

MANUAL PARA A OBTENÇÃO E APLICAÇÃO DE VARIÁVEIS HIDROMETEOROLÓGICAS A PARTIR DE BASES DE DADOS GLOBAIS PARA A GESTÃO DAS SECAS

MANUAL PARA A OBTENÇÃO E APLICAÇÃO DE VARIÁVEIS
HIDROMETEOROLÓGICAS A PARTIR DE BASES DE DADOS
GLOBAIS PARA A GESTÃO DAS SECAS

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.9788497498340>

N.º de páxinas: 74
Índice: p. 5

ISBN: 978-84-9749-834-0
Depósito legal: C 223-2022

Desenho e maquetación: Tórculo Comunicación Gráfica, S. A.

Este traballo é complementado polos recursos necesarios para o desenvolvemento do manual, que son descritos nos Capítulos I e II deste documento (p. 27). Podem ser encontrados no persistente URL <http://hdl.handle.net/2183/29801> (Repositorio institucional da UDC, RUC).



PROXECTO REALIZADO COM A COLABORACIÓN DE



EDIÇÃO

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

© Editora: Servizo de Publicacións. Universidade da Coruña
© Conteúdo e imáxens: dos autores



This book is released under a Creative Commons license Attribution-Non-Commercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). / Este libro é lanzado sob uma licena Creative Commons CC BY-NC-SA (Atribuioo - No-Comercial - Partilhar Igual) 4.0 International.

MANUAL

PARA A OBTENÇÃO E APLICAÇÃO DE VARIÁVEIS HIDROMETEOROLÓGICAS A PARTIR DE BASES DE DADOS GLOBAIS PARA A GESTÃO DAS SECAS

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 geral

Glossário de siglas e abreviaturas	7	7	Objectivos específicos	26
I Introdução	9	8	Seguimiento del manual	26
1 Objectivo Geral	9	II Obtenção de dados	29	
2 Escopo e limitações do manual	9	9	Área de estudo	29
3 Organização do documento	10	10	Precipitação: CHIRPS	32
4 Gestão da seca	10	11	índices de vegetação do NOAA STAR	38
4.1 Ciclo hidrológico	10	12	Varios: TerraClimate	47
4.2 Equilíbrio hídrico	11	13	Solução para unificar escalas espaciais	54
4.3 Definição e classificação de seca	12	III Aplicações à gestão de secas	57	
4.4 Índices de seca	14	14	Aplicações do balanço hídrico	57
4.5 Características das secas	16	15	Calculo e aplicação dos índices de seca do SPI e SPEI	61
5 Requisitos	18	16	Caracterização de secas usando o SPI	67
5.1 Sistemas de Informação Geográfica (SIG)	18	Conclusões	71	
5.2 Sistemas de Computação Numérica (SCN)	20	Referências	73	
6 Bases de datos globais	22			
6.1 Bases de datos complementares	22			
6.2 Bases de datos de variáveis hidrometeorológicas	23			

Índice de figuras

Fig. 1. O ciclo da água. Fonte: U.S. Geological Survey (USGS)	11
Fig. 2. Sequência de ocorrência dos tipos de seca e seus impactos ..	13
Fig. 3. Características das secas de acordo com a “run-theory”	18
Fig. 4. Camadas temáticas disponíveis em um SIG e comparação entre dados raster e vetoriais	19
Fig. 5. Ambiente RStudio	21
Fig. 6. Localização e representação dos limites de Moçambique e da bacia do rio Licungo	31

Índice de tabelas

Tabela 1. Categorias de secas de acordo com os valores dos índices .	17
--	----



Glossário de siglas e abreviaturas

ETP	Evapotranspiração potencial
ETr	Evapotranspiração real
CHIRPS	Estimativas de precipitação a partir de pluviômetro e observações satelitais <i>(Rainfall Estimates from Rain Gauge and Satellite Observations)</i>
STAR	O Centro de Aplicações e Pesquisas Satelitais <i>(The Center for Satellite Applications and Research)</i>
NOAA	A Administração Nacional dos Oceanos e Atmosfera <i>(The National Oceanic and Atmospheric Administration)</i>
VCI	Índice de condição de vegetação <i>(Vegetation Condition Index)</i>
TCI	Índice de condição da temperatura <i>(Temperature Condition Index)</i>
VHI	Índice de Saúde Vegetal <i>(Vegetation Health Index)</i>
NDVI	Índice de vegetação por diferença normalizada <i>(Normalised Difference Vegetation Index)</i>
BT	Temperatura do Brilho <i>(Brightness Temperature)</i>
SPI	Índice de Precipitação Padronizado <i>(Standardized Precipitation Index)</i>
SPEI	Índice padronizado de precipitação e evapotranspiração <i>(Standardized Precipitation and Evapotranspiration Index)</i>
SIG	Sistema de Informação Geográfica
SCN	Sistemas de Computação Numérica

I. Introdução

1. Objectivo Geral

Este manual pretende ser uma ferramenta introdutória básica para a gestão da seca. Este fenómeno hidrológico extremo é um dos mais importantes na gestão dos recursos hídricos, pois causam anualmente imensas perdas económicas e ecológicas em todo o mundo que condicionam a segurança alimentar e a qualidade de vida das pessoas.

Devido às mudanças climáticas, os eventos da seca se tornarão mais frequentes e severos no futuro. Portanto, entender, caracterizar e monitorar esse fenómeno é essencial para elaborar planos de contingência que reduzam o risco e sejam mais resilientes às secas. Para alcançar este objectivo são necessários dados, ferramentas e procedimentos que nos permitam compreender e modelar o fenómeno. No entanto, as realidades socioeconómicas amplamente variáveis em cada região do mundo significam que o acesso a certos dados de uma qualidade suficiente é limitado em algumas regiões e países.

Este manual é uma ferramenta que permite resolver esta problemática através de ferramentas criadas para obter de forma rápida e eficiente os dados necessários à gestão e monitoria das secas em qualquer região. Estes dados são produzidos por bancos de dados globais livres e gratuitos para todo o mundo pelo que estes dados podem ser usados em entornos administrativos (países, cidades, etc.) ou naturais (bacias hidrográficas, reservas naturais, etc.). Além disso, o manual apresenta diversos usos e aplicações desses dados como ferramentas específicas para a gestão das secas. Como parte do procedimento, é feita uma breve introdução aos diferentes aspectos, conceitos e requisitos necessários para a utilização do manual.

2. Escopo e limitações do manual

As informações conceituais e práticas apresentadas neste manual pretendem ser uma ferramenta introdutória para a obtenção e aplicação de dados na monitoria e caracterização de secas. No manual são apresentadas as informações necessárias para a compreensão do fenómeno e seus componentes de forma geral pelo que as ferramentas e resultados práticos se limitam a estes conceitos.

Os conceitos-chave sintetizados são de uso geral por especialistas e organizações internacionais, entretanto, se limitam aos objectivos do manual, podendo ser expandidos e até apresentar diferenças em relação a outras fontes bibliográficas. Da mesma forma, as ferramentas utilizadas, os procedimentos e interpretações aplicados para os exercícios práticos, podem conter erros e ser melhorados ou aplicados em outros cenários. Todas as ferramentas informáticas utilizadas são de uso público e gratuitas, podendo ser substituídas por outras que apresentem fins semelhantes.

3. Organização do documento

Este documento consiste em três capítulos. O primeiro capítulo é uma introdução aos conceitos, ferramentas e requisitos necessários para o uso do manual. Em primeiro lugar, se compilam os conceitos associados ao estudo e gestão das secas, sua classificação e caracterização. Em seguida, se explica de forma geral os requisitos técnicos e computacionais necessários. A continuação, se apresenta uma introdução às várias bases de dados globais que existem e aos dados necessários para a gestão da seca que ofertam, enfatizando aqueles que se usan no manual. Finalmente, se explicam os objectivos específicos que são buscados neste trabalho. No segundo capítulo, se detalham e explicam os passos necessários para obter a série de diferentes variáveis que são necessárias para a gestão da seca. No terceiro capítulo, se faz uso dos dados obtidos no Capítulo II e se expoem as etapas para a correcta aplicação na caracterização e monitoria das secas numa região específica. Nos anexos se entregam três roteiros que são o produto final complementar para seguir o manual. Como estudos de caso escolhemos Moçambique localizado no cone sul africano e a bacia hidrográfica do rio Licungo na parte centro-norte deste país.

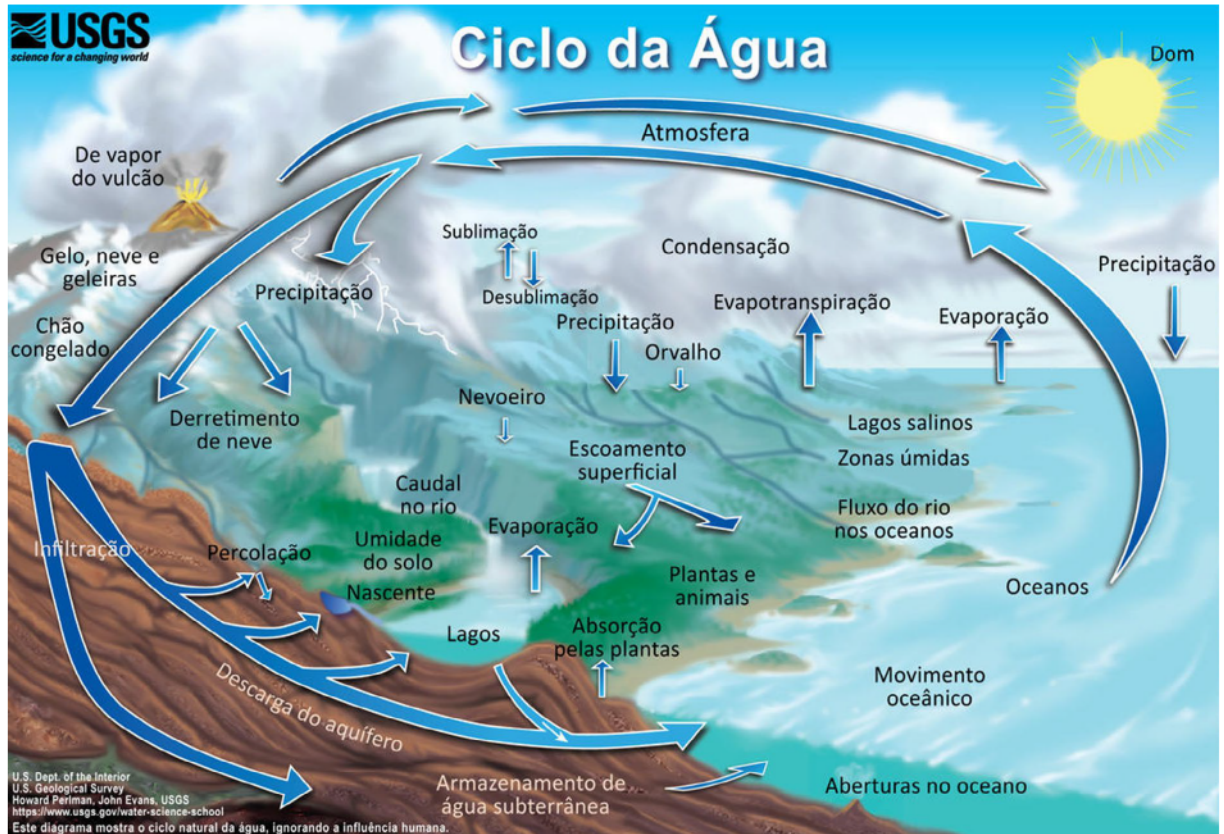
4. Gestão da seca

4.1. Ciclo hidrológico

O ciclo hidrológico ou ciclo da água é o processo de circulação da água entre os diferentes compartimentos da hidrosfera (Fig. 1). A água da terra e dos oceanos entra na atmosfera por evaporação ou sublimação, onde se condensa em nuvens e cai como chuva ou neve. A água que cai na forma de precipitação entra nos corpos de água doce ou se infiltra no solo. O ciclo se completa quando as águas superficiais ou subterrâneas voltam a entrar no oceano. Os principais processos envolvidos no ciclo da água e que são variáveis importantes na análise das secas são:

- Precipitação (P): A atmosfera perde água por condensação (chuva e orvalho) ou sublimação inversa (neve e geada) que passam, dependendo do caso, para o solo, para a superfície do mar ou para a bolsa de gelo (camada de gelo sobre os oceanos).
- Evapotranspiração (ET): A água evapora-se na superfície oceânica, no solo e também pelos organismos, no fenómeno da transpiração, de forma que ET se refere ao fenómeno combinado. A evapotranspiração potencial (ETP) é a quantidade máxima de água que pode evaporar de um solo totalmente coberto por vegetação, que se desenvolve em condições ótimas e, no suposto caso, sem limitações na disponibilidade de água. A evapotranspiração real ou efetiva (ETr) é aquela que realmente ocorre nas condições existentes em cada caso.
- Infiltração: O fenómeno ocorre quando a água que chega ao solo penetra pelos poros e passa para o subsolo. A proporção de água que infiltra e que circula na superfície (escoamento) depende da permeabilidade do substrato, da encosta e da cobertura vegetal. A quantidade de água contida num determinado momento é a umidade do solo e pode ser medida em diferentes profundidades.
- Escoamento: este termo se refere aos vários meios pelos quais a água líquida desliza colina abaixo pela superfície do solo. Ao atingir um corpo d'água (rios, lagos, reservatórios, etc.) o volume que por ele passa por um determinado período é chamado fluxo.

Fig. 1
O ciclo da água. Fonte: U.S.
Geological Survey (USGS)



- Circulação subterrânea: ocorre a favor da gravidade, como o escoamento superficial. Ocorre nos aquíferos através dos poros de uma rocha permeável, intervindo fenômenos como pressão e capilaridade.

4.2. Equilíbrio hídrico

O ciclo da água pode ser formulado matematicamente a partir do balanço hídrico, que é derivado do conceito de conservação da matéria, ou seja, é o equilíbrio entre todos os recursos hídricos que entram e saem dum sistema (bacia hidrográfica, país, região, etc.), em determinado intervalo de tempo:

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

Onde, as entradas de água são a precipitação (P), que pode ocorrer na forma de chuva, orvalho, neve ou condensação; contribuição da água subterrânea adjacente ao sistema; e, transferências de água de outros sistemas, como descarga de barragens ou esgoto. As saídas podem ser evapotranspiração; infiltrações profundas que alimentam os aquíferos; a água é derivada para sistemas externos adjacentes, consumo humano e indústria; e as que saem do sistema para um receptor externo, como um grande mar ou o mar. O estado se refere ao volume contido no sistema após adicionar e subtrair as entradas e saídas ao volume de água na etapa de tempo anterior.

Em geral, para cada intervalo de tempo a eq.1 pode ser formulada da seguinte forma:

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

Onde P é a precipitação, I a infiltração, U os diferentes usos da água (no mesmo ou em outro sistema), ETr a evapotranspiração real, Q a vazão de saída do sistema e ΔV a variação do volume de água no sistema. Ao longo prazo, a variação tende a zero, porém, a cada intervalo de tempo pode haver momentos em que o sistema apresenta déficit ou superávit. Cada um dos processos pode ser medido e o equilíbrio pode ser concluído. Ao simplificar o equilíbrio, podemos detectar meses úmidos e secos (3); e calcular o volume de água disponível (4):

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

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

Onde, na eq. 3 a variação ΔV que define um mês seco ou déficit ($\Delta V < 0$) ou um mês úmido com superávit ($\Delta V > 0$) no sistema num determinado período de tempo. Enquanto isso, na eq. 4 a segunda parte da igualdade é chamada precipitação útil e é a água disponível para recarregar o solo, rios e outros corpos de água existentes.

O balanço hídrico também nos permite calcular o déficit hídrico (D):

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

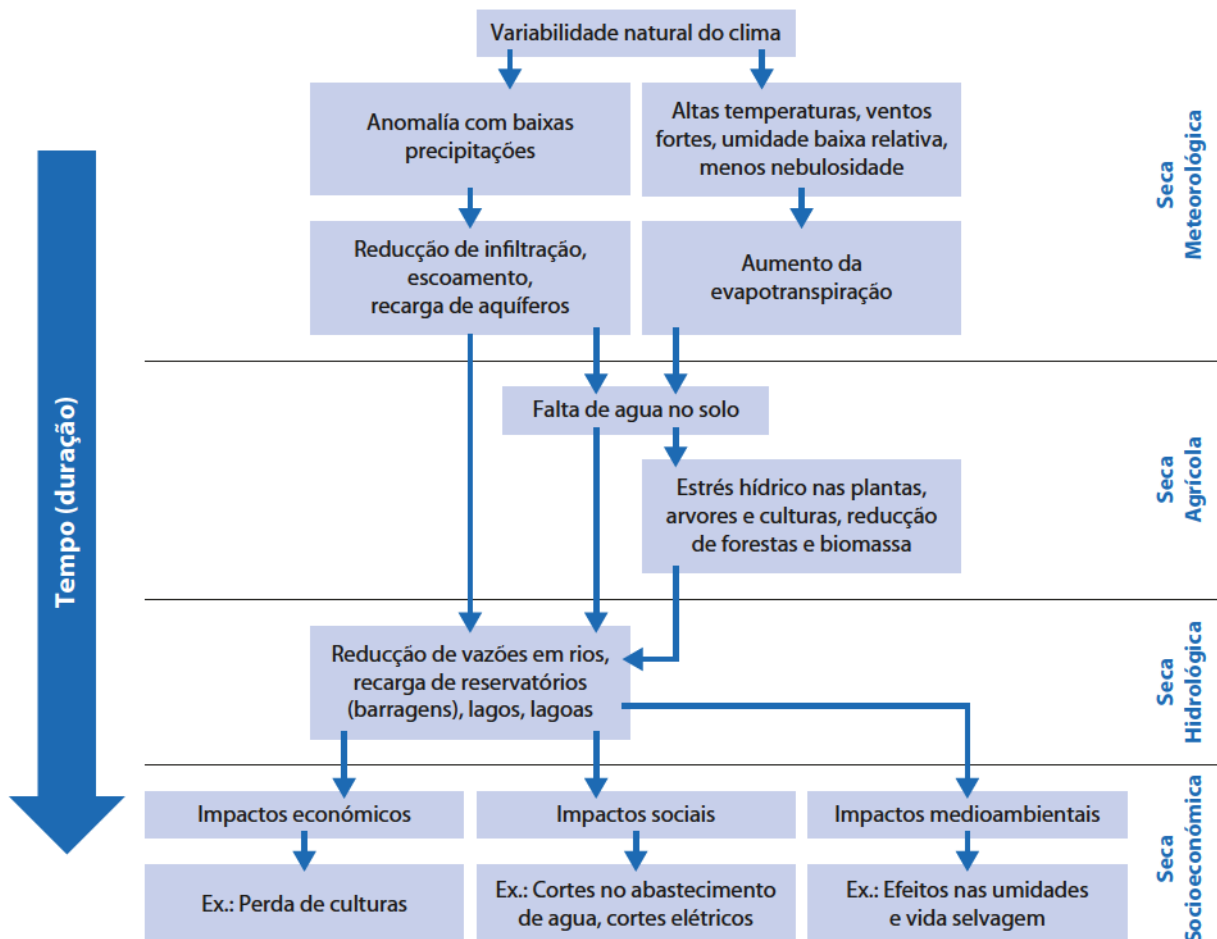
Onde, um déficit ($D > 0$) é a quantidade de água que falta para cobrir as necessidades potenciais de água para evaporar e transpirar.

