamente extraídos:



- 11) Se extrae las coordenadas de cada píxel con información de precipitación dentro del shapefile de Mozambique. Seguimos los pasos del script y se obtienen una matriz con dos columnas de pares ordenados longitud y latitud.

```
#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
```

Para saber cuántos píxeles hay se usa la función `dim` (26868 píxeles), y se grafican para observar en el viewer si se extrajo correctamente los pares ordenados:



Esta matriz (denominada `coordinates`) se guarda en una carpeta de interés en formato `.txt`. Este archivo es muy importante, pues contiene la localización de las celdas contenidas en el shapefile de interés y el orden en que se guarde no deberá cambiarse posteriormente (paso siguiente).

```
#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) Durante el proceso de extracción de coordenadas (usando la función “`extract`”), además de las coordenadas, se extraen los datos de precipitación asociada a cada coordenada. En este paso se lanza un bucle donde “`i`” varía desde 1 hasta el número de iteraciones de tiempo que contiene el archivo “`precip`”. Se repite el proceso de extracción de datos para cada uno de los pasos de tiempo y se almacenan en una matriz. Este proceso tarda dependiendo del número de píxeles, es decir del tamaño de la región bajo estudio. Al final del bucle se revisa su dimensión, que deberá tener una fila de datos por cada paso de tiempo y una columna de izquierda a derecha correspondiente a la serie de la coordenada ordenada de arriba abajo del paso anterior. Es muy importante mantener el archivo de coordenadas y de datos de precipitación juntos, pues su orden es primordial para el correcto uso de los datos.

```

#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) Las series de datos mensuales se guardan en formato .RDS, que es un formato compatible en RStudio que es más eficiente que un texto cuando las matrices son de gran tamaño. Se puede cargar nuevamente la matriz (función “readRDS”) y con el uso de la función “head” se observa la composición superior izquierda de datos de la matriz y se comprueba que el almacenaje de los datos fue correcto:

```

#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) Se agregan los datos mensuales a anuales mediante un bucle. El código filtra la información de cada píxel asociada a cada año (12 meses), la suma y la acumula en una nueva matriz. Esta nueva matriz tendrá dimensiones de número de años por número de píxeles. Esta nueva variable se guarda como archivo .RDS o .txt.

```

#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) Desde la matriz de datos anuales se puede estimar la precipitación media anual para cada píxel. Se la computa generándose una matriz de 3 columnas. Las dos primeras con el par ordenado de las coordenadas a la que pertenece cada dato y una tercera con el vector resultado de promediar las precipitaciones anuales. Este archivo se guarda en una carpeta de conveniencia.

```

#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) Para graficar de manera elegante y profesional las precipitaciones medias anuales, se abre el archivo con dicha información. Se convierte la matriz en un ráster (ver código con las funciones utilizadas), y se usa la función “levelplot” para generar un gráfico de alta calidad. Se pueden manipular la gama de colores, los valores que se muestran en la leyenda, mostrar o no cuadrantes geográficos, marcos, etc. Este gráfico se puede guardar en el formato que se desee.

```
#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_l<-rasterFromXYZ(dat_grafico)#The information is rasterised
class(raster_l); mode(raster_l)#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 ""

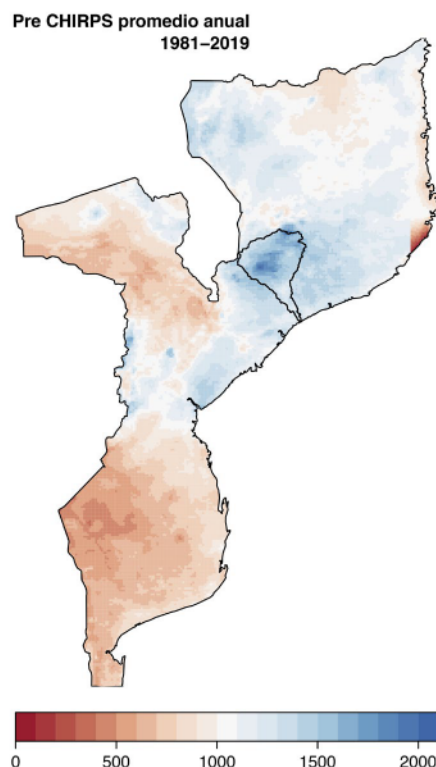
#The data is plotted in an elegant way
professional_graphics<-levelplot(raster_l,main="Pre CHIRPS promedio anual
1981-2019",ylab="",xlab="",#Choose the raster to plot

margin=FALSE, #The title can be changed
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_profesional2.png",height = 2700,width = 4000,res=900)
professional_graphics #Assigned name of the graphic to be exported as png
dev.off() #Export closes

#The graphic is exported as pdf, the destination folder and the size can be modified.
pdf("E:/19_MANUAL_data_global/Graphics/pre_anual_CHIRPS_profesional2.pdf", width = 5.25, height = 6.75, onefile = FALSE)
professional_graphics #Assigned name of the graphic to be exported as pdf
dev.off() #Export closes
```

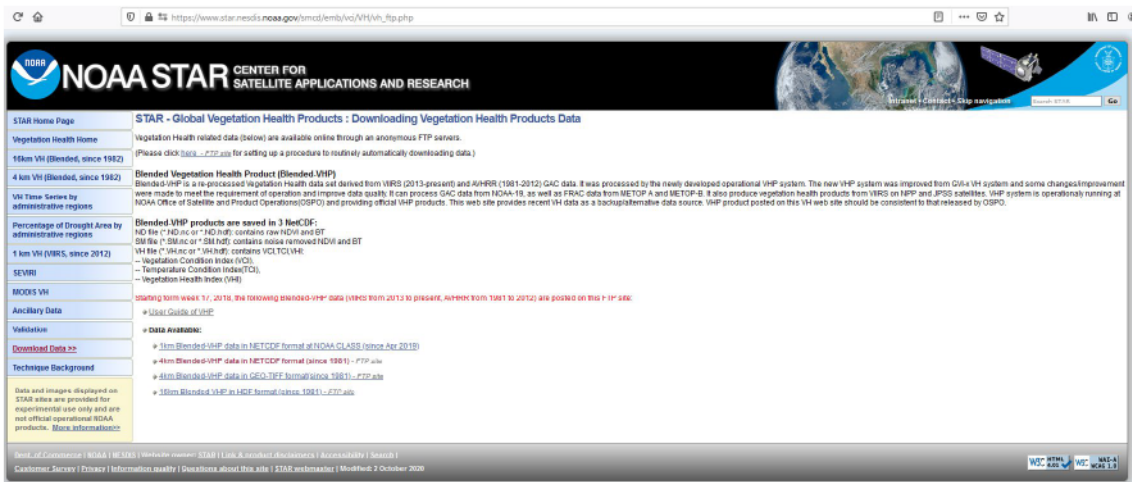
El resultado de este gráfico se verá así, en el que además hemos incluido el shapefile de la cuenca del río Licungo:



11. Índices de vegetación desde NOAA STAR

En esta sección se descargan para Mozambique los índices de sequía VCI, TCI y VHI desde la NOAA STAR. El proceso puede aplicarse a otros productos como el NDVI o el BT. Se utiliza otra estrategia de descarga de datos diferente a la usada desde el CHIRPS. Las series de datos vienen en escala semanal, por lo que se agregan a escala mensual para ser comparables con la precipitación u otros índices de sequía. Se grafica el estado mensual de 6 meses consecutivos para analizar su cambio.

- 1) Los archivos que contienen los productos de NOAA STAR se descargan desde la página web: https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. En la que se puede observar la siguiente información:



Donde, se especifica que los productos están guardados en 3 archivos NetCDF (extensión .nc). Se indican cuales archivos contienen los distintos datos disponibles y se tiene acceso a un manual de usuario el cual es recomendado revisar. Por lo tanto, aquí se requiere los "VH files" que contienen los índices VCI, TCI y VHI. Se entra dentro de la opción para descargar datos a una escala de 4 km.



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

Se descarga uno a uno todos los archivos VH y se guardan en una carpeta de interés. Cada uno de estos archivos presenta las series de los índices VCI, VHI y TCI en todo el mundo a una escala de 4km para cada semana desde 1981. Al ser una malla muy pequeña, la NOAA STARR ha decidido ofrecer un archivo por cada semana. Si bien existen mecanismos de programación para realizar la descarga de forma directa, en este manual de iniciación lo haremos manualmente.

Adicionalmente en la carpeta donde guardemos los archivos (1946 archivos desde la semana 35 de 1981 hasta la semana 52 de 2019), se requiere un archivo tipo texto (.txt), que contenga los nombres de cada uno de los archivos descargados.

- 2) Una vez en RStudio, se instalan y cargan los paquetes necesarios.
- 3) Se lee el shapefile de Mozambique, lo hemos asignado con el nombre mz.
- 4) Se lee el archivo que contiene los nombres de los archivos VH. Se extraen de cada uno de ellos los datos en la zona que nos interese.

```
#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) Al igual que con los archivos de datos CHIRPS, se carga uno de los archivos VH y se imprimen para ver su configuración. Aquí se usa la lista de nombres de los archivos en .txt para leer los archivos.

```
#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
```

Lo primero que se observa es el número de variables que contiene cada uno de estos archivos:

```
> print(aux)#The file is displayed and its composition is shown
File E:/data_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):
```

De las 4 variables o productos, el VCI, TCI y VHI se llaman de esa manera y vienen con una configuración que depende de las variables “WIDTH” y “HEIGHT”. Varían de entre 0 y 100 y se encuentran sobre el sistema cartesiano.

```
  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 probably 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 probably fail at this point"
```

Lo siguiente que se puede ver es el tamaño de las dimensiones “WIDTH” y “HEIGHT”. Esto significa que cada matriz que contiene cada variable tiene 10.000 × 3.616 celdas. Además, se ve que no tiene dimensiones, esto significa que no es posible extraer directamente los datos utilizando la estrategia usada en CHIRPS. Se explica lo que se hace en estos casos más adelante.

```
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

```

Finalmente se muestran otros varios atributos del archivo, como son el instrumento de medición, el satélite, la fecha en que la misión de toma de datos se desarrolló, el año y la semana a los que corresponden los datos contenidos, entre otros. Datos relevantes son el rango en las que cada píxel se localizan geográficamente.

