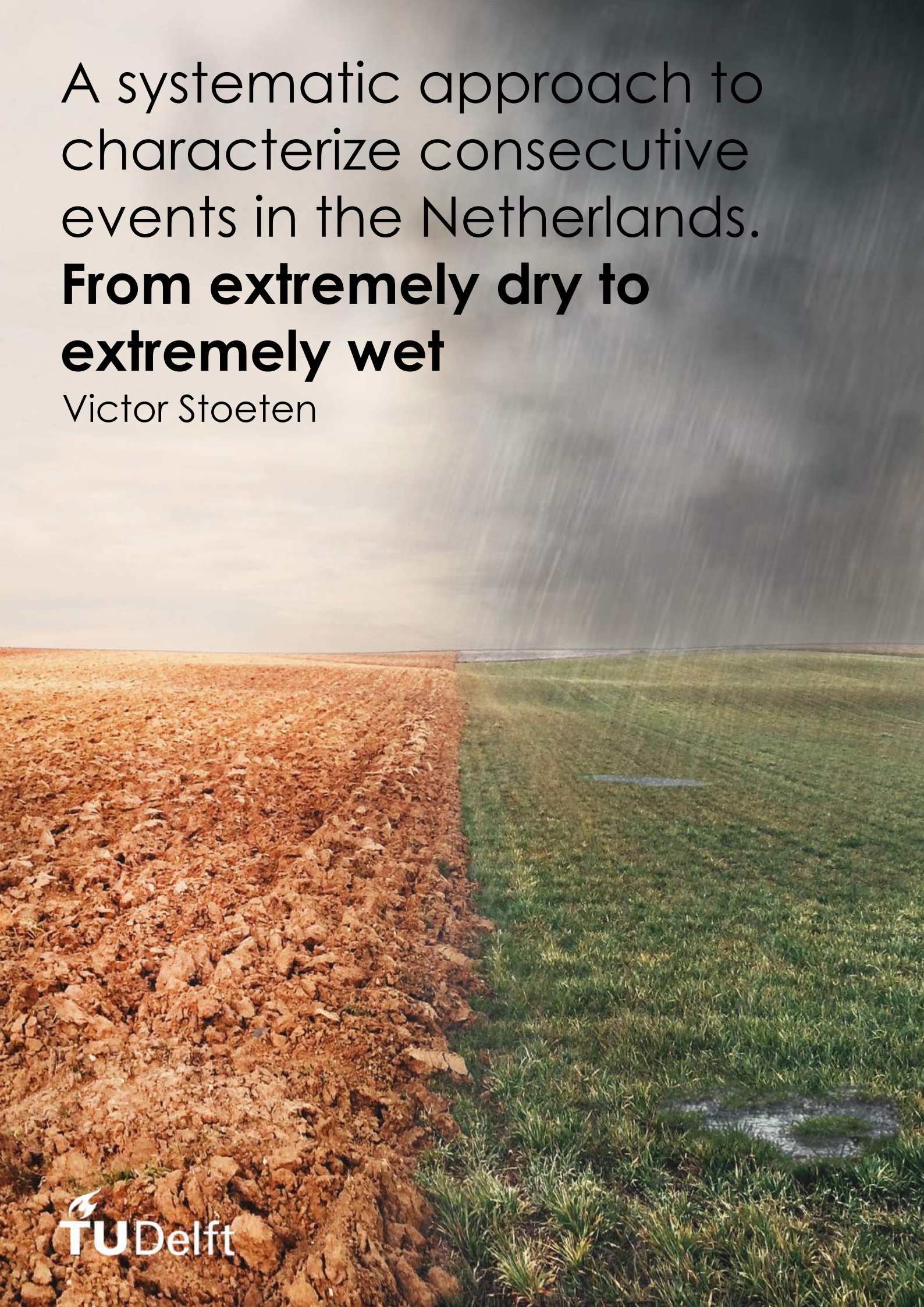
# A systematic approach to characterize consecutive events in the Netherlands. **From extremely dry to extremely wet**

Victor Stoeten



TU Delft

# A systematic approach to characterize consecutive events in the Netherlands. From extremely dry to extremely wet.

by

Victor Stoeten

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on 11-03-2022 at 10:45.


| | | |
|---|---|---|
| Student number: | 4313542 | |
| Project duration: | April 14, 2021 – Month Day, 2022 | |
| Thesis committee: | Dr. ir. E. Ragno, | TU Delft - CiTG (Chair) |
| | Dr. ir. G.F. Nane, | TU Delft - EWI |
| | Dr. ir. O. Morales Nápoles, | TU Delft - CiTG |
| | Dr. ir. M. Hrachowitz, | TU Delft - CiTG |


Photo front page: Adapted from Elisabeth Lies / Unsplash
An electronic version of this thesis is available at http://repository.tudelft.nl/.

# Preface

When I started this thesis, little did I know about this subject. I began to explore and dived further into details. Slowly but surely I lost myself in the multitude of elements the subject lead me to. This experience of loosing myself has been a beautiful experience. With the knowledge acquired over time, I started to see the world through a different lens. I even developed some frequency illusion as I started to see phenomena researched on places other than my laptop screen. For example, I was reading a novel while trying to escape from my thesis. Words and sentences related to the topic struck me. Normally, I would have never noticed them out of the ordinary. My perception was changed. This thesis taught me that consecutive events aren't necessarily about one event following another. Moreover, it is about the possible interactions between those events and recognizing their patterns. This realisation has lead me to approach hydraulic engineering challenges with a more comprehensive mindset. A lesson learned that I will take with me into the future.

I would like to express my sincere gratitude to all the people who were involved in this thesis. Elisa Ragno, chair of this thesis, for introducing me to this subject, your effort and many invested hours. Tina Nane, for sharing your knowledge on Non-parametric Bayesian Networks with great enthusiasm. To Oswaldo Morales Nápoles, for strong gate-keeping on do's and don'ts for this thesis. To Markus Hrachowitz, for your help during the last part of this thesis.

Additionally I would like to express my gratitude to my girlfriend, for her unparalleled love and support past year. To my family, for their support despite difficult times. To my roommates and friends, for keeping an eye on me.

Victor Stoeten
*Delft, March 2022*

# Abstract

Floods and droughts, also known as hydro-hazards, are phenomena that generally involve detrimental consequences to society and environment. Traditional practices for risk assessment consider flood and drought independently. However, they are two opposite extremes of the same hydrological cycle. Omitting their interaction might lead to an under- or overestimation of the current and future risks associated with such natural hazards. In history, a number of drought-flood interactions have been observed in various parts of the world. Research to these drought to flood interactions is still in its infancy. Therefore, this research explores the concept of consecutive dry and wet (CDW) events in the Netherlands. The aim of this research is two sided. First, the Consecutive Events Graph (CEG) is introduced. This is a radar chart type of graph used to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons. These can be used to identify hot-spots prone to opposite extremes. Second, a fully probabilistic framework based on Non-parametric Bayesian Network is developed to model the dependence between dry and wet indicators. Such model can be used to infer expected wet conditions in a given region when dry conditions are known.

For the CEG and probabilistic model a number of settings were introduced to quantify meteorological dry and wet extremes and to couple them spatial-temporally. First, a number of indicators were selected to quantify both type of extremes. The so called dry indicators are the three months Standardized Precipitation Index (SPI-3), maximum potential rainfall deficit (RDx), summer days (SD25), consecutive dry days (CDD) and the heat score (Hscore). The so called wet indicators are the total precipitation (Ptot), precipitation exceeding 20mm per day (P20), simple precipitation intensity index (SDII) and maximum 1-day, 3-day and 5-day precipitation (R1D, R3D and R5D). Second, the dry period is defined in summer (June, July and August) and consecutive wet period in fall (September, October and November). Third, the Netherlands was subdivided in five homogeneous regions such that both wet and dry indicators were characterized on a regional scale. Maximum values of the indicators in its corresponding period were calculated for each single region and for every single year between 1965 to 2020. This resulted in a dataset consisting quantities for dry extremes in summer and wet extremes in fall for 5 unique regions over 56 years. Application to the CEG shows potential to identify and quantify CDW extremes. Region-to-region and year-to-year comparison is possible to quantify changes between years or regions. Application to the NPBNs disclosed limited interdependencies across the dry and wet indicators. Using the NPBNs for precise forecasting of expected wet conditions is deemed unsuitable as of low precision. Making inference of wet indicators based on hypothetical mild to extremely dry indicators revealed multiple trends of those wet indicators. These trends are increasing for short term precipitation indicators (R1D, R3D, and R5D) and simple precipitation intensity index (SDII) and are mildly decreasing for the total precipitation (Ptot).

Extreme dry events, extreme wet events and consecutive occurrences of these events are inevitable. It is expected that these phenomena will occur more frequently and become more severe due to a changing climate. A number of recommendations for future research is proposed. Findings from this thesis will help to smooth the path towards better understanding of the identification, quantification and interaction of CDW events or multi-hazard events in general.

*"Generally speaking, whether something is logical or isn't, what's meaningful about it are the effects. Effects are there for anyone to see, and can have a real influence. But pinpointing the cause that produced the effect isn't easy. It's even harder to show people something concrete that caused it, in a "Look, see?" kind of way. Of course there is a cause somewhere. There can't be an effect without a cause. You can't make an omelet without breaking some eggs. Like falling dominoes, one domino (cause) knocks over the adjacent domino (cause), which then knocks over the domino (cause) next to it. As this sequence continues on and on, you no longer know what was the original cause."*

– Haruki Murakami, *Killing Commendatore*

# Contents

# List of Figures

# List of Tables

# Abbreviations and symbols

The following list describes abbreviations and symbols that are used within this dissertation

## Abbreviations

| | |
|---|---|
| BN | Bayesian network |
| CDW | Consecutive Dry to Wet |
| KNMI | Royal Dutch Meteorological Institute |
| KNMI-A | Automatic weather stations of the KNMI |
| KNMI-M | Manual rain gauge stations of the KNMI |
| LCW | The National Coordination Committee for Water Allocation |
| NDVI | Normalized Difference Vegetation Index |
| NPBN | Non-parametric Bayesian network |
| RIVM | The National Institute for Public Health and the Environment |
| rv | Random variable |
| rvs | Random variables |
| ETC-CDIs | Expert Team on Climate Change Detection and Indices |

## Symbols

| | |
|---|---|
| $C_n$ | Empirical copula |
| $C_{\theta_n}$ | Parametric copula |
| $d_H$ | Heillinger distance |
| $d_{cal}$ | d-calibration score |
| $EV24$ | Daily evapotranspiration [mm/day] |
| $P$ | Daily precipitation [mm/day] |
| $S_n$ | Cramer Von Mises score |
| $TG$ | Daily mean temperature [C°] |
| $TX$ | Daily maximum temperature [C°] |
| | |
| $CDD$ | Conecutive Dry Days [days] |
| $Hscore$ | Heat score [C°] |
| $RDx$ | Maximum Cumulative Potential Precipitation Deficit [mm] |

| | |
|---|---|
| $SD25$ | Summer days [days] |
| $SPI$ | Standardized Precipitation Deficit [-] |
| | |
| $P20$ | Precipitation Exceeding 20mm per day [days] |
| $Ptot$ | Total precipitation [mm/ time period] |
| $R1D$ | Maximum one day precipitation [mm/day] |
| $R3D$ | Maximum three day precipitation [mm/ 3 days ] |
| $R5D$ | Maximum five day precipitation [mm/ 5 days ] |
| $SDII$ | Simple precipitation Intensity Index [mm/ wet days ] |

# 1 Introduction

Floods and droughts are natural phenomena causing severe impacts to society and environment. Both hazards accounted for 64% of the 1000 most disastrous natural disaster events recorded globally between 1900 and 2006 (Adikari and Yoshitani, 2009) and combined affected over three billion people worldwide in the past two decades (United Nations, 2015, WHO, 2021). It is expected that these phenomena will occur more frequently and become more severe in multiple parts of the world due to a changing climate (Alfieri et al., 2015, IPCC, 2012, Visser-Quinn et al., 2019). At the same time, population growth and urban expansion will lead to the increase of exposure (Güneralp et al., 2015) and vulnerability (Jahn, 2015) over time. Disasters occur where these elements of risk (i.e. hazards, exposure and vulnerability) interact with each other. Examples for drought-flood interactions have been observed in history. For example, the California drought from 2012 to 2016 and consecutive extreme precipitation events in 2017. During the height of the drought a high-pressure ridge was stationed off the west coast amplifying dry conditions, while in the winter of 2016-2017 a low-pressure trough appeared in the same region causing precipitation extremes (Simon Wang et al., 2017). Similarly, the severe drought of 2012 in England and Wales that was followed by one of the wettest April-July for these countries switching concerns from drought impacts to flood risk (Parry et al., 2013). For the Huaihe River basin in China, sudden drought to flood alternations are a reoccurring phenomena and show an upward trend in frequency of occurrence from 1961 to 2007 (Sheng et al., 2009).

Traditional practices for risk assessment consider flood and drought independently however, they are two opposite extreme of the same hydrological cycle (Krysanova et al., 2008, Ward et al., 2020). Hence, omitting their interaction might lead to an under- or over-estimation of the current and future risks associated with such natural hazards (Gill and Malamud, 2014, Hao et al., 2018, Leonard et al., 2014). For example, flood management practices adopted in one location might exacerbate the vulnerability of that location against drought (Di Baldassarre et al., 2017, Kreibich et al., 2019, Ward et al., 2020). The main challenge in analysing opposite extremes resides in defining a methodology able to assimilate the different temporal and spatial scales of the two phenomena whereas different methods to identify the dry and wet spells are used on various timescales. Research to these drought to flood interactions is still in its infancy. Those available are either based on an event ((Hanwei and Shixuan, 2012, Parry et al., 2013) on a regional scale (Ji et al., 2018, Parry et al., 2013, Swain et al., 2018, Zhao et al., 2020) and global scale (He and Sheffield, 2020, Rashid and Wahl, 2021)) or related to global warming (Madakumbura et al., 2019).

The interaction between flood and drought can be addressed via a multi-hazard approach.

This approach reverts to the early 1990s (Ward et al., 2021). The UNDRR (2020) defines the concept as *"(1) the selection of multiple major hazards that the country faces, and (2) the specific contexts where hazardous events may occur simultaneously, cascadingly or cumulatively over time, and taking into account the potential interrelated effects"*. Recent years show an increasing amount of attention to this approach, which is becoming more important to the Disaster Risk Management (DRM) and resilience debate (Pescaroli and Alexander, 2018, Tilloy et al., 2019, Ward et al., 2021) especially in light of a changing climate. It is possible that these multi-hazard events have larger impacts than any single-hazard event occurring in isolation (Aghakouchak et al., 2020, Hao et al., 2018, Raymond et al., 2020, Wahl et al., 2015, Ward et al., 2021, 2020).

The Netherlands recently experienced a number of dry summer periods: 2018, 2019 and 2020 (Daniëls, 2021). These dry conditions may impose preconditions on the system making it more vulnerable against a flood event. For example, dry conditions may decrease the stability of an embankment, which exacerbates the risk of breaching due to high water levels resulting from an extreme precipitation event (Figure 1.1). Dutch flood protection systems are able to withstand extreme events (e.g. high discharges or low discharges), however these standards are generally based on processes from a single event and do not considered preconditions of the system, e.g. as a consequence of dry period (STOWA, 2019). Recent dry years did not result in severe impacts as in figure 1.1. However, it is not excluded that this could happen, especially with an expected increase in severity and frequency of hydro-hazards due to a changing climate (Alfieri et al., 2015, IPCC, 2012, Visser-Quinn et al., 2019). In this research, the focus is on the interaction between dry and wet events, i.e. periods of water scarcity and water abundance. The aim of this research is then two fold. First, radar charts are introduced to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons. These can be used to identify hot-spots prone to opposite extremes. Second, a fully probabilistic framework based on Non-parametric Bayesian Network is developed to model the dependence between dry and wet indicators. Such model can be used to infer expected wet conditions in a given region when dry conditions are known.

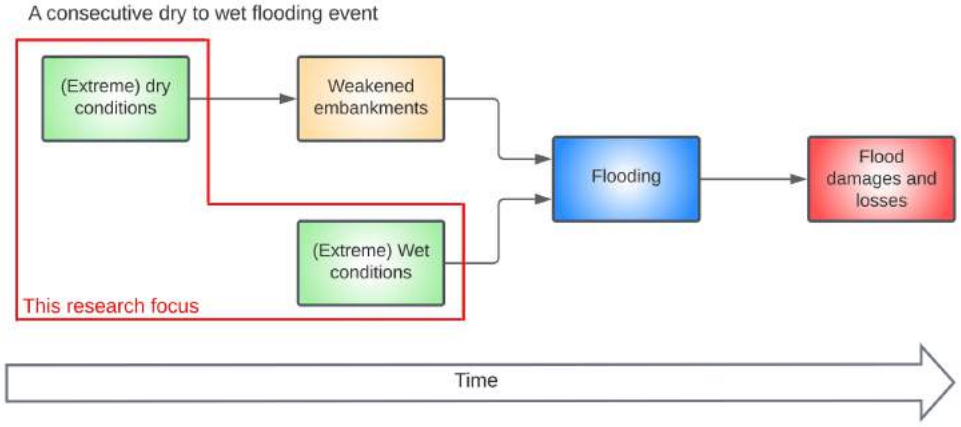



Figure 1.1: The motivating concept of a preconditioned disastrous event based on typology from Zscheischler et al. (2020). Here applied with sources of a consecutive dry wet event. This research only focuses on consecutive dry wet events in the Netherlands.

## 1.1. Outline of this thesis

This research focuses on the identification and quantification of consecutive extremely dry and extremely wet events and their interaction in the Netherlands. In *chapter 2* we perform a literature review on the concepts of extremely dry (drought) and extremely wet events (floods) in general and in relation to the Netherlands. Furthermore, we explain the method for the statistical model used in this research. *Chapter 3* explains the process carried out on how we identify and quantify the consecutive extremes and apply to a novel framework and the probabilistic model (Steps 1 to 4 in Figure 1.2). In *chapter 4* the results obtained from the novel framework and probabilistic model are presented (Steps 3 and 4 in Figure 1.2). The results are then discussed in *chapter 5* followed by the conclusion in *chapter 6* (Step 5 in Figure 1.2).