4.3. Definição e classificação de seca

A seca é um fenômeno hidrológico extremo cujos limites geográficos e temporais são difíceis de determinar, sendo um evento natural especialmente perigoso quando não existe uma boa gestão dos recursos hídricos. A seca é entendida como uma anomalia temporária de precipitação mais ou menos prolongada, caracterizada por um período de tempo com valores pluviométricos inferiores ao normal. Pode ou não causar danos às culturas, degradação do solo e/ou situação de abastecimento insuficiente de água, dependendo do nível de demanda e das características dos sistemas de exploração de recursos hídricos.

Fig. 2
Sequência de ocorrência dos
tipos de seca e seus impactos

Fonte: Centro Nacional de Mitigação de Secas da Universidade de Nebraska-Lincoln



Nos últimos anos, muitas classificações de seca foram desenvolvidas. Porém, a mais aceita e disseminada é aquela que define a seca meteorológica, agrícola, hidrológica e socioeconômica. As três primeiras medem a seca como um fenômeno físico, enquanto a última trata da seca em termos de oferta e demanda, baseada no seguimento dos efeitos do déficit hídrico conforme se agita nos sistemas socioeconômicos. Cada uma delas está relacionada a uma parte do ciclo hidrológico e se desenvolvem uma após a outra, a partir das anomalias climáticas que fazem parte da variabilidade natural do clima (Fig. 2).

A seca meteorológica ocorre quando anomalias de ausência de chuva dominam uma área. Essas anomalias podem ser acompanhadas por um aumento nas temperaturas, que se traduz em aumento da evapotranspiração. A seca agrícola ocorre quando as culturas são afetadas pela falta de chuvas e pela baixa umidade do solo. A seca hidrológica ocorre quando o abastecimento Baixo de água se torna aparente, especialmente em riachos, reservatórios e níveis de água subterrânea, geralmente após muitos meses de seca meteorológica. A seca socioeconômica relaciona a oferta

e a demanda de diversos produtos básicos com a seca (água para consumo humano, irrigação, indústrias, hidrelétricas, etc.). A seca meteorológica pode começar e terminar rapidamente, enquanto a seca hidrológica leva muito mais tempo para se desenvolver e se recuperar.

A seca deve ser diferenciada de outros fenômenos relacionados como a escassez e a aridez. A escassez representa uma situação de déficit permanente (calculado com o balanço hídrico) em relação à demanda por água num sistema regional de recursos, caracterizado por um clima árido ou por um rápido crescimento da demanda. A aridez é uma situação estrutural natural de uma região e, portanto, permanente. Mesmo nessas circunstâncias, não deveria haver déficit se os sistemas de exploração fossem adequadamente planejados e operados e as demandas fossem mantidas dentro de limites razoáveis, de acordo com as características climáticas da região. Isso requer ações planejadas a médio e longo prazo.

4.4. Índices de seca

Os índices são representações numéricas que mostram informações associadas à severidade das secas, determinada por dados climáticos ou hidrometeorológicos. São usados para quantificar a gravidade, localização, desenvolvimento cronológico e duração dos episódios de seca. A gravidade refere-se ao desvio da normalidade de um índice com base num limite previamente estabelecido. Desta forma, podemos identificar quando uma seca começou, quando terminou e a área geográfica afetada. Localização refere-se à área geográfica em que as condições de seca são registradas. As datas aproximadas de início e fim definem a duração do evento de seca. Existem índices que também podem servir a outra função essencial, no sentido de que podem fornecer uma referência histórica para planejadores ou tomadores de decisão. Esta referência fornece aos usuários informações sobre a probabilidade de ocorrência ou recorrência de secas de gravidade variável. No entanto, deve-se notar que as mudanças climáticas começarão a alterar as tendências históricas. Recomenda-se representar o fenômeno em várias escalas de tempo, uma vez que reconhece períodos chuvosos de curto prazo dentro de secas de longo prazo ou períodos de seca de curto prazo dentro de períodos chuvosos de longo prazo.

Dado que a classificação da seca responde à análise de anomalias em cada parte do ciclo hidrológico, existem índices concebidos para caracterizar cada tipo de seca baseados em variáveis específicas, por exemplo, precipitação para seca meteorológica ou vazões para seca hidrológica. Nos últimos anos, foram desenvolvidos índices mistos que tentam quantificar as secas em todo o ciclo da água ou em etapas dele. No entanto, usar vários índices com base em várias variáveis é a melhor maneira de caracterizar o estado de seca em todo o sistema.

Devido à complexidade do fenômeno, dezenas de índices foram propostos nas últimas décadas. Esses índices podem ser calculados com uma ou mais variáveis de input hidrometeorológicas em diferentes escalas de tempo (diária, semanal ou mensal), sendo a mais comum e aceita a escala mensal. Neste manual, explicaremos em detalhes apenas 5 índices, que foram extensivamente avaliados e validados em todo o mundo. Para obter mais informações sobre alguns índices de seca, pode consultar a bibliografia adicional, como o manual de Indicadores de Seca (Svodova et al., 2016), onde mais de 100 índices de seca são explicados.

4.4.1. Índice Precipitação Padronizado (SPI)

O SPI foi introduzido McKee et al., (1993) e é definido como um valor que representa o número de desvios-padrão da precipitação ao longo dum período de acumulação definido, em relação à média (padronização), uma vez que a distribuição da precipitação original foi transformada para uma distribuição normal. Como exemplo, pode ser indicado que um valor SPI para 3 meses de -2,3 indicaria que a quantidade de precipitação que foi registrada ao longo dum período foi 2,3 vezes o desvio padrão abaixo do valor médio. O passo a passo para o cálculo completo pode ser revisto em diversos documentos, por exemplo, em Kumar et al., (2009). Atualmente este é recomendado pela Organização Meteorológica Mundial para monitorar o estado das secas. É muito flexível, fácil de interpretar, aplicável e comparável em diferentes zonas climáticas e depende de uma única variável, a precipitação.

Através da utilização do índice SPI é possível quantificar e comparar as intensidades dos déficits de precipitação entre zonas com climas muito diferentes e tem a propriedade de ser integrado em qualquer escala de tempo, o que significa que pode ser utilizado como um indicador de diferentes tipos de seca. Escalas temporárias de curta duração (2 - 6 meses) reproduzem secas principalmente nos setores agrícola e florestal, sendo úteis para secas meteorológicas e agrícolas, enquanto as de longa duração (12-24 meses) são associadas as secas hidrológicas já que representam efeitos de seca nos reservatórios, rios e aquíferos.

O cálculo do SPI requer iniciar a partir de séries históricas de precipitações mensais de pelo menos 30 anos. Séries de precipitação em agregações diárias, semanais ou outras são admissíveis no cálculo do SPI; no entanto, recomenda-se seguir as escalas em que normalmente se desenvolvem as secas, sendo esta comumente mensal, sazonal ou interanual. Uma vez que o cálculo do SPI foi realizado numa série de tempo de precipitação, os valores mensais do SPI podem ser avaliados de acordo com a Tabela 1.

4.4.2. Índice padronizado de precipitação e evapotranspiração (SPEI)

O SPEI é um índice de seca com propriedades semelhantes ao SPI, com a diferença de que leva em consideração a temperatura por evapotranspiração, no cálculo, além da precipitação. Essa diferença o torna adequado para analisar tendências e mudanças nas características das secas devido às mudanças climáticas. Seu cálculo requer séries históricas mensais da variação da água no sistema ($\Delta V = P - ETP$), onde o ETP inclui as informações de temperatura. O processo de padronização que o SPEI segue é o mesmo do SPI, com a diferença que a série ΔV é previamente transformada numa distribuição log-normal. Da mesma forma que o SPI, o SPEI pode ser calculado para várias escalas de tempo, permitindo avaliar as secas em várias etapas do ciclo da água e relacioná-las com os diferentes tipos de seca. Sua avaliação segue a Tabela 1.

4.4.3. Índice de condição de vegetação (VCI)

O VCI é um índice utilizado para a detecção e monitoria da seca agrícola. Permite avaliar a duração, a área coberta, a intensidade e os impactos que provoca uma seca na vegetação. Para um determinado ponto ou região, e para cada etapa de tempo (mensal no nosso caso), o VCI é derivado do Índice de Vegetação por Diferença Normalizada (NDVI – Normalized Difference Vegetation Index), obtido a partir da série histórica de imagens de satélite obtidas por AVHRR (Advance Very High Resolution Radiometer) em operação desde 1984:

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

Onde, o max e o min são os valores máximo e mínimo de NDVI presentes nas séries temporais analisadas e o NDVI é o valor de cada intervalo de tempo. O NDVI é, por si só, um índice de vegetação que permite estimar a quantidade, qualidade e desenvolvimento da vegetação com base na medição de sensores de satélite da intensidade da radiação de certas bandas do espectro eletromagnético que a vegetação emite ou reflete. A avaliação do grau de intensidade da seca segundo o VCI é apresentada na Tabela 1.

4.4.4. Índice de Condição de la Temperatura (TCI)

Semelhante ao VCI, o TCI é um índice derivado de medições do sensor satelital AVHRR, neste caso da temperatura superficial. Seu uso está relacionado a monitoria de secas meteorológicas e agrícolas. Sua estimativa usa o brilho estimado pela banda infravermelha térmica (10,3-11,3 μm) detectado pelos sensores do espaço. De acordo com o brilho medido, é estimada a Brightness Temperature (BT), que por si só é um indicador diretamente correlacionado com a temperatura da superfície terrestre. Sua formulação é:

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

Onde, dentro de um período de tempo, BT é o valor em cada passo de tempo (mensal no nosso caso) e o min e max se referem aos valores máximo e mínimo de BT dentro desse período. Sua avaliação segue a Tabela 1.

4.4.5. Índice de Saúde Vegetal (VHI)

O VHI é calculado pela combinação ponderada de duas anomalias já descritas: o VCI e o TCI, ambos derivados de observações obtidas por sensores satelitais. A ideia básica do indicador é a seguinte: quanto mais baixo for o VCI observado (vegetação verde relativamente pobre) e quanto mais baixo for o TCI observado (temperatura relativamente quente), menor é o VHI. Os valores baixos de VHI são indicadores de seca, especialmente quando persistem por longos períodos de tempo.

A equação usada é:

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

Onde w é o peso ponderado, cujo valor está entre 0,0 e 1,0, sendo geralmente 0,5. A avaliação da intensidade da seca segundo o VHI é apresentada na Tabela 1.

4.5. Características das secas

A caracterização das secas é fundamental para a gestão da seca. Definir e quantificar as características de cada evento de seca, permite a planificação de campanhas de monitoria contínua e aplicação de modelos de previsão. Consiste em definir a gravidade de um determinado evento para que um gestor ou técnico possa tomar decisões e aplicar planos de mitigação adequados para

Tabela 1
Categorias de secas
de acordo com os
valores dos índices

Categorias	SPI/SPEI	VCI/TCI/VHI
Úmido	> 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
Seca moderada	-1.00 - -1.50	25 - 40
Seca severa	-1.50 - -2.00	10 - 25
Seca extrema	< -2.00	0 - 10

evitar um grave déficit hídrico, portanto, esse componente metodológico é fundamental para as partes interessadas.

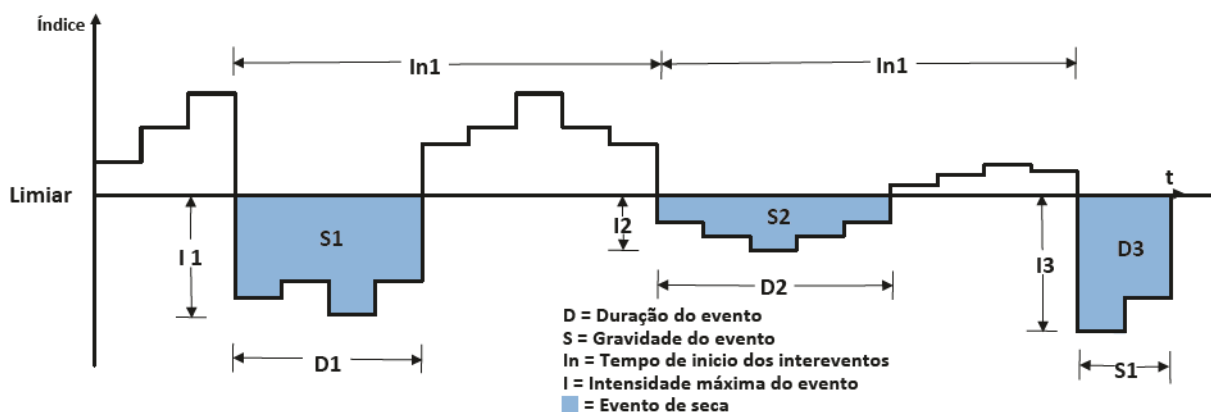
Dentre os diferentes métodos propostos para caracterizar as secas, o método da “run-theory” descrito em Yevjevich, (1969) tornou-se generalizado, devido à objetividade na definição de seca. O método permite uma derivação analítica das distribuições de probabilidade das características da seca, apresentando um desempenho melhor do que outras abordagens. O método de execução pode ser aplicado a uma série temporal de interesse, que no nosso caso serão os índices de seca, assumindo um limiar representativo da gravidade do evento. Os limites são definidos de acordo com a categoria de seca que precisa ser caracterizada.

Um evento seco é definido por sua duração, intensidade, severidade e extensão geográfica. A duração é o período de tempo em semanas ou meses que dura a seca, entendido como o tempo em que o índice esteve consecutivamente abaixo de um limiar, definindo o início e o fim do evento de seca ao ultrapassar esse limiar no tempo. A intensidade máxima ou média é o valor absoluto máximo e a média entre os valores que um índice assume durante a duração do evento. A severidade é definida como a soma das intensidades ou valores tomados pelo índice durante cada período que durou a seca. Essas três características são indicadas na Fig. 3 para três eventos de seca. A extensão geográfica é a área ou porcentagem de área dentro de uma região ou zona em estudo que é afetada por secas numa ou mais categorias de intensidade durante cada intervalo de tempo.

Para quantificar essas características, é necessário calcular os índices de seca sobre vários pontos do território em estudo. Essa distribuição espacial deve ser a mais densa possível, recomendando em Organização Meteorológica Mundial, (2008) pelo menos uma estação ou ponto geográfico onde o índice seja avaliado a cada 250 km².

Os índices de seca adequados são selecionados com base no tipo de seca de interesse. Os índices podem ser considerados gerais para todo o sistema ou específicos, dependendo da parte do ciclo hidrológico para os que foram desenhados. Fazer esta distinção nem sempre é fácil, pelo que, recomenda-se o uso de vários índices para obter uma caracterização completa das secas numa região. Aqui, usaremos os índices expostos anteriormente.

Fig. 3
Características das secas de acordo com a “run-theory”



5. Requisitos

Para a utilização e aplicação deste manual, além das bases teóricas associadas à gestão da seca, é necessário saber que ferramentas computacionais que iram se utilizar. Os dois grupos de ferramentas requeridas são os Sistemas de Informação Geográfica (SIG) e Sistemas de Computação Numérica (SCN). Embora existam muitos softwares semelhantes aos que se explicam a seguir, os conceitos em torno a arquitetura destes programas são os mesmos que os seguidos por outras ferramentas destinadas a cumprir funções semelhantes.

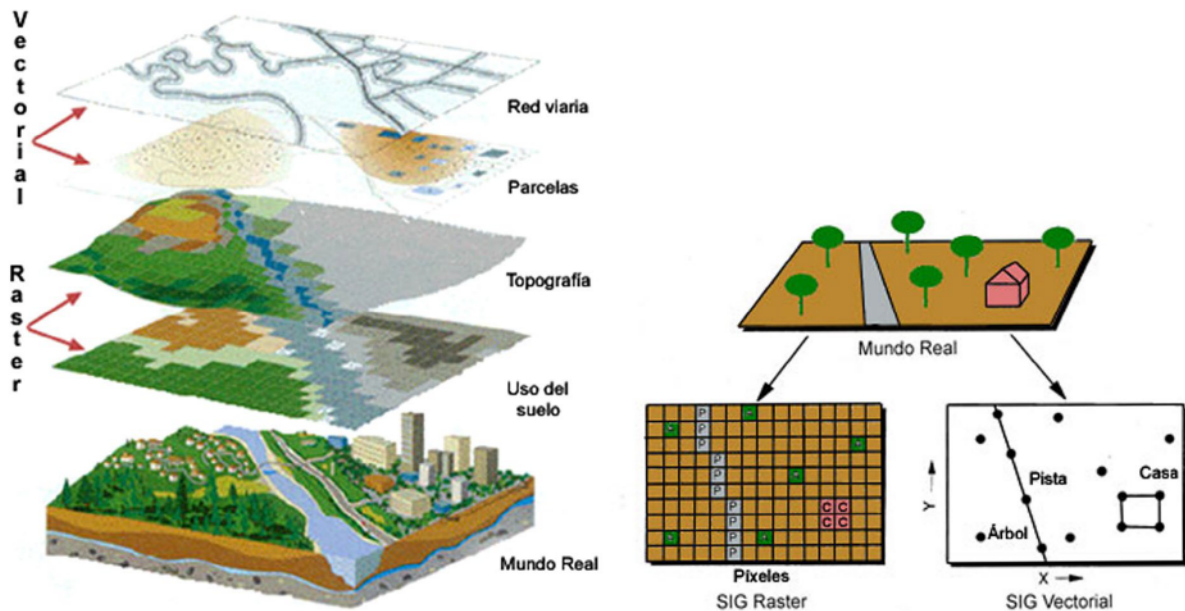
5.1. Sistemas de Informação Geográfica (SIG)

SIG são programas informáticos (softwares) que fornecem ferramentas para recolha, processamento, gestão, análise e representação de dados com componente cartográfico (informação geográfica) em formato digital. Eles nos permitem representar o espaço físico em que vivemos em mapas e tabelas por meio de camadas temáticas que podem ser sobrepostas. Qualquer registro de informação com referência geográfica, como endereço, nome de cidade, identificador de parcela de terreno, coordenadas GPS, etc., pode ser localizado e disponibilizado num mapa (Fig. 4).

O GIS pode ser classificado em dois grupos principais exemplificados na Fig. 4:

Vetor GIS: utilizam vetores para descrever objetos geográficos. São comumente apresentados como shapefiles (extensões comuns: .SHP, .SHX, .DBF, .PRJ, etc.). Os elementos geográficos são representados a partir de três estruturas básicas: pontos, linhas e polígonos. As retas são representadas por segmentos que unem pontos e estes, por sua vez, se fecham para formar polígonos. Como resultado, obtemos três naturezas diferentes de shapefile. Cada um deles possui qualidades específicas de forma que não é possível realizar determinados processos com eles. Desta forma, não é impossível calcular superfícies para pontos, mas é impossível para polígonos. Ou

Fig. 4
Camadas temáticas disponíveis em um SIG
e comparação entre dados raster e vetoriais
Fonte: <https://geopaisa.blog/2017/03/08/que-es-un-sig/>



não é possível calcular comprimentos de polígonos, mas podemos fragmentar polígonos em várias partes.