- 6) En este paso se extrae a modo de ejemplo desde este archivo los valores de VCI. Se hace esto para observar su composición y saber cómo tratar los datos más adelante. Se guarda en una variable (vci en este caso) la matriz de dimensión 10.000 x 3.616 (lo revisamos con la función “dim”). Se revisa además la clase y modo de la variable, que deberá ser una matriz con valores numéricos.

```

#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) Terminada la inspección de los datos, se procede a crear dos vectores: uno de longitud y latitud que correspondan a las dimensiones “WIDTH” y “HEIGHT”. Dado que no vienen georreferenciadas tenemos que crearlas manualmente utilizando la información extraída en el paso 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) Las coordenadas longitudinales y latitudinales de todos los píxeles que conforman la base datos se acotan según los límites que marquen un rectángulo que contenga el shapefile de interés. De esta manera la matriz que contenga las coordenadas de la cual se extraerán las coordenadas contenidas en el shapefile de interés no será excesivamente grande. Se utilizan como límites longitudinales y latitudinales los que observamos en la fig. 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

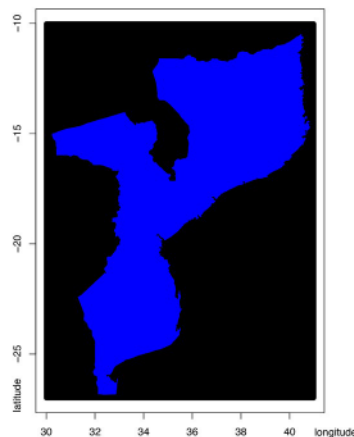
lat_a<-lat[lat<=lat2]
lat_b<-data.frame(lat_a[lat_a>=lat1])#Latitudes are cut off within the boundaries

```

- 9) Utilizando los vectores de longitud y latitud ya acotados a los límites del shapefile (lon_b y lat_b), se crea una matriz con los pares ordenados correspondientes a todas las coordenadas contenidas en los límites propuestos. A esta matriz se la grafica y superpone el shapefile de interés de manera que comprobemos de manera visual que las coordenadas creadas abarquen todo el territorio de interés.

```
#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) El siguiente paso es extraer las coordenadas que se encuentren contenidas en el shapefile. Se puede hacer el uso un SIG o en directamente R, que es como se lo realiza aquí. Primero, se transforma la matriz de coordenadas a un archivo de clase espacial (puntos con información geográfica) donde la proyección debe ser definida, en este caso se usa la proyección geográfica WGS84. Luego, se extraen las coordenadas contenidas en el shapefile de interés, en este caso “mz”. Finalmente, se transforma nuevamente el archivo a una matriz con datos numéricos (sin información geográfica), se revisa la dimensión de este archivo, se grafica para observar si se ha realizado de manera correcta la operación y se guarda la matriz en una carpeta conocida.

```
#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) Dado que las variables a extraer dependen de las dimensiones “WIDTH” y “HEIGHT”, y estas varían desde 1 a 1000 y 1 a 3616, respectivamente, se busca la posición a la que corresponde cada coordenada contenida en el shapefile. Para esta búsqueda, se redondean los decimales de las coordenadas de la matriz objetivo (contenidas en mz) y las de los vectores que contienen todas las longitudes y latitudes posibles. El resultado será una matriz con un número de filas igual al número de coordenadas contenidas en mz y 2 columnas correspondientes a las posiciones de “WIDTH” y “HEIGHT”, se han denominado “lol”.

```
#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){
    lol<-cbind(which(r2[,1]==r1[i,1]),which(r3[,1]==r1[i,2]))
  }
  if (i!=1){
    lol<-rbind(lol,cbind(which(r2[,1]==r1[i,1]),which(r3[,1]==r1[i,2])))
  }
}
#lol is the variable containing the positions within the files to extract for each coordinate
colnames(lol)<-c("longitude", "latitude")#Name the position columns
dim(lol)#Check dimension, must be equal to that of coordinates
head(lol)#Matrix organisation is checked, must be equal to that of coordinates
```

- 12) Se lanza un bucle con un número de iteraciones igual al número de archivos semanales a extraer la información. Cada iteración sigue el orden del vector que contiene los nombres de los archivos, por lo que se debe tener cuidado de escribir correctamente el nombre de las carpetas donde están guardado los archivos .nc originales. Los resultados serán 3 matrices, correspondientes a los valores de VCI, TCI y VHI de todas las celdas contenidas en el shapefile. Su dimensión será el número de filas igual al número de semanas con datos y el número de columnas igual al número de píxeles. Se debe tener en cuenta que el orden que siguen las coordenadas de izquierda a derecha es coincidente con el orden de arriba a abajo de las coordenadas obtenidas en el paso 10.

Este paso puede tardar varias horas, dependiendo del tamaño de la región de interés.

```
#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_NOASTAR/doc_pag/web/",vh[i,1]))#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(lol)){#The number of steps in the loop is the number of cells we will extract values from
    if (j==1){
      vcil<- data.frame(year,month,week,vci[lol[j,1],lol[j,2]])
      tcil<- data.frame(year,month,week,tci[lol[j,1],lol[j,2]])
      vhil<- data.frame(year,month,week,vhi[lol[j,1],lol[j,2]])
    }
    if (j!=1){
      vcil<- cbind(vcil,vci[lol[j,1],lol[j,2]])#The vector contains the VCI data of the corresponding week for each cell
      tcil<- cbind(tcil,tci[lol[j,1],lol[j,2]])#The vector contains the TCI data of the corresponding week for each cell
      vhil<- cbind(vhil,vhi[lol[j,1],lol[j,2]])#The vector contains the VHI data of the corresponding week for each cell
    }
    print(paste0(j,"de",nrow(lol)))#Step counter
  }
  colnames(vcil)<-c("year","month","week",1:(ncol(vcil)-3))#The columns are named
  colnames(tcil)<-c("year","month","week",1:(ncol(tcil)-3))#The columns are named
  colnames(vhil)<-c("year","month","week",1:(ncol(vhil)-3))#The columns are named
  if (i==1){
    vci2<-vcil
    tci2<-tcil
    vhi2<-vhil
  }
  if (i!=1){
    vci2<-rbind(vci2,vcil)#The matrix accumulates the corresponding weekly VCI series for each cell
    tci2<-rbind(tci2,tcil)#The matrix accumulates the corresponding weekly TCI series for each cell
    vhi2<-rbind(vhi2,vhil)#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(vci2)#Check dimension of matrix, must be number of weeks x number of cells +3 (year, month and week)
dim(tcil2)#Check dimension of the matrix, must be number of weeks x number of cells +3 (year, month and week)
dim(vhi2)#Check dimension of the matrix, must be number of weeks x number of cells +3 (year, month and week)
```

- 13) Se guardan las matrices creadas en archivos con formato .RDS. Se recomienda este tipo de archivos frente a los .txt ya que comprimen la información resultando en variables más fáciles y rápidas de leer en RStudio. Se pueden cargar nuevamente las matrices, observar la composición superior izquierda de datos y comprobar que el almacenaje de los datos fue correcto:

```
#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) El siguiente paso será agregar los datos semanales en datos mensuales mediante un bucle. Primero se leen los datos semanales y se reemplaza cualquier valor extraño que pueda haberse generado por NA. Un NA es un valor no medido, pero con el cual se puede operar, mientras otros como los NAN representan un error con el cual no se puede realizar cálculos.

```
#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
```

El código filtra la información de cada píxel asociado a las semanas de cada mes (12 meses), la promedia y la guarda en una nueva matriz. Esta nueva matriz tendrá dimensiones de número de meses por número de píxeles. Esta nueva variable se guarda como archivo .RDS o .txt.