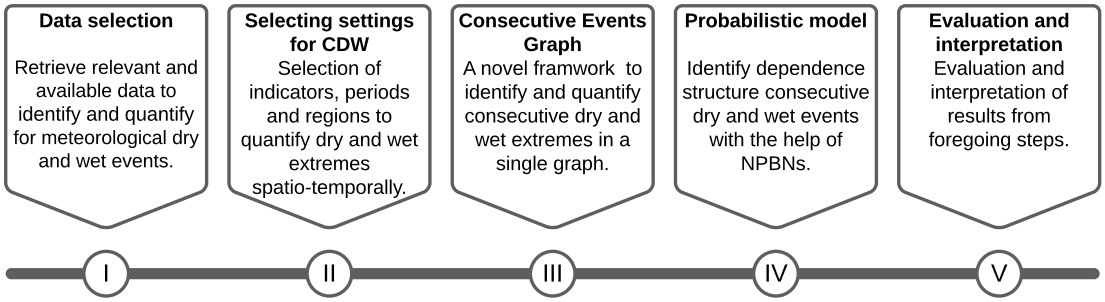



Figure 1.2: Outline of this thesis. The steps performed in chronological order are reading from left to right.

# 2
# Literature review

## 2.1. Drought

Drought is a natural hazard, just like floods, earthquakes or cyclones. However, despite being in the same category, drought differs in various respects (Wilhite, 2000). First, it can be considered as a creeping phenomenon as it does not show a as predictable and distinct onset and ending. Second, the impact of drought is less visible when compared to many other natural hazards. Instead, the impact of drought has a relative slow development and can linger on in the period after the drought took place. As a result, quantifying damages and providing help for drought issues is much more complex when compared to the other natural disasters (Wilhite, 2000).

Commonly drought is referred to as a shortage of water or a moisture deficiency and is often relative to a long term average condition or climate. This definition is rather generic, whereas other definitions for drought are frequently deduced by its impact where answers are to be found on socio-economic issues resulting from drought (Smakhtin and Schipper, 2008). For many decades, the search for a generalized objective definition of drought has been a curse for the the field of water resources. There are a large number of evaluations of definitions on drought across literature (Dracup et al., 1980, Lloyd-Hughes, 2014, Mishra and Singh, 2010, Sheffield and Wood, 2011, Tallaksen and Lanen, 2004, Wilhite and Glantz, 1985, Yevjevich, 1969). However, the spatio-temporal variation of the hydrological, meteorological and socio-economic conditions makes it difficult to move towards such generalized objective definition (Lloyd-Hughes, 2014, Mishra and Singh, 2010). For this reason Lloyd-Hughes (Lloyd-Hughes, 2014) states it we cannot expect and should not want to have a generalized objective definition of drought.

Table 2.1: A non exhaustive list of drought definitions in common literature.

| Definition | Source |
|---|---|
| "An exceptional period of water shortage for existing ecosystems and the human population (due to low rainfall, high temperature, and/or wind)." | IPCC (2021) |
| "Drought is a sustained period of below-normal water availability. It is a recurring and worldwide phenomenon, with spatial and temporal characteristics that vary significantly from one region to another" | Tallaksen and Lanen (2004) |
| "Drought are periods of time when natural or managed water systems do not provide enough water to meet established human and environmental uses because of natural shortfalls in precipitation or stream-flow." | Werick and Whipple (1994) |
| "Drought means a sustained, extended deficiency in precipitation." | WMO (1987) |
| "Drought is an interval of time, generally of the order of months or years in duration, during which the actual moisture supply at a given place rather consistently falls short of the climatically expected or climatically appropriate moisture supply." | Palmer (1965) |

Drought definitions can be subdivided into conceptual and operational definitions (Wilhite and Glantz, 1985). Conceptual definitions of drought involve a more general description and presents the boundaries of the concept. The rather generic definition in the previous paragraph can be classified as such conceptual definition, as holds for the definitions found in most dictionaries. Literature shows a general consensus on the presence of "a moisture deficit on a temporary scale" in the conceptual definition of drought (Table 2.1). Albeit differences exist in the nature of the water deficit (e.g. meteorological and/or hydrological), affected subject or activity (e.g. humans, environment and/or agriculture) and sometimes a cause of the drought (e.g. natural and/or anthropogenic). A poor conceptual understanding of drought can lead to bad decision making and improper measures when facing issues related to drought (Smakhtin and Schipper, 2008, Wilhite and Glantz, 1985). Furthermore, it may result in the misclassification of non-drought related issues to drought (Glantz and Katz, 1977, Smakhtin and Schipper, 2008).

Operational definitions of drought attempt to pinpoint the onset, degree of severity and ending of drought periods (Wilhite and Glantz, 1985). They are more or less related to quantification of the characteristics of drought, existing of a time dimension (e.g. duration, frequency, trend), space dimension (e.g. spatial extent) and severity. The quantification of these characteristics require data on hourly to seasonal scale, which are commonly aggregated to drought indicators. Most drought indicators introduced in literature focus on the severity of drought alone, in specific the moisture deficiency. It should be noted that both the duration, and to a lesser extent, the area play a role in the severity of drought. Additional literature of this century shows more interest in the space-time dimension of drought to get a better understanding of e.g. drought propagation (Herrera-Estrada et al., 2017, Lloyd-Hughes and Saunders, 2002, Vicente-Serrano, 2006, Zhou et al., 2019). But, as of its complex nature, most research reduce the 3-dimensional space-time drought structure (severity, area and duration) to a subspace of lower order (Lloyd-Hughes, 2012). Drought indicators are frequently utilized in a variety of applications for communication on drought, although usually not designed for widespread use (Quiring, 2009). Such use requires carefulness as drought characteristics vary on the type or impact of drought considered and by differences in conditions such as hydrometeorological

conditions, type of climate, environmental variation, socioeconomic variation and anthropogenic influences (Leelaruban and Padmanabhan, 2017, Lloyd-Hughes, 2014, Mishra and Singh, 2010, Satoh et al., 2021). This leads us to the types of drought commonly found in literature on drought which are to be explained in next paragraphs.

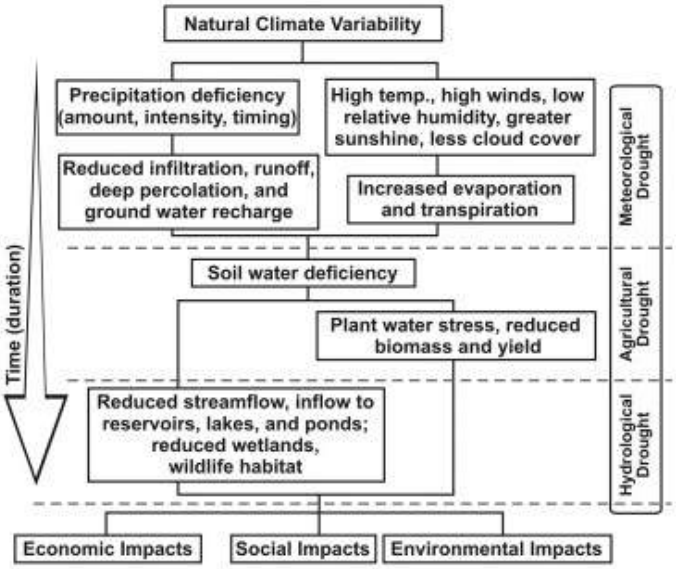



Figure 2.1: Schematization of drought types, their chronological order and causal factors generally found in literature. Dependent on the duration (vertical) drought will move to a next phase or drought type. Starting with meteorological drought having anomalies in weather e.g. lack of precipitation or increased temperature. As a result, a deficiency in soil water can arise (phase two: Agricultural drought). Finally, this can result in a deficiency for open water or ground water, meaning hydrological drought. Source: NDMC.

### 2.1.1. Drought topology

Wilhite and Glantz (Wilhite and Glantz, 1985) defined four categories for drought which are widely adopted in literature on drought (Figure 2.1). These categories are meteorological drought, agriculutral drought, hydrological drought and socio-economic drought. The categories can be deviated by either their nature of deficit (i.e. meteorological and hydrological drought) or impact (i.e. agricultural and socio-economic drought). Although widely used in literature, agricultural drought can be replaced with soil-moisture drought. Soil-moisture drought strongly relates to crop failure, which is the main focus of agricultural drought, but encompasses all impact by soil moisture deficits which makes it a definition by nature of deficit (van Loon, 2015). In the following paragraphs we will elaborate on the droughts types based on the nature of deficit and several drought types based on impact.

#### Drought topology by nature of deficit

The drought definitions based on nature of deficit are meteorological drought, soil moisture drought and hydrological drought. The boundaries separating these categories are defined as following.

(i) Meteorological drought is defined as a deficiency in precipitation extending over a region and period of time (Mishra and Singh, 2010). Additionally, it can be in combination with an increased potential evapotranspiration (van Loon, 2015). This type of drought is found most frequently in literature. A vast number of indicators to quantify and monitor meteorological drought are available, whereof precipitation is the most common input variable (Mishra and

Singh, 2010, Wilhite and Glantz, 1985). Some definitions involve other parameters such as humidity, temperature, wind or vapor pressure. Meteorological drought indicators should be independent of physical properties of the site it is measured and purely defined by meteorological variables (Wanders et al., 2010). Although site specific, meteorological drought impact may be neglected by buffer systems such as open waters, but a continuous call made on the moisture reserves stored in a hydrological system can lead to deficiencies within the next types of drought.

(ii) Soil-moisture drought refers to deficits in water stored in the upper layer of the soil. Soil-moisture drought may constrain the supply of moisture to vegetation and affect natural ecosystems and infrastructure (van Loon, 2015). Multiple elements from the water cycle play a role in the extent of this type of drought, whereof antecedent precipitation and evapotranspiration are the most important variables (Manning et al., 2018). Soil-moisture is commonly quantified by indirect methods as the number of in-situ observations are limited.

(iii) Hydrological drought is defined by a deficit in open surface waters and subsurface water. It is sometimes subdivided into a separate definition for streamflow drought and groundwater drought (van Loon, 2015). Hydrological drought is often caused by sustaining meteorological and soil-moisture drought, but is generally out of phase with these two types of drought (Wilhite and Glantz, 1985). Water Authorities In the Netherlands mostly refer to this type of drought as it can be noticed by below-normal water levels in lakes, rivers or groundwater.

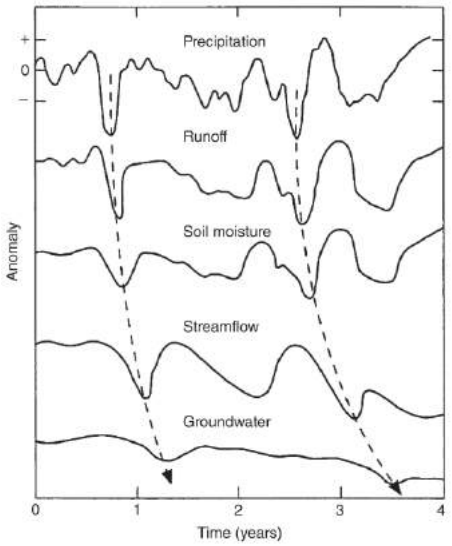



Figure 2.2: Theoretical propagation of drought through the terrestrial part of the hydrological cycle. Source: van Loon (van Loon, 2015)

The three types explained above generally occur in a chronological order. For example, a drought in precipitation (i.e. meteorological drought) can have an effect on all parts of the hydrological cycle and by means propagate through another part of the terrestrial part (Figure 2.1.1) (Mishra and Singh, 2010, Tallaksen and Lanen, 2004, van Loon, 2015). Although this may not be the only process triggering drought type lower in the propagation scheme. For example, the lack of snow melt or ice melt may trigger a hydrological drought (Van Loon and Van Lanen, 2012). Additionally, human interference in the hydrological system can trigger or amplify the types of droughts above (Van Loon et al., 2016). For example, land use, irrigation, dams and water abstraction may influence at the layers of soil moisture drought and hydrological drought. Therefore, many droughts in environments dominated by humans may not considered as natural only.

### Drought topology by impact

Each kind of the drought types by nature of deficit can lead environmental impact, but may not always do so. Environmental impact studies generally subdivide the environmental impact into the socio-economic realm and biophysical realm (Chadwick et al., 2005). This subdivision can also be applied to the typification of drought by impact:

(i) Socio-economic drought is one of the most common types of drought based on impact found in literature (Mishra and Singh, 2010, Tallaksen and Lanen, 2004, van Loon, 2015, Wilhite and Glantz, 1985). Economic goods depends on (fresh) water supply e.g. fish, hydroelectric power, navigation, foods and tap water. Drought will have socio-economic impact once a deficit in (fresh) water is not able to fulfil the demand of water to produce these type of goods or carry out these services anymore. This type of drought can incorporate many features of the types of drought mentioned above. As stated earlier, agricultural drought is a type of drought that has an impact on agricultural activities e.g. crop loss or yield loss. This type of drought is often the result of a period of declining soil moisture, which does not meet the requirements of the crops anymore. Which can be (and often is) a result of sustaining meteorological drought. However, a plants requirements for water differ per specie, growth stage, soil type, and meteorological conditions. For example, a plant that is deeply rooted will be able to subtract water from the soil more deeply than a young and shallow rooted plant. This also means that agricultural drought can be the result of wrongdoing by human action (Van Loon et al., 2016). Often economic loss is used to quantify the impact of drought to the agricultural sector (Jeuken et al., 2012).

(ii) Ecological drought is a type of drought that can be classified within the biophysical realm. The United States Geological Survey (USGS) defines this type of drought as a loss of ecosystem services. Ecological drought is similar to agricultural drought, but this time the impact is not on agricultural activities but on ecosystems. This can be loss of plants and trees, but also fish mortality as a result of declining oxygen in rivers and lakes.

### 2.1.2. Drought in the Netherlands

The Netherlands has a long history and relationship with water. Often this is related to historical flood events, whereas droughts - situated at the other side of the hydrological spectrum - is less associated with this country. However, recent years of (extreme) drought (e.g. 2018) impacted the Netherlands in various sectors such as agriculture, navigation, energy and ecology (Rijkswaterstaat and Unie van Waterschappen, 2019). The recent experience with this natural hazard increased the attention to drought effects and mitigation of those effects (Beleidstafel Droogte, 2019, Didde, 2021). This is not a new topic, as concerns for the effects of drought on the longer term dates back to 1988. The National Institute for Public Health and the Environment (RIVM) (RIVM, 1988) published their report about environmental outlooks for the Netherlands: "Zorgen voor Morgen" (in English: Concerns for Tomorrow). This report covers similar concerns on the effects of drought as today's concerns, and already introduces source and effect-based measures to tackle this problem on the long run. Even though these concrete goals, concrete actions were lacking in the years following the publication (Didde, 2021).

Public Works and Water Management (in Dutch: Rijkswaterstaat) consider the months from April to September as the dry season. It is to a large extent linked to the growing season for plants as transpiration of plants barely takes place outside this period. Within this season the daily differences in Makkink's reference evaporation (Makkink, 1957) and precipitation are calculated and documented cumulatively per season (Figure 3.5). The Netherlands counts some historical extremely dry years such as 1976, 2003 and more recent 2018 (Rijkswaterstaat and Unie van Waterschappen, 2019).

The water system in the Netherlands is highly controlled with the help of pumps, sluices and weirs. All water flowing in via the Rhine river at Lobith can be directed to various parts in the Netherlands (e.g. IJsselmeer or Haringvliet). In times of drought this requires deliberate choice where water will flow to. A certain discharge is required for the Rhine river to host its vessels for transport. The agricultural sector requires fresh water for irrigation and salinization in coastal areas has to be prevented. The National Coordination Committee for Water Allocation (LCW) is the governing body for these choices on behalf of parties as e.g. regional water authorities, Public Works and Water Management and drinking water companies. This committee allocates water based on priority of water requirement (i.e. a displacement series). There are four categories within the displacement series in order of decreasing priority: Category one prioritizes large risk entities e.g. the stability of water retaining structures, preventing soil subsidence and preventing irreversible environmental damages. Category two covers utility services which for example require cooling water for their provision of energy. Category three prioritizes category four covers smaller risk entities (e.g. maritime transport, agriculture and industry) (Rijkswaterstaat and Unie van Waterschappen, 2019).

With the additional increase in population, the demand for fresh water, and the trend within the agricultural sector to more capital intensive crops the stress on fresh water increases. Furthermore, the climate is changing, there is an increase in extreme rainfall events and rising sea level hitherto. This increasing stress on fresh water and the increased discussion due to recent drought years lead to new goals to inventory the effects of extreme drought and how to tackle future extreme droughts for the Netherlands (Smit, 2021).

## 2.2. Floods

Floods are the kind of natural disaster with the largest socio-economic impact (Jonkman, 2005, Munich Re, 2007), and apart from the ice sheets, floods are quasi-total present around the world (Vinet, 2011). The European Environment Agency (European Union, 2007) defines a flood as a temporary covering of a certain area by water that is normally not covered by water. This can be water originating from rivers, sea or extreme precipitation, but may exclude floods from sewerage systems. Not all floods are disastrous events (e.g. from 100-year or larger return period events), whereas the more frequent floods are from small-magnitude events resulting in so called nuisance floods. These floods are characterized by low levels of inundations which do not threaten public safety, do not result in great damage propagation and create minor disruptions for e.g. routine activities (Moftakhari et al., 2018).

To understand floods, one needs to understand what type of floods can be faced. There are several kinds of floods bearing different characteristics such as temporal extent, spatial extent, the types of protection suitable, the damage it can cause and forecastability. Generally floods result from an extreme hydro-meteorological event originating from a water body (e.g. sea, ocean or river) or precipitation. This is also the reason why research does the typification of flood by the source of the water. These types most commonly include fluvial floods, coastal or seawater floods, groundwater floods, pluvial floods and flash floods (Nicklin et al., 2019, Nixon, 2015, Vojinovic and Huang, 2015). In some case non-natural sources are included, for example artificial water bearing infrastructure like dams (Federal Emergency Management Agency, 1999). In general fluvial and coastal floods receive most attention as they are typically the flood types with the largest spatio-temporal extent (Nicklin et al., 2019, Nixon, 2015). However, recent flood events originating from extreme precipitation events demonstrate that pluvial and flash floods may not be neglected (Spekkers et al., 2017). For example, the summer floods in Western Europe of 2021 causing at least 122 deaths and over twenty billion euros in damages (Task Force Fact-finding hoogwater 2021, 2021). Szewranski et al. (Szewrański et al., 2018) even suggest that the cumulative pluvial flood damages may exceed the damages from river and coastal floods, as pluvial flood events are more frequent than fluvial and coastal flood events (Van Ootegem et al., 2015). To have a correct understanding of the flood types we describe them as following (Prokić et al., 2019, Zurich Insurance, 2020):

(i) Fluvial floods are floods that occur when the water level in rivers, streams or lakes exceed the crest elevation of their banks and overflow its neighbouring land. They generally originate from precipitation events or snow melt, where large volumes of water flow into river systems causing increasing water levels. Characteristics of this type of flood events are dependent on the surrounding of the river system. Where mountainous or hilly areas generally will experience effects on a shorter term than flatter areas, as water moves quicker towards a river system. This is also the reason why this type of flooding can be relative predictable in downstream areas, when heavy precipitation events upstream of a river catchment result in increased water levels downstream of that river days later. The timescale of this type of flood event can differ from days to weeks.