SIG Raster: Os Sistemas de Informação Raster baseiam sua funcionalidade em formatos gráficos de bitmap (extensões comuns: .ECW, .JPG, .GRID, .TIF, .GIF, .SID). Sua maneira de armazenar informações é dividindo o espaço numa retícula ou matriz (geralmente) regular de pequenas células chamadas pixels e atribuir um valor numérico a cada célula como uma representação de uma qualidade ou propriedade espacial física (altitude, temperatura, distância, inclinação, etc.). Se pegamos a posição das coordenadas do centro de uma das células, pode-se dizer que todos os pixels são georreferenciados.

Dois dos softwares GIS mais comuns e desenvolvidos são o ArcGIS e o QGIS. Ambos são amplamente utilizados e cobrem a maioria das necessidades dos usuários. A principal diferença é que QGIS é um software de código aberto gratuito, enquanto ArcGIS é um programa de pago comercializado pela ESRI. No entanto, a informação geográfica que pode ser representada, analisada e apresentada em forma de mapas no referido software, também pode ser manipulada através de SCNs que oferecem estas opções nas suas características.

Os dados globais analisados neste manual estão em formatos raster, enquanto que para a extracção em áreas de estudo de interesse usam-se arquivos vetoriais que definem um área (país ou bacia hidrográfica). Esta informação pode ser baixada de bancos de dados gratuitos.

5.2. Sistemas de C omputa  o Num erica (SCN)

SCNs s o softwares ou ferramentas destinadas a resolver problemas num ericos complexos e c alculos por meio do uso de m etodos num ericos. O procedimento geralmente consiste em introduzir dados ou vari aveis de entrada a uma fun  o existente no software, o que desenvolve c alculos previamente programados e que dar  um resultado. Os SCNs t em seu pr oprio ambiente e linguagem de programa  o que permite aos usu arios usar as ferramentas e fun  es j  existentes no programa (opera  es simples, como adicionar ou subtrair as complexas, como resolver integrais ou equa  es diferenciais) e criar novas fun  es (programa  o espec fica e personalizada).

Existem muitos SCN aplic aveis   engenharia h drica e, portanto, a gest o de secas. Como exemplo, entre os mais difundidos em uso est o o MATLAB da empresa Mathworks, Python, R, C++, entre outros. Cada um deles   orientado e especializado na solu  o de problemas com bases num ericas espec ficas. Neste manual, se usa R em seu ambiente mais amig vel RStudio, que   especializado na computa  o estat stica e cria  o de gr ficos.

5.2.1. R e RStudio

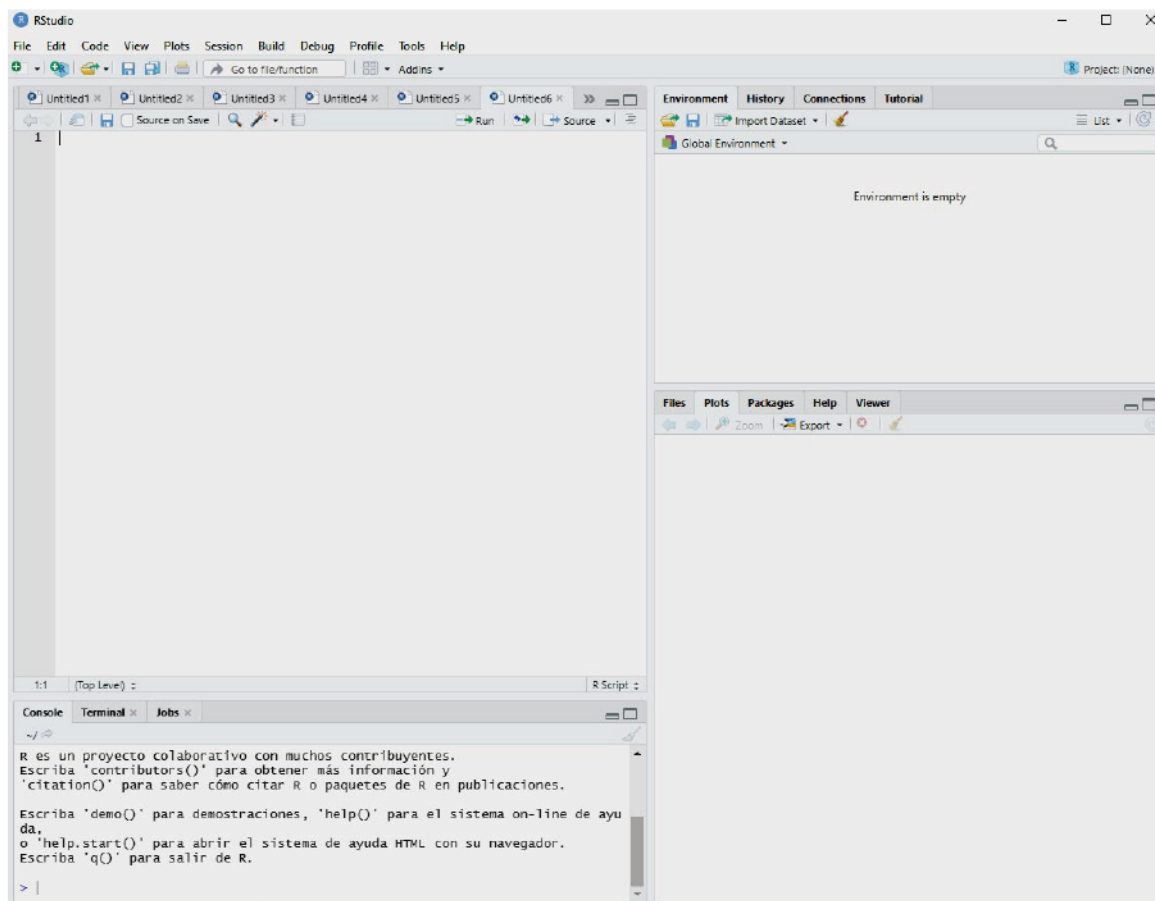
R   um ambiente de programa  o livre, de c odigo aberto e gratuito conceitualizado para o processamento e an lise de dados estat sticos (Dur n, 2000).   constitu do por um conjunto de ferramentas muito flex veis que podem ser facilmente expandidas atrav s de pacotes e bibliotecas de livre acesso ou programando nossas pr prias fun  es.

Os pacotes em R s o cole  es de fun  es e conjuntos de ferramentas desenvolvidas pela comunidade de usuarios. Estes aumentam a potencialidade de R melhorando as funcionalidades b sicas j  existentes, adicionando umas novas. Os milhares de pacotes que existem s o encontrados em reposit rios (CRAN) dos quais seus manuais de usu rio, sua arquitetura e conceitua  o podem ser baixados. CRAN (Comprehensive R Archive Network)   o site oficial do software onde se oferecem diversos recursos para o uso do programa: manuais do usu rio, cursos online, informa  es gerais, descarga de pacotes, informa  es sobre os pacotes instalados, etc.

Para instalar o R,   preciso baixar o instalador conforme ao sistema operacional que possa ser usada desde seu site oficial (<https://cran.rstudio.com/>). Al m disso,   aconselh vel instalar o RStudio (RStudio Team, 20016), um ambiente mais amig vel que permite realizar processos computacionais baseados na linguagem R. A vers o gratuita pode ser baixada do site do programa (<https://www.rstudio.com/products/rstudio/download/>).

RStudio   um ambiente de desenvolvimento integrado (EDI) para a linguagem de programa  o R, dedicado   computa  o estat stica e gr fica. Inclui um console, editor de sintaxe que oferece suporte   execu  o de c odigo, bem como ferramentas para plotar, depurar e gerir o espa o de trabalho.   composto por 4 janelas (Fig. 5): Janela superior esquerda:   o editor de sintaxe,   o local onde editamos a sintaxe para depois execut -la. Ao escrever nada acontecer , a menos que voc  pressione um bot o para executar os comandos. Nesta janela se escrevem novos scripts e ser o abertos os entregues neste manual. Janela superior direita:   o ambiente de trabalho do programa, neste local s o mostrados o conjunto de dados e os objetos (resultados, vari veis, gr ficos, etc.) que s o armazenados durante a execu  o de diferentes an lises. Janela inferior direita: possui v rias subguias: (i) a guia arquivos permite visualizar o hist rico dos arquivos trabalhados com o programa; (ii) a guia plotagens permite que visualizar os gr ficos gerados; (iii) a guia pacotes permite visualizar os pacotes baixados e guardados no disco duro, bem como

Fig. 5
Ambiente RStudio



gerir a sua instalação ou actualização; (iv) a janela de ajuda permite o acesso ao CRAN; (v) a janela do visualizador mostra os resultados ao construir relatórios. Janela inferior esquerda: é o console, corresponde ao que o software R seria na versão básica. Nesse local, o software executa as operações realizadas a partir do editor de sintaxe.

Dos milhares de pacotes existentes para R, existem vários desenvolvidos especificamente á gestão de recursos hídricos e secas (Slater et al., 2019). Alguns dos pacotes requeridos nos scripts anexos são os “SPEI” desenvolvidos por Begueria e Vicente-Serrano, (2017), que contém funções que permitem calcular diretamente os índices de seca SPI e SPEI; pacote “ggplot2” permite fazer gráficos de uma forma profissional, e permite inclusive representar graficamente shapefiles e rasters; o pacote “tmap” permite ao usuário R manipular informações geográficas, normalmente manipuladas por meio de GIS com R.

Nos scripts apresentados neste manual (ANEXOS: Scripts em R e RStudio para gestão de secas) que se analisam e desenvolvem passo a passo nos capítulos II e III. A informação específica requerida para cada um deles se obtém a partir de bancos de dados gratuitos. Os scripts

estão escritos de forma lógica e ordenada para poder serem usados, replicados, manipulados, melhorados e complementados pelos usuários deste documento.

Como informação geral R permite trabalhar com todos os tipos de dados: matrizes, vetores, informações geográficas vetoriais e ráster, etc. Cada arquivo pode ser lido e atribuído um nome (usando o símbolo <-). Podem ser criados diferentes loops e algoritmos existentes (por exemplo, funções seno, cosseno). Os gráficos podem ser gerados e visualizados no visualizador e exportados para qualquer pasta no Windows Explorer em diferentes formatos (.pdf, .jpg, png, etc.). Recomenda-se o seguinte manual para iniciar com a programação em R: R para Iniciantes (Ahumada, 2003). No entanto, existem centenas de manuais e blogs online que podem ser usados como fonte de orientação.

6. Bases de datos globais

Como afirmado anteriormente, para a gestão ideal da seca são necessárias séries temporais de dados hidrometeorológicos correspondentes a cada componente do ciclo da água. Esses dados devem ser de longa duração (pelo menos 30 anos) e ter sido medidos com suficiente distribuição sobre o território natural (uma estação meteorológica a cada 250km²). No entanto, a medição dessas variáveis com estas condições é uma tarefa complexa. O principal motivo geralmente é o alto custo de instalação e manutenção, motivo pelo qual em muitas áreas do mundo a disponibilidade de dados é extremamente limitada.

Para obter dados como precipitação, umidade do solo, taxas de fluxo, etc. diversas instituições desenvolveram bancos de dados globais nos quais oferecem séries temporais gratuitas e abertas de variáveis de interesse hidrológico e climático. Esses productos estão baseados em dados pontuais medidos que foram interpolados por meio de técnicas estatísticas, em modelos climáticos e hidrológicos globais e em informações baseadas provenientes de radares satelitais.

Além de bases de dados específicas para gestão de água e secas, se apresentam alternativas para a obtenção das informações geográficas necessárias também para a representação de variáveis hidrológicas, mapas e resultados. Esses dados são arquivos vetoriais correspondentes aos limites administrativos dos países e aos limites das principais bacias hidrográficas do mundo. Existem centenas de fontes de dados geográficos, productos ofertados por organizações governamentais e ambientais, entre outras, que oferecem este tipo de informação de forma gratuita e aberta. Também se podem encontrar arquivos raster com informações topográficas e climáticas, entre outras.

6.1. Bases de datos complementares

Os arquivos vetoriais (shapefiles) dos limites de países ou bacias hidrográficas georeferenciados se descarregam desde os bancos de dados geográficos públicos GADM e FAO (Food and Agricultural Organization).

O GADM é uma base de dados de alta resolução das áreas administrativas em todo o mundo, oferecido para todos os países, todos os níveis administrativos e em qualquer período de tempo. O GADM criou dados espaciais para muitos países a partir de bancos de dados espaciais fornecidos por governos nacionais, ONGs e/ou de mapas e listas de nomes disponíveis na Internet. O banco de dados está disponível em vários formatos de exportação, incluindo shapefiles usados na maioria dos aplicativos GIS. Estes arquivos também podem ser usados com a linguagem de

análise de dados R. Os shapefiles delimitam as áreas para a descarga a análise de dados, além de ser necessários para a criação de gráficos de dados descritivos que incluam mapas geográficos. O endereço web é <https://gadm.org/>. Possui um ambiente muito intuitivo onde dentro da guia Dados se escolhe o país de interesse e se baixam os arquivos vetoriais distribuídos em níveis de 0 a 5 onde 0 é o limite nacional e 5 é a menor divisão que cada país possui (nem todos têm 5 níveis).

A FAO na sua secção “geonetwork” oferece uma ampla variedade de mapas interativos, conjuntos de dados GIS, imagens de satélite e aplicativos relacionados ao seu trabalho. Possui uma grande variedade de informação espacial global básica desenvolvida para apoiar a toma de decisões, promover abordagens multidisciplinares para o desenvolvimento sustentável e melhorar a compreensão dos benefícios da informação geográfica. GeoNetwork é uma fonte de dados aberta que facilita a partilha de informações temáticas referenciadas geograficamente entre diferentes organizações. O endereço web é <http://www.fao.org/geonetwork/srv/en/main.home>; no qual é possível acessar a guia “Hydrology and Water Resources” e obter arquivos vetoriais georreferenciados correspondentes aos limites de bacias hidrográficas ao redor do mundo em diferentes níveis (bacias e sub-bacias), leitos de rios e outros.

6.2. Bases de dados de variáveis hidrometeorológicas

6.2.1. Tipos de bases de dados

Neste manual se mostra como baixar, extrair e aplicar séries temporais das seguintes variáveis: precipitação em mm/mês, temperatura média em °C por mês, ETP em mm/mês e ETr em mm/mês, além dos índices de seca mensais VCI, TCI e VHI. Cada uma dessas variáveis é medida ou estimada de forma diferente, portanto, os bancos de dados que oferecem estes productos provem de fontes diferentes. Em geral, os dados existentes são obtidos a partir de 4 técnicas: a partir de medições in situ, dados obtidos desde medições de satélite, dados modelados numericamente mediante algoritmos e dados gerados a partir de técnicas de reanálise. Alguns bancos de dados integram estas técnicas para melhorar a precisão dos productos.

A forma tradicional de medir a precipitação, e a temperatura (variáveis mais comuns e básicas) é usando pluviômetros e termômetros. Cada uma dessas estações está localizada num ponto específico do território. Nos últimos anos, essas variáveis também foram estimadas por meio de detecção remota desde radares infravermelhos de satélites que orbitam a Terra. Esta última metodologia é o mecanismo pelo qual se estimam as variáveis das quais os índices de seca VCI, TCI e VHI são derivados.

ETr e ETP, umidade do solo e vazões (escoamento) são outras variáveis importantes no ciclo da água e, portanto, para estudos de seca. Medi-los requer lisímetros, perfurações no solo e estações de medição de vazão em diferentes locais ao longo dos rios. No entanto, a aplicação desses métodos costuma ser cara para instalar e manter, portanto, essas variáveis não são menos comuns de medir, especialmente em países e regiões com limitações socioeconômicas. A solução mais comum para obter esses dados é por meio de modelos hidrológicos/climáticos globais (grande escala) ou locais (escala pontual). Os modelos utilizam as variáveis de entrada medidas (precipitação e temperatura) para estimar as variáveis perdidas. Outra forma de estimativa é por meio de formulações matemáticas e estatísticas com base nas relações empíricas entre as variáveis medidas e estimadas. Um exemplo claro é o cálculo do ETP usando formulações como a de Thornthwaite (Thornthwaite, 1948) que estima a ETP com base na temperatura e precipitação.

A grande variedade de metodologías para estimar as variáveis hidrológicas faz com que existam diferentes bancos de dados geralmente pertencentes a grandes instituições científicas (NASA, Universidades, governos, etc.). Os productos oferecem dados medidos de forma tradicional (através de estações de medição insitu) são geralmente intermitentes (com lacunas) e contêm erros resultantes de seu processo de medição. Além disso, a cobertura global é muito diferente em diferentes partes do globo. Para superar essa limitação, as instituições automatizaram a análise de vários dados pontuais e calcularam uma distribuição desses valores no espaço com o uso de modelos de interpolação, redução da escala e técnicas de reanálise. A aplicação dessas técnicas tem sido apoiada por dados obtidos de modelos numéricos e de medições de satélite.

Dependendo da base de dados, pode-se encontrar vários formatos de descarga. Embora geralmente, por serem parâmetros variáveis no tempo, as informações espaciais e temporais são unificadas num único arquivo. Essa união é realizada em camadas, como uma matriz tridimensional (ou multimatriz). Por exemplo, num único arquivo pode-se encontrar uma malha de dados de precipitação para o mês de janeiro, outra malha de dados para o mês de fevereiro, outra para o mês de março, etc., salvando 12 malhas (correspondentes aos 12 meses de um ano), um após o outro. As grades da malha estão associadas a um valor (de precipitação, temperatura, etc.) comparável ao de uma estação localizada no centroide de cada uma, de forma que várias células podem ser adicionadas dentro de uma área ou região para formar uma única série de média temporal.

Devido à grande quantidade de informações que cada banco de dados apresenta, os arquivos que contêm os dados costumam se fragmentar em vários arquivos correspondentes aos dados de cada ano, mês ou outro do território. Embora existam vários tipos de arquivos, os mais comuns que encontraremos e os que mais utilizaremos neste manual são os netCDF (formulário de dados comuns online). Este é um formato de arquivo destinado a armazenar dados científicos multidimensionais (neste caso variáveis de interesse hidrológico). Cada uma dessas variáveis pode ser exibida numa dimensão (por exemplo, tempo). Esses arquivos (extensão comum: .NC) são comparáveis a arquivos raster e são tratados com o pacote R “ncdf4”.

Para outras extensões, no entanto, é necessário extrair uma camada específica e convertê-la ao formato desejado (recortando as informações correspondentes à nossa área de estudo), exigindo o uso de algoritmos específicos. O processamento de dados requer a consulta da documentação específica de cada banco de dados para aprender como explorar as diferentes camadas de informação. Os gestores das bases de dados tendem cada vez mais a facilitar essa tarefa e até mesmo códigos de leitura de arquivos são publicados.

Neste manual nos concentremos em dados históricos, porém, existem diversos bancos de dados que oferecem variáveis hidrológicas projetadas até o ano 2100, que podem ser obtidas de maneira similar. Essas projeções são suportadas em modelos climáticos globais que, a partir de medições históricas, projectam as estimativas de acordo com as diferentes ações do ser humano frente as mudanças climáticas. Como exemplo, citaremos o “Climate Change Knowledge Portal” (<https://climateknowledgeportal.worldbank.org/download-data>) que oferece uma ampla gama de variáveis de interesse hidrológico que podem ser baixadas e utilizadas de forma semelhante ao que explicaremos nos capítulos II e III.