```
#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[,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[,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[,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 dimension 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) En la última sección se muestra como generar varios gráficos correspondientes al estado de una variable sobre todo el territorio de interés, juntarlos en una matriz y exportarlos. Se usa el VHI mensual como ejemplo. Se convierte la matriz VHI en un ráster (ver código con las funciones utilizadas), y se usa la función “levelplot” para generar un gráfico de alta calidad. Se repite el proceso para 12 filas consecutivas, correspondientes a los meses de octubre de 1992 a septiembre de 1993, época en la cual se puede observar el desarrollo de una importante sequía ocurrida en Mozambique.

```
#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,v(vhi[98,-c(1:2)]))#A matrix containing the coordinates of each
#pixel and its corresponding annual average
#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,v(vhi[99,-c(1:2)]))
dat_grafico3<-data.frame(coordinates,v(vhi[100,-c(1:2)]))
dat_grafico4<-data.frame(coordinates,v(vhi[101,-c(1:2)]))
dat_grafico5<-data.frame(coordinates,v(vhi[102,-c(1:2)]))
dat_grafico6<-data.frame(coordinates,v(vhi[103,-c(1:2)]))
dat_grafico7<-data.frame(coordinates,v(vhi[104,-c(1:2)]))
dat_grafico8<-data.frame(coordinates,v(vhi[105,-c(1:2)]))
dat_grafico9<-data.frame(coordinates,v(vhi[106,-c(1:2)]))
dat_grafico10<-data.frame(coordinates,v(vhi[107,-c(1:2)]))
dat_grafico11<-data.frame(coordinates,v(vhi[108,-c(1:2)]))
dat_grafico12<-data.frame(coordinates,v(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

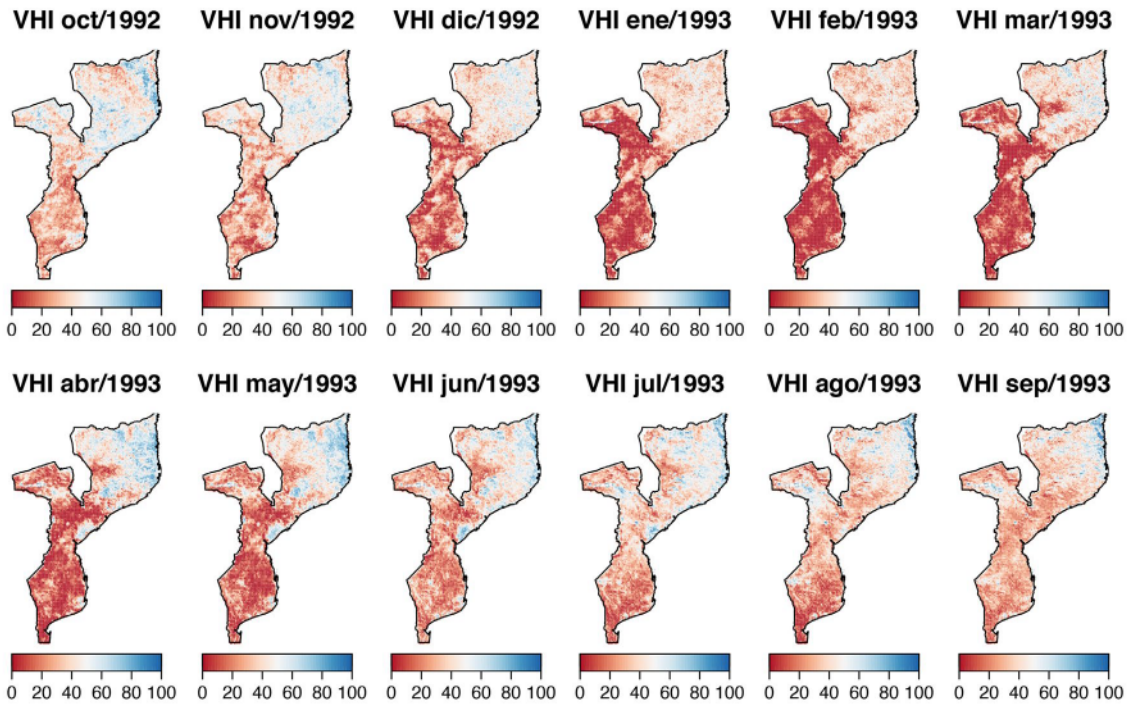
max(dat_grafico[,3],na.rm = T)#Maximum value, 3 is the column where the data of interest is located
min(dat_grafico[,3],na.rm = T)#Minimum value, 3 is the column where the data of interest is located

#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 puede 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(ms))#Limits of interest are attached
```

Finalmente para imprimir los 12 gráficos generados se usa la función “grid.arrange”, en la cual es necesario definir como ordenar los gráficos. Aquí se define una matriz de 2 x 6. Este arreglo debe definirse según el tamaño del gráfico que se desee. El código necesario para realizar este paso y el resultado final es:

```
#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. Varios: TerraClimate

En esta sección se descargan los datos de precipitación, ETr y ETP desde la base de datos TerraClimate sobre la cuenca de río Licungo. Utilizaremos una estrategia mixta entre las dos aplicadas anteriormente. Las series de datos vienen en escala mensual. Aunque solo se muestra cómo obtener estas tres variables mencionadas, el proceso será idéntico para todos los productos que ofrece esta base de datos.

- 1) Los archivos de TerraClimate se descargan desde la página web: https://climate.northwestknowledge.net/TERRACLIMATE/index_directDownloads.php. En la que se puede observar la siguiente información:

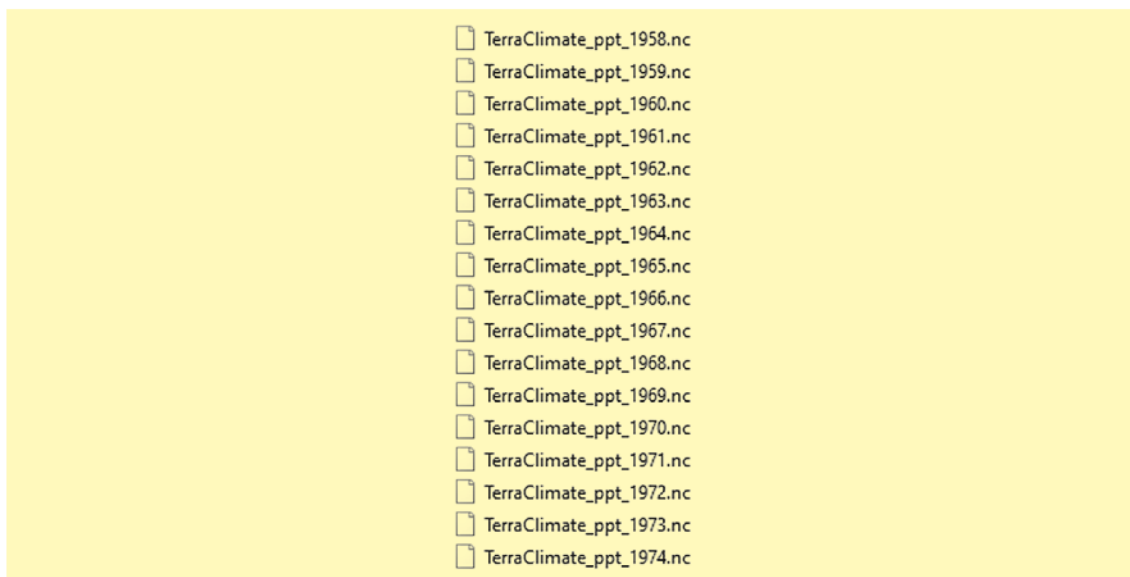
Filter the directory listings:

Select a Year: ▾

Select a Variable:

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

Se elige la variable de interés y se descarga uno a uno los archivos anuales que se requieran. Aquí se muestra la extracción de datos desde 1958 hasta 2019. Se guardan los archivos en una carpeta de destino. Los nombres de los archivos varían solamente en el número del año al que corresponden.



- 2) Se instalan y cargan los paquetes necesarios.
- 3) Se lee el shapefile de la cuenca del río Licungo, asignado con el nombre cl.
- 4) Al igual que con las otras fuentes de información, se lee un archivo y se analiza su composición.

- 5) Las características del archivo se muestran en la consola aplicando la función “print” a la variable “aux”. Al igual que para todo archivo .nc se analiza toda la información necesaria para comprender como están almacenados los archivos y extraerlos en la zona de interés correctamente.

En la primera sección se muestra el número de variables que contiene el archivo. En este caso contiene una sola variable denominada “ppt” y que almacena sus datos en una matriz tridimensional basada en longitud, latitud y tiempo. La unidad de cada dato es mm y el paso de tiempo mensual:

```
#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
```

```
> print(aux)  
File E:/data_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,EPSSG:4326  
    scale_factor: 1  
    add_offset: 0  
    unsigned: true
```

La siguiente sección informa las características de las dimensiones en la que el archivo guarda la información. En este caso divide al mundo en dirección longitudinal en 8640 partes desde el este y en dirección latitudinal en 4320 partes desde el norte. Estas matrices existen para 12 pasos de tiempo (meses) que se denominan con un número de días desde el 1 de enero de 1900 (es decir por ejemplo que el número 29 se refiere al 29 de enero de 1900). La dimensión “crs” nos indica la proyección con la que se han guardado los datos.

```
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
```

Finalmente, el archivo contiene varios atributos relativos a la fuente de información.

- 6) Lo siguiente es analizar las variables que se encuentran contenidas en cada uno de los archivos anuales. Se extrae la información de precipitación como ejemplo. Aquí se lee nuevamente el archivo de TerraClimate con la particularidad de leer directamente el dato “ppt” que se analizó previamente. A esta variable se le examina la clase y modo, en la que se comprueba que sea tipo “RasterBrick” y “S4” que corresponde a información ráster. Como se conoce que la dimensión de cada ráster mundial contiene 12 pasos de tiempo, se extrae uno de ellos (el primero en el script) para analizar su composición. Al extraer esta información se verifica su dimensión, la cual debe ser de 2 dimensiones con valores iguales a al número de divisiones longitudinales y latitudinales en las que la base de datos ofrece sus productos.

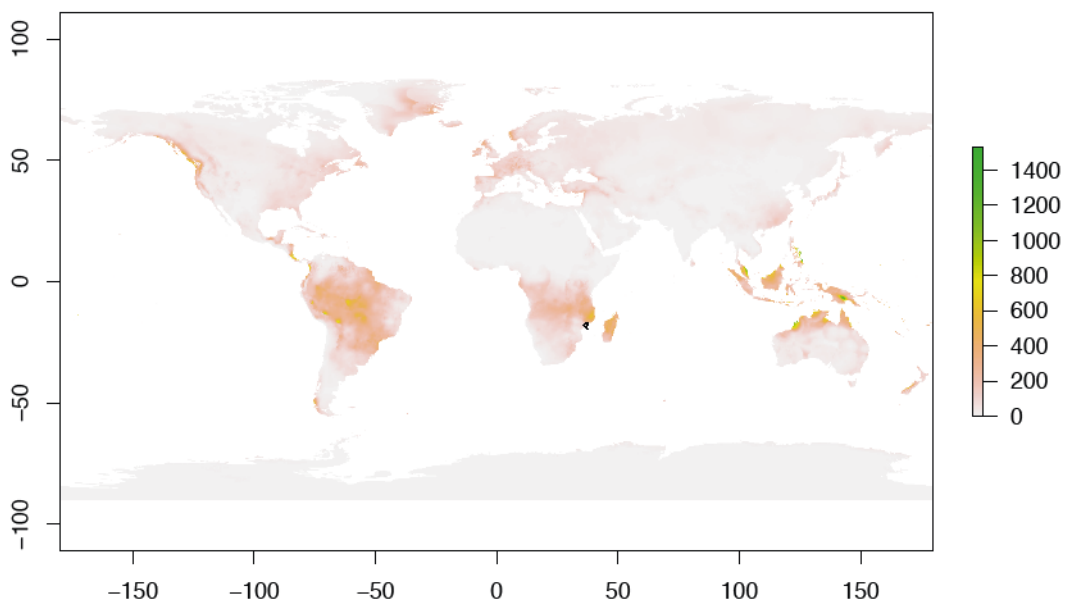
```
#A variable is extracted to analyse its composition