(ii) Coastal or seawater floods is when seawater comes ashore in coastal regions and causes the inundation of land. They are commonly caused by high tide and storm surges where high water levels and waves attack an unprotected or protected coastline and overflow or overtop the crest elevation of that coastline. For sandy coastlines this process results in erosion and possible breaching of coastal defence works. An additional cause of coastal flooding are tsunami waves, which generally cause large destruction due to the height and velocity of these waves.

(iii) Groundwater floods are the inundation of land as a result of a rising water table. In such case, the water table reaches a higher level than ground level, resulting in seepage of water from the ground. This saturation of the soil may find its cause by prolonged precipitation, which could have taken weeks before this type of flooding to occur. This is because water flow slowly through the ground.

(iv) Pluvial floods is the result of precipitation not being able to runoff to watercourses, drainage systems or sewers creating ponds or overland flow. This can only occur when the rate of precipitation exceeds the rate of evacuation of water by drains and infiltration in the ground. It is therefore dependent on the precipitation intensity and duration, and the hydrological characteristics of the basin such as area, soil type and run off magnitude (Prokić et al., 2019). The inundation depth is generally low for pluvial floods and can cause simple fluid nuisance to significant economic damage.

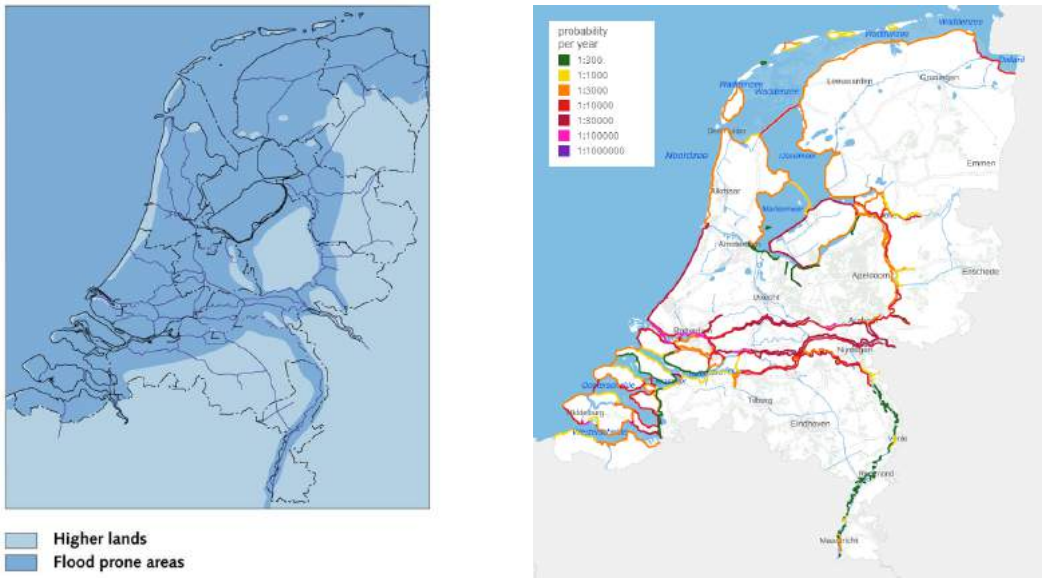
(v) Flash floods are caused by extreme precipitation events or heavy downpour where the run off volumes concentrate fast and channel up towards lower lying areas (Kuksina and Golosov, 2020, Montz and Gruntfest, 2002). They generally develop fast, are of short duration, cover a small area and are commonly associated with other events such as mudslides and fluvial floods (Montz and Gruntfest, 2002). Damages may not only be from water. Debris that is being picked up by the flow of water can damage objects or create additional obstructions when caught by underpasses and bridges which results in additional damming of water and intensifying flooding.

### 2.2.1. Coastal and fluvial floods in the Netherlands

Since the Middle Ages flood protections have been a prerequisite for settlement in low-lying areas in the Netherlands. During these centuries, defensive strategies were required as of the rising sea level and subsiding (peat) soils. Temporal flooding were part of life for those living in the flood prone areas. In more recent years, additional flood defence measures were realised in response to the 1953 flood disaster. Large scale measures were implemented nation wide and the Delta Committee was put into practice. This committee set new standards for flood protection which were based on cost-benefit analyses of each single dike ring by weighing the costs of the reinforcement measure against the reduction of flood risk achieved by the reinforcement measure. Little knowledge was available on flood defence bursts, so the approach for flood defence failure was based on normative water levels. These standards were used for decades, however not impeccable for riverine floods in more rural areas. In December 1993 and January 1995, two flood disasters happened in Limburg as a result of river bank overflow along the Meuse river. Though being of smaller scale when compared to the 1953 floods and no casualties were reported, these floods had significant impact on the for riverine flood risk perception.

Up until today, flood prevention is the dominant strategy in the Netherlands (Hegger et al., 2014). Flood risks have been reduced by a flood defence system consisting of (storm surge) barriers, dunes and embankments. A strategy high on the agenda, because the country is located in a low-lying delta with an extensive coastline and a large number of rivers, canals and lakes; an approximate of 55% of the Netherlands is prone to flooding whereof 29% at risk due to river flooding and 26% at risk from sea as it is below sea level (Figure 2.3a). Also in the Netherlands there is an increase in population over the past decades, with a resulting urbanisation of areas below sea level. Furthermore, there is an increase in extreme rainfall events and rising sea level due to a changing climate. If flood defences remain unchanged, the flood impacts will increase as a result in increased loads and exposure. The 2017, new flood

protection standards were adopted to safeguard these risks (STOWA, 2019).




(a) (b)

Figure 2.3: Left, the flood prone areas in the Netherlands covering approximately 55% of the Netherlands. Right, flood safety standards from 2017 - present. Presented probabilities are the flooding probabilities (various colors) per embankment trajectory (each single continuous and colored line) per year. From `www.pbl.nl`.

The new 2017 flood protection standards are anchored in the Dutch law, Water Act (Figure 2.3b). One of the important changes is the replacement of the design water levels for embankments by a more comprehensive flooding probability. The design water level translates to the probability of exceedance of a set water level at a water retaining structure. Whereas the flooding probability translates to the probability of losing the water retaining capacity for a dike trajectory, thus including the probability of exceedance of a set water level and breaching.

The new standards were based on three criteria: efficiency, equity and societal disruption. To account for equity, a minimum flood protection level was defined as the probability to die due to a flood event, i.e., Local Individual Risk (LIR). Every single individual in the Netherlands who is situated behind a primary flood defence has at least a level of protection of the LIR. This probability for loss of life as a consequence of flooding defines at a maximum 1/100.000 per year. Additional protection is applied to embankment stretches where failure may result in large number of fatalities or large economic damages and/or damage to or failure of vital and vulnerable infrastructure.

### 2.2.2. Pluvial and flash floods in the Netherlands

As stated before, generally fluvial and coastal floods receive most attention as they are typically the flood types with the largest spatio-temporal extent (Nicklin et al., 2019, Nixon, 2015). This is also true for the Netherlands. As presented in section 2.2.1 the Netherlands holds concrete safety standards for dikes and embankments that are anchored in Dutch law, while no legal standards apply for pluvial floods (Gilissen, 2014). There is, however, still a governmental responsibility to prevent and limit impact of this type of flooding, although the interpretation and concretization of this responsibility may vary per municipality (Dai et al., 2018). An increasing number of pluvial flood events in the public space in the past decade have led to the

encouragement of climate adaptation in urban areas which are to decrease pluvial flood impacts (Dai et al., 2018). Cities are especially sensitive to pluvial flooding due to the large proportion of impermeable surfaces that prevents excess water to drain in a natural way. Instead, water has to be retained or drained via local sewage systems which has a limited capacity. Pluvial floods can have a large socio-economic impact, not only due to damages to or destruction of building and interiors, but also due to secondary damages when pluvial floods affects power, telecom, data and public infrastructure (Runhaar et al., 2016). It is expected that this type of flooding is to occur more frequently and with larger impact due to the combined effect of climate change and urbanization (Rosenzweig et al., 2018). Especially for traditional water-draining infrastructure in cities like Amsterdam and Utrecht that were designed hundreds of years ago. The ability to cope with an increase of precipitation extremes is limited for those systems (Dai et al., 2018).

In the past century, the annual amount of precipitation in the Netherlands increased by approximately 26% as of a changing climate (Figure 2.4) (Koninklijk Nederlands Metereologisch Instituut, 2015). The largest percentile increase is during Winter (43%) and smallest during Summer (17%) (KNMI, 2021c). The main factors for these increases are changes in atmospheric circulation and an increase in sea surface temperature respectively (van Haren et al., 2013). The intensification of heavy precipitation events may be an additional factor fueling the total precipitation (Eden et al., 2018). This intensification is expected to be a result of an increase in temperature. Lochbihler et al. (Lochbihler et al., 2019) showed that an increase in temperature is accompanied by larger amounts of water vapour in the atmosphere, leading to larger and more intense extreme precipitation events. Extreme precipitation event are expected to increase up to 14% with each degrees Celcius. These trends are already visible in the Netherlands and additional increase should be accounted for in future expected precipitation events.

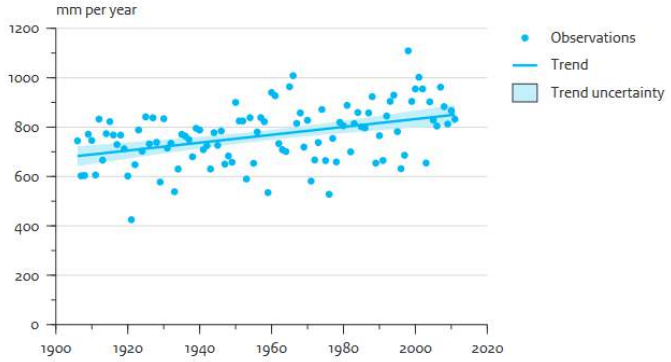



Figure 2.4: Increase in precipitation past century in the Netherlands due to a changing climate. From: KNMI (2018) (KNMI, 2018)

So what makes precipitation events extreme? Precipitation can be characterized by three main elements: the depth or intensity, area affected and duration of the event. It has been a common way for decades to perform a so called Depth-Area-Duration analyses on precipitation data (WMO, 1969). This means that extreme precipitation can take different shapes. For example, a short but high intensity local precipitation event could lead to urban flooding or flash flooding. Or low intensity precipitation, but with a long duration (e.g. a week) and a large catchment area affected could lead to high discharges downstream. The Dutch Meteorological Institute (KNMI) defines extreme precipitation in multiple manners. A local precipitation event

of over 25 millimeters in an hour is considered a downpour. Whereas a 50 millimeters of precipitation in a day is considered as an heavy precipitation event (KNMI, 2021c). More extreme precipitation events exist, but generally have a return period of approximately 100 years or more. Working with return periods is a general way of comparing precipitation events. In the current climate, precipitation extremes with a duration smaller than 12 hours do not show large spatial differences in the Netherlands. However, for precipitation extremes with a larger duration, spatial differences are present (Beersma et al., 2019). For example, the South-East of the Netherlands (Limburg province) and West of the Netherlands (Coastal regions) are regions with larger intensities compared to the rest of the Netherlands (KNMI, 2021c).

## 2.3. Multi-hazard approach

The concept of a multi-hazard approach was first introduced in the early 1990s (Ward et al., 2021). The UNDRR (2020) defines this concept as *"(1) the selection of multiple major hazards that the country faces, and (2) the specific contexts where hazardous events may occur simultaneously, cascadingly or cumulatively over time, and taking into account the potential interrelated effects"*. Several types of multi-hazard events are defined through literature, such as cascading events (AghaKouchak, 2018, May, 2007), compound events (Aghakouchak et al., 2020, Leonard et al., 2014, Zhang et al., 2021b, Zscheischler et al., 2019, 2018) and consecutive events (de Ruiter et al., 2020). The large number of alternating names present[1] lead to the fragmentation of literature (Kappes et al., 2012, Leonard et al., 2014, Pescaroli and Alexander, 2015). Additionally, the use these terms as synonyms causes confusion and redundancy (Pescaroli and Alexander, 2018). This is one of the reasons for reviewing the multiple facets of the multi-hazard approach available in literature and to work towards a more comprehensive methodology (Pescaroli and Alexander, 2018, Tilloy et al., 2019, Ward et al., 2021). Below the line, the major part of these aforementioned studies aim for a more comprehensive and holistic approach on one or more elements of risk assessment (Table 2.2). Whereas understanding the interaction of multiple hazards, such as consecutive dry and wet (CDW) events, will allow risk management and disaster preparedness to be more effective (Brunner et al., 2021, Kreibich et al., 2019, Pescaroli and Alexander, 2018, Tilloy et al., 2019, Ward et al., 2021).

Table 2.2: Comparison of Current and Future Practice in Modeling Climate Extremes from Leonard et al. (Leonard et al., 2014)

| Item | Current Practice (2014) | Future Practice |
|---|---|---|
| 1 | Poor mapping of inputs to impact | Well understood mapping of variables |
| 2 | Poor communication of model options | Influence diagrams convey model options |
| 3 | Less precise definitions of impact events | Greater precision in defining impact events |
| 4 | Causal reasoning focused on process | Diagnostic reasoning starting with stakeholder |
| 5 | Compound events are exceptional | Compound events are normal |
| 6 | Extreme inputs | Extreme impacts |
| 7 | Univariate analysis of extremes | Multivariate analysis of extremes |
| 8 | Analysis of outputs (climate incompatible) | Conditioned on inputs (climate compatible) |
| 9 | Silos of knowledge | Cross-disciplinary collaboration |

[1] with additional terms such as domino, chaining, interconnected and interacting events

## 2.4. Statistical modelling

Zhang et al. (2021a) summarized common statistical methods employed across literature on hydrological multi-hazard events[2] used to investigate multivariate dependencies. The methods vary from simply counting the events (empirical approach) to quantifying dependence between drivers with multivariate distributions or complex networks. Among the network approaches are Bayesian networks. Bayesian networks are becoming increasingly popular in dependency analysis and multi-hazard studies (Tilloy et al., 2019). They are a powerful tool, well suitable when dealing with a great number of variables while remaining user-friendly due to its compact and graphical approach. Bayesian networks, compared to other statistical modelling methods such as Fault Trees, Petri Nets and Markov Chains, are capable of making both predictions and diagnostics, quantify the likelihood of occurrences of events and make inference of variables based on conditioning variables and the structure of the Bayesian Network (Weber et al., 2012).

### 2.4.1. Bayesian networks

Pearl first introduced the concept of Bayesian networks to model uncertainty in AI networks in 1986 (Pearl, 1986). A Bayesian Network (BN) is a causal model for a set of variables that are defined over a directed acyclic graph (DAG) $G = (\mathbf{V}, A)$, where $\mathbf{V}$ is the set of nodes and $A$ is the set of arcs. Each node represents a random variable and the arcs connecting nodes represent causal or statistical dependency between nodes. This relationship is directed from a parent node to a child node, which are the two types of nodes present in a BN. As of its acyclic characteristic, arcs may not form a directed loop or have nodes that are self connected. This methodology for visualization of relationships between different random variables is relatively straightforward and therefore appealing for understanding complex networks (Figure 2.5).

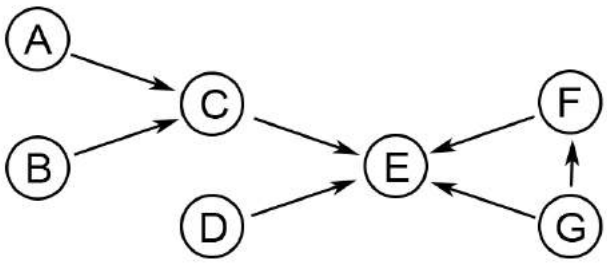



Figure 2.5: A directed acyclic graph (DAG) with seven variables. Nodes A and B are parent node of C, where C is the child of A and B. The relationships between nodes E, F and G would not be legitimate if the directional relationship from G to E was inverted. This would create a directional loop.

The joint probability distribution of $\mathbf{V}$ factorizes with respect to the DAG into a set of local probability distributions given by the Markov property of Bayesian networks,

$$f(v_1, \ldots, v_n) = \prod_{i=1}^{n} f(v_i|\text{parents}(v_i)) \tag{2.1}$$

Bayesian networks offers the means to model complex networks as of its robustness and flexibility as explained in section 2.4. Aguilera et al. subdivides Bayesian networks into a qualitative

[2]Instead of using multi-hazard events or a like, Zhang et al. (2021a) refers to them as compound extremes

component and a quantitative component (Aguilera et al., 2011):

- The qualitative component describing the DAG. Where the statistical dependence structure of the network is described in a qualitative sense.
- The quantitative component consists of the individual relationship of each child node (or variable) and its preceding parent nodes. This is necessary to understand the strength of a relationship between variables described by the model.

BNs are broadly used in a variety of fields, but most of these applications make use of discrete BBNs (Hanea et al., 2015). Here, parent nodes are marginal distributed and child nodes are compromised from conditional probability tables. While merely built up from discrete variables, this type of network suffers from multiple impracticalities such as subjectively assessing the discretization of continuous variables (Hanea et al., 2015, 2006). Alternatively, there are Gaussian BBNs which allow to deal with continuous variables, but assumes joint normality. As a consequence each marginal or conditional is also Gaussian. This method can also be combined with the discrete model, resulting in a discrete-continuous model. In such model the parents of a continuous node can be discrete but the children not. Additionally, Kurowicka and Cooke (2003) introduced Non-parametric Bayesian networks (NPBNs). In this type of model, no joint distribution is assumed, hence the name non-paramtetric (Hanea et al., 2006). For NPBNs, the nodes are described by continuous variables with invertible distribution functions. The arcs are described by (conditional) rank correlations obtained with a copula. More information on copulas can be found in appendix A.