A seguir, se apresentam as bases de dados das quais se obtêm as variáveis hidrológicas. Se explica de onde vêm a informação, que variáveis e em quais escalas de tempo e espaço as oferecem.

6.2.2. CHIRPS

CHIRPS é a sigla para Rainfall Estimates from Rain Gauge and Satellite Observations. CHIRPS, da Universidade da Califórnia em Santa Bárbara (Funk et al., 2015) oferece séries de precipitação diária e mensal numa escala de $0,05^\circ$ ($\approx 5,6$ km no equador) desde 1981. Os produtos são baseados em medições de observações de satélite NOAA (National Oceanic and Atmospheric Administration) e NASA (National Aeronautics and Space Administration). Posteriormente, eles são calibrados e validados a partir de um grande número de estações pluviométricas localizadas em toda a superfície terrestre.

Aquí se utilizará a versão V2.0 (<https://www.chc.ucsb.edu/data/chirps>). A sua precisão espacial motiva a utilização desta base de dados em estudos hidrológicos e gestão de secas uma vez que é particularmente útil em estudos em pequenas áreas ou fossos, podendo ser adicionada a qualquer superfície (por exemplo, K. Wang, Li e Wei 2019). Embora o escopo seja indicado como global, os dados estão disponíveis apenas para a superfície localizada a menos de 50° ao norte e ao sul do plano equatorial. As atualizações são feitas todas as semanas, permitindo uma monitoria em tempo quase instantâneo.

6.2.3. TerraClimate

TerraClimate (<http://www.climatologylab.org/terraclimate.html>) procedente do Climatology Lab da Universidade de Idaho ofereceu vários produtos a escala mensais desde 1958 (Abatzoglou et al., 2018). Usando técnicas de interpolação, redução de escala e reanálise criou malhas com uma resolução espacial de $1/24^\circ$ ($\approx 4,6$ km no equador) de dados de precipitação, temperatura (máxima e mínima), pressão de vapor, velocidade do vento e radiação de ondas curtas de superfície em todo o mundo. Com essas informações usando o modelo de balanço hídrico global Thornthwaite-Mather (WBM) (Willmott et al., 1985), o ETP (seguindo a formulação de Penman-Montieth), o fluxo ou escoamento, ETr, déficit hídrico (D), umidade do solo e água da neve equivalente em cada célula.

As variáveis são atualizadas anualmente, portanto seu uso como fonte de dados para monitoria de curto prazo não é possível. Porém, devido à sua curta resolução espacial, pode ser utilizado na análise de secas ou balanços hídricos em regiões de todos os tamanhos. Aquí vão se baixar e usar a precipitação, a temperatura média (derivada da temperatura máxima e mínima) e o déficit hídrico.

6.2.4. NOAA STAR

O Center for Satellite Applications and Research (STAR) e a National Oceanic and Atmospheric Administration (NOAA) da NASA fornecem uma variedade de produtos de sensores de satélite globais relacionados à vegetação (Kogan, 1995). As principais variáveis são o NDVI e o BT (Kogan, 1995), dos quais se podem derivar o VCI, TCI e VHI, que também já foram calculados. Esses produtos foram usados em vários estudos relacionados ao teor de água do solo, rendimentos agrícolas e secas, entre outros (por exemplo, Agutu et al. 2017; Bento et al. 2018).

Os dados podem ser baixados gratuitamente em resolução espacial de malha de 4×4 km e resolução temporal semanal, 1983 até hoje em https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. Suas atualizações são feitas cada uma ou duas semanas, permitindo uma monitoria quase em tempo real. Faremos o download dos índices VCI, TCI e VHI semanais e os adicionaremos mensalmente.

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 se encuentran 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. Objectivos específicos

Depois de fazer uma breve introdução aos conceitos associados à gestão da seca e, de expor as bases de dados de interesse, nos capítulos II e III seguiremos os seguintes objectivos específicos:

- a) Descarregar e representar graficamente os arquivos vetoriais com os limites de Moçambique e a bacia hidrográfica do rio Licungo.
- b) Descarregar dados mensais de precipitação do CHIRPS, obtê-los para Moçambique e plotar as suas médias anuais. Calcule o SPI e representar graficamente suas características.
- c) Descarregar os dados semanas de VCI, TCI e VHI do NOAA STAR, obtenha-os em Moçambique, agregue-os numa escala mensal e represente graficamente o seu desenvolvimento durante um período de interesse.
- d) Descarregar os dados de precipitação mensal, ETr e ETP do TerraClimate, obtê-los dados da bacia do rio Licungo, calcule o balanço e o déficit hídrico mensal e representar graficamente suas médias anuais. Calcule o SPEI e tabule suas características.
- e) Traçar o estado de seca/não seca com base no SPI em Moçambique.
- f) Representar graficamente a área afetada pelas secas de acordo com o SPEI na bacia do rio Licungo.
- g) Representar graficamente as características espaciais das secas em Moçambique.

Para atingir cada um desses objectivos se apresentam 3 scripts disponíveis nos ANEXOS, os mesmos que se explicam passo a passo nos capítulos I e II.

8. Seguimiento del manual

O ambiente de trabalho é projetado para o sistema operacional Windows. Para seguir o manual é preciso instalar o R e o RStudio, a partir dos quais podem se seguir os passos detalhados nos capítulos II e III de acordo com os scripts fornecidos nos anexos e que estão organizados da seguinte forma:

- a) O ANEXO I apresenta o código necessário para extrair as informações das diferentes bases de dados. Numa pasta de destino, os dados serão guardados, arrumados e prontos para uso.
- b) O ANEXO II contém as ferramentas necessárias para a leitura das informações baixadas e ordenadas obtidas no Anexo I e utiliza-las tanto para caracterizar secas por meio de índices como para realizar balanços hídricos.
- c) O ANEXO III apresenta os scripts necessários para a leitura da informação gerada nos Anexos I e II, representá-la graficamente e exportá-la em diferentes formatos.

Ao início de cada script se descrevem os pacotes que devem ser baixados (apenas uma vez por CPU utilizada) e ativados (a cada vez que uma nova sessão do RStudio é aberta) para usar esses códigos. R possui funções internas que permitem baixar e ativar pacotes diretamente do ambiente de trabalho. Para usar os códigos é necessário abrir um novo arquivo “R Script” no RStudio e copiar os códigos dos anexos. Para executar cada linha de código, é necessário posicionar-se no início de cada linha e clicar na opção “executar” no canto superior direito da janela de trabalho. Tanto no manual como no código, o passo a passo é especificado.

Dentro de cada script existem várias partes que correspondem a cada um dos subcapítulos do Manual (os pacotes são necessários para todo o código). Se quer seguir um único processo de descarga de dados, calculando índices e gráficos para qualquer região, se poderia copiar cada parte que se requer dos 3 scripts apresentados e ordená-los num novo arquivo. Este manual pretende ser um guia a partir do qual seu conteúdo pode ser expandido e modificado. À medida que avança no manual, se encontram etapas e seções, os mesmos que serão explicadas em detalhe a primeira vez que apareçam.

Se requer ter descarregado do GADM e da FAO e guardado numa pasta dentro do explorer Windows os arquivos vetoriais (.shp) de Moçambique e da bacia do rio Licungo. Vamos salvá-los numa pasta dentro do Windows Explorer. A pasta nos scripts se denomina “Layer_Mozambique”, e cada um dos arquivos “mozambique.shp” e “basin_Licungo.shp”. Porém, os nomes da pastas e dos arquivos podem ser modificados por escolha do usuário.



II. Obtenção de dados

9. Área de estudo

Nesta seção se segue o script gráfico no qual simplesmente se verifica se os arquivos contendo os contornos de interesse estão localizados geograficamente corretamente. Para isto se requer um shapefile de todos os países do mundo baixado do GADM (https://gadm.org/download_world.html).

- 1) Os pacotes necessários são instalados e carregados. Destes pacotes se usam algumas funções, embora para profundizar nelas ou conhecer todas as funções que os pacotes oferecem, se podem pesquisar no CRAN os manuais oferecidos pelos programadores destes pacotes.

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

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

- 2) Os shapefiles de Moçambique e da bacia do rio Lacungo são lidos, e denominados mz e cl, respectivamente. O arquivo também é lido com os limites de todos os países do mundo. Leve em consideração a direção onde a pasta está localizada e a direção da linha 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) Os shapefiles são plotados de maneira básica para verificar se os arquivos que carreguei estão corretos. Vemos os resultados no visualizador.

```
#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 produz um gráfico e se importam os arquivos de interesse, como .pdf e .png, para uma pasta onde se verifica se a localização geográfica é apropriada. Leve em consideração a direção onde a pasta está localizada e a direção da linha inclinada. Neste caso, se verifica que os arquivos correspondem aos limites exigidos.



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

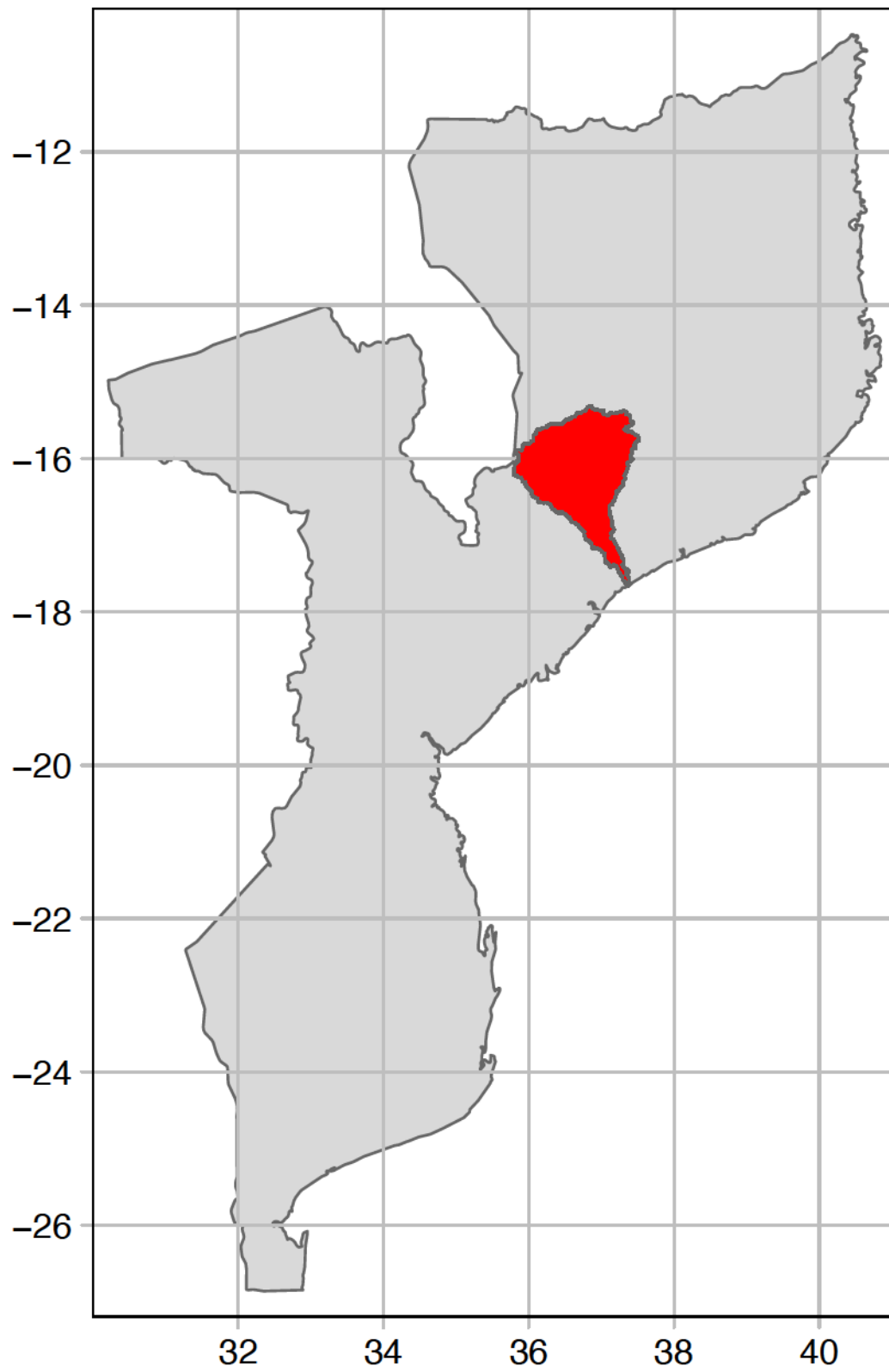
#The graphic is exported as png, the destination folder and the size can be modified
png("E:/19_MANUAL_data_global/Graphics/ubicacion_mz_cl.png",height = 270,width = 400)
graph #Assigned name of the graphic to be exported as png
dev.off() #Export closed

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

Caso os arquivos exijam alterações de projeção no globo terrestre ou não estejam localizados corretamente, um software GIS deve ser utilizado para corrigir os erros.



Fig. 6
Localização e representação
dos limites de Moçambique
e da bacia do rio Licungo



10. Precipitação: CHIRPS

Nesta seção, se baixam os arquivos contendo as informações de precipitação do site do CHIRPS. Após, a informação mensal de todos os pixels que estão dentro de Moçambique é obtida. Se agregam os dados numa escala anual e se representam graficamente de forma simples, elegante e com alta qualidade. As etapas a seguir são:

- 1) Na página CHIRPS https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_monthly/netcdf/ baixe o arquivo: `chirps-v2.0.monthly.nc`. Este arquivo pesa cerca de 6 Gb e contém as informações globais de precipitação mensal desde 1981. Ele é mantido numa pasta conhecida. Nessa página da web, você pode baixar outros arquivos, como dados de precipitação diários e trimestrais, entre outros.



- 2) Os pacotes necessários são instalados e carregados.
- 3) O shapefile de Moçambique é lido e atribuído com o nome `mz`.
- 4) O arquivo `chirps-v2.0.monthly.nc` é lido, seguindo o script (a variável chamada “`chi`” que pode ser alterada para o que você quiser).

```
#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) A variável “`chi`” é exibida no console por meio da função de impressão que permite ver as características do arquivo. Aqui se extraem todas as informações necessárias para entender como os arquivos são armazenados e desta forma extraí-los corretamente na área de interesse. Esta etapa será comum para todos os arquivos com a extensão `.nc`.

A primeira seção mostra o número de variáveis que o arquivo contém. Nesse caso, ele contém uma única variável chamada “`precip`” e armazena seus dados numa matriz tridimensional baseada na longitude, latitude e tempo. A unidade de cada dado é mm/mês, o intervalo de tempo mensal e para uma célula com dados perdidos o valor de -9999 é atribuído:

```
> 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
```


A próxima seção relata as características das dimensões nas quais o arquivo armazena as informações. Nesse caso, ele divide o mundo longitudinalmente em 7.000 partes desde o leste e latitudinalmente em 2.000 partes desde o norte. Essas matrizes existem para 468 etapas de tempo (meses) que são nomeadas com um número de dias desde 1 de janeiro de 1980 (ou seja, por exemplo, o número 29 refere-se a 29 de janeiro 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, o arquivo contém diversos atributos relacionados à fonte da informação. Nesse caso, são 15 atributos, incluindo a versão, o título, um e-mail de contato para o criador, a data de atualização, a instituição do desenvolvedor, 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) A variável “tempo” é extraída do arquivo “chi” (tempo) para criar um vetor com as datas às quais pertence cada passo de tempo. Vetores adicionais de ano e mês são gerados na 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) Ao final da análise das informações e extração das datas a que pertencem os dados, o arquivo é encerrado. Recomenda-se não deixá-lo aberto, pois normalmente essas variáveis ocupam muito espaço na memória do disco, podendo retardar o programa ou impedir a criação de novas variáveis.

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

- 8) As informações de precipitação são extraídas. Aqui se abre o arquivo CHIRPS novamente com a particularidade de ler diretamente a variável “precip” analisada previamente. À variável a que foi atribuída esta informação, examinam-se a classe e o modo, no qual se verifica que é o tipo “RasterBrick” e “S4” que corresponde à informação raster.

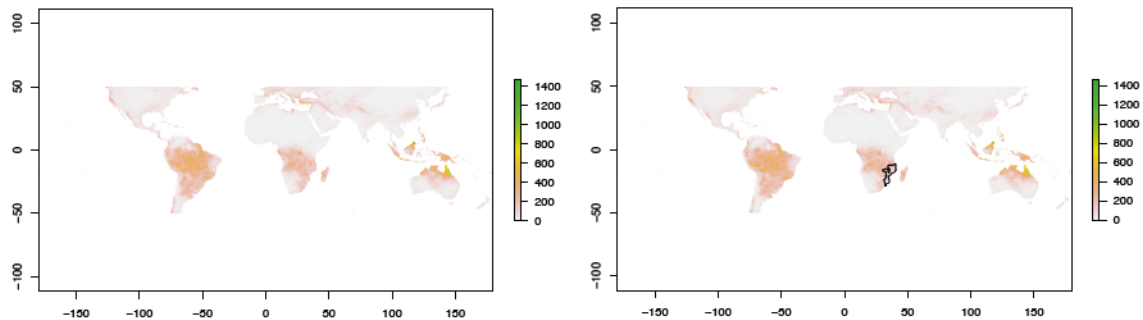
```
#Precipitation information is extracted
precip<-brick("E:/19_MANUAL_data_global/CHIRPS/chirps-v2.0.monthly.nc", varname = "precip")#Precipitation is extracted
class(precip)#It shows what kind of variable it is
mode(precip)#It is shown whether the variable information is numeric or text
dim(precip)#The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

- 9) Como a dimensão de cada raster possui 468 passos de tempo, um deles (o primeiro do script) é extraído para análise de sua composição. Ao extrair esta informação se verifica sua dimensão. Além disso, é representado graficamente primeiro por si próprio e depois com o shapefile da área de interesse (Moçambique) para verificar se está georreferenciado corretamente.