# 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
```

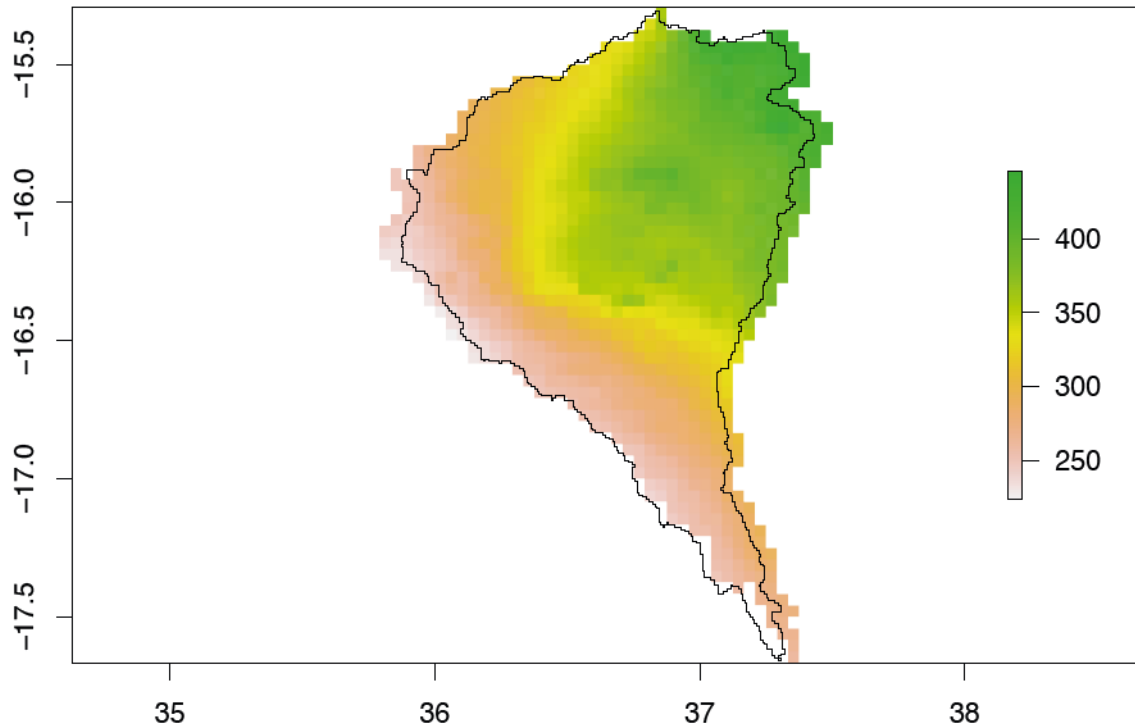
Para comprobar la localización de la información y del shapefile de interés, se grafica el ráster “precip_1” con el shapefile del área de interés (cuenca del río Licungo). En el viewer se obtiene algo así:



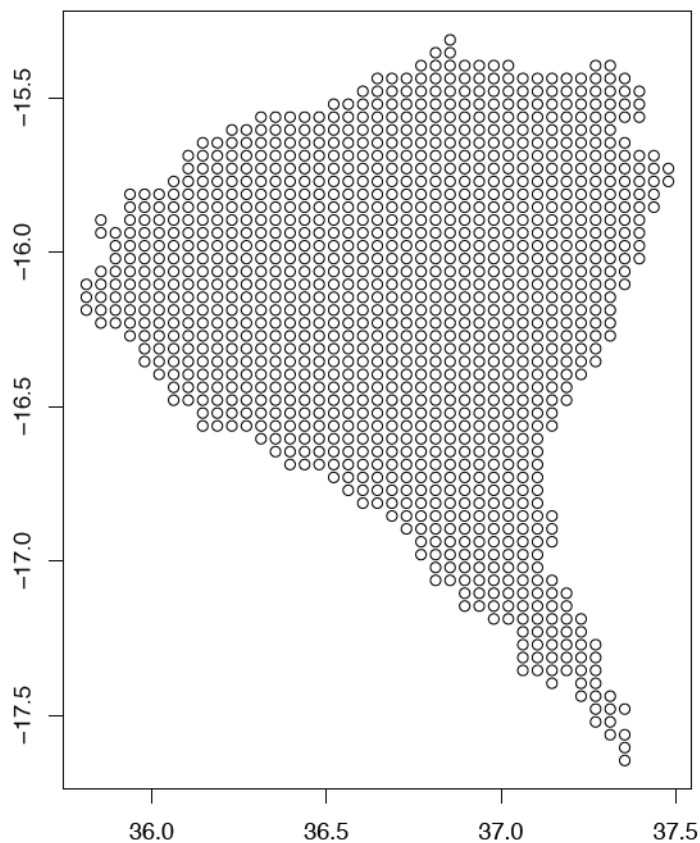
- 7) Una vez comprobado la información ráster está correctamente georreferenciada, se procede a extraer solamente los datos dentro del área de interés, siguiendo los pasos similares a los usados en la extracción de datos desde el CHIRPS. El procedimiento se realiza en tres pasos:

```
#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) Se extraen las coordenadas de cada píxel con información de precipitación dentro del shapefile de la cuenca del río Licungo. Siguiendo los pasos del script se obtienen una matriz con dos columnas de pares ordenados longitud y latitud. Mediante el uso de la función `dim` se obtiene el número de celdas contenidas en el área (1119 píxeles), y se grafican para comprobar en el viewer si los pares ordenados están correctos:



Esta matriz (denominada `coordinates`) se guarda en una carpeta de interés en formato `.txt`. Este archivo es muy importante, pues contiene la localización de las celdas contenidas en el `shapefile` de interés y el orden en que se guarde no deberá cambiarse posteriormente (paso 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

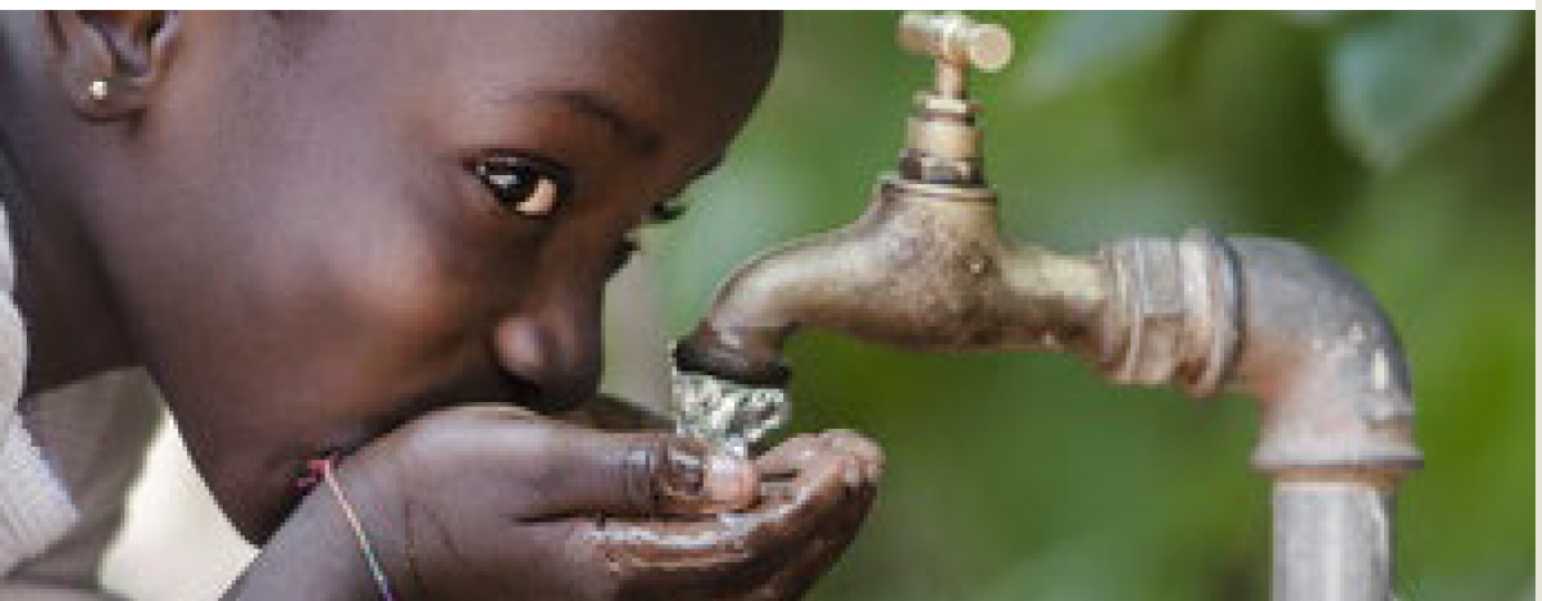
#The coordinates are saved as a .txt file
write.table(coordinates,"E:/19_MANUAL_data_global/TerraClimate/coordendas_TC_Licungo.txt",col.names = T,row.names = F)
#Change the desired location
```

- 9) Dado que cada archivo `.nc` contiene 12 ráster, que contienen la información de las variables de cada mes, se crea un vector de fechas, años y meses correspondientes al período de interés. Aquí se descargan todos los archivos desde 1958 hasta el 2019.

```
#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) Se lanza un bucle con un número de iteraciones igual al número de archivos anuales que queramos extraer la información. La variable "k" varía según el año correspondiente a cada archivo, por lo que cada iteración seguirá el orden anual de los archivos según sus nombres. Se debe tener cuidado de escribir correctamente el nombre de las carpetas donde están guardados los archivos `.nc` originales. En cada iteración del bucle, se revisa la composición de los datos de cada archivo y se almacena la información de la variable dentro del `shapefile` de interés en una matriz. El resultado serán 3 matrices, correspondientes a los valores de precipitación, ETP y ETr de todas las celdas contenidas en el `shapefile`. La dimensión de cada matriz será de un número de filas igual al número de meses con datos y de un número





de columnas igual al número de píxeles. Se debe tener en cuenta que el orden que siguen las coordenadas de izquierda a derecha es coincidente con el orden de las coordenadas de arriba a abajo obtenidas en el paso 8.