### 2.4.2. NPBNs in hydraulic engineering

Although the combination of copulas and Bayesian networks is a dynamic area of study, the application of them to natural hazards is still limited across literature (Tilloy et al., 2019). Hanea et al. (2015) touches on several of the application of these NPBNs. One of the papers featured in this research is on dam safety by Morales-Nápoles et al. (2014a). In this paper a NPBN is created to model the safety of an earth dam in the State of Mexico. This model contains parent nodes such as loads (i.e. seismic frequency and precipitation rate) and activities (i.e. maintenance frequency) that are considered contributing factors for failure modes (i.e. overtopping, loss of global stability and piping) leading to dam failure. Possible impacts by such events are also captured in this model. By conditioning nodes (e.g. the loads and activities) one can simulate the effect of this precondition to quantify its possible risks. The application of NPBNs for this use-case has proven itself to be successful.

More recent research, by Ragno et al. (2021), presents an application of NPBNs to estimate the monthly maximum river discharge by investigating 240 catchments across the United States. Again all available and possible contributing factors to this discharge event are implemented in the model to simulate estimates. The estimates are then compared to the observed discharge event with the help of statistical tests. It is shown that the use of NPBNs is suitable to deduce a good estimate of the discharge event when remaining variables are known, but suggest some challenges in defining a suitable NPBN.

Another paper introduces the use of NPBNs for modelling vegetated hydrodynamic systems such as mangroves and salt marshes (Niazi et al., 2021). To create a probabilistic description of these systems, they were schematized to deduce all contributing factors. From there NPBNs were set up and Monte Carlo sampled for multiple types of vegetated hydrodynamic systems to create a better understanding of these systems.

Application of NPBNs based on the normal copula assumption is made possible with the help software like UNINET (UNINET Team, 2019) and Banshee for Matlab (Paprotny et al., 2020) and Python (Mendoza Lugo et al., 2021).

# 3
# Data and methodology

This chapter describes the methodology and steps performed. Section 3.1 describes what data is used for the analysis. Sections 3.2, 3.3 and 3.4 build upon the knowledge of droughts and floods in the literature review and describe the settings to create a dataset for spatio-temporal coupled CDW extremes. These sections describe the methods to (i) identify and quantify dry and wet extremes separately (with the help of indicators), (ii) select the consecutive periods for dry and wet extremes and (iii) present the regions to link the dry and wet extremes spatially. Section 3.5 describes the resulting dataset of spatio-temporal coupled CDW extremes. This dataset is applied in two fold, inline with the aim of the research. First, in section 3.6, the Consecutive Events Graph (CEG) is introduced. This is a practical framework to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons. Section 3.7 describes the second application. Here NPBNs are developed, tested and applied. This application aims to quantify the interaction of the consecutive dry and wet extremes and infer wet extremes based on hypothetical or observed dry events.

Generally all analyses are performed in Python (Van Rossum and Drake, 2009), unless stated differently. All codes, data and figures can be acquired from: thesis repository.

## 3.1. Data selection

### 3.1.1. Dry event and wet events definition

The definition of drought varies from one region to another and can be referred to by the impact of the lack of water or the lack of water itself (Smakhtin and Schipper, 2008). This research focuses on the lack of water based on meteorological variables only. The definition of drought for this research is therefore meteorological drought (Wanders et al., 2010). The wet events are also defined by its meteorological aspect only (i.e. precipitation). Other variables such as Normalized Difference Vegetation Index (NDVI), discharge and groundwater levels were considered. Although these either had a limited temporal extent, a limited spatial extent, poor data availability or complicated the translation to the statistical model.

All daily historical weather information is retrieved from the KNMI open archive. This archive compromised weather stations (KNMI-A) and the precipitation stations (KNMI-M) (Figure 3.1.1).

### 3.1.2. KNMI-A stations

KNMI-A stations measure a large number of variables at a 10 minute interval e.g. air temperature, precipitation and wind speed. They are most relevant for real-time weather monitoring, weather forecasting and scientific research. Daily data from these stations are retrieved (KNMI,

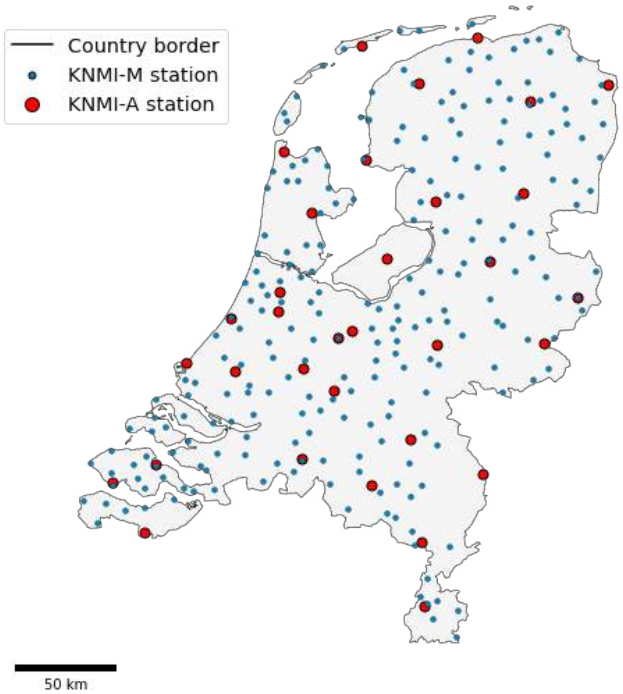



Figure 3.1: KNMI-A (red) and KNMI-M (blue) stations considered in this research. Locations retrieved from the KNMI manual of observations (KNMI, 2000).

2021a). The various number of variables allows calculation of multiple drought or heat related indicators, from hereon referred to as dry indicators. The indicators are elaborated in section 3.2.1 and are summarized in table 3.1. The climate variables used from these stations are daily maximum temperature ($T_X$), daily average temperature ($T_G$), daily precipitation ($P$) and daily evapotranspiration ($EV24$). The evapotranspiration retrieved from the KNMI-A stations is also known as the Makkink reference evaporation. It is not measured, but calculated on the basis of the daily average temperature ($T_G$) and the daily average radiation ($Q$) (KNMI, 2000):

$$EV24 = \frac{1000 \cdot 0.65 \cdot \delta(T_G)}{\{\delta(T_G) + \gamma(T_G)\} \cdot \rho \cdot \lambda(T_G)} \cdot Q \quad [\text{mm/day}] \tag{3.1}$$

Where $\delta(T_G)$ is the saturated vapor pressure gradient relative to water, $\gamma(T_G)$ the psychometric constant, $\rho$ is the density of water and $\lambda(T_G)$ the heat of evaporation of water. Generally, the daily global radiation is the limiting factor for the temporal extent of the variables. KNMI-A stations with a long historical record (i.e. older than 1980) are generally affected by this (Appendix B).

### 3.1.3. KNMI-M stations

The KNMI-M network is a precipitation station network run by volunteers who measure precipitation amounts manually on a daily base. The total of 240 stations are relevant for calibration of historical weather data and scientific research. Daily homogenized precipitation data ($P$) of these stations are retrieved from KNMI Climate Explorer portal (KNMI, 2021b). All stations have an equal temporal extent starting from 1955 until now. The density of the precipitation

stations allows determination of local precipitation events.

## 3.2. Identifying and quantifying dry and wet extremes

To identify and quantify dry and wet extremes indicators are selected that portray the characteristics of these extremes. First, the indicators for dry extremes are described (from hereon dry indicators). Followed by the indicators for wet extremes (from hereon wet indicators).

### 3.2.1. Dry indicators

In section 2.1 meteorological drought is defined as a deficiency in precipitation and can be combined with evapotranspiration. A number of five indicators were selected that either highlight one or multiple elements of these characteristics. These indicators are the three months Standardized Precipitation Index, maximum potential rainfall deficit, summer days, consecutive dry days and the heat score (Table 3.1). Given the required variables for the dry indicators, only KNMI-A stations are used for calculation of these dry indicators. Each indicator is further explained in next subsections, along with a figure that describes the number of occurrences or propagation of the considered indicator during the growing season for station "de Bilt". This figure will help to identify the months where the largest number of occurrences or increase are observed for considered dry indicators.

Table 3.1: The indicators for dry events, short description, required variables, unit and source used in this research.

| Dry indicator | Short description | Input variables | Units | Source |
|---|---|---|---|---|
| SPI-3: Standardized Precipitation Index | Quantifies precipitation anomalies on a 3-months time scale | P | - | McKee et al. (1993) |
| CDD: Consecutive Dry Days | Maximum number of consecutive days with no precipitation for a time period. | P | days | Tank et al. (2011), Zhang et al. (2011) |
| HS: Heat Score | Cumulative amount of $°C$ above 18 ° per day for a time period. | $T_G$ | $°C$ | KNMI (2021d) |
| SD25: Summer Days | Number of days with an air temperature above 25 ° for a time period. | $T_X$ | days | Tank et al. (2011), Zhang et al. (2011) |
| RDx: Maximum Potential Precipitation Deficit | Cumulative difference in potential evaporation and precipitation for a time period. | EV24, P | mm | KNMI (2000) |

#### Three months Standardized Precipitation Index (SPI-3)

The Standardized Precipitation Index (SPI) is designed to quantify precipitation anomalies based on probability such that it can be compared across regions with different climates (Guttman, 1999, McKee et al., 1993). Preferably a continuous period of 30 years of data should be used, but its robustness increases with longer time frames. Timescale variation is possible, where 1-, 3-, 6-, 12- and 24-month moving precipitation totals are commonly used periods. Although it only requires precipitation data as input, precipitation deficits at larger time scales can relate to effects further down the drought propagation order (Figure 2.1.1).

For the first step in the calculation procedure of the SPI one determines the distribution that describes each set of precipitation time series for the timescale of choice. In other words, a SPI-3 at the end of August requires the total precipitation of the months of June, July and August of every year for fitting the distribution. Next this relationship of probability to precipitation is transformed into a standardized normal distribution with a mean of zero and standard deviation

of one. The corresponding value on this normal distribution for any particular precipitation value is the SPI value indicating the anomaly (Table 3.2).

In this research the SPI-3 for the summer months of the Netherlands is used (i.e. June, July and August). The SPI at this timescale reflects on short- and medium-term moisture conditions (WMO, 1987). The gamma distribution was used for fitting, which common for this location on the northern hemisphere and short timescales like 3 months (Guenang and Mkankam Kamga, 2014).

Table 3.2: Standardized Precipitation Index classification and corresponding theoretical probabilistic occurrence (WMO, 1987)

| SPI Value | Class | Probability |
|---|---|---|
| 2.0 or more | Extremely wet | 0.023 |
| 1.5 to 2.0 | Severely wet | 0.044 |
| 1.0 to 1.5 | Moderately wet | 0.092 |
| 1.0 to -1.0 | Normal | 0.682 |
| -1.0 to -1.5 | Moderate drought | 0.092 |
| -1.5 to -2.0 | Severe drought | 0.044 |
| -2.0 or less | Extreme drought | 0.023 |

### Consecutive Dry Days (CDD)

Consecutive Dry Days is a climate indicator that quantifies the maximum number of consecutive no precipitation days per time period ($j$) (Tank et al., 2011, Zhang et al., 2011). No precipitation days are days ($i$) where no measurable precipitation takes place, but is commonly defined as days with less than one millimeter of precipitation. The indicator does not show a clear month to be more favourable in observing a high number of dry days (Figure 3.2). Nevertheless, a large number of this indicator corresponds with a extensive period of low precipitation which can be favouring drought conditions. A long-term record is required for a good estimate range of expected values.

To calculate the CDD, one has to count the largest number of consecutive days where:

$$P_{ij} < 1mm \tag{3.2}$$

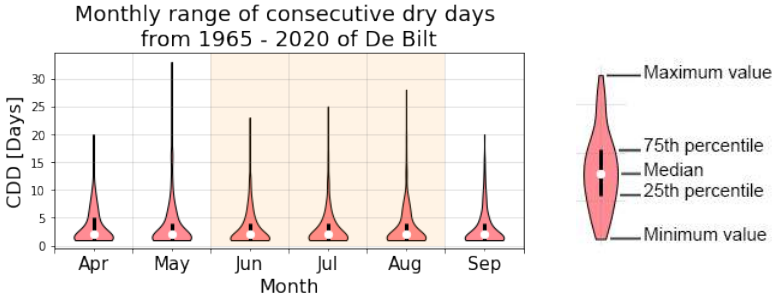



Figure 3.2: Violin plots of monthly number of the maximum consecutive dry days in De Bilt over the period from 1965 until 2020. Description of the violin plots on the right. The summer period is annotated in yellow.

### Heat score (Hscore)

The heat score is an indicator used by the KNMI to classify the heat of the summer period (KNMI, 2021d). The heat score is calculated by taking the difference between the daily average temperature ($TG$) and a threshold of 18°C and only accounts for positive values (i.e. $TG > 18^{\circ}C$). The score is calculated cumulatively for the same period as the growing season i the Netherlands, i.e. from April 1 to September 30. The indicator finds its largest increase during the summer months June, July and August (Figure 3.3).

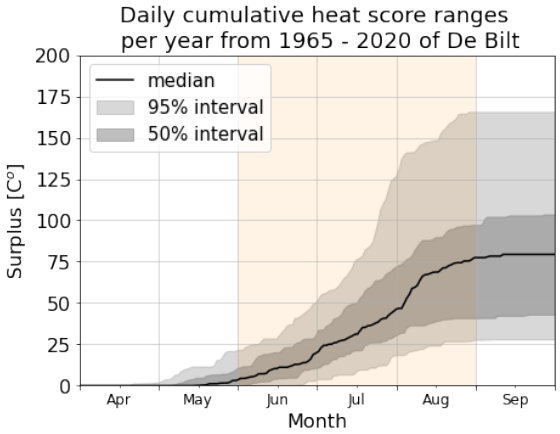



Figure 3.3: Daily cumulative heat score ranges per year for De Bilt. Largest increases can be observed during the months June, July and August (in yellow). Largest scores are observed in September as the indicators is calculated cumulatively.

### Summers Days (SD25)

Summers Days is defined by the Expert Team on Climate Change Detection and Indices (ETCCDIs) as a climate extremes indicator (Tank et al., 2011, Zhang et al., 2011). It quantifies the number of days where the maximum air temperature ($TX$) is above 25°C per time period. For this study the number of Summers Days in the summer period is used as indicator. A long-term record of is required for a good estimate range of expected values. The largest number of summer days are generally found in July, August and June (Figure 3.4).

To calculate the SD25, one has to count the number of days where:

$$T_X > 25^{\circ} \tag{3.3}$$

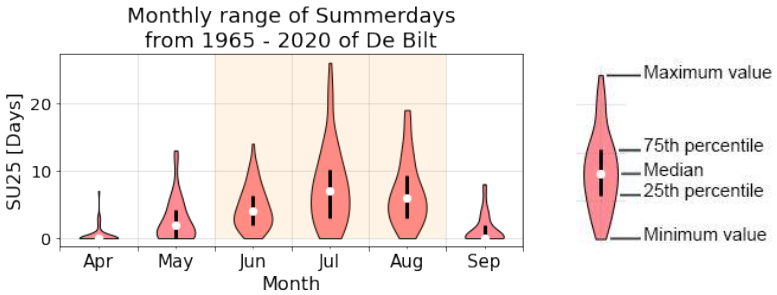



Figure 3.4: Violin plots of monthly number of summer days in De Bilt over the period from 1965 until 2020. Description of the violin plots on the right. The summer period is annotated in yellow.

### Maximum Cumulative Potential Precipitation Deficit (RDx)

The potential precipitation deficit ($RDx$) is the difference between the potential evaporation ($EV24$) and amount of precipitation ($P$) on a single day or a period of multiple days (cumulative). In the Netherlands, typically the growing season for plants is used for the cumulative potential precipitation deficit i.e. from April 1 to September 30 (Wolters et al., 2011). This is the period where the daily potential evaporation is generally larger than the average daily precipitation. When negative ($\sum EV24 < \sum P$), the potential precipitation deficit is reset to zero. Negative values can only exist if retention measures are present, otherwise foregoing surpluses will not make up for consecutive deficits due to runoff. The EV24 represents the evaporation of short grass in optimal watered conditions (KNMI, 2000). Although this is not the actual condition for the greater part of the Netherlands during summer, it appears to be attractive for practical application (De Bruin and Stricker, 2000). Which is also the reason the KNMI uses the cumulative potential precipitation deficit as indicator for drought, such as for the analysis of drought year 2018 (Sluijter et al., 2018). The RDx is calculated cumulatively for the same period as the heat score (i.e. April - September) (KNMI, 2000). Generally, largest values and increases are observed in months June, July and August (Figure 3.5).

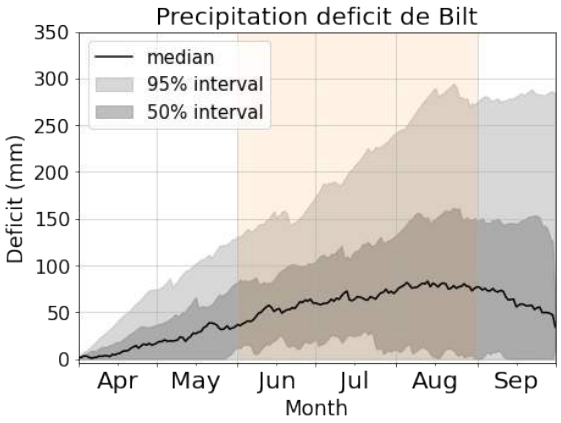



Figure 3.5: Precipitation deficit yearly ranges from 1965 - 2020. The summer period is annotated in yellow.