```
#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
```

No visualizador, vai se obter algo assim:



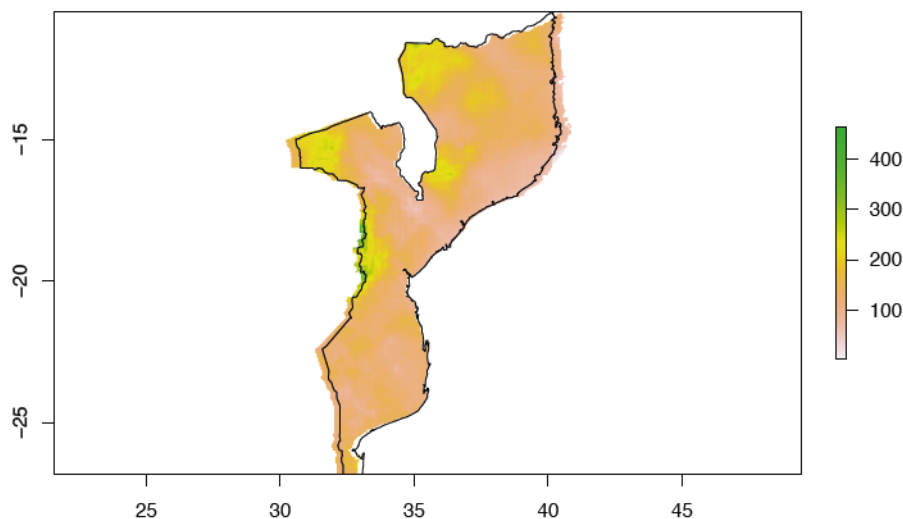
Onde se pode confirmar a coincidência espacial correta da informação raster de precipitação com o shapefile “mz”. Também pode se ver como a precipitação mundial para este intervalo de tempo específico (janeiro de 1981) varia de 0 a um pouco mais de 1400 mm.

- 10) Uma vez verificado que a localização da informação raster coincide com os shapefiles, os dados dentro da área moçambicana são extraídos. O procedimento se realiza em três etapas:

```
#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)
```

Os gráficos permitem verificar que os dados foram extraídos corretamente:

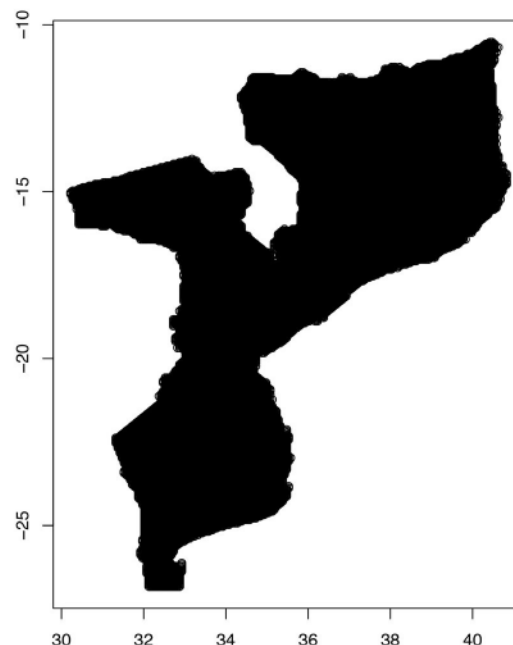


- 11) As coordenadas de cada pixel com informações de precipitação são extraídas dentro do shapefile de Moçambique. Seguimos os passos do script e obtemos uma matriz com duas colunas de pares ordenados de longitude e latitude.

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

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

Para saber quantos pixels existem, a função `dim` (26868 pixels) é usada, e eles são representados graficamente para observar no visualizador se os pares ordenados foram extraídos corretamente:



Esta matriz (chamada de `coordinates`) é guardada numa pasta de interesse no formato `.txt`. Este arquivo é muito importante, pois contém a localização das células contidas no shapefile de interesse e a ordem em que se guarde não deve ser alterada posteriormente (próximo passo).

```
#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 o processo de extração de coordenadas (usando a função “`extract`”), além das coordenadas, são extraídos os dados de precipitação associados a cada coordenada. Nesta etapa, se lança um loop em que “`i`” varia de 1 até o número de etapas de tempo contidas no arquivo “`precip`”. Se repete o processo de extração de dados para cada uma das etapas de tempo e se armazenam numa matriz. Esse processo demora dependendo da quantidade de pixels, ou seja, do tamanho da região em estudo. No final do loop, se revisa sua dimensão, que deverá ter uma linha de dados para cada etapa de tempo e uma coluna da esquerda a direita correspondendo à série da coordenada, arrumada de cima para baixo na etapa anterior. É muito importante manter o arquivo de coordenadas e dados de precipitação juntos, pois sua ordem é essencial para o uso correto dos dados.

```

#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) As séries de dados mensais são guardadas no formato .RDS, que é um formato compatível no RStudio e mais eficiente que um texto quando as matrizes são grandes. A matriz (função “readRDS”) pode ser carregada novamente e com o uso da função “head”, observa-se a composição superior esquerda dos dados da matriz e verifica-se que o armazenamento dos dados está correto:

```

#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) Os dados mensais são agregados aos anuais por meio de um loop. O código filtra as informações de cada pixel associado a cada ano (12 meses), adiciona e acumula numa nova matriz. Esta nova matriz terá dimensões de número de anos por número de pixels. Esta nova variável se guarda como um arquivo .RDS ou .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) A partir da matriz de dados anual, a precipitação média anual para cada pixel pode ser estimada. Se calcula gerando uma matriz de 3 colunas. As duas primeiras com o par arrumado de coordenadas ao qual cada dado pertence e um terceiro com o vetor resultante da média da precipitação anual. Este arquivo se guarda numa pasta conveniente.

```

#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 elaborar um gráfico de maneira elegante e profissional da precipitação anual, abra o arquivo com esta informação. A matriz é convertida num raster (veja o código com as funções utilizadas), e a função “levelplot” é utilizada para gerar um gráfico de alta qualidade. Pode se manipular a gama de cores, os valores mostrados na legenda, mostrar ou não os quadrantes geográficos, quadros, etc. Este gráfico pode se guardar em qualquer formato que desejar.

```
#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
class(raster_1); mode(raster_1)#The variable is checked to ensure that it is of raster type
dim(dat_grafico)#Their dimensions are revised
head(dat_grafico)#The composition of data is reviewed

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

#The boundaries within which the information is found are sought
max(dat_grafico[,3],na.rm = T)#Maximum value, 3 is the column where the data of interest is located
min(dat_grafico[,3],na.rm = T)#Minimum value, 3 is the column where the data of interest is located

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

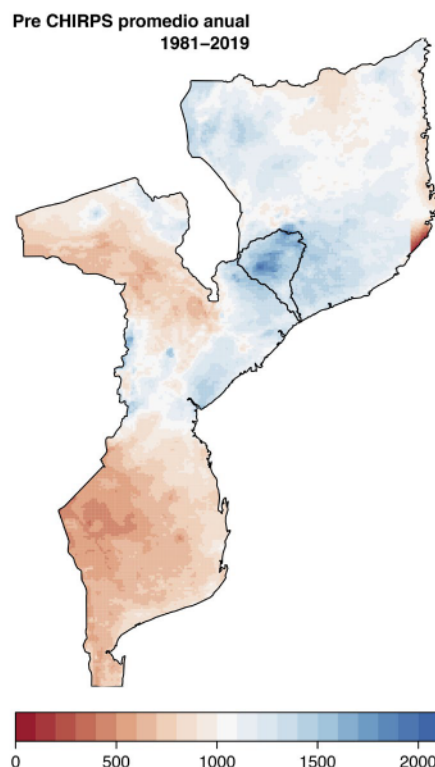
#The data is plotted in an elegant way
professional_graphics<-levelplot(raster_1,main="Pre CHIRPS promedio anual
1981-2019",ylab="",xlab="",#Choose the raster to plot

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 = 6.25, height = 6.75, onefile = FALSE)
professional_graphics #Assigned name of the graphic to be exported as pdf
dev.off() #Export closes
```

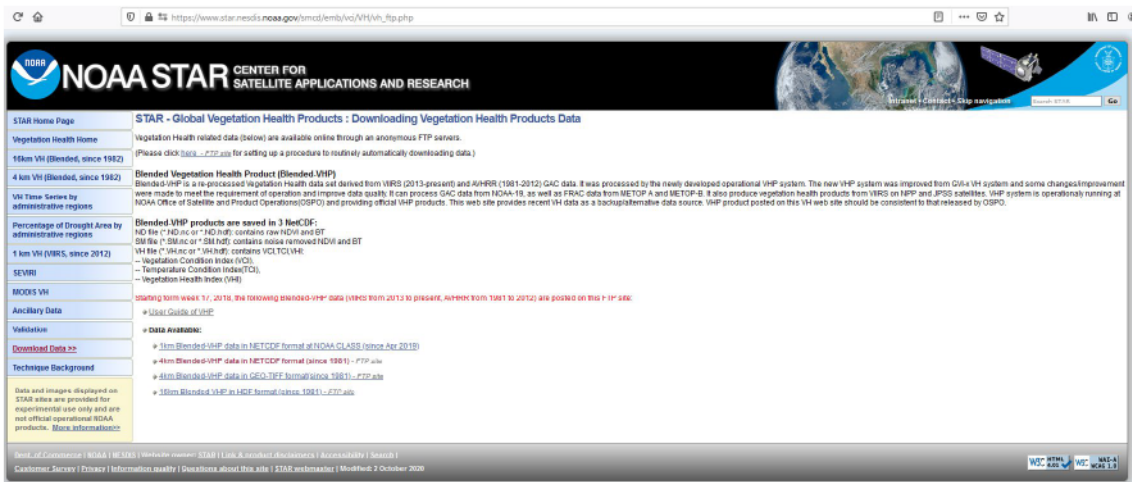
O resultado deste gráfico será semelhante a este, no qual também se inclui o shapefile da bacia do rio Licungo:



11. Índices de vegetação do NOAA STAR

Nesta seção, vão ser descarregados os índices de seca VCI, TCI e VHI para Moçambique desde o NOAA STAR. O processo pode ser aplicado a outros produtos, como NDVI ou BT. Se utiliza uma estratégia de download de dados diferente da usada no CHIRPS. As séries de dados vêm numa escala semanal, portanto, devem se agregar numa escala mensal para serem comparáveis com a precipitação ou outros índices de seca. Se faz um gráfico do status mensal de 6 meses consecutivos para analisar sua alteração.

- 1) Os arquivos que contem os productos do NOAA STAR são baixados do site: https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_ftp.php. Em que as seguintes informações podem ser observadas:



Onde, se especifica que os produtos são armazenados em 3 arquivos NetCDF (extensão .nc). Se indicam que arquivos contêm os diferentes dados disponíveis e se tem acesso a um manual do usuário o qual se recomenda revisar. Portanto, aqui se exigem os “arquivos VH” que contêm os índices VCI, TCI e VHI. Se entra na opção de baixar dados numa 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

Todos os arquivos VH são baixados um por um e guardados numa pasta de interesse. Cada um desses arquivos apresenta a série dos índices VCI, VHI e TCI em todo o mundo numa escala de 4km para cada semana desde 1981. Por ser uma malha muito pequena, o NOAA STARR decidiu oferecer um arquivo para cada semana. Embora existam mecanismos de programação para realizar o download diretamente, neste manual de iniciação faremos isso manualmente.

Além disso, na pasta onde guardamos os arquivos (1946 arquivos da semana 35 de 1981 à semana 52 de 2019), é requerido um arquivo do tipo texto (.txt), que contém os nomes de cada um dos arquivos baixados.

- 2) Uma vez no RStudio, os pacotes necessários são instalados e carregados.
- 3) O shapefile de Moçambique é lido e atribuído o nome mz.
- 4) O arquivo contendo os nomes dos arquivos VH é lido. Os dados da área de interesse são extraídos de cada um deles.

```
#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) Tal como acontece com os arquivos de dados CHIRPS, se carrega um dos arquivos VH e se imprimem para ver suas configurações. Aquí se usa a lista de nomes dos arquivos em .txt para ler os arquivos.

```
#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
```

A primeira coisa que se pode ver é o número de variáveis que cada um desses arquivos contém:

```
> 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):
```

Das 4 variáveis ou produtos, o VCI, TCI e VHI se identificam com esse mesmo nome e vêm com uma configuração que depende das variáveis “WIDTH” e “HEIGHT”. Eles variam de 0 a 100 e estão no 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"
```

A seguinte que se pode ver é o tamanho das dimensões “WIDTH” e “HEIGHT”. Isso significa que cada matriz que contém cada variável tem 10.000 x 3616 células. Além disso, verifica-se que não possui dimensões, o que significa que não é possível extrair os dados diretamente utilizando a estratégia utilizada no CHIRPS. O que é feito nesses casos é explicado posteriormente.

```
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

```

Por fim, são apresentados vários outros atributos do ficheiro, como o instrumento de medição, o satélite, a data em que foi efectuada a missão de recolha de dados, o ano e a semana a que correspondem os dados contidos, entre outros. Os dados relevantes são o intervalo em que cada pixel está localizado geograficamente.

- 6) Nesta etapa, os valores VCI são extraídos deste arquivo como exemplo. Isso é feito para observar sua composição e saber como tratar os dados posteriormente. A matriz de dimensão 10000 x 3616 é salva em uma variável (vci neste caso) (verificamos com a função dim). Também é revisada a classe e o modo da variável, que deve ser uma matriz com 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) Após a inspeção dos dados, dois vetores são criados: um de longitude e latitude que correspondem às dimensões “LARGURA” e “ALTURA”. Como não são georreferenciados, temos que criá-los manualmente usando as informações extraídas na etapa 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) As coordenadas longitudinais e latitudinais de todos os pixels que compõem o banco de dados são dimensionadas de acordo com os limites que marcam um retângulo que contém o shapefile de interesse. Desta forma, a matriz contendo as coordenadas das quais as coordenadas contidas no shapefile de interesse serão extraídas não será excessivamente grande. Se usam como limites longitudinais e latitudinais aqueles observados na 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

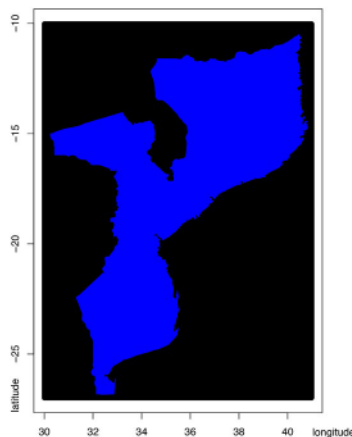
lat_a<-lat[lat<=lat2]
lat_b<-data.frame(lat_a[lat_a>=lat1])#Latitudes are cut off within the boundaries

```

- 9) Utilizando os vetores longitude e latitude já delimitados aos limites do shapefile (lon_b e lat_b), é criada uma matriz com os pares ordenados correspondendo a todas as coordenadas contidas nos limites propostos. Esta matriz é representada graficamente e sobreposta no shapefile de interesse para que seja possível verificar visualmente se as coordenadas criadas cobrem todo o território de interesse.

```
#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) O próximo passo é extrair as coordenadas que estão contidas no shapefile. Se pode usar um GIS ou R diretamente, que é como se realiza aqui. Primeiramente, a matriz de coordenadas é transformada em um arquivo de classe espacial (pontos com informação geográfica) onde a projeção deve ser definida, neste caso se usa uma projeção geográfica WGS84. Em seguida, são extraídas as coordenadas contidas no shapefile de interesse, neste caso “mz”. Por fim, o arquivo é transformado novamente em uma matriz com dados numéricos (sem informação geográfica), a dimensão deste arquivo é revisada, é representado um gráfico para ver se a operação foi realizada corretamente e se guarda a matriz numa pasta conhecida.

```
#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 as variáveis a extrair dependem das dimensões “WIDTH” e “HEIGHT”, e estas variam de 1 a 1000 e 1 a 3616, respectivamente, procura-se a posição a que corresponde cada coordenada contida no shapefile. Para esta pesquisa, são arredondados os números decimais das coordenadas da matriz alvo (contidas em mz) e as dos vetores que contêm todas as longitudes e latitudes possíveis. O resultado será uma matriz com um número de linhas igual ao número de coordenadas contidas em mz e 2 colunas correspondentes às posições de “WIDTH” e “HEIGHT”, chamadas “lo1”.

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

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

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

- 12) Um loop é iniciado com um número de etapas igual ao número de arquivos semanais dos quais queremos extrair as informações. Cada etapa seguirá a ordem do vetor que contém os nomes dos arquivos, portanto, deve-se tomar cuidado para digitar corretamente os nomes das pastas onde os arquivos .nc originais estão armazenados. Os resultados são 3 matrizes, correspondendo aos valores VCI, TCI e VHI de todas as células contidas no shapefile. Sua dimensão será o número de linhas igual ao número de semanas com dados e o número de colunas ao número de pixels. Deve-se observar que a ordem seguida pelas coordenadas da esquerda para a direita é a mesma que a ordem de cima para baixo das coordenadas obtidas na etapa 10.

Essa etapa pode levar várias horas, dependendo do tamanho da região de interesse.

```
#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(lo1)){#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[lo1[j,1],lo1[j,2]])
      tcil<- data.frame(year,month,week,tci[lo1[j,1],lo1[j,2]])
      vhil<- data.frame(year,month,week,vhi[lo1[j,1],lo1[j,2]])
    }
    if (j!=1){
      vcil<- cbind(vcil,vci[lo1[j,1],lo1[j,2]])#The vector contains the VCI data of the corresponding week for each cell
      tcil<- cbind(tcil,tci[lo1[j,1],lo1[j,2]])#The vector contains the TCI data of the corresponding week for each cell
      vhil<- cbind(vhil,vhi[lo1[j,1],lo1[j,2]])#The vector contains the VHI data of the corresponding week for each cell
    }
    print(paste0(j,"de",nrow(lo1)))#Step counter
  }
  colnames(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) As matrizes criadas são salvas em arquivos no formato .RDS. Este tipo de arquivo é recomendado em vez de .txt, pois comprime as informações resultando em variáveis que são mais fáceis e rápidas de ler no RStudio. Você pode recarregar as matrizes, observar a composição superior esquerda dos dados e verificar se o armazenamento de dados está correto:

```
#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) A próxima etapa será agregar os dados semanais em dados mensais usando um loop. Primeiramente são lidos os dados semanais e se substitui quaisquer valores estranhos que possam ter sido gerados por NA. Um NA é um valor não medido, mas com o qual pode ser operado, enquanto outros, como NANs, representam um erro com o qual os cálculos não podem ser realizados.

```
#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
```