Este paso puede tardar varias horas, dependiendo del tamaño de la región de interés.

```
#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
pa11 <- subset(aa1,j)#The raster of the month varying from 1 to 12 is extracted when "j" varies
pa12 <- subset(aa2,j)
pa13 <- subset(aa3,j)

step1_1 <- crop(pa11, cl) #Intersect the raster with the shapefile of interest
step1_2 <- crop(pa12, cl)
step1_3 <- crop(pa13, 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) A las tres matrices que contienen datos se le agregan los 3 vectores de fechas, años y meses y se los guarda en formato .RDS.

```
#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")
```

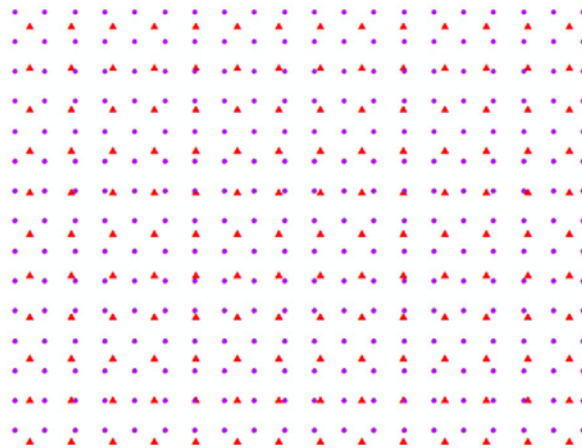
En la consola se puede observar los resultados de aplicar la función head a las tres matrices, donde se puede mirar y comprobar la composición final de las matrices.

```
> 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. Solución para unificar escalas espaciales:

En esta sección se explicarán soluciones sencillas en caso de requerir que dos variables tengan exactamente las mismas coordenadas. Se analizan dos opciones, la primera será el caso de requerir de una sola serie temporal representativa de una zona específica; y, la otra cuando necesitamos que todas las variables, aunque provengan de diferentes bases de datos tengan el mismo número de píxeles.

Para ejemplificar estos casos haremos uso de las series de precipitación y el VHI mensual obtenidas para Mozambique desde el CHIRPS y el NOAA STAR. En el siguiente gráfico se muestra un acercamiento a zona de 60 x 60 km aleatorio dentro de Mozambique, en donde los puntos azules indican coordenadas del NOAA STAR y los triángulos rojos coordenadas del CHIRPS.



- 1) Se leen las coordenadas de las bases de datos CHIRPS y NOAA STAR. Como variables ejemplo, leeremos también la precipitación y VHI. Las dimensiones de las coordenadas son distintas entre las dos bases de datos puesto que el tamaño de píxel es distinto.

```
#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) En este paso se promedian los valores de precipitación de todo Mozambique. Siguiendo el código se obtiene una matriz con el mismo número de filas (correspondiente al número de meses), y solamente columnas correspondientes a las fechas y a la precipitación promedio de todo el país. Si se deseara promediar varias subsecciones del país se debería filtrar las coordenadas que interesen y utilizar la misma función. Este trabajo no se lo muestra aquí, pero puede realizarse utilizando varias herramientas mostradas en el 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) A partir de esta sección se explicará un artificio para obtener datos de una variable en las coordenadas de otra base de datos. Existen muchas técnicas que pueden usarse como es la interpolación lineal o bilineal. Sin embargo, cuando el tamaño de las celdas es relativamente pequeño frente al área de interés, este proceso no genera una precisión significativa, pero sí demanda de mucho tiempo y carga computacional. Aquí se buscará la información (VHI) de la coordenada más cercana a la de interés (CHIRPS) y se asignará ese vector de información a la coordenada CHIRPS. Como ejemplo, primero se define una matriz de coordenadas objetivo, en este caso serán las del CHIRPS y, la matriz con coordenadas iniciales, que será las del NOAA STAR. Se calcula la mínima distancia “d” de la primera coordenada inicial hacia todas las coordenadas objetivo, luego la posición de la menor distancia será la posición de la coordenada inicial de la cual se extraerá la información.

```
#It shows how to find the position of the nearest coordinate with an example
objetivo<-data.frame(coor_chi[,1:2])#Target coordinates are defined
iniciales<-data.frame(coor_vhi)#The coordinates to be searched for are defined
d<-data.frame(which(data.frame(pointDistance(objetivo[1,],iniciales,lonlat = T))
  ==min(data.frame(pointDistance(objetivo[1,],iniciales,lonlat = T)))))
#d is the position of the nearest coordinate
```

Se muestra en la consola la coordenada objetivo y la inicial. Se puede ver que están muy cerca en latitud y longitud.

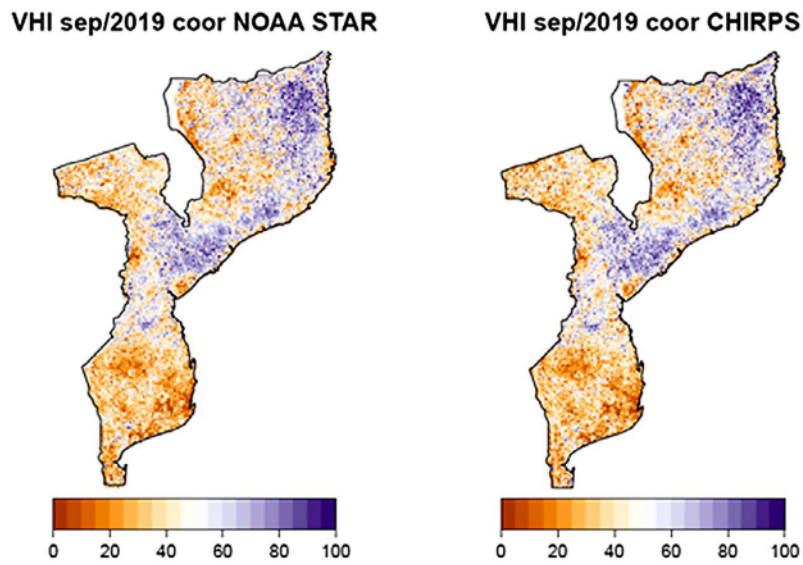
```
> #Nearby coordinates are displayed on the console
> objetivo[1,]
  long lat
1 40.425 -10.525
> iniciales[d[1,1],]
  longitud latitud
51085 40.41 -10.53
```

- 4) Se repite este proceso para todas las coordenadas iniciales (VHI) con un bucle que varía desde la primera hasta la última coordenada. En cada paso se extrae el vector de VHI perteneciente a la coordenada más cercana y se guarda en una nueva matriz. Esta matriz (denominada “dil”) deberá tener el mismo número de filas inicial de los datos VHI, pero tendrá el mismo número de columnas que el número de coordenadas CHIRPS. El orden de las coordenadas CHIRPS de arriba abajo indican el orden de vectores VHI de izquierda a derecha. Finalmente, se guarda la matriz en una carpeta conocida en formato .RDS.

```
#The nearest coordinates of the whole CHIRPS coordinate matrix are searched
objetivo<-data.frame(coor_chi[,1:2])#Target coordinates are defined
iniciales<-coor_vhi#The coordinates to be searched for are defined
for(i in 1:nrow(coor_chi)){
  d<-data.frame(which(data.frame(pointDistance(objetivo[i,],iniciales,lonlat = T))
    ==min(data.frame(pointDistance(objetivo[i,],iniciales,lonlat = T)))))
  if(i==1){
    dil<-data.frame(vhi[,1:2],data.frame(vhi[-c(1:2)]),d[1,1])
  }
  if(i!=1){
    dil<-cbind(dil,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(dil)<-c("year","month",c(1:(ncol(dil)-2)))#The columns of the new matrix are renamed
dim(dil)#Dimension must be equal to CHIRPS data
head(dil[,1:15])

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

5) Con el objetivo de comparar la nueva versión de datos VHI con la original, se grafican para una fecha aleatoria (sep/2019 en este caso) ambas bases de datos. Se sigue el proceso para graficar los ráster en el espacio siguiendo el código. El resultado será:



En el que se puede apreciar como la resolución original de los datos VHI es más fina que la de CHIRPS. Sin embargo, la distribución espacial de la información no ha perdido detalle significativo.

III. Aplicaciones a gestión sequías

14. Aplicaciones del balance hídrico

Esta sección está destinada a calcular y representar en mapas los meses secos (eq. 3), el déficit hídrico (eq. 4) y el volumen de agua disponible (eq. 5) sobre la cuenca del río Licungo. Utilizaremos las variables de precipitación, ETr y ETP extraídas desde el TerraClimate. Se muestran tres aplicaciones diferentes que podrían usarse de otras maneras de acuerdo con las necesidades de los usuarios.

- 1) Se instalan y cargan los paquetes necesarios.
- 2) Se lee el shapefile de la cuenca del río Licungo, se la ha asignado con el nombre cl.
- 3) Se lee el archivo con las coordenadas del TerraClimate que están dentro de la cuenca del río Licungo.

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

- 4) Se leen los archivos necesarios para los cálculos, estos son las matrices de precipitación, ETP y ETr. Se revisan sus dimensiones y su composición.

```
#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) Los meses con déficit o superhábit en cada una de las celdas se la calcula restando a la matriz de precipitación, la matriz de ETP. Se he denominado a esta variable “secos”.

```
#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) El volumen de agua disponible por celda se la calcula restando a la matriz de precipitación, la matriz de ETr. Se he denominado a esta variable “disponible”.

```
#The volume of available water is calculated (Pre-ETr)
available<-data.frame(pre[,1:3],pre[,4:ncol(pre)]-etr[,4:ncol(etr)])
```

- 7) El déficit hídrico por celda se la calcula restando a la matriz de ETP, la matriz de ETr. Se le denominó a esta variable “déficit”.

```
#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) A las 3 matrices resultantes, se revisan las dimensiones que deben coincidir con las de precipitación, ETP y ETr originales, y se las guarda como archivos .RDS.