### 3.2.2. Wet indicators

Following from section 2.2, precipitation may result in flooding due to a sudden release of large amount of water (e.g. a cloud burst) or a sustained wet period (e.g. large scale and long duration precipitation extreme). A total of six indicators were selected to quantify these characteristics. These are the total precipitation, precipitation exceeding 20mm per day, simple precipitation intensity index, maximum 1-day, 3-day and 5-day precipitation (Table 3.3). The KNMI-M stations were used to calculate these indicators, as all indicators require daily precipitation or precipitation on a larger window for calculation. Similar to the dry indicators, the wet indicators are further explained in following subsections. To quantify regional and seasonal differences of the indicators, the explanations are assisted with a seasonal graph of average values of the indicator per KNMI-A station.

Table 3.3: Wet indicators, required input variables, unit and source used in this research.

| Wet indicator | Short description | Unit | Source |
|---|---|---|---|
| PRCPTOT: Total precipitation | Total sum of precipitation for a time period | mm | (Tank et al., 2011, Zhang et al., 2011) |
| R20mm: Precipitation exceeding 20mm per day | Number of days with 20mm precipitation for a time period | days | (Tank et al., 2011, Zhang et al., 2011) |
| SDII: Simple precipitation intensity index | Average rainfall intensity on wet days for a time period | mm /day | (Tank et al., 2011, Zhang et al., 2011) |
| Rx1day: Maximum 1-day precipitation | Maximum precipitation in a single day for a time period | mm | (Tank et al., 2011, Zhang et al., 2011) |
| Rx3day: Maximum 3-day precipitation | Maximum precipitation in three consecutive days for a time period | mm | (Tank et al., 2011, Zhang et al., 2011) |
| Rx5day: Maximum 5-day precipitation | Maximum precipitation in five consecutive days for a time period | mm | (Tank et al., 2011, Zhang et al., 2011) |

#### Total precipitation (Ptot)

The total precipitation ($Ptot$) is the sum of precipitation of days with precipitation ($i$) for a period ($j$) of choice. Generally this indicator is used for a one year period, as defined by the ETC-CDIs (Tank et al., 2011, Zhang et al., 2011). Highest values occur in both summer and fall (Figure 3.6). In fall higher values are observed in coastal regions compared to inland regions. The value can be calculated by

$$Ptot_j = \sum J_{i=1}(P_{ij}) \tag{3.4}$$

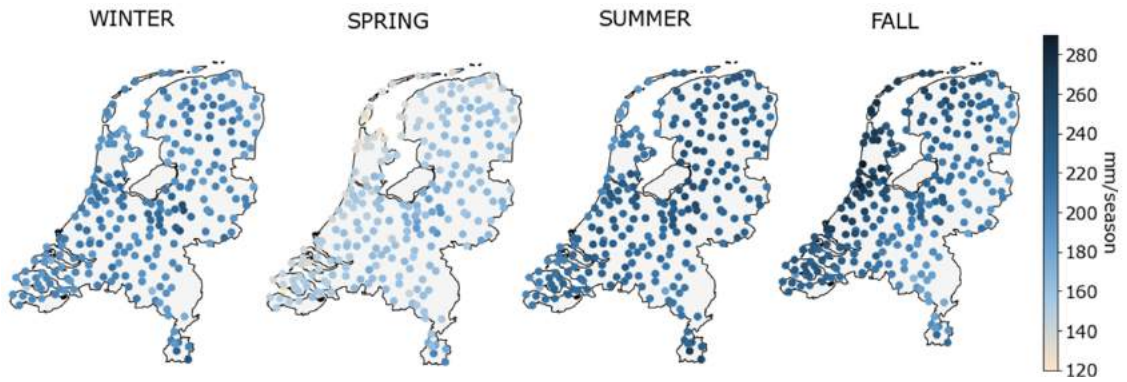



Figure 3.6: Average seasonal total precipitation in years 1965 - 2020.

### Precipitation Exceeding 20mm Per Day (P20)

The Precipitation Exceeding 20mm Per Day is an indicator that counts the number of days with a precipitation amount above 20 mm. It is an indicator used by the ETC-CDIs (Tank et al., 2011, Zhang et al., 2011). Highest number of days are observed during summer and fall period (Figure 3.7). During fall period these number of days are generally higher in coastal regions than inland regions. Calculation of the value can be done by counting the number of days where

$$P_{ij} > 20mm \tag{3.5}$$

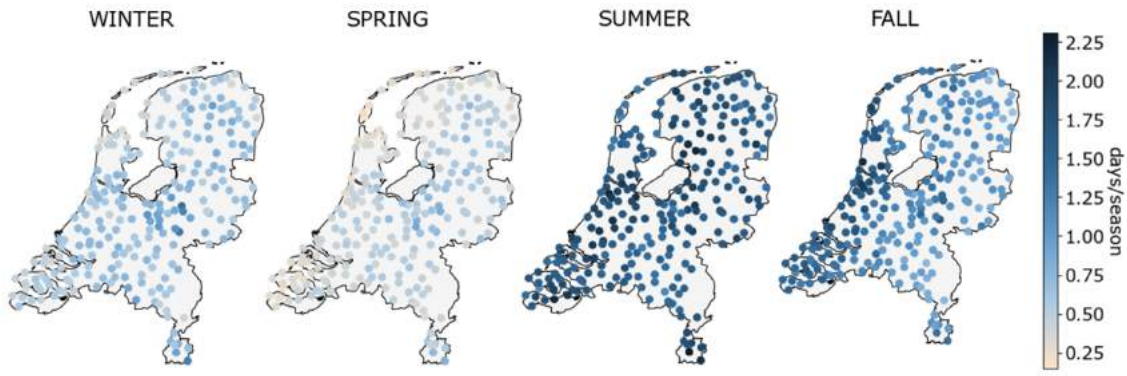



Figure 3.7: Average seasonal P20 days in years 1965 - 2020.

### Simple precipitation Intensity Index (SDII)

Simple precipitation Intensity Index (SDII) indicates the average rainfall in a period (Tank et al., 2011, Zhang et al., 2011). It is calculated by averaging the precipitation of all rain days ($w$) over the number of wet days ($W$) for a period ($j$) of choice. Highest values are generally observed during summer followed by fall. Coastal and inland differences exist in fall period. This indicator can be denoted as

$$SDII_j = \sum W_{w=1} \frac{P_{wj}}{W} \tag{3.6}$$

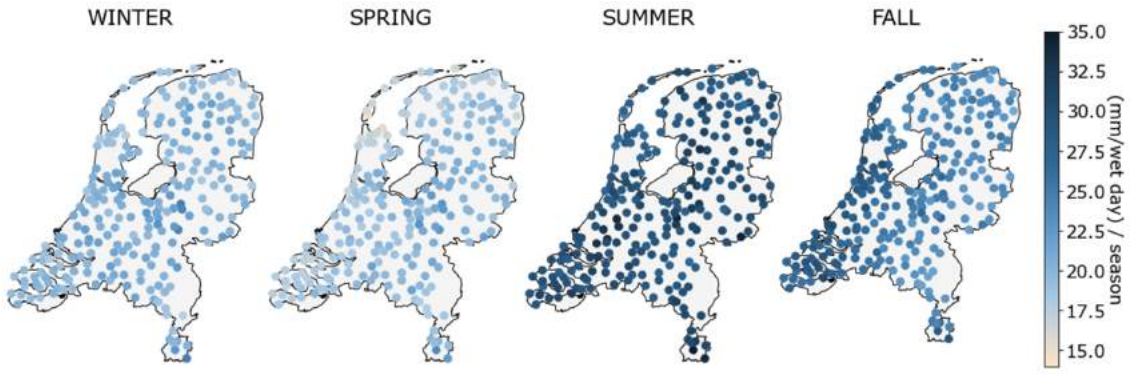



Figure 3.8: Average seasonal maximum SDII in years 1965 - 2020.

### Maximum 1-day precipitation (R1D)

Maximum 1-day precipitation is the highest precipitation amount on a single day $(i)$ in a period $(j)$ (Tank et al., 2011, Zhang et al., 2011). Highest values are generally observed during summer followed by fall. Coastal and inland differences exist in fall period.

$$R1D_j = max(P_{ij}) \tag{3.7}$$

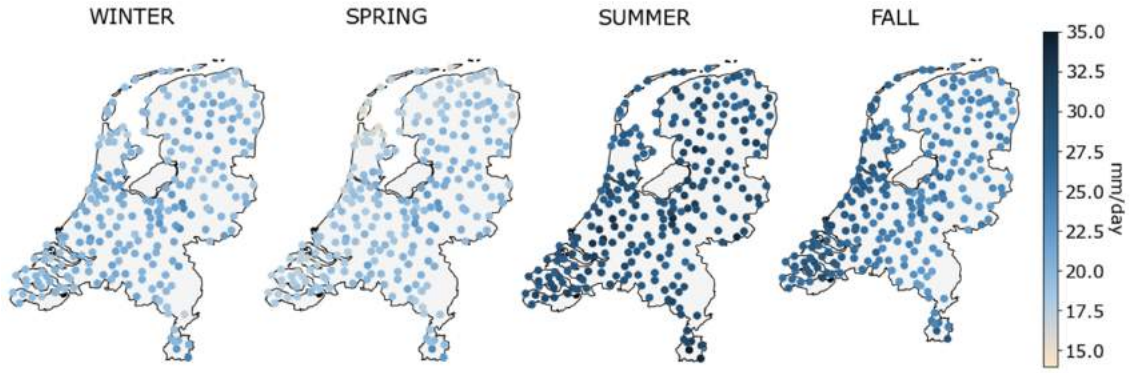



Figure 3.9: Average seasonal maximum R1D in years 1965 - 2020.

### Monthly maximum 3-day precipitation (R3D)

Maximum 3-day precipitation highest precipitation amount on three consecutive days $(m)$ in a period $(j)$ (Tank et al., 2011, Zhang et al., 2011). This indicator captures extreme rainfall with a longer lasting period than just a single day extreme event.

$$R3D_j = max(P_{mj}) \tag{3.8}$$

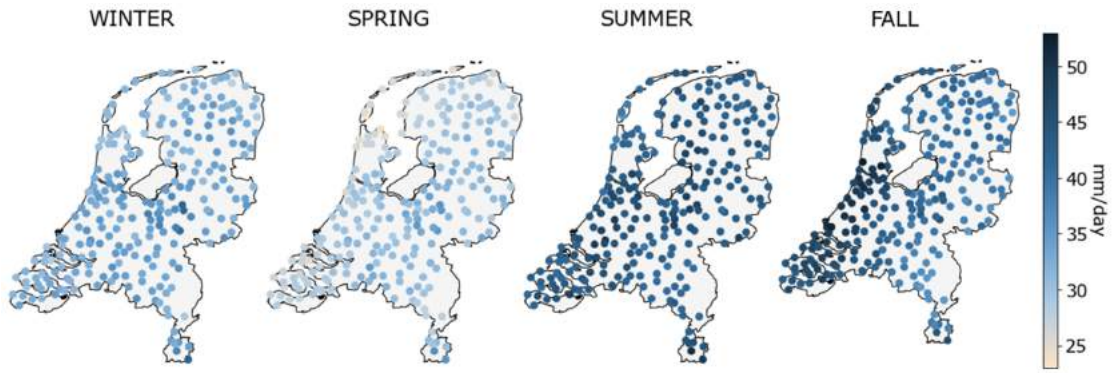



Figure 3.10: Average seasonal maximum R3D in years 1965 - 2020.

### Monthly maximum 5-day precipitation (R5D)

Maximum 5-day precipitation highest precipitation amount on five consecutive days ($n$) in a period ($j$) (Tank et al., 2011, Zhang et al., 2011). This indicator captures extreme rainfall with a longer lasting period than just a single day extreme event. For this study this is the highest precipitation amount for three days during the wet period annually.

$$R5D_j = max(P_{nj}) \tag{3.9}$$

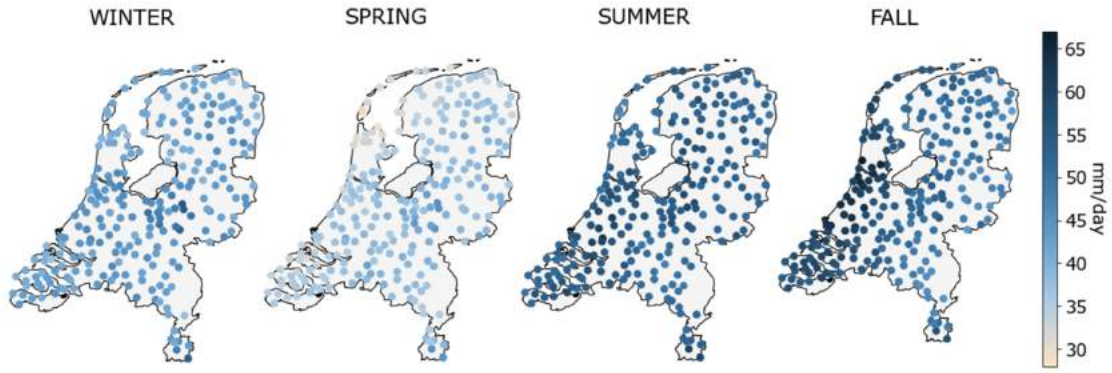



Figure 3.11: Average seasonal maximum R5D in years 1965 - 2020.

## 3.3. Dry and wet period

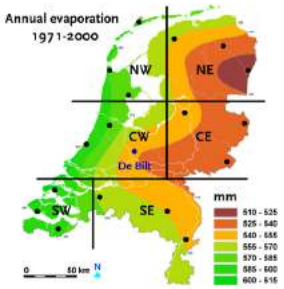
For selection of the dry and wet extremes the dry period is required to precede the wet period. The figures describing the single indicators of section 3.2 were used to select these periods. In general, most indicators find their maximum values or largest increases during the months June, July and August (i.e. summer; annotated by the yellow background in the figures). Furthermore, the precipitation deficit (RDx) and summer days (SD25) show a downward trend in the month September. Therefore, the dry period was defined by months June, July and August. For the wet period, the largest values are observed in both summer and fall. Both winter and spring show smaller average values of the indicators. Therefore, the wet period is defined by months September, October and November (i.e. fall).

## 3.4. Region selection

To link the dry and wet extremes spatially, the Netherlands was subdivided in homogeneous regions where both wet and dry events were characterized on a regional scale. This was required as dry events generally occur on a regional to national scale, while for wet events this can be more locally (Lloyd-Hughes, 2012). It is common to make a subdivision for drought regions in the Netherlands (Figure 3.12). There are multiple organisations or sources of literature that do so. For example, the LCW its Regional Drought Consultation (RDO) regions defines its regions by the water authority administrative borders and associated water systems. The Netherlands can also be subdivided based on its drought problems (Rijkswaterstaat and Unie van Waterschappen, 2019). Differences exist as demarcation on the action taking authority can be more prominent than the expected type of drought problems faced in an area. Furthermore, research on regional frequency analysis with precipitation deficits presents alternative subdivisions for drought regions (Beersma and Buishand, 2007).


(a) The 6 RDO regions defined by the LCW based on waterboard administrative borders and associated water systems. Source: (Rijkswaterstaat and Unie van Waterschappen, 2019)

(b) The 11 regions defined by drought problems e.g. salinisation, water depletion, presence of inclines for water delivery. Source: (Rijkswaterstaat and Unie van Waterschappen, 2019).




(c) Disctricts defined in precipitation deficit study by Beersma and Buishand. Source: (Beersma and Buishand, 2007).

(d) The 21 regions defined by waterboard administrative borders.

Figure 3.12: Regions defined by other drought related organisations, literature or research.

The KNMI-M stations used for calculation of the wet indicators all have a temporal extent of 66 years starting from 1955. The KNMI-A stations used for calculation of the dry indicators have a much smaller and unequal temporal extent. For example, only five stations have the temporal extent of at least 56 years starting from 1965. Fourteen stations have a minimum temporal extent of 33 years and up to 30 stations have a minimum temporal extent of 20 years (Appendix B). The long term record of 55 years was preferred, as this contributes to retrieving a better picture of the synoptic dry extremes per region. Furthermore, for the application to the

probabilistic model a as long as possible record was preferred. The corresponding five KNMI-A stations are well distributed in the Netherlands (Figure 3.4). Using these five stations was therefore considered suitable for quantifying the dry indicators on a regional scale.

For this research the region selection is based on the following criteria:

- Applicability: A region requires one (or more) KNMI-A stations to define the dry conditions for that region.
- Action: A district its borders may be defined by its active authority i.e. multiple waterboards.
- Coastal/inland: Synoptic meteorological conditions differs between coastal and inland regions in the Netherlands (Beersma and Buishand, 2007, Philip et al., 2020).
- Distance: The distance from a region its borders to a KNMI-A station should be limited.
- Similarity: Comparing to related regions defined by other drought related organizations or research (Figure 3.12)

Resulting regions are visualized in figure 3.4. Each KNMI-A station related region contains a single KNMI-A station, but multiple KNMI-M stations for quantifying wet extremes. Wet indicators calculated from these KNMI-M stations were therefore spatially linked to the dry indicators from each KNMI-A station. This means that largest calculated value of every unique wet indicator from all KNMI-M stations within a region is paired with the only calculated value of a dry indicator of the KNMI-A station of that same region.

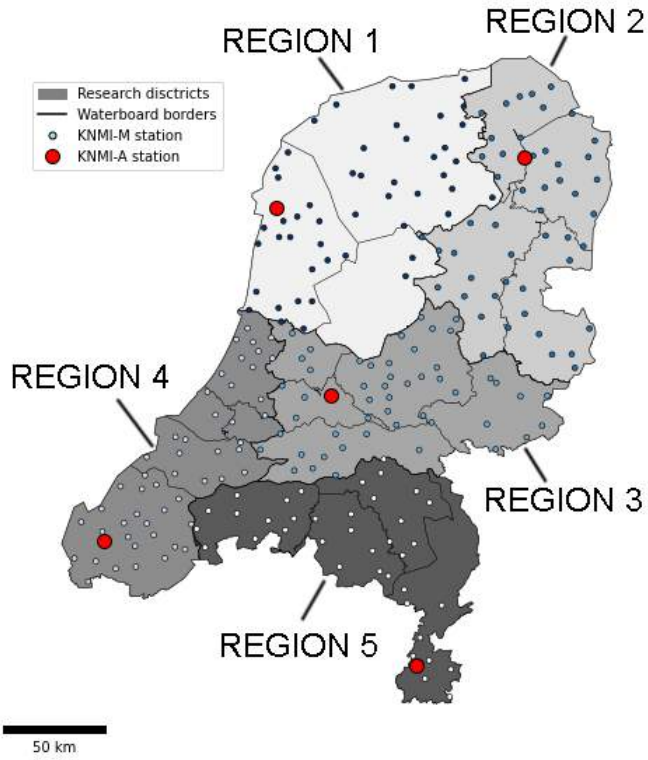



Figure 3.13: The five created districts (gray shaded areas), waterboards (black borders) and regarding KNMI-A and KNMI-M stations.