O código filtra as informações de cada pixel associado às semanas de cada mes (12 meses), faz a média e guarda numa nova matriz. Esta nova matriz terá dimensões de número de meses por número de pixels. Esta nova variável se guarda como um arquivo .RDS ou .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,auxil2[,1]==i & auxil2[,2]==j)#An auxiliary matrix is created with the weekly VCI
    #values for the year and month according to the loops
    fir3<-filter(auxil3,auxil3[,1]==i & auxil3[,2]==j)#An auxiliary matrix is created with the weekly TCI
    #values for the year and month according to the loops
    fir4<-filter(auxil4,auxil4[,1]==i & auxil4[,2]==j)#An auxiliary matrix is created with the weekly VHI
    #values for the year and month according to the loops
    if (j==1){
      gir2<-data.frame(year=i,month=j,vci=t(data.frame(colMeans(fir2,na.rm = T))))
      gir3<-data.frame(year=i,month=j,tci=t(data.frame(colMeans(fir3,na.rm = T))))
      gir4<-data.frame(year=i,month=j,vhi=t(data.frame(colMeans(fir4,na.rm = T))))
    }
    if (j!=1){
      gir2<-rbind(gir2,data.frame(year=i,month=j,vci=t(data.frame(colMeans(fir2,na.rm = T)))))#CA monthly VCI matrix is created
    }
    for each year of the loop
    gir3<-rbind(gir3,data.frame(year=i,month=j,tci=t(data.frame(colMeans(fir3,na.rm = T)))))#A monthly TCI matrix is created
    for each year of the loop
    gir4<-rbind(gir4,data.frame(year=i,month=j,vhi=t(data.frame(colMeans(fir4,na.rm = T)))))#A monthly VHI matrix is created
    for each year of the loop
  }
}
if (i==1981){
  vci_m<-gir2
  tci_m<-gir3
  vhi_m<-gir4
}
if (i!=1981){
  vci_m<-rbind(vci_m,gir2)#Monthly VCI matrices are collected at each step of the loop
  tci_m<-rbind(tci_m,gir3)#Monthly TCI matrices are collected at each step of the loop
  vhi_m<-rbind(vhi_m,gir4)#Monthly VHI matrices are collected at each step of the loop
}
print(paste(i,"de", "2019"))#Loop step counter
}
dim(vhi_m)#dimension of the monthly VCI matrix is revised, should be num. of months x num. of cells +2 (year and month)
dim(vci_m)#Check the 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) A última secção mostra como gerar diversos gráficos correspondentes ao estado de uma variável sobre todo o território de interesse, agrupá-los em uma matriz e exportá-los. O VHI mensal é usado como exemplo. A matriz VHI é convertida em um raster (veja o código com as funções utilizadas), e a função “levelplot” é utilizada para gerar um gráfico de alta qualidade. Se repete o processo com 12 linhas consecutivas, correspondentes aos meses de outubro de 1992 a setembro de 1993, época em que se observa o desenvolvimento de uma grande seca em Moçambique.

```
#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

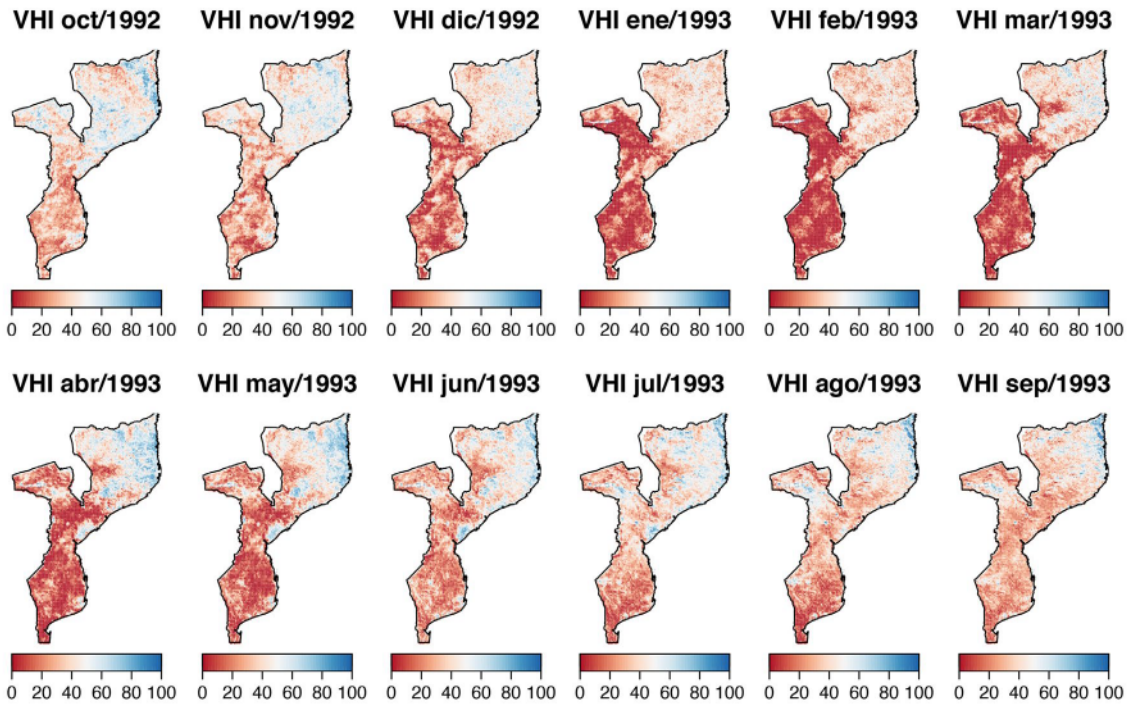
max(dat_grafico[,3],na.rm = T)#Maximum value, 3 is the column where the data of interest is located
min(dat_grafico[,3],na.rm = T)#Minimum value, 3 is the column where the data of interest is located

#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
```

Por fim, para imprimir os 12 gráficos gerados, é utilizada a função “grid.arrange”, na qual é necessário definir a forma de ordenação dos gráficos. Aqui é definida uma matriz 2 x 6. Este arranjo deve se definir com base no tamanho do gráfico deseado. O código necessário para realizar esta etapa e o resultado final é:

```
#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

Nesta secção descarregamos dados de precipitação, ETr e ETP da base de dados TerraClimate sobre a bacia do rio Licungo. Utilizaremos uma estratégia mista entre as duas aplicadas anteriormente. As séries de dados são mensais. Embora apenas estas três variáveis sejam mostradas como exemplo, o processo será idêntico para todos os produtos oferecidos por esta base de dados.

- 1) Os arquivos Terraclimate serão baixados do site: https://climate.northwestknowledge.net/TERRACLIMATE/index_directDownloads.php. Em que as seguintes informações podem ser observadas:

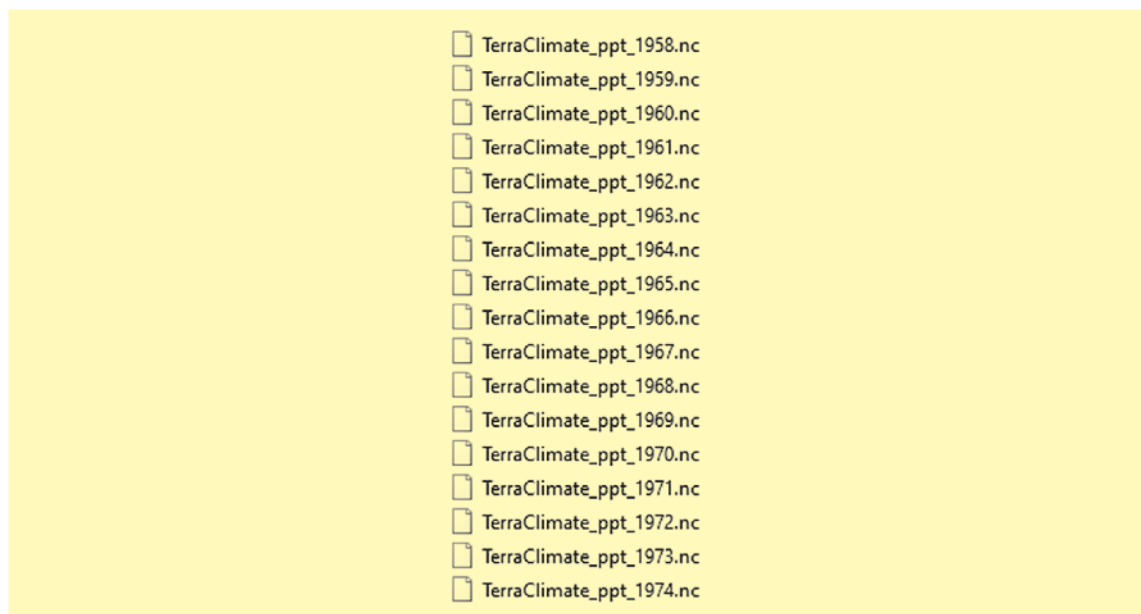
Filter the directory listings:

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)

A variável de interesse é escolhida e os arquivos anuais requeridos são baixados um a um. Aqui está a extração de dados de 1958 a 2019. Os arquivos se guardam numa pasta de destino. Os nomes dos arquivos variam apenas no número do ano a que correspondem.



- 2) Os pacotes necessários são instalados e carregados.
- 3) O shapefile da bacia do rio Licungo é lido e atribuído com o nome cl.
- 4) Da mesma forma que com as demais fontes de informação, um arquivo é lido e a sua composição é analisada.

- 5) As características do arquivo são mostradas no console aplicando a função “print” à variável “aux”. Como acontece com qualquer arquivo .nc, todas as informações necessárias são analisadas para entender como os arquivos são armazenados e extraí-los corretamente na área de interesse.

A primeira seção mostra o número de variáveis que o arquivo contém. Nesse caso, ele contém uma única variável chamada “ppt” e que armazena seus dados numa matriz tridimensional baseada em longitude, latitude e tempo. A unidade de cada dado é mm e o intervalo de tempo mensal:

```
#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
```

A próxima seção relata as características das dimensões nas quais o arquivo armazena as informações. Nesse caso, ele divide o mundo longitudinalmente em 8640 partes do leste e 4320 partes latitudinalmente do norte. Essas matrizes existem para 12 etapas de tempo (meses) que são nomeadas com um número de dias desde 1º de janeiro de 1900 (ou seja, por exemplo, o número 29 refere-se a 29 de janeiro de 1900). A dimensão “crs” indica a projeção com a qual os dados foram salvos.

```
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
```

Por fim, o arquivo contém diversos atributos relacionados à fonte da informação.

- 6) O seguinte é analisar as variáveis que estão contidas em cada um dos arquivos anuais. As informações de precipitação são extraídas como exemplo. Aqui o arquivo TerraClimate é lido novamente com a particularidade de ler directamente o dado “ppt” analisado previamente. A esta variável examina-se a classe e o modo, no qual se verifica que é o tipo “RasterBrick” e “S4” que corresponde à informação raster. Como é conhecido que a dimensão de cada raster mundial contem 12 passos de tempo, um deles (o primeiro do script) é extraído para analisar sua composição. Ao extrair esta informação se verifica a sua dimensão, que deve ser 2 dimensões com valores iguais ao numero de divisões longitudinais e latitudinais em que o banco de dados oferece seus productos.

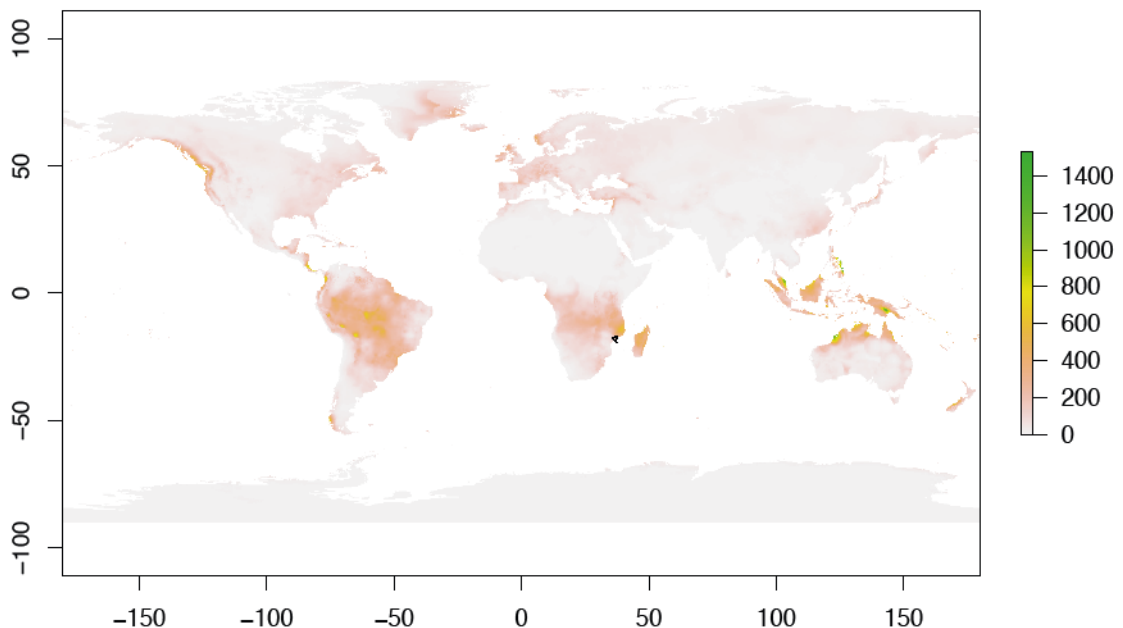
```
#A variable is extracted to analyse its composition

# 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(c1, add=T)#Plot the raster and check if the shapefile of interest (c1) is located where it is supposed to be
```

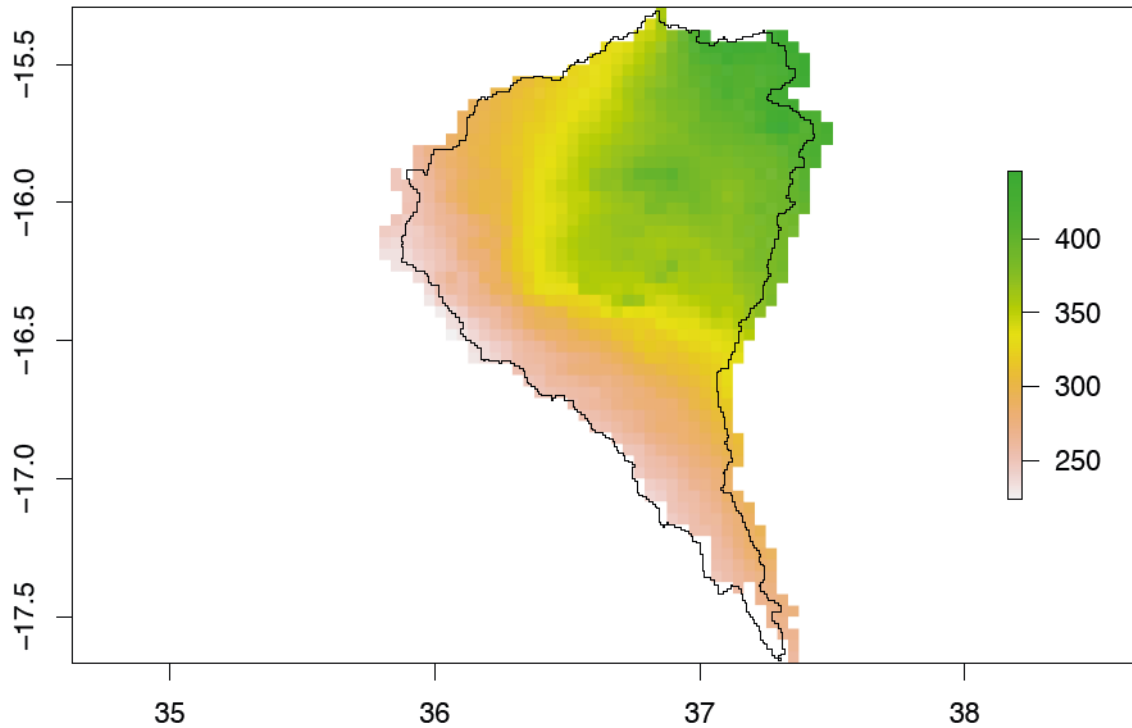
Para verificar a localização das informações e do shapefile de interesse, o raster “precip_1” é representado graficamente com o shapefile da área de interesse (bacia do rio Licungo). No visualizador, se obtém algo assim:



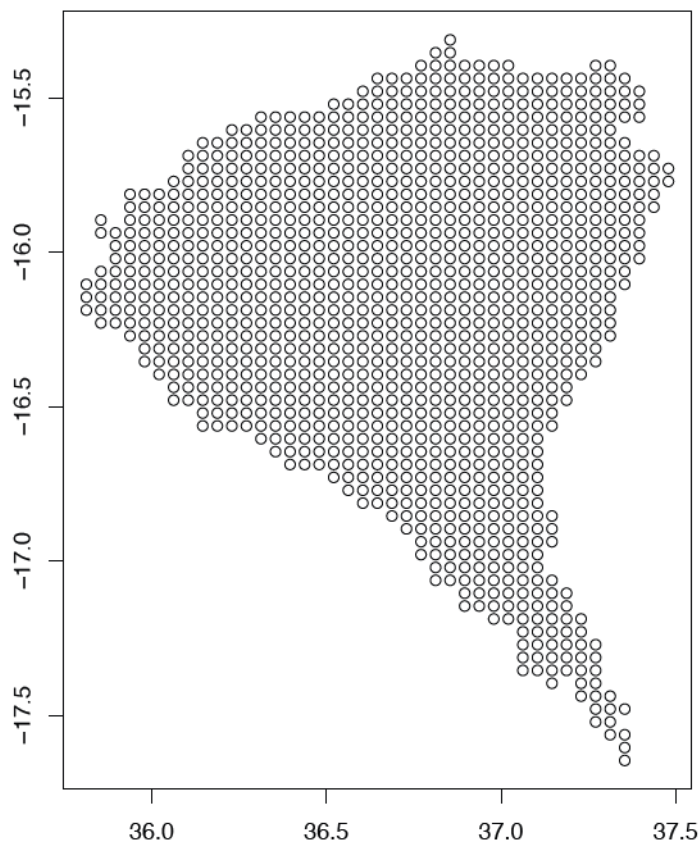
- 7) Uma vez verificada que a informação raster está correctamente georeferenciada, se procede à extração apenas dos dados dentro da área de interesse, seguindo os passos semelhantes aos usados na extração de dados do CHIRPS. O procedimento é feito em três etapas:

```
#Precipitation information is extracted within the shapefile of interest (c1)
step1 <- crop(precip_1, c1) #Intersect the raster with the area of interest
step2 <- rasterize(c1, step1) #Rasterize the shapefile
final <- step1*step2 #Create final product

plot(final)#Plot the raster of interest
plot(c1, add=TRUE)
```



- 8) As coordenadas de cada pixel são extraídas com informações de precipitação dentro do shapefile da bacia do rio Licungo. Seguindo os passos do script se obtém uma matriz com duas colunas de pares ordenados de longitude e latitude. Usando a função `dim` se obtém o número de células contidas na área (1119 pixels), e são representados graficamente para verificar no visualizador se os pares ordenados estão correctos:



Esta matriz (chamada de coordinates) se guarda numa pasta de interesse no formato .txt. Este arquivo é muito importante, pois contém a localização das células contidas no shapefile de interesse e a ordem em que é salvo não deve ser alterada posteriormente (etapa 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 arquivo .nc contém 12 rasters, que contém a informação das variáveis de cada mês, se cria um vetor de datas, anos e meses correspondentes ao período de interesse. Neste caso, se baixam todos os arquivos de 1958 a 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) Um loop é iniciado com um número de passos igual ao número de arquivos anuais que queremos extrair as informações. A variável “k” varia de acordo com o ano correspondente a cada arquivo, então cada etapa seguirá a ordem anual dos arquivos de acordo com seus nomes. Deve-se tomar cuidado para nomear corretamente as pastas onde os arquivos .nc originais estão armazenados. Se repete o processo usado para revisar a composição dos dados em cada arquivo e armazenar as informações variáveis dentro do shapefile de interesse numa matriz. O resultado serão 3 matrizes, correspondendo aos valores de precipitação, ETP e ETr de todas as células contidas no shapefile. A dimensão de cada matriz será de um número de linhas igual ao número de meses com dados e de um número de colunas igual





ao número de pixels. Deve-se notar que a ordem seguida pelas coordenadas de esquerda a direita é a mesma que a ordem das coordenadas de cima para baixo obtidas na etapa 8.

Essa etapa pode levar várias horas, dependendo do tamanho da região de interesse.

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

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

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

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

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

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

if(j==1){
cor_data1<-data.frame(t(data.frame(ext1)[,2]))
cor_data2<-data.frame(t(data.frame(ext2)[,2]))
cor_data3<-data.frame(t(data.frame(ext3)[,2]))
}
if(j!=1){
cor_data1<-rbind(cor_data1,data.frame(t(data.frame(ext1)[,2])))
cor_data2<-rbind(cor_data2,data.frame(t(data.frame(ext2)[,2])))
cor_data3<-rbind(cor_data3,data.frame(t(data.frame(ext3)[,2])))
}
print(paste(j, "de", 12, "en", k))#Contador de pasos
}

aa1<-0#The auxiliary variable is replaced to reduce memory usage
aa2<-0#The auxiliary variable is replaced to reduce memory usage
aa3<-0#The auxiliary variable is replaced to reduce memory usage

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



- 11) Às três matrizes que contêm dados, os 3 vetores de datas, anos e meses são adicionados e guardados no 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")
```

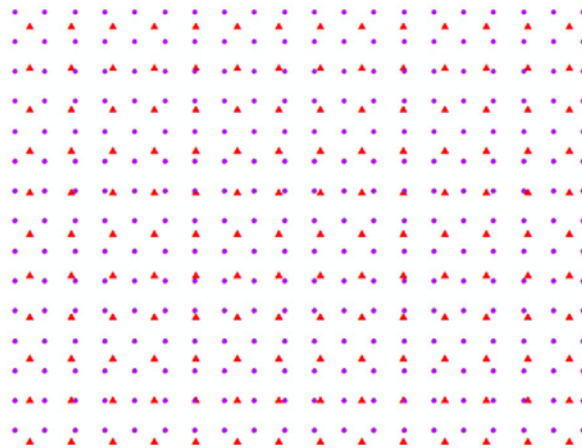
No console é possível ver os resultados da aplicação da função head às três matrizes, onde a composição final das matrizes é visualizada e verificada.

```
> 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. Solução para unificar escalas espaciais

Nesta secção, se explicam soluções simples no caso de duas variáveis precisarem ter exatamente as mesmas coordenadas. Duas opções são analisadas: a primeira será o caso de exigir uma única série temporal representativa para uma área específica; e a outra quando precisamos que todas as variáveis, mesmo que venham de bancos de dados diferentes, tenham o mesmo número de pixels.

Para exemplificar estes casos, usaremos a série de precipitação e o VHI mensal obtido para Moçambique do CHIRPS e NOAA STAR. O gráfico a seguir mostra uma abordagem para a área de 60 x 60 km., aleatorio dentro de Moçambique, onde os pontos azuis indicam as coordenadas NOAA STAR e os triângulos vermelhos indicam as coordenadas CHIRPS.



- 1) As coordenadas dos bancos de dados CHIRPS e NOAA STAR são lidas. Como variáveis de exemplo, também leremos a precipitação e VHI. As dimensões das coordenadas são diferentes entre os dois bancos de dados, pois o tamanho do pixel é diferente.