```
#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) Con la matriz de meses secos, se calcula el porcentaje del tiempo (número de meses dentro del período de análisis) que ha presentado meses secos. A este vector se le agregan las coordenadas correspondientes y se guarda la matriz en una carpeta conocida. Siempre se debe revisar la dimensión y composición de la variable antes de guardar.

```
#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) A la matriz de agua disponible, se filtran los datos correspondientes a cada mes (enero, febrero, etc.) y se promedian todos los valores (de 62 años en este caso). Se agregan los 12 vectores, uno de cada mes, con sus respectivas coordenadas y se guardan las matrices en una carpeta conocida. Siempre se debe revisar la dimensión y composición de la variable antes de guardar.

```
#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 1.0032258 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) A la matriz de déficit, se la transforma de mm/mes a km³. Para este procedimiento se calcula un área por celda, utilizando el tamaño de celda al que corresponde cada píxel, en el caso del TerraClimate es de ≈ 21.43 km². Posteriormente, se calcula el promedio de todas las celdas y se obtiene un vector en km³ del déficit promedio en cada uno de los pasos de tiempo mensual. A este vector se le agregan las coordenadas correspondientes y se guardan la matriz final. Siempre se debe revisar la dimensión y composición de la variable antes de guardar.

```

#A constant is defined which transforms the mm of cell to km3
area_celda_km3<-(((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) Para graficar se leen el shapefile de la cuenca del río Licungo y las matrices de porcentaje de meses secos, agua disponible y déficit en km³.

```

#The necessary geographical information is read from the folders where they are stored
c1<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/basin_Licungo.shp")#The Mozambique shapefile is read
plot(c1)

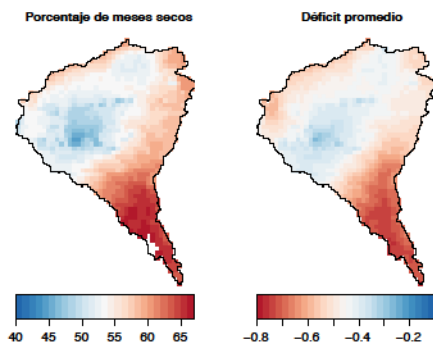
#The percentage of dry months in the Licungo river basin is read as follows
dry_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_dry.txt", header=T)
dim(dry_percentage2)
head(dry_percentage2)

#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)

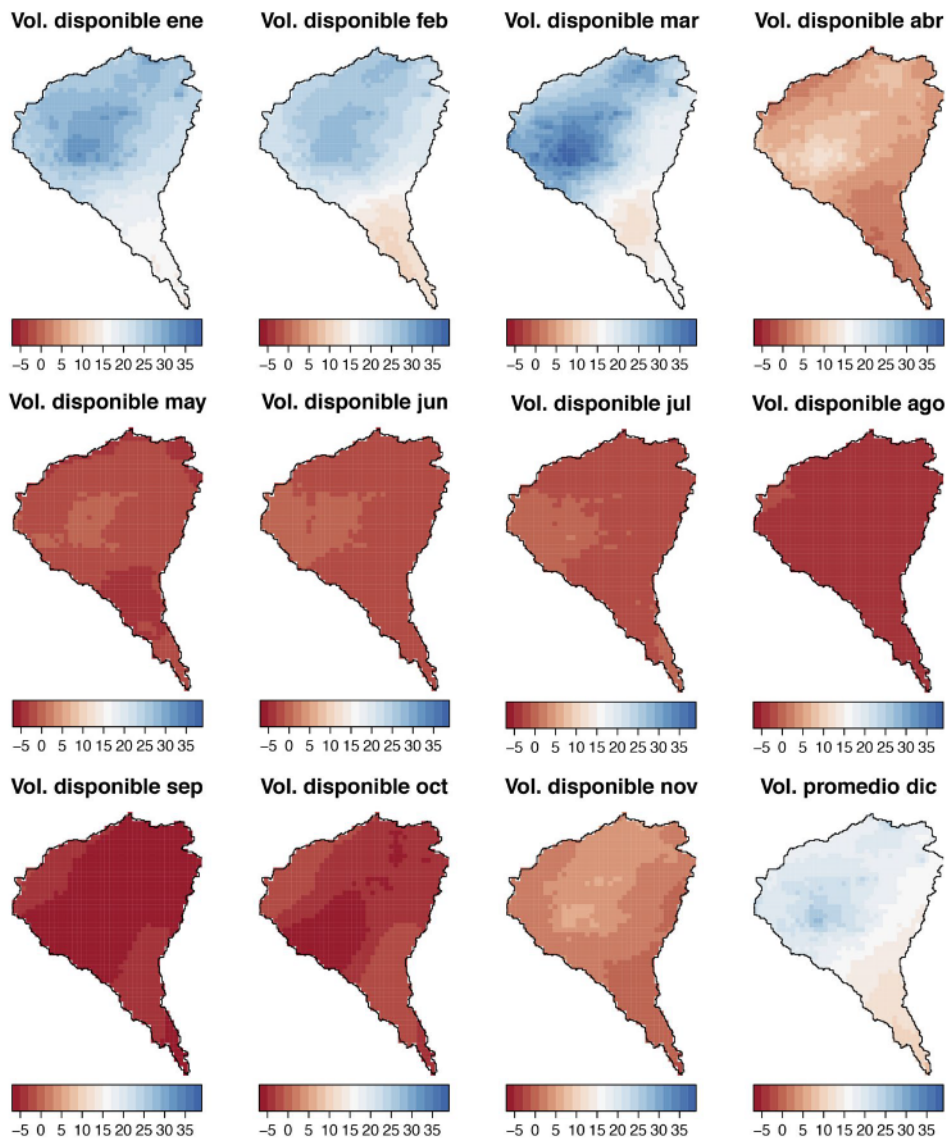
#The percentage of deficit in km3 in the Licungo river basin is read as follows
stress_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_stress.txt", header=T)
dim(stress_percentage2)
head(stress_percentage2)

```

13) Se siguen los pasos de graficas utilizadas con anterioridad (ejemplo paso 16 en la sección 10) para graficar las matrices de porcentaje de meses secos y de déficit en km^3 . El resultado será algo así:



14) Siguiendo los pasos de las gráficas definidas para el VHI (ejemplo paso 15 de la sección 11), se grafican los volúmenes promedio mensuales en km^3 de cada mes en la cuenca del río Licungo. El resultado será algo así:



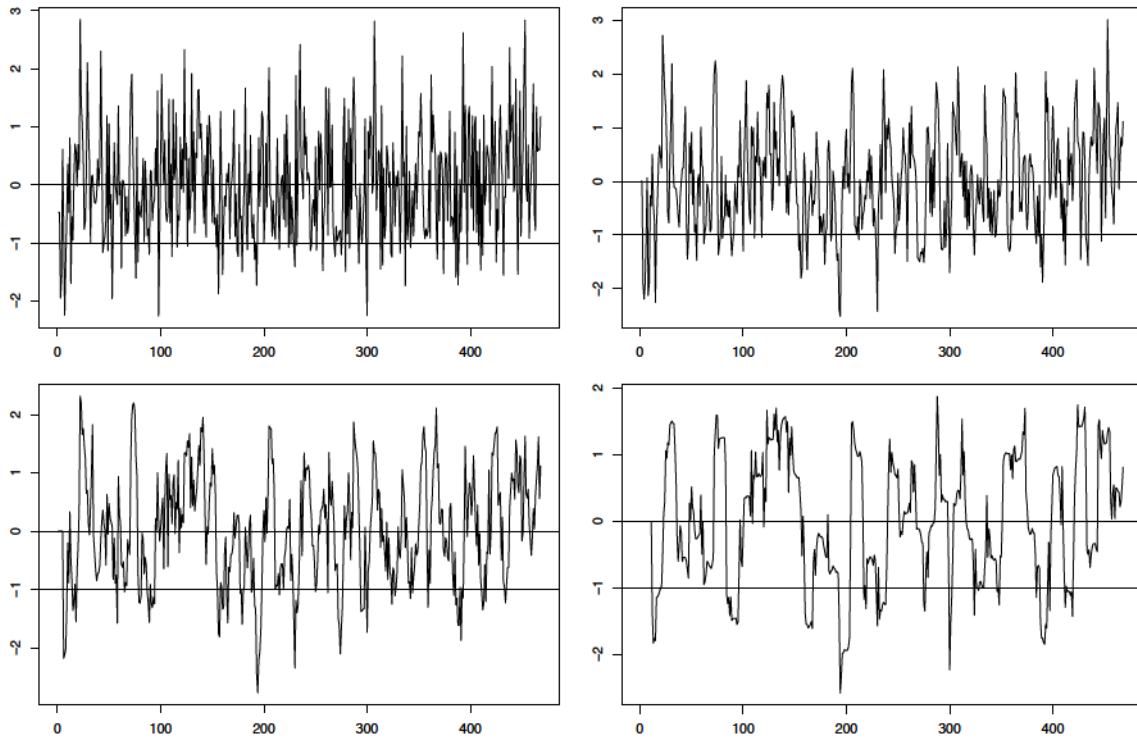
15. Cálculo y aplicación de los índices de sequía SPI y SPEI

En esta sección se explica paso a paso el cálculo de los índices de sequía SPI y SPEI. Se hará uso de los datos de precipitación del CHIRPS obtenidas sobre Mozambique para el SPI y, los datos de precipitación menos ETP (meses secos y húmedos) procedentes del TerraClimate sobre la cuenca del río Licungo para el SPEI. Primero se muestra el proceso para realizar el cómputo con una sola serie temporal en cada caso y posteriormente para toda la matriz de píxeles contenidas en las áreas de interés. Se culmina esta sección mostrando dos aplicaciones comunes y muy útiles de los índices. La primera es el cálculo de porcentaje de área afectada por sequías en cada paso de tiempo y la segunda, gráficos espaciales donde se puede mostrar estados de sequía/no sequía para un período específico.

- 1) Se instalan y cargan los paquetes necesarios. El paquete principal será el “SPEI”, en el que se ha programado funciones necesarias para el cómputo directo del SPI y el SPEI.
- 2) Se leen los shapefiles de la cuenca del río Licungo y de Mozambique, se han asignado con el nombre cl y mz respectivamente.
- 3) Se leen las coordenadas de los datos del CHIRPS localizadas dentro de Mozambique y de los datos del Terraclimate pertenecientes a la cuenca del río Licungo.
- 4) Para un vector aleatorio de la matriz de precipitación CHIRPS, entendida como una serie temporal mensual en un punto dentro de Mozambique, se calcula el SPI. Primero se define un vector que contenga las escalas temporales que se desee calcular, aquí se ha planteado a modo de ejemplo las escalas de 1, 3, 6 y 12 meses. El resultado será un vector de los SPI-1, -3, -6 y -12 meses correspondientes al vector de precipitación original. Se grafican los 4 vectores para comparar la diferencia, además se delimita el umbral de 0 y -1 (sequía moderada).