## 3.5. Resulting data set

Based on the defined indicators and regions a dataset can be created containing the extreme values of each indicator per region and year (Table 3.4). For every region a number 56 years (1965 - 2020) of extreme values are calculated. The variation of each indicator per region is visualized in figure 3.14 for dry indicators and figure 3.15 for wet indicators. This resulting dataset is used for the analysis in the following sections.

Table 3.4: Snapshot of the resulting dataset used for further analysis in this research. For every region (column = 'Region') a number 55 years (column = 'Year') of extreme values are present. All dry indicators are labeled have a "D_" prefix, whereas all wet indicators have a "W_" prefix.

| | Year | Region | D_SPI_03 | D_CDD | D_Hscore | D_SD25 | D_P_def | W_Ptot | W_P20 | W_SDII | W_R1D | W_R3D | W_R5D |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1965 | 1 | 0.76 | 11 | 1.4 | 2 | 82.0 | 254.1 | 2 | 7.28 | 26.5 | 50.7 | 63.7 |
| 1 | 1966 | 1 | 1.06 | 10 | 9.5 | 0 | 78.0 | 301.5 | 4 | 7.99 | 30.11 | 50.7 | 57.2 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 278 | 2019 | 5 | -0.3 | 22 | 191.8 | 37 | 268.0 | 307.9 | 3 | 7.23 | 35.9 | 54.5 | 69.6 |
| 279 | 2020 | 5 | -0.23 | 16 | 180.1 | 36 | 322.6 | 234.0 | 2 | 6.25 | 37.6 | 63.6 | 77.5 |

.

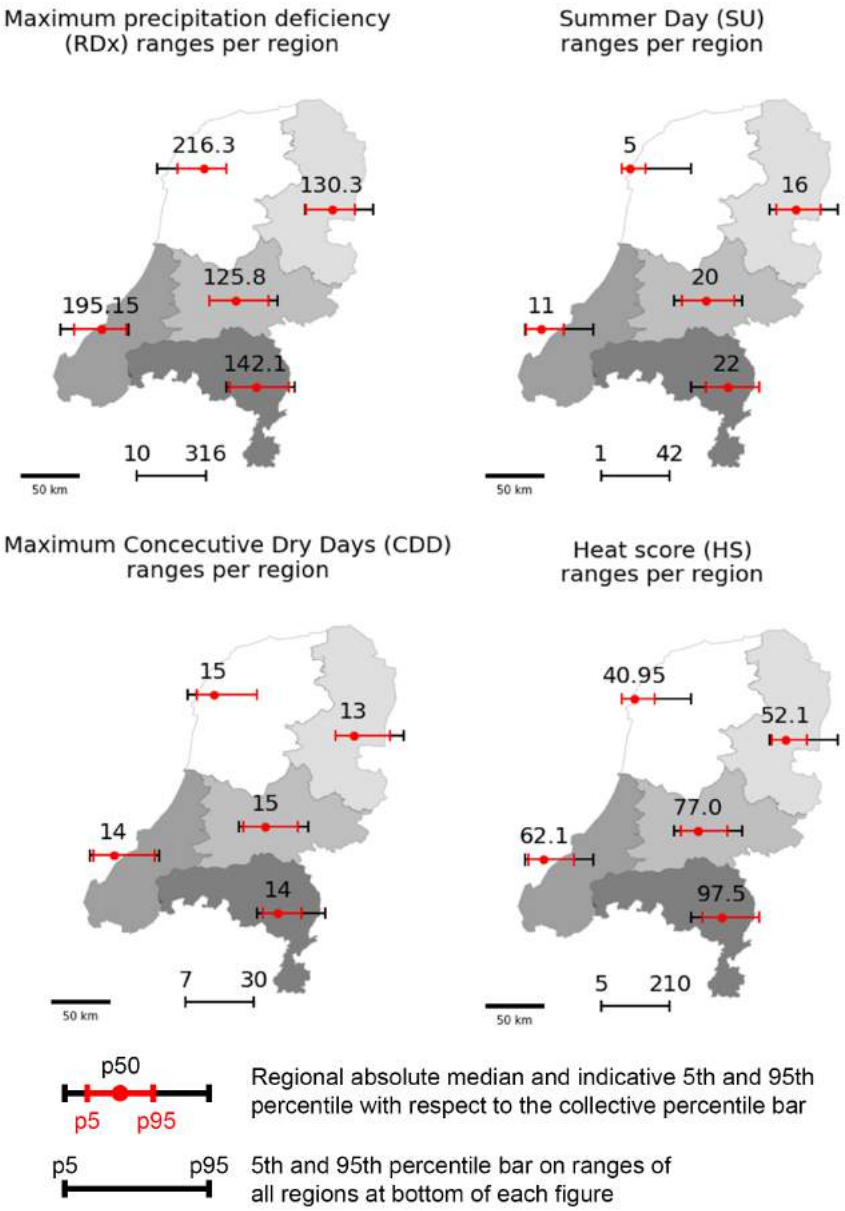



Figure 3.14: Ranges of the dry indicators per region (red) on a all region scale bar (black) in summer (June, July and August). The median is annotated and presented by the red dot. The minimum and maximum p5 and p95 are annotated by the black scale bar at the bottom of each figure. Note that the SPI-3 indicator is not presented as this indicator is normalized and would result in near equal ranges (mean of 0 and standard deviation of 1).

.

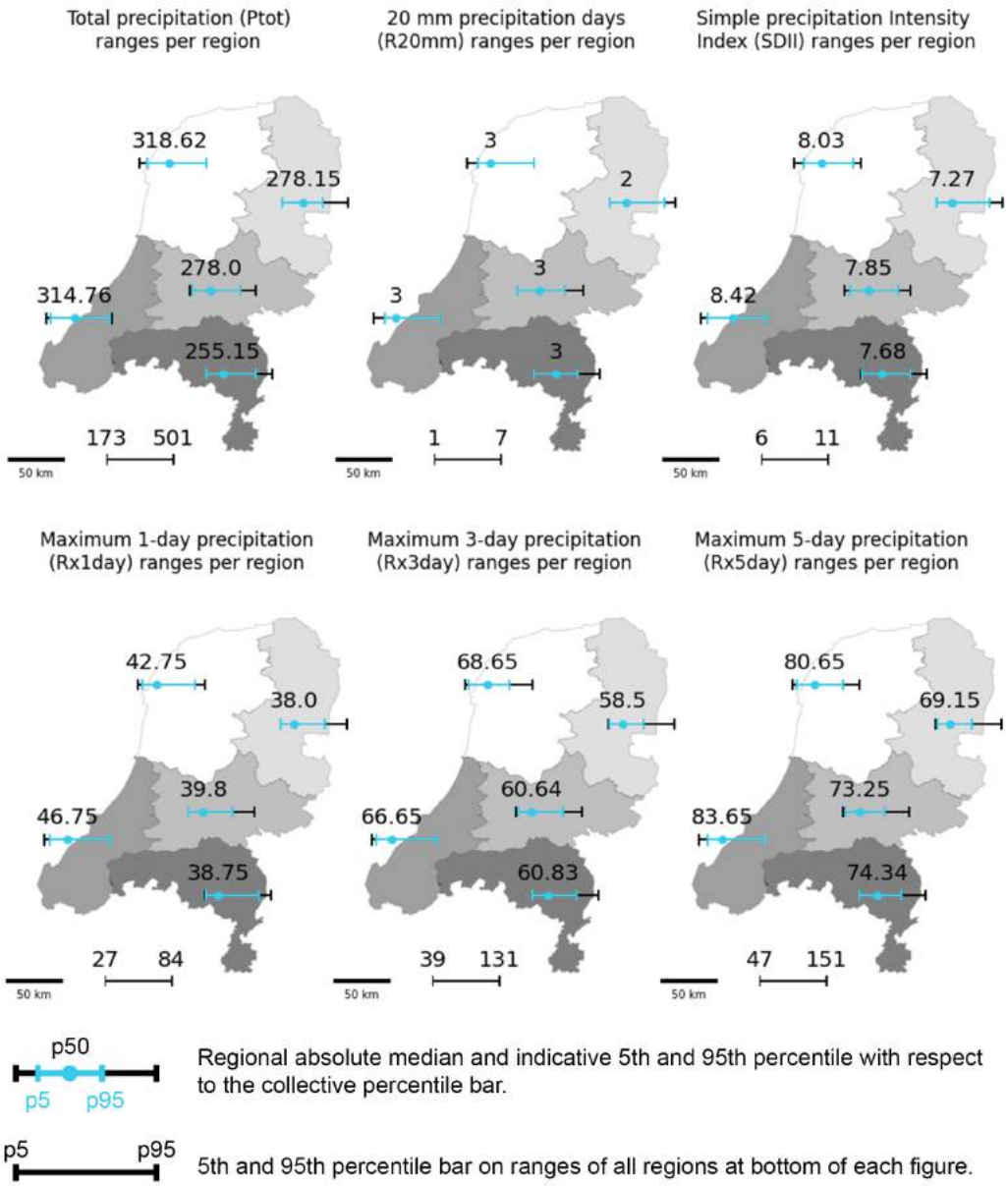



Figure 3.15: Ranges of the wet indicators per region (cyan) on a all region scale bar (black) in fall (September, October and November). The median is annotated and presented by the cyan dot. The minimum and maximum p5 and p95 are annotated by the black scale bar at the bottom of each figure.

.

## 3.6. Consecutive Event Graph (CEG)

The Consecutive Event Graph (CEG) is a practical framework introduced to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons (Figure 3.16). To place them in chronological order, all dry indicators are presented on the left half of the figure and all wet indicators on the right half of the figure. A vertical black line divides these wet and dry indicators. The dataset of section 3.5 has CDW extremes indexed per region and year. To present changes of CDW extremes from year to year in a single region (Y2Y) or region to region in a single year (R2R) the dataset is adapted. The quantities of indicators are placed in historical perspective per region by transforming to a standard uniform distribution. To do so, the empirical cdf of each indicator per region is calculated defined by

$$\hat{F}_n(x) = \frac{1}{n+1} \sum_{i=1}^{n} 1\{X_i \leq x\} \tag{3.10}$$

where

$$1\{X_i \leq x\} = \begin{cases} 1, & \text{if } X_i \leq x \\ 0, & \text{if otherwise} \end{cases}. \tag{3.11}$$

There is one exception, the SPI-3 is already standardized per region such that its computed cdf can be used to retrieve a normalized value. A small value for the SPI-3 (e.g. -2.00) corresponds with an extreme dry event, but is situated at the lower end of the cdf curve (i.e. near 0). Therefore the corresponding value of the SPI-3 on this cdf curve $F(x_i)_{SPI-3}$ is inverted

$$z_{i_{SPI-3}} = 1 - F(x_i)_{SPI-3}. \tag{3.12}$$

With the chosen transformation method, any observation on the axis of an indicator corresponds to the cumulative probability for a specific region i.e. the probability of occurrence of values less than or equal to the observation. These characteristics largely increase the readability of the figure, which is additionally strengthened by the indication of common percentiles in the figure (gray lines and black dot on axes).

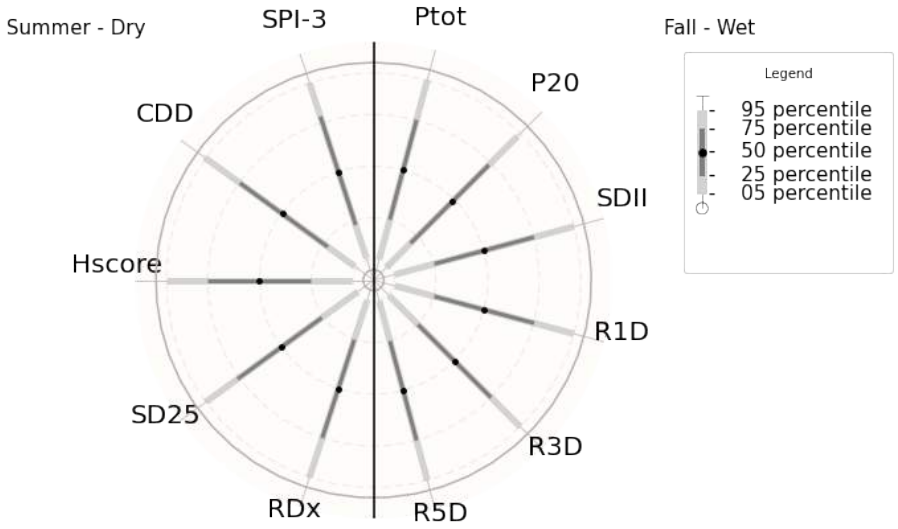



Figure 3.16: Base figure of the CEG created to visualize the quantities of dry and wet indicators. All dry and wet indicators are placed on a polar chart, whereof each indicator has its unique axis going from the center of the graph to the edge of the graph. The left half of the figure holds all dry indicators. Split up by a vertical black line, the right half of the figure holds all wet indicators. All indicators are transoformed to a standard uniform distribution by computing the empirical cdf of each indicator per region. A value close to 0 would represents a low extreme and will be observed at the inner circle. A value close to 1 represents a large extreme and will be observed at the outer edge. All values in between can be valuated with the help of the percentile bars present on each axis. The values indicated are found in the legend. As input, the figure requires seasonal maxima per each indicator.

## 3.7. Non Parametric Bayesian Networks

This study aims to improve the understanding and quantification of the interaction of CDW extremes by developing NPBNs. NPBNs are a powerful tool to model complex networks with a great number of variables and support probabilistic inference of unknown variables based on preconditioning of known variables (see section 2.4). Characteristics desirable considering the large number of indicators and aim of this study. Successful application of NPBNs requires the selection of network configuration(s) and testing of these networks. Hereafter the network can be used for making inference. These steps are elaborated in the following subsections.

### 3.7.1. Network selection

Selection of the network is a challenging task due to the large number of structures possible. The number of structures ($q_n$) grow super-exponentially with increasing number of variables ($n$) used for the Bayesian network, denoted as (Robinson-Garcia et al., 2020)

$$q_n = \sum_{k=1}^{n} (-1)^{k+1} \left(\frac{n}{k}\right) 2^{k(n-k)} q_{n-k} \tag{3.13}$$

Imposing the network by testing all network configurations (i.e. a "brute force" method) is possible, but would require too many resources to actually compute. Other than this, there are three options to learn the structure of a Bayesian network. The first option is from expert input, where the expert explains how the data are related to each other. The second option is to learn it from the data. In this case, one can search for dependencies manually or with the help of

learning algorithms. The third option is a combination of both fore-mentioned options.

For this study, we a priori selected one network configuration and impose a second network configuration by learning the network from data. The first network is a saturated network (SN). In this network configuration all nodes are interconnected by arcs, thus interdependence of all variables will be accounted for. For the second network we apply algorithm based structure learning to create a restricted network (RN). Learning algorithms for the second network are broadly subdivided into three categories (Marco Scutari, 2010):

- Constraint-based algorithms are algorithms that single out conditional independence with the help of statistical tests using the Markov property of the BN (Eq 2.1). The nodes that are found to be non-independent are linked if satisfying its corresponding direction-dependent separation statements. The methods are generally based on the Inductive Causation algorithm (Pearl, 1991).

- Score-based algorithms are algorithms that assign a score to every candidate DAG. This score is maximized with the help of a heuristic search algorithm (e.g. tabu search and hill-climbing).

- Hybrid algorithms are a combination of the above. Here the candidate DAG space for applying a score-based algorithm is restricted by an initial constraint-based algorithm.

The algorithm based structure learning is conducted using the *bnlearn* package v4.8 (Marco Scutari, 2010) in *R* v1.4.1106 (R Core Team, 2019). For the continuous setting, *bnlearn* makes the assumption of a multivariate Gaussian Bayesian network. This assumption could inflict faulty arcs, as not all indicators show Gaussian behaviour (Appendix C). However, a better alternative to learning the network did not exist (apart from the brute force method).

Learning the network is performed with the hill-climbing algorithm, which is a score-based algorithm. For the mixed approach (i.e. expert input and learning from the data) one can add a whitelist and blacklist of arcs to either force inclusion or exclusion of these arcs in the network, respectively. No expert was consulted to include or exclude arcs by expert judgement. However, arcs from indicators of wet events pointed towards indicators of dry events were blacklisted, as the past is independent of the future. One could debate an arc directed from any indicator towards the "region"-node, as no dry or wet event causes the location one may be. Nevertheless, these directions were not blacklisted. Arguments such as `restart` and `perturb` are used to increase the chance of reaching the global maximum score of possible network configurations. The `restart` argument defines the number of random restarts and the `perturb` argument defines the number of attempts adding, removing or reversing an arc per random restart. From here a single network configuration is obtained.

The robustness of the resulting network configuration is checked by performing a strength test procedure and a bootstrap procedure. The strength test procedure quantifies the strength of the probabilistic relationships of the arcs based on the Bayesian Information Criterion (bic). The bootstrap procedure creates a number of bootstrap replicates of equal size as the input dataset, where the hill-climbing algorithm with equal settings is used to impose the network form each bootstrap. The empirical frequency of each arc is calculated over all bootstraps. The robustness tests are used purely informative and no arcs were added or removed based on these test results.

### 3.7.2. Validation of NPBNs

Both the SN and RN were implemented as a NPBN. Validating these NPBNs involved two steps to make sure computations from the network are of value (Hanea, 2008). First, when fitting a copula to data one has to question whether this copula prescribes the dependence structure of the multivariate data well enough. Secondly, one has to question whether the network is an adequate model of the saturated graph given its conditional independence relations.

For the first validation question, Genest et al. (2007) introduced the application of the *Cramer-Von Mises* test on copulas by calculation of the distance between the empirical copula ($C_n$) and a parametric copula ($C_{\theta_n}$)

$$S_n = \sum_{i=1}^{n} \left( C_n(u_1, \dots, u_i) - C_{\theta_n}(u_1, \dots, u_i) \right)^2 \tag{3.14}$$

The lower the value of $S_n$, the better the fit of the parametric copula to the data. A value of zero would be a perfect fit. In this research we only test parametric copulas of the the Gaussian, Frank, Clayton and Gumbel families. These families are widely used in hydrological studies (Ragno et al., 2021). Explained in section A.0.1 these copulas show different behaviour, whereof Gaussian and Frank show fairly similar behaviour. The Clayton and Gumbel copulas show stronger tail dependencies of which the lower tail and upper tail respectively.