```
#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) Neste passo, os valores de precipitação para todo o Moçambique são calculados. Seguindo o código, obtém-se uma matriz com o mesmo número de linhas (correspondente ao número de meses), e apenas colunas correspondentes às datas e à precipitação média de todo o país. Se você deseja calcular a média de várias subsecções do país, deve filtrar as coordenadas que lhe interessam e usar a mesma função. Este trabalho não é mostrado aqui, mas pode ser feito usando várias ferramentas mostradas no 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 desta seção, será explicado um artifício para obter dados de uma variável nas coordenadas de outro banco de dados. Existem muitas técnicas que podem ser utilizadas, como interpolação linear ou bilinear. No entanto, quando o tamanho das células é relativamente pequeno em comparação com a área de interesse, esse processo não gera uma precisão significativa, mas sim demanda muito tempo e carga computacional. Aqui a informação (VHI) da coordenada mais próxima daquela de interesse (CHIRPS) será pesquisada e esse vetor de informação será atribuído à coordenada CHIRPS. A título de exemplo, defina primeiro uma matriz de coordenadas alvo, neste caso serão as do CHIRPS e a matriz com as coordenadas iniciais, que serão as do NOAA STAR. A distância mínima “d” da primeira coordenada inicial a todas as coordenadas do alvo é calculada, então a posição da menor distância será a posição da coordenada inicial da qual a informação será extraída.

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

O alvo e as coordenadas iniciais aparecem no console. Pode ver que eles estão muito próximos em latitude e longitude.

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

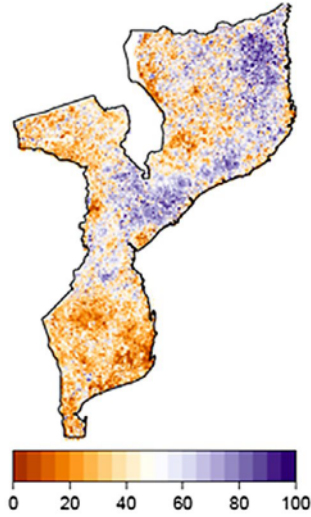
- 4) Este processo é repetido para todas as coordenadas iniciais (VHI) com um loop variando da primeira à última coordenada. Em cada etapa, o vetor VHI pertencente à coordenada mais próxima é extraído e armazenado numa nova matriz. Esta matriz (denominada “di1”) deve ter o mesmo número inicial de linhas que os dados VHI, mas terá o mesmo número de colunas que o número de coordenadas CHIRPS. A ordem das coordenadas CHIRPS de cima para baixo indica a ordem dos vetores VHI da esquerda para a direita. Finalmente, a matriz se guarda numa pasta conhecida no formato .RDS.

```
#The nearest coordinates of the whole CHIRPS coordinate matrix are searched
objetive<-data.frame(coor_chi[,1:2])#Target coordinates are defined
initials<-coor_vhi#The coordinates to be searched for are defined
for(i in 1:nrow(coor_chi)){
  d<-data.frame(which(data.frame(pointDistance(objetive[i,],initials,lonlat = T))
    ==min(data.frame(pointDistance(objetive[i,],initials,lonlat = T))))))
  if(i==1){
    dil<-data.frame(vhi[,1:2],data.frame(vhi[, -c(1:2)])[i,d[1,1]])
  }
  if(i!=1){
    dil<-cbind(dil,data.frame(data.frame(vhi[, -c(1:2)])[i,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])

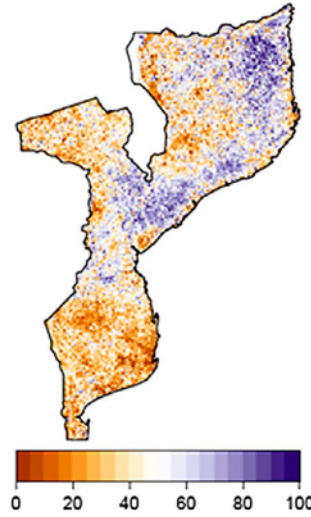
#The new VHI data matrix is saved but corresponding to the CHIRPS coordinates
saveRDS(dil,"E:/19_MANUAL_data_global/change_coordinates/vhi_coordinates_chirps.txt")
```

- 5) Para comparar a nova versão dos dados do VHI com o original, se realizam graficos a partir das duas bases de dados para uma data aleatória (set/2019 neste caso). Se seguiu o processo para representar graficamente os rasters no espaço seguindo o código. O resultado será:

VHI sep/2019 coor NOAA STAR



VHI sep/2019 coor CHIRPS



Se pode ver como a resolução original dos dados VHI é mais precisa do que a do CHIRPS. No entanto, a distribuição espacial da informação não perdeu detalhe significativo.

III. Aplicações à gestão de secas

14. Aplicações do balanço hídrico

Esta seção tem como objectivo calcular e representar em mapas os meses de seca (eq. 3), o déficit hídrico (eq. 4) e o volume de água disponível (eq. 5) na bacia do rio Licungo. Usaremos as variáveis de precipitação, ETr e ETP extraídas do TerraClimate. Três aplicações diferentes serão mostradas que podem ser usadas de outras maneiras conforme as necessidades dos usuarios.

- 1) Os pacotes necessários são instalados e carregados.
- 2) O shapefile da bacia do rio Licungo é lido, foi-lhe atribuído o nome de cl.
- 3) O arquivo é lido com as coordenadas do Terraclimate que estão dentro da bacia do rio 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) São lidos os arquivos necessários para os cálculos, são as matrizes de precipitação, ETP e ETr. Suas dimensões e composição são revistas.

```
#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) Os meses com déficit ou superhabit em cada uma das células são calculados subtraindo a matriz ETP da matriz de precipitação. Essa variável foi denominada “seca”.

```
#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) O volume de água disponível por célula é calculado subtraindo a matriz ETr da matriz de precipitação. Essa variável foi denominada “disponível”.

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

- 7) O déficit de água por célula é calculado subtraindo a matriz ETr da matriz ETP. Essa variável foi denominada “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) Às 3 matrizes resultantes, se revisam as dimensões que devem coincidir com a precipitação original, ETP e ETr, e estas são guardadas como arquivos .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) Com a matriz de meses secos, calcula-se a porcentagem de tempo (número de meses dentro do período de análise) que apresentou meses secos. As coordenadas correspondentes são adicionadas a este vetor e a matriz se guarda numa pasta conhecida. Você deve sempre verificar a dimensão e composição da variável antes.

```
#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) Para a matriz de água disponível, os dados correspondentes a cada mês (janeiro, fevereiro, etc.) são filtrados e todos os valores são calculados em média (de 62 anos neste caso). Os 12 vetores são somados, um de cada mês, com suas respectivas coordenadas e as matrizes se guardam numa pasta conhecida. Deve se verificar sempre a dimensão e composição da variável antes de salvar.

```
#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 matriz de déficit é transformada de mm/mês para km³. Para este procedimento, é calculada uma área por célula, utilizando o tamanho da célula a que corresponde cada pixel, no caso do TerraClimate é de $\approx 21.43 \text{ km}^2$. Posteriormente, a média de todas as células é calculada e se obtém um vetor em km³ do déficit médio em cada uma das etapas de tempo mensais. As coordenadas correspondentes são adicionadas a este vetor e a matriz final é guardada. Deve se verificar sempre a dimensão e composição da variável 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 realizar a representação gráfica, se tem o shapefile da bacia do rio Licungo e as matrizes de porcentagem de meses secos, a água disponível e o déficit em km³.

```

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

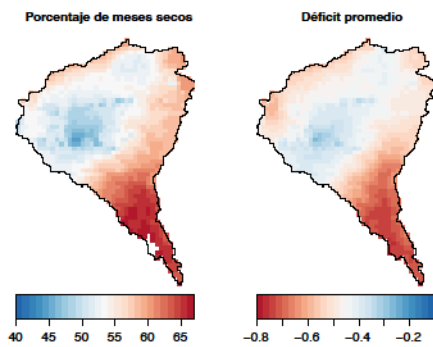
#The percentage of dry months in the Licungo river basin is read as follows
dry_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_dry.txt", header=T)
dim(dry_percentage2)
head(dry_percentage2)

#The average volume per month in the Licungo river basin is read as follows
vol_promedio2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/volumen_promedio_months.txt", header=T)
dim(vol_promedio2)
head(vol_promedio2)

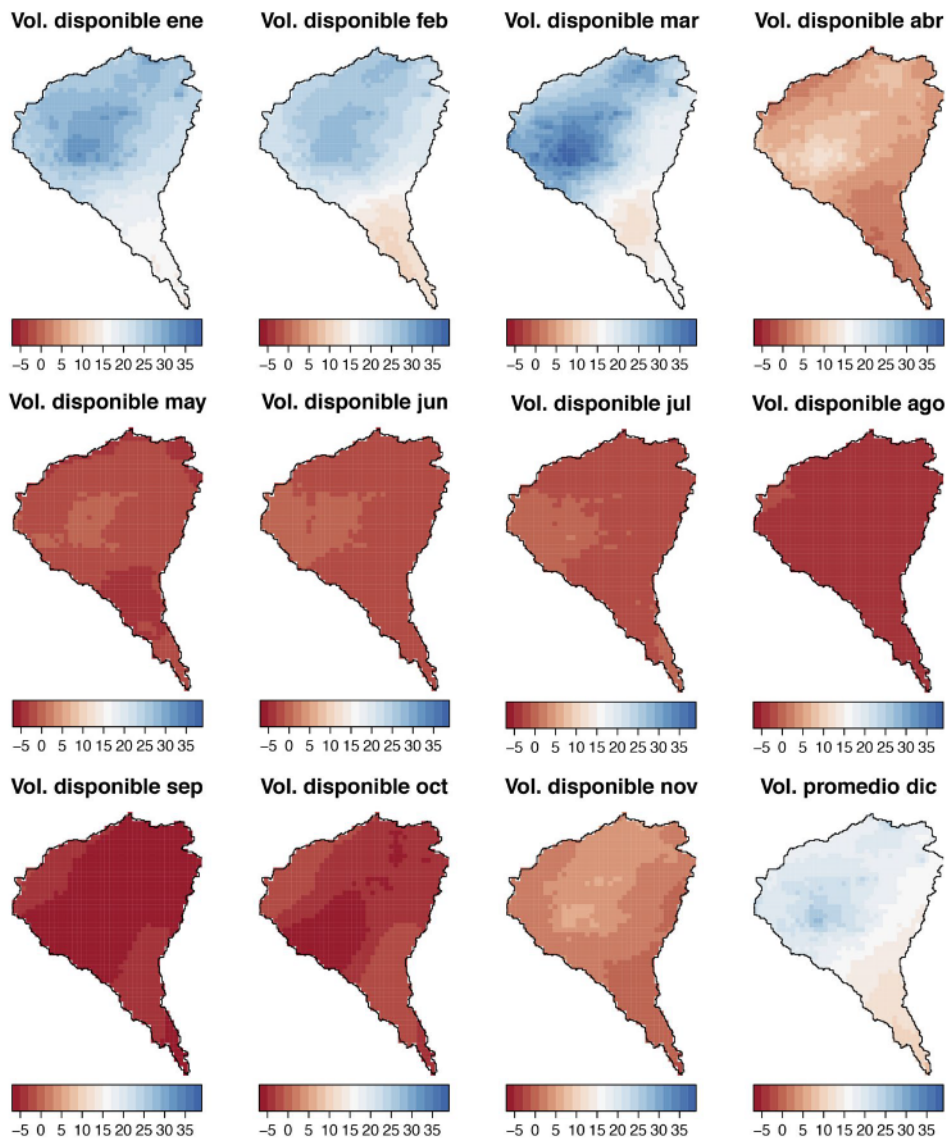
#The percentage of deficit in km3 in the Licungo river basin is read as follows
stress_percentage2<-read.table("E:/19_MANUAL_data_global/Balance_hidrico/percentage_months_stress.txt", header=T)
dim(stress_percentage2)
head(stress_percentage2)

```

13) As etapas de representação gráfica usadas anteriormente (exemplo da etapa 16 na seção 10) são seguidas para representar graficamente as matrizes de porcentagem de meses secos e de déficit em km^3 . O resultado será mais ou menos assim:



14) Seguindo os passos dos gráficos definidos para o VHI (exemplo da etapa 15 da seção 11), são traçados os volumes médios mensais em km^3 de cada mês na bacia do rio Licungo. O resultado será mais ou menos assim:



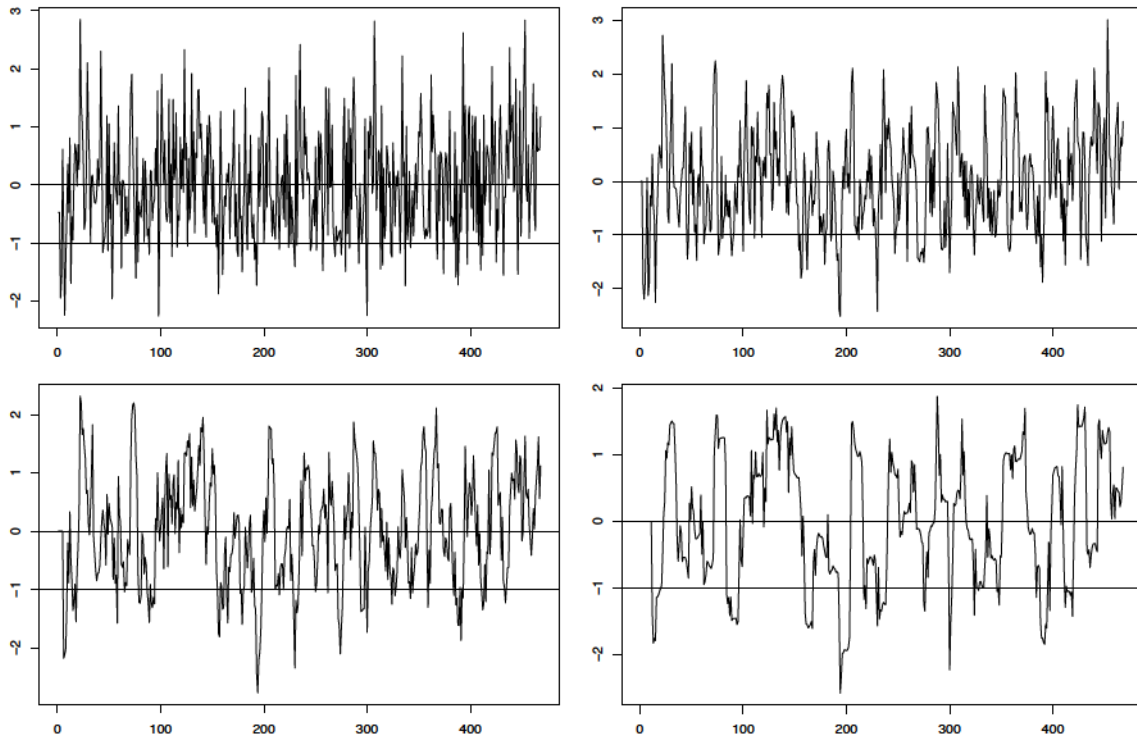
15. Cálculo e aplicação dos índices de seca do SPI e SPEI

Nesta secção, se explica passo a passo o cálculo dos índices de seca SPI e SPEI. Os dados de precipitação CHIRPS obtidos em Moçambique serão usados para o SPI e os dados de precipitação menos ETP (meses secos e úmidos) do TerraClimate para a bacia do rio Licungo para o SPEI. Primeiramente, é mostrado o processo para realizar o cálculo com uma única série temporal em cada caso e, posteriormente, para toda a matriz de pixels contida nas áreas de interesse. Esta secção termina mostrando duas aplicações comuns e muito úteis de índices. A primeira é o cálculo da percentagem da área afetada pelas secas em cada etapa de tempo e a segunda são gráficos espaciais onde os etapas da seca/sem seca para un período específico podem ser mostrados.

- 1) Os pacotes necessários são instalados e carregados. O pacote principal será o “SPEI”, no qual foram programadas as funções necessárias para o cálculo direto do SPI e do SPEI.
- 2) Os shapefiles da bacia do rio Licungo e Moçambique são lidos, foram atribuídos com o nome cl e mz respectivamente.
- 3) As coordenadas dos dados CHIRPS localizados em Moçambique e os dados Terraclimate pertencentes à bacia do rio Licungo são lidos.
- 4) Para um vetor aleatório da matriz de precipitação CHIRPS, entendido como uma série temporal mensal num ponto dentro de Moçambique, o SPI é calculado. Primeiro é definido um vetor que contém as escalas de tempo que queremos calcular, aqui as escalas de 1, 3, 6 e 12 meses foram propostas como exemplo. O resultado será um vetor do SPI-1, -3, -6 e -12 meses correspondente ao vetor de precipitação original. Se realiza um gráfico dos 4 vetores para comparar a diferença, além disso, o limite de 0 e -1 (seca moderada) é delimitado.

```
#SPI is calculated for a random cell (35)
vectorSPISPEI<-matrix(c(1,3,6,12)) #Vector with the required time scales to be calculated
SPI <- matrix(nrow=nrow(pre_chi),ncol = nrow(vectorSPISPEI))#Empty matrix to be filled with SPIs
for (j in vectorSPISPEI){ #4 SPIs will be calculated depending on the predefined vector
  spii<- spi(pre_chi[,35],j)#The function "spi" is applied and the input variables are a precipitation
  #vector and the time scale at which it is required to be calculated
  tr <- as.numeric(fitted(spii))#The spi data is extracted from the class spi
  tr[is.na(tr)] <- 0#NAs that can be generated are changed to zeros
  SPI[,which(vectorSPISPEI==j)] <- tr #The spi vector is stored in the SPI array
  colnames(SPI)[vectorSPISPEI==j] <- j #Each column is named with the number of the time scale
}
#SPI is reviewed results
head(SPI)
dim(SPI)

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



- 5) Para um vetor aleatório da matriz de meses secos e úmidos, entendido como uma série temporal mensal num ponto da bacia do rio Licungo, é calculado o SPEI. Quanto ao SPI, primeiro é definido um vetor que contém as escalas de tempo a serem calculadas, aqui as escalas de 1, 3, 6 e 12 meses foram propostas como exemplo. O resultado será um vetor dos SPEI a escalas de -1, -3, -6 e -12 meses correspondente ao vetor do déficit D original. Se realiza um gráfico (não mostrado) dos 4 vetores para comparar a diferença, além disso, o limite de 0 e -1 (seca moderada) é delimitado.