```
#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],"1")#SPI1
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,2],"1")#SPI3
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,3],"1")#SPI6
abline(h=c(0,-1))#Crosses the threshold 0 y -1
plot(c(1:nrow(SPI)),SPI[,4],"1")#SPI12
abline(h=c(0,-1))#Crosses the threshold 0 y -1
```



- 5) Para un vector aleatorio de la matriz de meses secos y húmedos, entendida como una serie temporal mensual en un punto dentro de la cuenca del río Licungo, se calcula el SPEI. Al igual que para el SPI, primero se define un vector que contenga las escalas temporales que se desee calcular, aquí se ha planteado a modo de ejemplo las escalas de 1, 3, 6 y 12 meses. El resultado será un vector de los SPEI a escalas de -1, -3, -6 y -12 meses correspondientes al vector de al déficit D original. Se grafican (no se muestran) los 4 vectores para comparar la diferencia, además se delimita el umbral de 0 y -1 (sequía moderada).

```
#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)))
colnames(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[,1:50])
head(SPI12[,1:50])

#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")
```

- 6) Mediante un bucle que varíe para todas las columnas de la matriz de precipitación (todas las series temporales de todas las coordenadas), se realiza el mismo procedimiento de cálculo de SPI para cada una de ellas. Las matrices resultantes al seguir el código serán 4 (en este caso), una por cada escala temporal que se haya calculado el índice. Si se sigue el código se generarán en un tiempo relativamente corto (variable al número de coordenadas) matrices que deben ser de igual dimensión que la matriz de precipitaciones CHIRPS originales. Las matrices se guardan como archivos .RDS en carpetas conocidas.

```
#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 1: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==1){
    SPEI1 <- data.frame(dry[,1:3],SPEI[,1])
    SPEI3 <- data.frame(dry[,1:3],SPEI[,3])
    SPEI6 <- data.frame(dry[,1:3],SPEI[,6])
    SPEI12 <- data.frame(dry[,1:3],SPEI[,12])
  }
  if (i!=1){
    SPEI1 <- data.frame(SPEI1,data.frame(SPEI[,1]))
    SPEI3 <- data.frame(SPEI3,data.frame(SPEI[,3]))
    SPEI6 <- data.frame(SPEI6,data.frame(SPEI[,6]))
    SPEI12 <- data.frame(SPEI12,data.frame(SPEI[,12]))
  }
  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 SPEIs 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")
```

- 7) Mediante un bucle que varíe para todas las columnas de la matriz de meses secos y húmedos déficits, (D), se realiza el mismo procedimiento de cálculo de SPEI para cada una de ellas. Las matrices resultantes al seguir el código serán 4 (en este caso), una por cada escala temporal que se haya calculado el índice. Si se sigue el código se generarán en un tiempo relativamente corto (variable al número de coordenadas) matrices que deben ser de igual dimensión que la matriz de déficits. Las matrices se guardan como archivos .RDS en carpetas conocidas.

```
#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 1: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==1){
    SPEI1 <- data.frame(dry[,1:3],SPEI[,1])
    SPEI3 <- data.frame(dry[,1:3],SPEI[,3])
    SPEI6 <- data.frame(dry[,1:3],SPEI[,6])
    SPEI12 <- data.frame(dry[,1:3],SPEI[,12])
  }
  if (i!=1){
    SPEI1 <- data.frame(SPEI1,data.frame(SPEI[,1]))
    SPEI3 <- data.frame(SPEI3,data.frame(SPEI[,3]))
    SPEI6 <- data.frame(SPEI6,data.frame(SPEI[,6]))
    SPEI12 <- data.frame(SPEI12,data.frame(SPEI[,12]))
  }
  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 SPEIs 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) En este paso se calcula el área de porcentaje de área afectada por sequías. Esto es cuantificar para cada paso de tiempo (cada mes) el número de píxeles que están en sequía y calcular su porcentaje frente al total. Se requiere un umbral de acuerdo con la intensidad de sequía que se desee estudiar. En el ejemplo hemos usado como umbral valores de sequía moderada (valores menores a -1), sequía severa (valores menores a -1.5) y sequía severa (valores menores a -2). Se muestra aquí la composición final de la matriz que se guarda. Se usa como ejemplo el SPI a una escala temporal de 6 meses sobre Mozambique.

```
#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(percentajes)
[1] 468 6
> head(percentajes)
      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) Se grafican los porcentajes como series temporales siguiendo el código del ANEXO III. En este caso se utiliza la función “ggplot”. Es necesario leer los porcentajes de área afectada por sequías, que aquí se definen como PAA. Si se quisiera mostrar las series temporales sin sombrearlos se utilizar el paso 4.

```
#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 Licungo basin
mz<-readShapePoly("E:/GIS_Mozambique/Layer_Mozambique/mozambique.shp")#The Mozambique shapefile is read
plot(mz)
plot(cl,add=T)

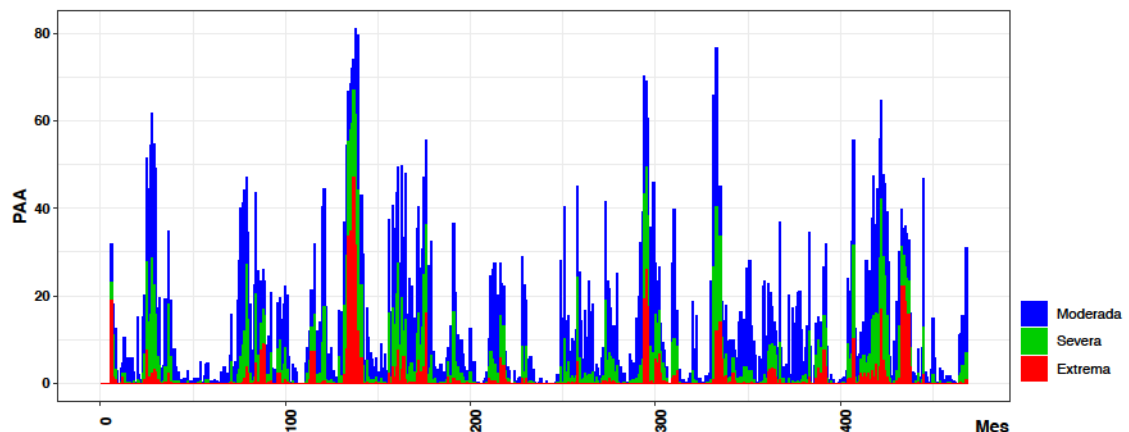
#Drought percentages are read for the different intensities
percentages<-read.table("E:/19_MANUAL_data_global/SPI_SPEI/percentages.txt",header=T)
dim(percentages)
head(percentages)

#The limits to be plotted are calculated
spil<-data.frame(c(1:nrow(percentages)), (percentages[,4]-percentages[,5]),
                (percentages[,5]-percentages[,6]),percentages[,c(6)])#Calculation
colnames(spil)<-c("date","Moderate","Severe","Extreme")

#The created time series are plotted
aux1<-melt(data = spil, id.vars = "date",value.name = "por")#"date" must be equal to the column name of spil
# "por" can be changed
head(aux1)
spl<-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)
spl
dev.off() #Export closed
```


El gráfico resultado será algo así:



- 10) En este paso se calcula el estado de sequía o no sequía sobre Mozambique. Esto es transformar todos los valores inferiores a un umbral de sequía como 1 y los mayores al umbral como 0. Se usa como ejemplo el SPI a una escala temporal de 6 meses sobre Mozambique y para una intensidad de sequía moderada, es decir un umbral valores de SPI menores a -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
coord_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordinate_CHIRPS.txt",header=T)
head(coord_chi)
dim(coord_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#All 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) Se siguen los pasos de gráficas utilizados con anterioridad (ejemplo paso 16 en la sección 10) adaptados para graficar los estados de sequía o no sequía. El objetivo es mostrar para una serie de meses las zonas donde hay sequías. Se toma como ejemplo la época de octubre de 1992 a septiembre de 1993 cuando hubo una importante sequía en Mozambique. Se usa el “ggplot” como función para crear gráficos siguiendo el código, donde se pueden modificar títulos, colores y otras características. A diferencia de las funciones que se han usado para otros gráficos (como el uso de la función “levelplot”), aquí es necesario transformar la información geográfica (shapefile mz) en un matriz de datos compatible con el método de “ggplot”; se lo hace mediante la función “fortify”.