Another goodness-of-fit measure for NPBNs is the *d-calibration* score ($d_{cal}$) introduced by Morales et al. (Morales-Nápoles et al., 2014b). This score is calculated by subtracting the Heillinger distance ($d_H$) from 1.

$$d_{cal} = 1 - d_H \tag{3.15}$$

Here, the Heillinger distance is a measure to asses the similarity between two probability distributions. For the Gaussian case this is the similarity between the correlation matrices. The *d-calibration* score takes on values between 0 and 1 inclusive. The higher the score the more similar the two correlation matrices are. This goodness-of-fit measure is able to give an indication to both the validation steps by comparing the normal rank correlation matrix (NRC; the model) with the empirical rank correlation matrix (ERC; the data) and the Bayesian network rank correlation matrix (BNRC; selected network). Comparison of the ERC and the NRC informs the user whether the Gaussian copula describes the bivariate dependence structure between variables well. Comparison of the BNRC and NRC informs the user whether the joint normal copula is valid for the specific Bayesian network (non-saturated).

### 3.7.3. Conditioning of the NPBN

Both SN and RN are used to used for making inference of wet indicators by conditioning dry indicators. Forecasting weather data for a forthcoming season (i.e. seasonal prediction) is extremely useful in helping to plan for possible hazards in a number of domains such as flood risk management, agriculture or industry (Quesada et al., 2012). Inferring the wet indicators is done in two ways.

The first method is by conditioning the dry indicators and region on the input data. For the input data, the corresponding values of the wet indicators are known (Section 3.5). The 50th

percentile is calculated from the inferred wet indicators to obtain a single value. This allows to compare an estimated value for the inferred wet indicator with the corresponding value of the wet indicator for that region and dry conditions. Additionally, to test the accuracy of the model, each known wet indicator value is tested to fall in a quantile range from its inferred distribution. The number of values observed within this range is divided by the total length of the dataset i.e 280 (Figure 3.17, vertical direction). This is done for an increasing quantile range, where the quantile range is expanding from the median in both directions (Figure 3.17, horizontal direction). In other words, a quantile range of 2 is equal to the range between q49 and q51, and a quantile range of 4 is equal to the range between q48 and q52. This processes is repeated up to the quantile range of 98, representing the range between q01 and q99. The corresponding share of known values to fall in each tested quantile range is plotted against the quantile range.

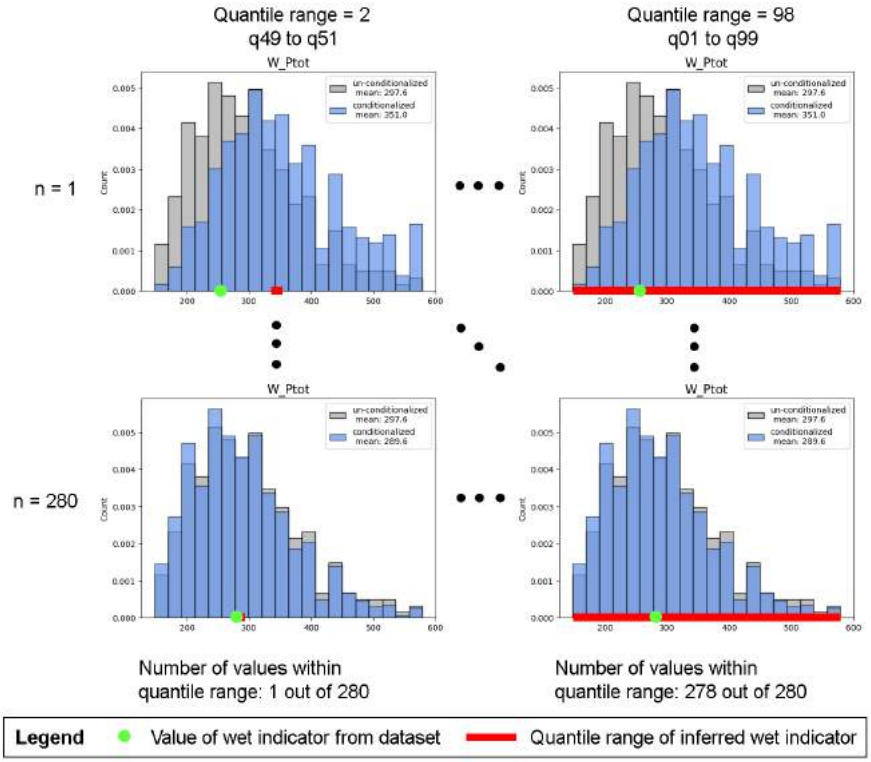



Figure 3.17: Visualization of the process to test each known wet indicator value to a quantile range from its inferred distribution applied to "W_Ptot". Here, $n$ is the length of the dataset. The green dot is the corresponding value of the wet indicator from the input dataset. The blue histogram is the distribution of the inferred indicator. The red bar is the considered quantile range and calculated from the inferred indicator. For all $n$ the quantile range differs as of different inferred wet indicator distribution. For all increasing quantile range, the number of values to fall within this range are expected to increase.

The second method is by conditioning the dry indicators on hypothetical extremely mild conditions to extremely dry conditions. The inferred distributions of each wet indicator are visualized by a box plots per dry conditions. This will allow comparison of changes in the distributions to detect any trend based on the changing dry conditions.

# 4
# Results

In this chapter the results related to the CEG and NPBNs are presented. First, section 4.1 presents application of the CEG relative to different years and different regions to understand changes in the wet and dry occurrences. Secondly, section 4.2 presents the results of the algorithm based structure learning, NPBNs testing and conditioning of the two different NPBNs.

## 4.1. Consecutive Events Graph (CEG)

Indicators for dry and wet extremes are used to quantify consecutive dry wet events in the CEG. The dry events are defined in summer period (June, July and August), where the dry indicators generally find their maximum values or largest increase (Section 3.2.1). Associated consecutive wet events are defined in fall period (September, October and November). This is the consecutive period where the largest values are observed for the wet indicators (Section 3.2.2). To understand and to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons, a year-to-year (Y2Y) comparison or region-to-region (R2R) comparison is presented. To allow such comparison the indicators were placed in historical perspective per region by transforming each indicator per region to a standard uniform distribution (section 3.6).

### 4.1.1. Year-to-Year comparison

#### Selecting years and region

To quantify changes in CDW extremes from year to year in a region, the CEG is suitable for comparison of multiple years in a single region. As an example, region 3 is selected which is expected to have an average climate relative to all other regions as it is located in the middle of the Netherlands(covering approximately provinces Utrecht and Overijssel).

A number of historical dry summers in the Netherlands based on rainfall deficit (RDx) is selected for comparison i.e. 1976, 2003, 2018 and 2020 (Jordi Huirne, 2020). The year 1976 was struck by a drought that covered over a million squared kilometers, where Netherlands was located close to its epicenter (KNMI, 1976). The year 2018 is commonly compared to the year 1976 in severity (Beleidstafel Droogte, 2019, Daniëls, 2021). Although of smaller areal size, it covered large parts of northwestern Europe and again its epicenter was located close to the Netherlands (Buras et al., 2020). The year 2003 was a drought year of lower severity for the Netherlands, here the epicenter was located more in the center of Europe (Buras et al., 2020). The year 2020 may not be classified as a drought year, but as a year with large water shortages in the Netherlands (IWR, 2021). There are several other years where dry conditions or water shortages

were present, however for the readability of the graph we limit ourselves to the four years stated above.

Out of years 2003, 2018 and 2020, KNMI (2022) only reports one extreme precipitation event on September 5 in 2018 located in The Green Heart (situated in the western part of region 3). Followed by two additional days of heavy precipitation events in Western Netherlands. Additional heavy precipitation events were observed in Northern Netherlands later that month. No report on precipitation events of Fall 1976 could be retrieved. In general, the fall period is characterized by large total precipitation differences across the Netherlands in this period KNMI (2022). The Y2Y comparison (Figure 4.1) allows relative comparison in the main figure (top) and absolute comparison by showing the corresponding values per indicator in a per year unique figure (bottom four figures). All on equal scale, with percentiles explained in the legend.

### Analysis of the Y2Y

All years show dry indicators in higher percentiles since such years were recorded as particularly dry. For the year 2020, the year where no drought was recorded, the SPI-3 shows a value approximately equal to the 50th percentile. This indicates that the water shortage was not driven by a lack of precipitation during the summer months, but other climatic factors. For example, the rainfall deficit (RDx; The difference between evapotranspiration and precipitation) shows a very high value for this year ($\approx$ 95%). Moreover, rainfall deficit was characterized by an early onset of in months April and May, compared to the other years where the largest proportion in increase of the RDx took place in the summer months (Figure 4.2, left). This could indicate high evapotranspiration rate, which, in this study, is estimated based on temperature. Information about temperature in the CEG are included in the indicators heat score (Hscore) and summer days (SD25). Both these indicators score relatively high ($\approx$ 90 and 85%, respectively) for this year. In more detail, the early increase of RDx is paired with a stagnation of Hscore in months April and May in 2020 (Figure 4.2). An increase of RDx can only be caused by evapotranspiration, which has a daily median value of 2.4 mm in the months April - September (Appendix D). Thus, the increase of RDx, of which the high values of evapotranspiration, took place irrespective of the Hscore in these months. On the other hand, for the parts where Hscore is increasing (generally in months June to August) one observes an increase in RDx as well (Figure 4.2).

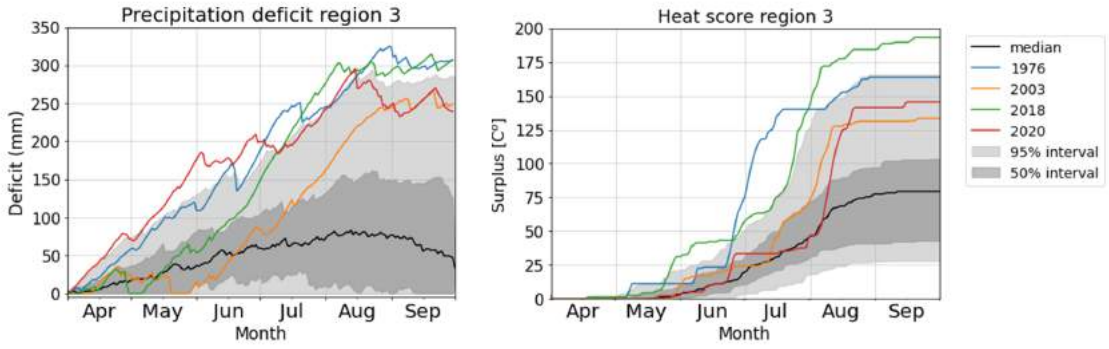



Figure 4.2: Propagation of precipitation deficit (left) and heat score (right) over the growing season for years 1976 blue), 2003 (yellow), 2018 (green) and 2020 (red) in region 3. In the background the indicator its median, 50% and 95% intervals from 1955 - 2020 in region 3 are visualized.

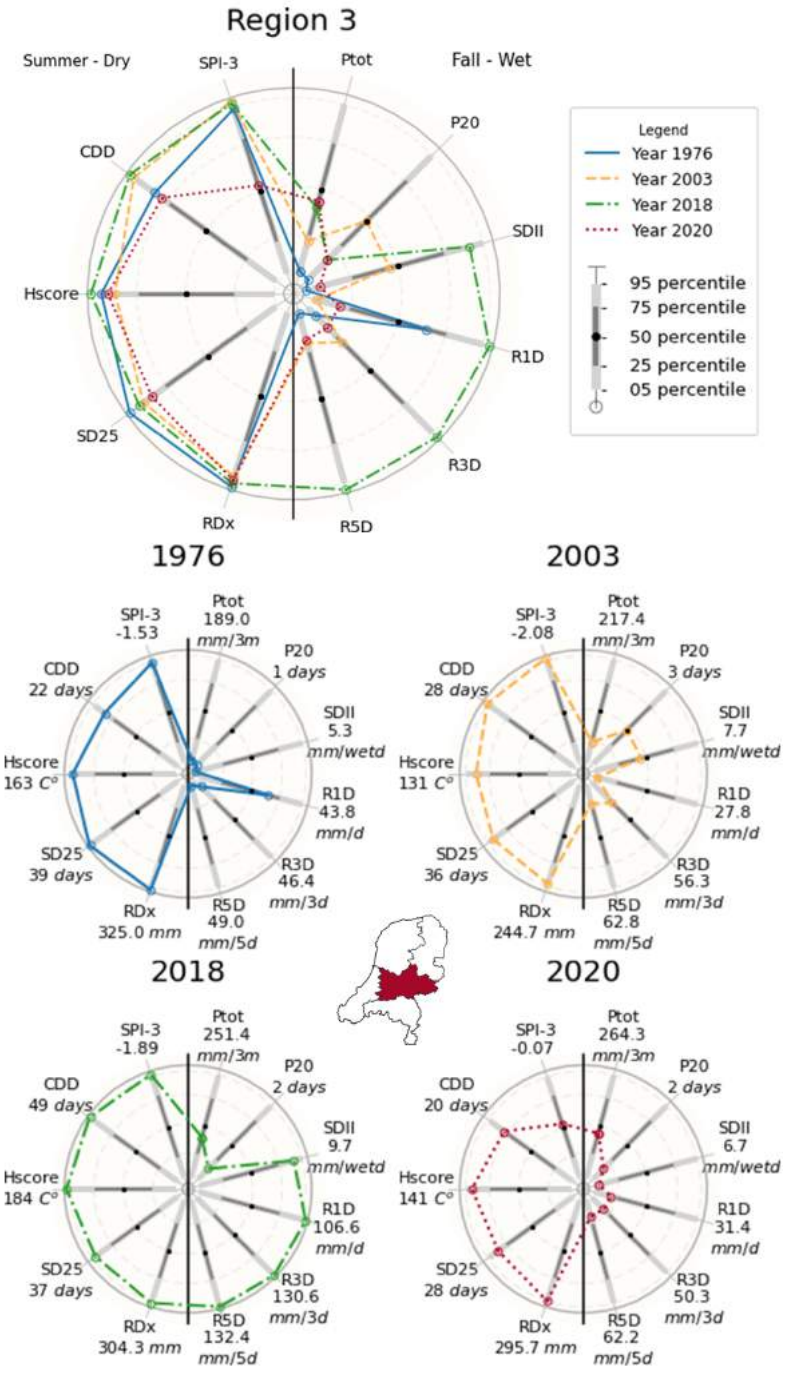



Figure 4.1: Year-to-Year comparison with CEG for dry and consecutive wet events. Application is performed on region 3, indicated in red on the small map. The selected years are 1976, 2003, 2018 and 2020 which are several historical dry years in the Netherlands. The large figure at the top allows comparison of the selected years. The four figures at the bottom present each year separately. The colors and line-styles are corresponding with the left figure and are unique per year. Additionally, the actual values are annotated on the single year figures, just below the indicator names.

In Appendix G the same years for all other regions are presented. Similar behaviour is observed for dry indicators across all regions in all four years. That means, for the drought years (1976, 2003 and 2018) high values for all dry indicators are observed. Across all regions a lower value is

observed for SPI-3 ($< 50\%$) and higher values for indicators heat score (Hscore), summer days (SD25) and precipitation deficit (RDx) (all $> 80\%$). The consecutive dry days (CDD) indicator in year 2020 shows a mixed picture, where values range between the 85 and 25 percentiles. High values ($> 80\%$) for wet indicator are observed in all years. In year 1976 the two northern regions show high wet indicators, where region 1 (indicators R1D, R3D and R5D) and region 2 (R1D). In year 2003 only in the north-eastern region i.e. region 2 (P20, SDII and R1D). In year 2018 the regions 3 and 4 (both SDII, R1D, R3D and R5D). In year 2020 the two western coastal regions show high wet indicators, where region 1 (R1D and R3D) and region 4 (R1D, R3D and R5D).

In table 4.1, the top five largest values of the dry indicators and their corresponding years are presented to provide a reference to the actual value of the indicators since indicators in CEG are normalized. In line with Jordi Huirne (2020), all years considered in figure 4.1 are in the top five of the RDx indicator. As expected, the highest indicators have been observed in the years 1976 and 2018, years in which the Netherlands experienced severe drought conditions (Hekman et al., 2019, Kramer et al., 2019).

Table 4.1: Top five largest values for dry indicators in region 3. All years used in figure 4.1 are in bold.

| Year | SPI-3 [–] | Year | RDx [$mm$] | Year | SD25 [$days$] | Year | Hscore [$^{\circ}C$] | Year | CDD [$days$] |
|---|---|---|---|---|---|---|---|---|---|
| **2003** | -2.08 | **1976** | 325.0 | **1976** | 39 | **2018** | 184.3 | **2018** | 49 |
| **2018** | -1.89 | **2018** | 314.6 | 1983 | 38 | 2006 | 176.7 | **2003** | 28 |
| 1983 | -1.86 | **2020** | 295.7 | 1995 | 38 | 1995 | 169.7 | 1995 | 28 |
| **1976** | -1.53 | **2003** | 266.5 | **2018** | 37 | **1976** | 163.8 | 1985 | 26 |
| 2013 | -1.44 | 1982 | 256.4 | 2006 | 37 | 1994 | 147.9 | 1970 | 25 |

Figure 4.1 shows an extremely dry and consecutive extremely wet for indicators SDII, R1D, R3D and R5D [1] for the year 2018. In table 4.2 the five largest values for indicators SDII, R1D, R3D and R5D for region 3 of year 2018 are presented including corresponding station ID's and date of occurrence. The R3D and R5D dates are moving sum values, where the date denoted in table 4.2 is that last day of the three or five days included. This means they include day 09-05 if the date for R3D is 09-07 and the date for R5D is 09-09. The SDII does not show its origin on a single date as it is calculated from all wet days (P>1mm). The high indicators values observed in 2018 (figure 4.1) include the same date (Date : 2018-09-05) and most probably from the reported storm conditions by KNMI (2022).

[1] Average precipitation on wet days, maximum one day precipitation, maximum three day precipitation, maximum five day precipitation respectively.