```
#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[,1:5])#The composition of the matrices is reviewed
head(SPI3[,1:5])
head(SPI6[,1:5])
head(SPI12[,1:5])

#The matrices corresponding to the SPIs are stored in a folder known as .RDS
saveRDS(SPI1,"E:/19_MANUAL_data_global/SPI_SPEI/SPI1_mz.rds")
saveRDS(SPI3,"E:/19_MANUAL_data_global/SPI_SPEI/SPI3_mz.rds")
saveRDS(SPI6,"E:/19_MANUAL_data_global/SPI_SPEI/SPI6_mz.rds")
saveRDS(SPI12,"E:/19_MANUAL_data_global/SPI_SPEI/SPI12_mz.rds")
```

- 6) Por meio de um loop que varia para todas as colunas da matriz de precipitação (todas as séries temporais de todas as coordenadas), o mesmo procedimento de cálculo do SPI é realizado para cada uma delas. As matrizes resultantes ao seguir o código serão 4 (neste caso), uma para cada escala de tempo em que o índice foi calculado. Se o código for seguido, as matrizes serão geradas num tempo relativamente curto (variável ao número de coordenadas) que devem ser da mesma dimensão da matriz de precipitação CHIRPS original. As matrizes serão guardadas como arquivos .RDS em pastas conhecidas.

```
#SPI 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 SPIs are stored in a folder known as .RDS
saveRDS(SPEI1,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI1_cl.rds")
saveRDS(SPEI3,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI3_cl.rds")
saveRDS(SPEI6,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI6_cl.rds")
saveRDS(SPEI12,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI12_cl.rds")
```

- 7) Por meio de um loop que varia para todas as colunas da matriz de déficits de meses secos e úmidos, (D), o mesmo procedimento de cálculo do SPEI é realizado para cada uma delas. As matrizes resultantes ao seguir o código serão 4 (neste caso), uma para cada escala de tempo em que o índice foi calculado. Se o código for seguido, matrizes que devem ter a mesma dimensão da matriz de déficit serão geradas num tempo relativamente curto (variável ao número de coordenadas). As matrizes serão guardadas como arquivos .RDS em pastas conhecidas.

```
#SPI 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 SPIs are stored in a folder known as .RDS
saveRDS(SPEI1,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI1_cl.rds")
saveRDS(SPEI3,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI3_cl.rds")
saveRDS(SPEI6,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI6_cl.rds")
saveRDS(SPEI12,"E:/19_MANUAL_data_global/SPI_SPEI/SPEI12_cl.rds")
```

- 8) Neste passo se calcula a área da porcentagem de área afetada por secas. Isso serve para quantificar para cada etapa de tempo (a cada mês) o número de pixels que estão em seca e calcular sua porcentagem em relação ao total. Um limite é necessário de acordo com a intensidade da seca a ser estudada. No exemplo, usamos como limites os valores de seca moderada (valores menores que -1), seca severa (valores menores que -1,5) e seca severa (valores menores que -2). A composição final da matriz guardada é mostrada aqui. O SPI numa escala de tempo de 6 meses é usado como exemplo para Moçambique.

```
#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[,-(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[,-(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[,-(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      0.000000  0.000000  0.000000
2 1981-01-01 1981 2      0.000000  0.000000  0.000000
3 1981-01-01 1981 3      0.000000  0.000000  0.000000
4 1981-01-01 1981 4      0.000000  0.000000  0.000000
5 1981-01-01 1981 5      0.000000  0.000000  0.000000
6 1981-01-01 1981 6     31.736644  23.031122  18.970522
```

- 9) As porcentagens são traçadas como séries temporais seguindo o código do ANEXO III. Nesse caso, se usa a função “ggplot”. É necessário ler as porcentagens da área atingida por secas, aqui definidos como PAA. Se for necessária mostrar as series temporais não sombreada se utiliza o passo 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)

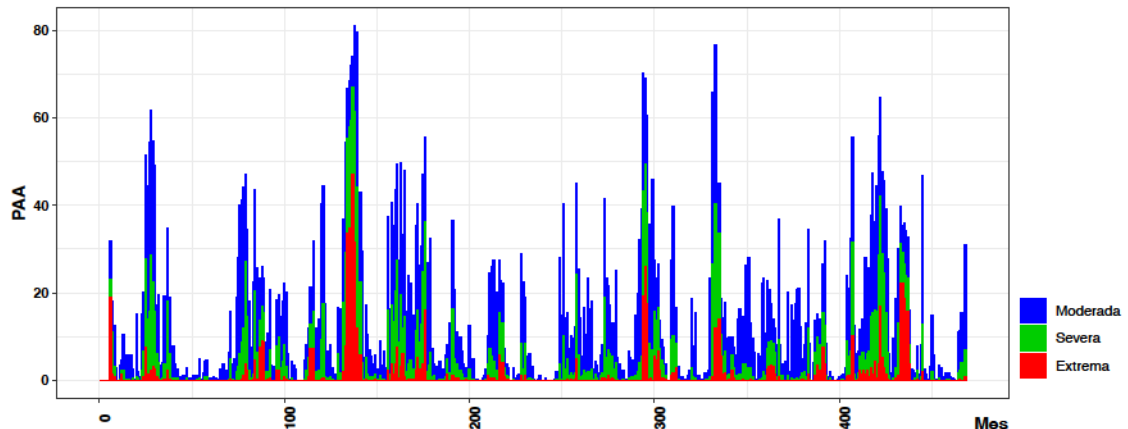
#Drought percentages are read for the different intensities
percentages<-read.table("E:/19_MANUAL_data_global/SPI_SPEI/percentages.txt",header=T)
dim(percentages)
head(percentages)

#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
```


O gráfico de resultado será algo assim:



- 10) Nesta etapa se calcula o estado de seca ou não seca em Moçambique. Isso transforma todos os valores abaixo de um limite de seca em 1 e aqueles acima do limite em 0. O SPI é usado como exemplo numa escala de tempo de 6 meses em Moçambique e para uma intensidade de seca moderada, ou seja, um limite de valores SPI menores que -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 siguem os passos dos gráficos usados anteriormente (exemplo da etapa 16 na seção 10) adaptados para representar graficamente os estados de seca ou não seca. O objetivo é mostrar as áreas onde ocorrem secas há vários meses. O período de outubro de 1992 a setembro de 1993 é usado como exemplo, quando houve uma seca significativa em Moçambique. O “ggplot” é usado como uma função para criar gráficos seguindo o código, onde títulos, cores e outras características podem ser modificados. Ao contrário das funções que têm sido utilizadas para outros gráficos (como o uso da função “levelplot”), aqui é necessário transformar a informação geográfica (shapefile mz) numa matriz de dados compatível com o método “ggplot”; isso é feito usando a função “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)

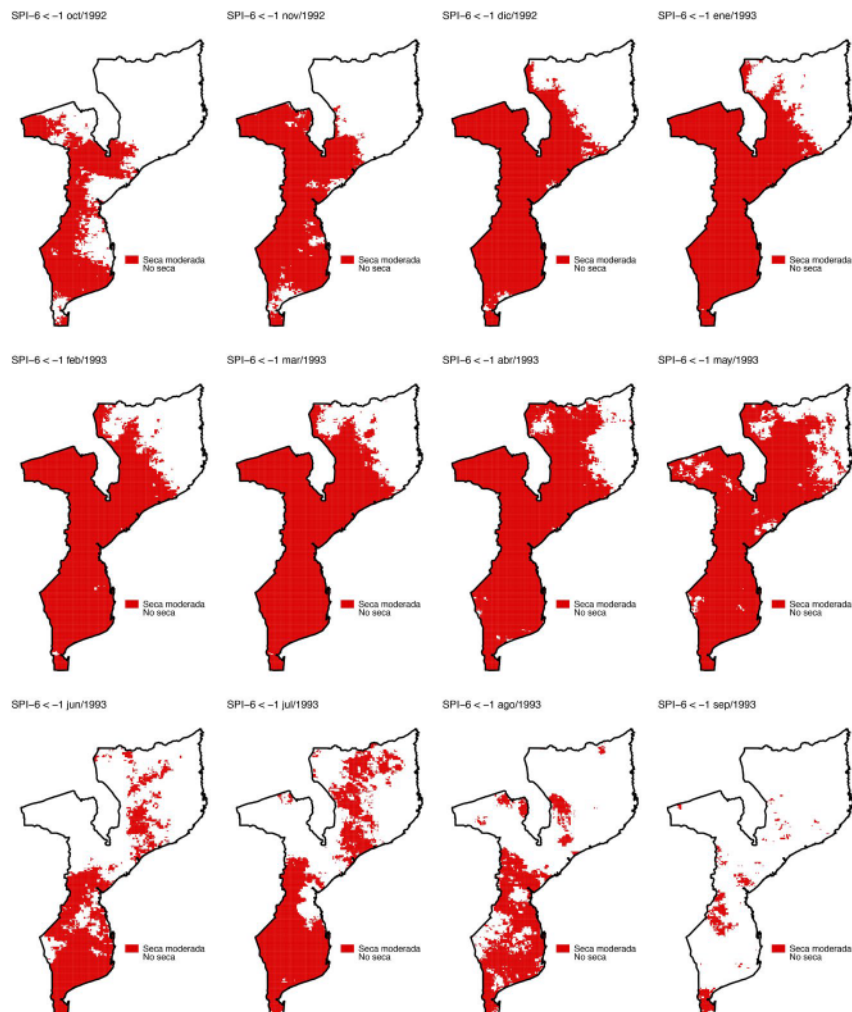
#The matrix is read with information of zeros (drought) and 1 (no drought)
spi6_b<-readRDS("E:/19_MANUAL_data_global/SPI_SPEI/seq_no_seq.rds")
dim(spi6_b)
head(spi6_b[,1:5])

#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()

```

Um gráfico é criado para cada mês e guardado como pdf (ou outro tipo de arquivo). O resultado será mais ou menos assim:



16. Caracterização de secas usando o SPI

Esta secção estabelece um procedimento para calcular as características médias das secas. Essas características são a duração, intensidade máxima e gravidade; calculando também o número de eventos ocorridos e a probabilidade de haver um mês de seca. Os eventos de seca são definidos quando pelo menos dois meses consecutivos encontram-se abaixo do limite de intensidade definido. O procedimento é realizado para uma série temporal SPI (SPI-6 em Moçambique como exemplo) da qual são extraídas as médias de cada característica dentro da série temporal. Os limites podem ser alterados e o código apresentado é usado para qualquer outro índice de seca que seja padronizado (média 0 e desvio padrão 1).

- 1) Os pacotes necessários são instalados e carregados.
- 2) É lida a matriz contendo as séries temporais do indicador a ser caracterizado. Neste caso, o SPI-6 de Moçambique foi calculado com a precipitação 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) As coordenadas dos dados CHIRPS localizados em Moçambique são lidas.

```
#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) O loop que calcula as características de todos os SPIs na matriz de destino tem muito poucas variáveis para modificar, se necessário. O limite é definido antes do lançamento do loop, é importante ter em conta que o loop apenas identifica as características de épocas secas não úmidas, pelo que os valores devem ser negativos. O início do loop segue a ordem da esquerda para a direita, onde a primeira coluna é o número a partir do qual varia a variável auxiliar "l", este valor deve ser modificado se a matriz de destino começar noutra coluna. Da mesma forma, o número inicial na parte final do código deve ser alterado seguindo as instruções mostradas nele. Se for preciso extrair as características duma única série temporal ou um número limitado, também se deve alterar o "ncol (spi6)" que indica o número de colunas da matriz chamada "spi6" para o da posição da última coluna da qual deseja obter as 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 o loop no código entregue nos 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

```

A matriz “dat” é o resultado do loop e contém as características médias de todos os eventos de seca para cada série temporal do SPI. Sua dimensão possui um número de linhas igual ao número de coordenadas CHIRPS e 5 colunas correspondentes a cada característica:

```

> head(dat[,1:5])
      Duracion Intensidad Severidad numero probabilidad
colMeans.cas1. 7.529412  -1.705905 -7.935980    19    1.608849
colMeans.cas1. 7.470588  -1.754820 -8.184736    18    1.596280
colMeans.cas1. 9.384615  -1.743732 -9.567796    14    2.005260
colMeans.cas1. 7.750000  -1.734323 -8.028418    18    1.655983
colMeans.cas1. 7.687500  -1.735632 -8.060110    18    1.642628
colMeans.cas1. 7.176471  -1.737271 -8.029506    18    1.533434
> |

```

As coordenadas correspondentes são adicionadas a esta matriz e são guardadas numa pasta de destino conhecida.

```

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

```

- 5) Para representar graficamente essas características médias se usa o procedimento semelhante ao usado para realizar os gráficos dos estados de seca ou não seca na seção 1. Lemos a informação geográfica de Moçambique e a matriz que contém as características das secas de acordo com o SPI6 (neste caso) e suas coordenadas. Apenas as durações e probabilidades são representadas num gráfico a modo de exemplo.

```

#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) Extraí-se o vetor da característica a ser representada no gráfico, onde o número dentro dos colchetes é a posição da coluna que contém a informação a ser representada no gráfico.

```

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

```

- 7) Este vetor é transformado em seções que estão dentro de seus limites máximo e mínimo. O número atribuído deve estar dentro dos limites 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 realiza a representação gráfica dos resultantes usando a função “ggplot”. Podem se modificar as cores, o nome da seção (previamente definido), entre outros. É importante definir os valores “x” e “y” como as coordenadas de longitude e latitude contidas na matriz de coordenadas. O gráfico é guardado com um nome para depois exporta-lo.

```
#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) O processo anterior é repetido com a matriz das 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

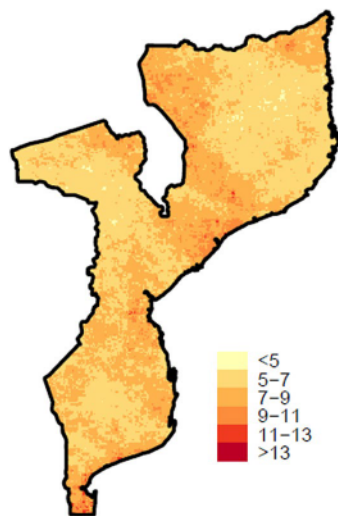
graphic_2<-ggplot(probabilities)+
  geom_tile(aes(x = carac[,1], y = carac[,2], fill=factor(probabilities[,1])))+#Here you change the name of the column
  theme_classic()+
  theme(legend.position = c(0.75, 0.25),legend.box="horizontal",legend.title = element_blank(),
        legend.text=element_text(size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

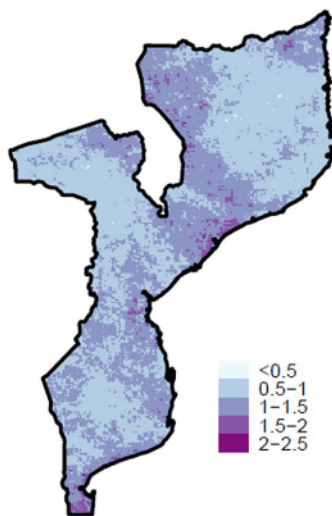
```
#The graphic is exported as pdf, the destination folder and the size can be modified
pdf("E:/19 MANUAL_data_global/Graphics/characteristics.pdf",height = 8,width = 6)
grid.arrange(graphico_1,graphic_2,
              nrow = 1) #The 2 graphs are arranged in a 1 x 2 matrix, can be changed
dev.off()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

O resultado será mais ou menos assim:

SPI-6 < -1 Duración



SPI-6 < -1 Probabilidades



Conclusões

Este documento é um manual introdutório para obter variáveis hidrometeorológicas em qualquer região e aplicá-las na gestão das secas. As secas são um fenômeno hidrológico extremo, que ocorre naturalmente em qualquer regime climático. Seus impactos no meio ambiente causam imensas perdas econômicas e ecológicas todos os anos, razão pela qual sua caracterização e monitoria se tornam cada vez mais importantes. No entanto, a caracterização das secas é um processo complexo, pois além de afetar diferentes partes do ciclo da água, não há uma indicação clara do início ou do fim. Soma-se a isso o problema de que as ferramentas necessárias para modelar o fenômeno, como dados de interesse hidrológico e climático (precipitação, temperatura, etc.), nem sempre atendem à qualidade exigida.

Este trabalho coleta uma série de conceitos de gestão de secas amplamente aceitos pela comunidade científica e técnica que são aplicados por meio de vários scripts (códigos de programação) codificados no Software R. Esses scripts divididos em 3 seções mostram o passo a passo para i) descarregar, extrair, manipular e guardar variáveis hidrometeorológicas de bancos de dados globais; ii) usar e aplicar essas variáveis para a gestão da seca; e, iii) apresentar os resultados por meio de gráficos em diferentes formatos.

Seguindo o passo a passo das ferramentas fornecidas neste manual, muitas informações úteis podem ser geradas para qualquer região do mundo. Os resultados que podem ser obtidos podem ser do interesse de gestores e técnicos, que nem sempre têm fácil acesso a este tipo de informação. Uma grande vantagem deste manual é a flexibilidade de suas ferramentas, podendo ser utilizadas em diferentes locais e permitindo a criação de novos scripts e a geração de novas formas de utilização das informações baixadas pelos usuários. As metodologias distribuídas neste documento podem ser expandidas em trabalhos futuros, tais como a obtenção de dados resultantes de modelos de projecção de mudanças climáticas, análise de tendências, representação de resultados em regiões dum país (cidades ou províncias) ou bacias hidrográficas (sub-bacias), entre outros.

Referências

- 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://tu cows.icm.edu.pl/packages/cran/web/packages/SPEI/SPEI.pdf>. A case study Birkoor Kortigiri Mandals. <https://doi.org/10.1175/2009JCLI2909.1>
- 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., PRODOSCIMI, I., VITOLO, C., SMITH, K., 2019. Using R in hydrology: a review of recent developments and future directions. *Hydrol. Earth Syst. Sci. Discuss.* 1-33. <https://doi.org/10.5194/hess-2019-50>
- 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 DIALY 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)

PROJETO REALIZADO COM A COLABORAÇÃO DE