```

#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(c1,add=T)
cd<-fortify(mz)

#CHIRPS coordinates are read inside Mozambique
coord_chi<-read.table("E:/19_MANUAL_data_global/CHIRPS/coordinates_CHIRPS.txt",header=T)
head(coord_chi)
dim(coord_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(coord_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 Moderada"))+#Legends and colours are
  changed
  coord_fixed()

```

Se crea un gráfico para cada mes y se los guarda como .pfd (u otro tipo de archivo). El resultado será algo así:



16. Caracterización sequías mediante el SPI

En este apartado se establece un procedimiento para calcular las características promedio de las sequías. Estas características son la duración, la intensidad máxima y la severidad; calculándose además el número de eventos ocurridos y la probabilidad de tener un mes de sequía. Los eventos de sequía se definen cuando al menos dos meses consecutivos estuvieron bajo el umbral de intensidad definido. El procedimiento se lo realiza para una serie temporal de SPI (SPI-6 en Mozambique como ejemplo) de donde se extraen los promedios de cada característica dentro de la serie temporal. Los umbrales pueden ser cambiados y el código presentado sirve para cualquier otro índice de sequía que esté estandarizado (media 0 y desviación estándar 1).

- 1) Se instalan y cargan los paquetes necesarios.
- 2) Se lee la matriz que contiene las series temporales del indicador a caracterizar. En este caso el SPI-6 de Mozambique computado con las precipitaciones del CHIRPS.

```
#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) Se leen las coordenadas de los datos del CHIRPS localizadas dentro de Mozambique.

```
#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) El bucle que calcula las características de todos los SPI de la matriz objetivo tiene muy pocas variables que modificar en caso de requerirse. El umbral se define previo al lanzamiento del bucle, se debe tener en cuenta que el bucle solo identifica las características de épocas secas no húmedas, por lo que los valores deben ser negativos. El inicio del bucle sigue el orden de izquierda a derecha, donde la primera columna es el número desde el cual varía la variable auxiliar "l", este valor debe ser modificado si la matriz objetivo empieza en otra columna. Así mismo se debe cambiar el número de inicio en la parte final del código siguiendo las instrucciones mostradas en el mismo. Si se necesitara extraer las características de una sola serie temporal o un número limitado, se debe cambiar también "ncol(spi6)" que indica el número de columnas de la matriz llamada "spi6" por el de la posición de la columna última que se desee obtener las características.

```
#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
```

... (ver todo el bucle en el código entregado en los anexos).

```

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
casil<-data.frame(du,int,sev,freq,prob)
if(nrow(data.frame(which(casil[,1]==1))) != 0){
  casil<-casil[-which(casil[,1]==1),]
}

if(nrow(casil)==0){
  casil<-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(casil)))
}
if (l!=4){#Change the 4 when you change the start of the initial loop
  dat<-rbind(dat,t(data.frame(colMeans(casil))))
}
print(paste(l,"de",ncol(spi6)))#Step counter, one for each column
} #fin de l
colnames(dat)<-c("Duration","Intensity","Severity","numero","probabilidad")
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

```

La matriz “dat” es el resultado del bucle y contiene las características promedio de todos los eventos de sequía de cada serie temporal SPI. Su dimensión tiene un número de filas igual que el número de coordenadas CHIRPS y 5 columnas correspondiente a cada característica:

```

> head(dat[,1:5])
  colMeans.casil. Duracion Intensidad Severidad numero probabilidad
colMeans.casil. 7.529412 -1.705905 -7.935980 19 1.608849
colMeans.casil. 7.470588 -1.754820 -8.184736 18 1.596280
colMeans.casil. 9.384615 -1.743732 -9.567796 14 2.005260
colMeans.casil. 7.750000 -1.734323 -8.028418 18 1.655983
colMeans.casil. 7.687500 -1.735632 -8.060110 18 1.642628
colMeans.casil. 7.176471 -1.737271 -8.029506 18 1.533434
> |

```

A esta matriz se le agrega las coordenadas correspondientes y se guarda en una carpeta de destino conocida.

```

#The respective coordinates are attached to their characteristics
carac<-data.frame(coor_chi,dat)
head(carac)
dim(carac)

```

- 5) Para graficar estas características promedio se utiliza el procedimiento similar al usado para graficar los estados de sequía o no sequía en la sección 1. Se lee la información geográfica de Mozambique y la matriz que contiene las características de sequías según el SPI6 (en este caso) y sus coordenadas. Se grafican solamente las duraciones y las probabilidades a modo de ejemplo.

```

#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) Se extrae el vector de la característica a graficar, donde el número dentro de los corchetes es la posición de la columna que contiene la información a graficar.

```

#The durations are extracted
duration<-data.frame(carac[,3])

```

- 7) A este vector se transforman en secciones que estén dentro de sus límites máximos y mínimos. El número que se le asigna debe estar dentro de los límites existentes.

```
#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) Se grafican los resultantes haciendo uso de la función “ggplot”. Se pueden modificar los colores, el nombre de la sección (antes definidas), entre otras. Es importante definir los valores de “x” y “y” como las coordenadas de longitud y latitud contenidas en la matriz de coordenadas. Al gráfico se le guarda con un nombre para después exportarlo.

```
#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("<0.5", "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) Se repite el proceso anterior con la matriz de probabilidades.

```
#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

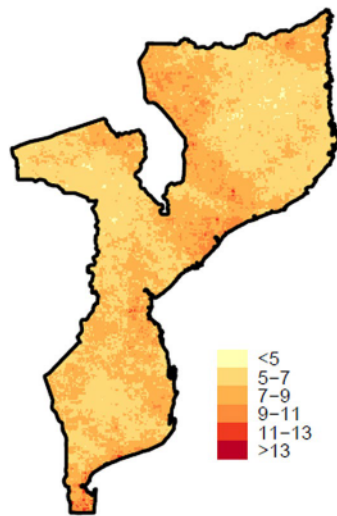
graphic_2<-ggplot(probabilities)+
  geom_tile(aes(x = carac[,1], y = carac[,2], fill=factor(probabilities[,1])))+#Here you change the name of the column
  theme_classic()+
  theme(legend.position = c(0.75, 0.25),legend.box="horizontal",legend.title = element_blank(),
        legend.text=element_text(size =
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) Se exportan los gráficos como .pdf.

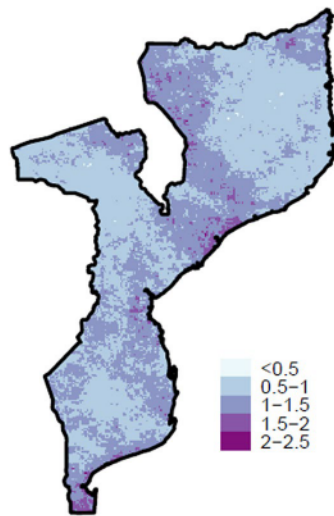
```
#The graphic is exported as pdf, the destination folder and the size can be modified
pdf("E:/19 MANUAL_data_global/Graphics/characteristics.pdf",height = 8,width = 6)
grid.arrange(graphico_1,graphic_2,
              nrow = 1) #The 2 graphs are arranged in a 1 x 2 matrix, can be changed
dev.off() #Export closed
```

El resultado será algo así:

SPI-6 < -1 Duración



SPI-6 < -1 Probabilidades

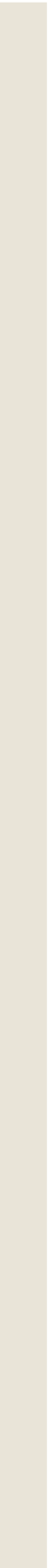


Conclusiones

Este documento es un manual introductorio para obtener variables hidrometeorológicas en cualquier región y aplicarlas a la gestión de sequías. Las sequías son un fenómeno hidrológico extremo, que ocurre de forma natural en cualquier régimen climático. Sus impactos sobre el medio ambiente causan inmensas pérdidas económicas y ecológicas todos los años, por lo que su caracterización y monitorización se ha vuelto cada vez más importante. Sin embargo, la caracterización de las sequías es un proceso complejo, pues además de afectar diferentes partes del ciclo del agua, no existe un indicativo claro de inicio o fin. A esto se le suma la problemática de que las herramientas que se requieren para modelar el fenómeno, como datos de interés hidrológico y climático (precipitación, temperatura, etc.), no siempre cumplen con la calidad requerida.

Este trabajo recoge una serie de conceptos de gestión de sequías ampliamente aceptados entre la comunidad científica y técnica que se los aplica a través de varios scripts (códigos de programación) codificados en el Software R. Estos scripts divididos en 3 secciones muestran el paso a paso para i) descargar, extraer, manipular y guardar variables hidrometeorológicas desde base de datos globales; ii) utilizar y aplicar estas variables para la gestión de sequías; y, iii) presentar resultados mediante gráficos en distintos formatos.

Al seguir el paso a paso de las herramientas entregadas en este manual, se puede generar mucha información útil para cualquier región del mundo. Los resultados que se pueden obtener pueden ser de interés para gestores y técnicos que no siempre encuentran acceso fácil a este tipo de información. Una gran ventaja de este manual es lo flexible de sus herramientas, pudiendo ser utilizadas en distintos lugares y permitiendo crear nuevos scripts y generar nuevas formas de utilizar la información descargada por los usuarios. Las metodologías repartidas en este documento pueden ampliarse en futuros trabajos, como por ejemplo con la obtención de datos resultado de modelos de proyección de cambio climático, análisis de tendencias, representación de resultados sobre regiones contenidas en un país (ciudades o provincias) o cuencas hidrográficas (subcuencas), entre otras.



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>
- 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.
- 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>
- 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>
- DURÁN, A., 2000. Agua en el suelo.
- 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>
- 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](https://doi.org/10.1016/0273-1177(95)00079-T)
- KUMAR, N. M., MURTHY, C. S., SSHA 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>
- 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>
- RSTUDIO TEAM, 2016. Integrated Development Environment for R. RStudio, Inc., Boston, MA. <http://www.rstudio.com/>
- SLATER, L. J., THIREL, G., HARRIGAN, S., DELAIGUE, O., HURLEY, A., KHOUAKHI, A., PRODOSCHMI, 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>
- 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>
- VERMOTE, E., WOLFE, R., 2015. MOD09A1 MODIS/Terra Surface Reflectance DAILY L2G Global 1km and 500 m SIN Grid V006 [Data set].
- 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>
- 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>
- 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>
- 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](https://doi.org/10.1016/0022-1694(69)90110-3)



PROYECTO REALIZADO CON LA COLABORACIÓN DE