Table 4.2: Top five largest values for region three in year 2018 and corresponding station ID. NaN values are present where the station ID is not in the top 5 largest values of the wet indicator. The rows are sorted on the SDII indicator. For the RxD indicators the total sum of precipitation (P) in millimeters and date of occurrence (Date;[mm-dd]) is shown. For SDII there does not exist a single date, as it is the average intensity of wet days. The values used in the CEG of 4.4 are in bold.

| Station ID | R1D | | R3D | | R5D | | SDII |
|---|---|---|---|---|---|---|---|
| | P | Date | P | Date | P | Date | |
| 561 | **106.6** | 09-05 | **130.6** | 09-07 | **132.4** | 09-09 | **9.71** |
| 470 | 71.6 | 09-05 | 122.0 | 09-07 | 126.1 | 09-09 | 9.13 |
| 465 | NaN | - | NaN | - | NaN | - | 8.07 |
| 559 | NaN | - | 55.1 | 09-07 | 57.4 | 09-09 | 7.17 |
| 548 | 27.9 | 09-05 | 61.4 | 09-07 | 67.2 | 09-09 | 7.15 |
| 563 | NaN | - | 50.0 | 09-08 | 53.5 | 09-09 | NaN |
| 840 | 32.6 | 09-06 | NaN | - | NaN | - | NaN |
| 836 | 25.8 | 09-06 | NaN | - | NaN | - | NaN |

### 4.1.2. Region-to-Region comparison

#### Selecting year

The CEG can be used for comparison of multiple regions to quantify changes in CDW extremes from one region to another in a single year. The indicators are transformed to standard uniform per region, whereas drought conditions are relative to the climatology of the region. As an example for the region-to-region (R2R) comparison, the year 2018 is used, which is considered to be one of the driest summers of the Netherlands (Sluijter et al., 2018) (Figure 4.4). This also means that the region three displays an equal CEG to the region three year 2018 figure of the Y2Y-comparison (Figure 4.1). All other years are visualized in Appendix G.

All dry indicators in all regions show high values for dry its dry indicators ($> 0.8$). Only two regions show a consecutive extremely wet event, i.e. region 3 and 4 located at mid-East and South-West of the Netherlands respectively. Region 3 was already denoted in the previous paragraph. To denote region four in a similar way, table 4.3 shows the five largest values for indicators SDII, R1D, R3D and R5D of region four in 2018 and corresponding station ID's. Tables 4.3 and 4.2 both show extremes to occur around the same date (2018-09-05). Their locations are also clustered in the same area (Figure 4.3). This gives the reason to believe that the extreme values originate from the same precipitation event(s). This also means that current use of regions for coupling dry and wet events is sensitive to acknowledging a single extreme wet event over two (or possibly more) regions. This may give the impression that two extreme wet events took place such as in figure 4.4.

Table 4.3: Top five largest values for region four in year 2018 and corresponding station ID. The rows are sorted on the SDII indicator. For the RxD indicators the total sum of precipitation (P) in millimeters and date of occurance (Date) is shown. For SDII there does not exist a single date, as it is the average intensity of wet days. The values presented in figure 4.4 are in bold.

| Station ID | R1D | | R3D | | R5D | | SDII |
|---|---|---|---|---|---|---|---|
| | P | Date | P | Date | P | Date | |
| 442 | **118.5** | 09-05 | **190.4** | 09-07 | **194.0** | 09-09 | **14.16** |
| 454 | 46.2 | 09-05 | 91.3 | 09-07 | 94.1 | 09-09 | 9.83 |
| 458 | 47.2 | 09-05 | 101.5 | 09-07 | 109.3 | 09-09 | 9.36 |
| 439 | 53.0 | 09-05 | 84.3 | 09-07 | 86.2 | 09-09 | 9.22 |
| 443 | 46.8 | 09-05 | 75.2 | 09-07 | 79.6 | 09-09 | 8.4 |

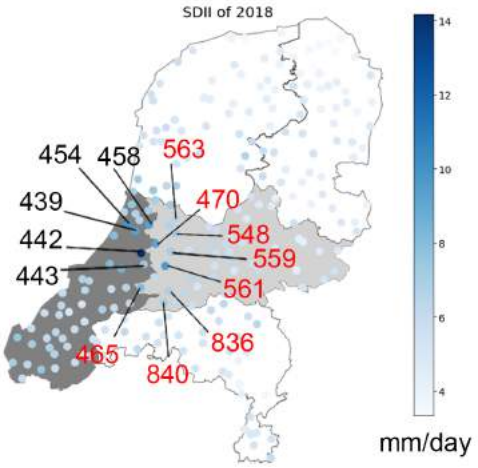



Figure 4.3: Visualization of the SDII indicator values of all KNMI-M stations for 2018. Region three and region four are highlighted in light gray and dark gray respectively. The location of stations stated in tables 4.2 and 4.3 are annotated, whereof stations for region three in red and stations of region four in black. The locations show to be clustered in the same area.

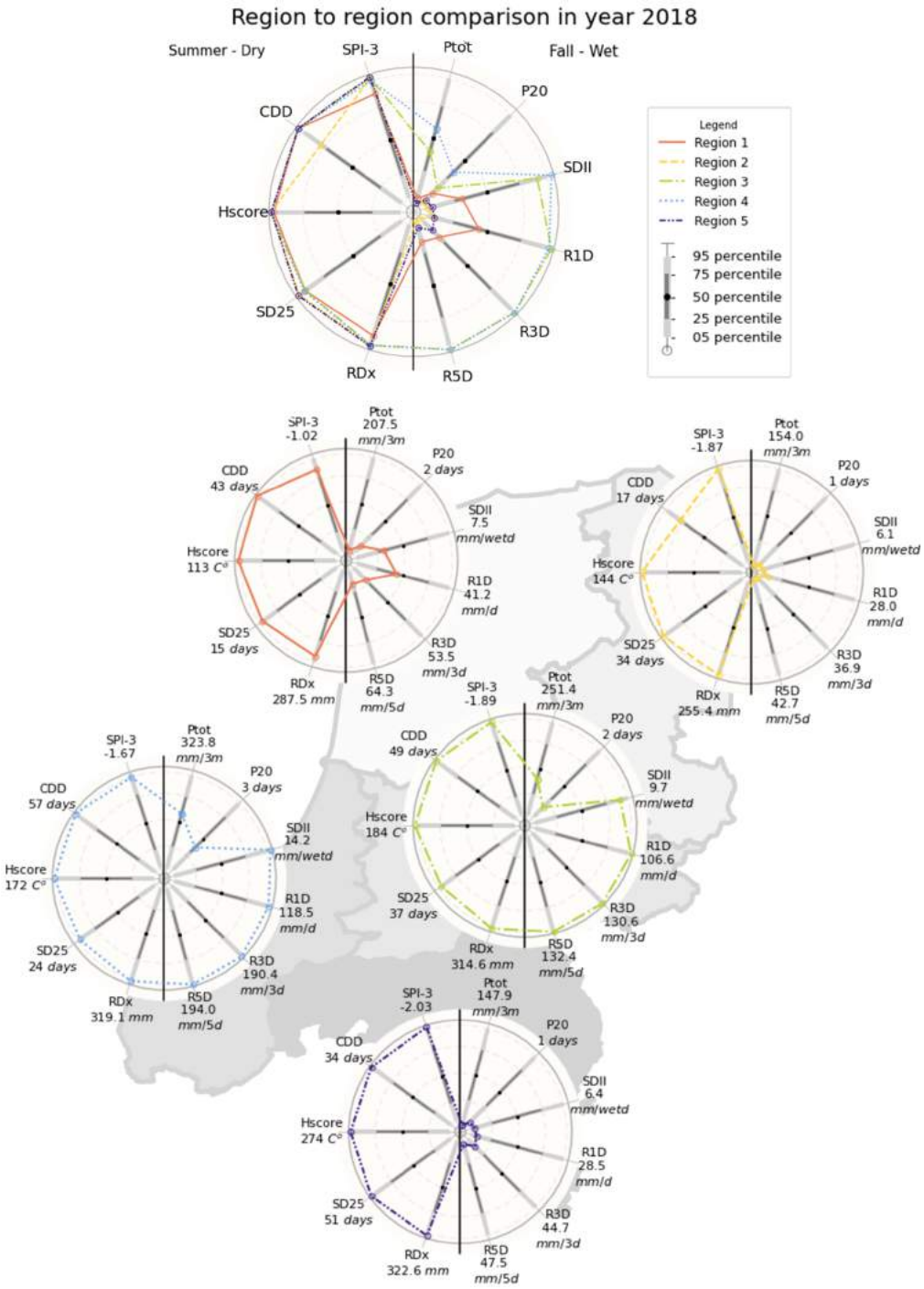



Figure 4.4: Region-to-Region comparison with CEG for dry and consecutive wet events. The year of comparison is 2018, which is considered as one of the driest years in the Netherlands (Sluijter et al., 2018). The large figure on the top the contains all regions for the desired comparison. The five figures on top of the map of the Netherlands presents each region separately and are located on their corresponding region in different gray-scales. The colors and line-styles are corresponding with the top figure and are unique per region. Additionally, the actual values are annotated on the single region figures, just below the indicator names.

## 4.2. Non-parametric Bayesian Networks

Non-parametric Bayesian Networks (NPBNs) are used to quantify the interaction between the dry and wet indicators and to assess expected wet conditions following a dry period. As we explore these interactions we introduce two networks. The first network is a saturated network (SN), which accounts for interdependence of all variables. The second network is a restricted network (RN) taught by a score-based structure learning algorithm. Here, each candidate network configuration is assigned a Bayesian Information Criterion score and with the help of a Hill-Climbing search algorithm this score is aimed to be maximized. The resulting RN shows a much smaller number of arcs compared to the SN (Figure 4.5). From here, first, the results from the strength test procedure and the bootstrap procedure are presented (subsection 4.2.1). Secondly, the results from validating the SN and RN as NPBNs are presented (subsection 4.2.2). Thirdly, the rank correlations of both networks are visualized (subsection 4.2.3). In the last place, the results of conditioning both the SN and RN on the input data and on hypothetical dry conditions are presented (subsection 4.2.4).

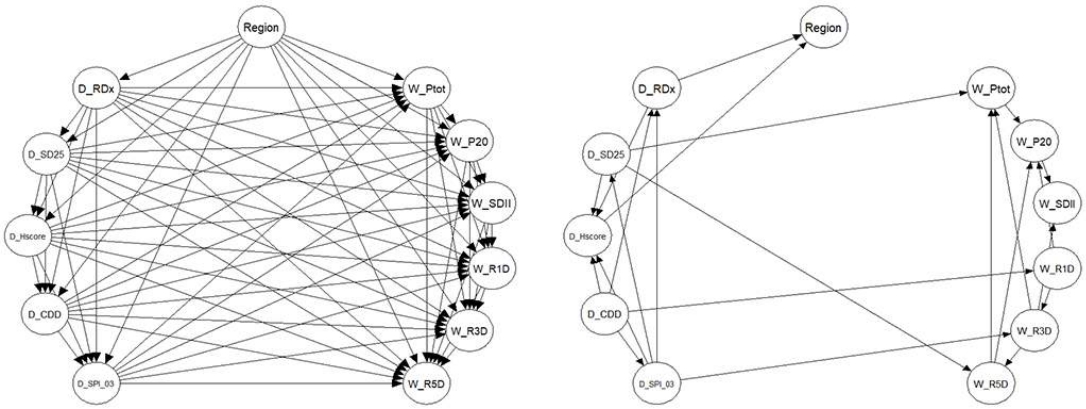



Figure 4.5: The saturated network (SN; left) and restricted network (RN; right).

### 4.2.1. Robustness of RN in *bnlearn*

The network resulting from the score-based structure learning algorithm on the input data shows a total of 22 arcs where all nodes are interconnected (Figure 4.5, right). Four arcs are directed from dry indicators to wet indicators. To construct this network, *bnlearn* makes the assumption of a Gaussian Bayesian network. The results of the strength test procedure and a bootstrap procedure are presented below. The tests are used to gain informative insight only.

Strength values of the probabilistic relationships represented by each arc in the RN network vary considerably (Figure 4.6). Values are negative, where smaller values represent greater dependence and vice versa. The smallest value observed of arcs between dry indicators is arc "D_SD25" > "D_Hscore" (-248.26). Both these dry indicators are calculated from temperature, which could explain the high dependency. All other arcs show values of -29.18 or higher. The largest value observed is arc "D_CDD" > "D_Hscore" (-3.55). The CDD is calculated from (a lack of) precipitation and Hscore on average temperature, which could explain to this limited dependency. The smallest values observed of arcs between wet indicators are "W_R1D" > "W_R3D" > "W_R5D" (-168.28; -313.42). These indicators are all based on short term precipitation values; one day, three days and five days respectively. All other arcs show values of -68.7 or higher. The largest value observed is arc "W_R3D" > "W_Ptot" (-1.5). Arcs crossing from dry to wet indicators generally show high to very high values, meaning their probabilistic

relationships is limited.

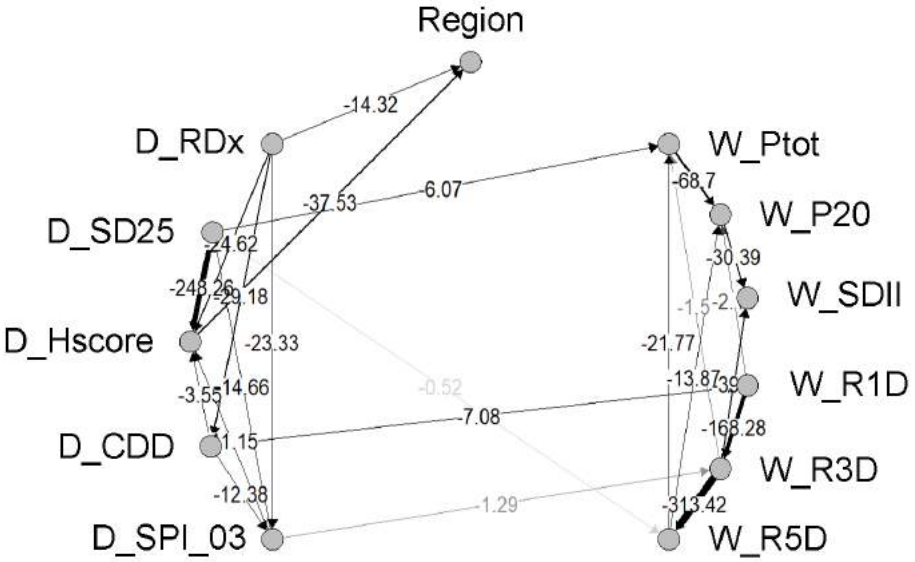



Figure 4.6: Arc strength values of network resulting from the hill-climbing greedy search algorithm. The values are indicated with a Bayesian Information Criterion Score. Scores are all negative, where a smaller score represents an greater dependence between variables. Greater dependence is represented with by increased line thickness and darker color.

Bootstrapping is performed to asses the stability of the network (Figure 4.7). Here all arcs with an empirical frequency larger than 0.3 are visible. Nodes may be connected by arcs in both directions (e.g. D_Hscore and Region) as this strength test runs the hill-climbing algorithm over $n$ bootstrap samples (in this study $n = 5000$). In each run an arc can be non-existent, existent and if existent aimed in either direction. All arcs between dry-dry and wet-wet indicators show high frequencies ($\geq 0.68$ and $\geq 0.54$ respectively). Of all arcs going from dry to wet indicators, the four arcs present in the RN have the largest empirical frequency ($\geq 0.54$). Generally, all arcs present in the RN have a frequency of 0.54 or higher. Only "D_Hscore" $>$ "Region" shows a lower frequency, but pairs with a counter directive arc of 0.42 meaning the arc is commonly present in either one of the directions. No frequencies of larger than 0.52 were observed for arcs not present in the RN. Comparing figures 4.6 and 4.7), some arcs find its way to have large empirical frequencies although limited dependence scores (e.g. "D_CDD" $>$ "D_Hscore": 0.75; -3.55).

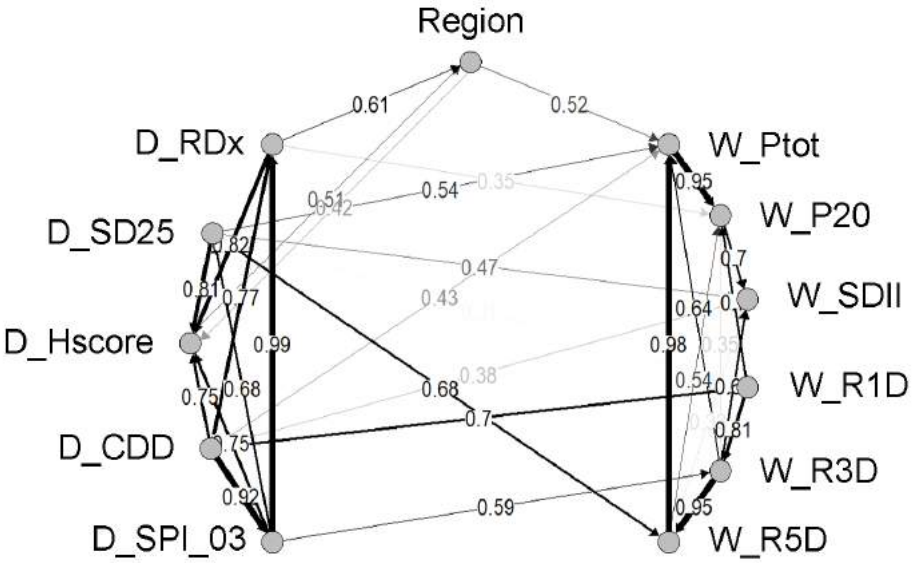



Figure 4.7: Empirical frequency of arcs for a set of networks learned from 5000 bootstrap replicates of same size as the original data. All arcs with a frequency smaller than 0.3 are removed. All arcs between the value of 0.3 and 0.6 are thin-lined and in gray scale. All arcs above 0.6 have an increasing line thickness and are in black.

### 4.2.2. Validation and verification of the NPBNs

To quantify the performance of both the SN and the RN as a NPBN the *Cramer-von-Mises* scores per pair of nodes and the *d-calibration* score are calculated. The CvM scores are generic to both the UN and RN. CvM statistics were calculated for the Clayton, Frank, Gaussian and Gumbel copulas following from equation 3.14. Figure 4.8 summarizes these statistics by presenting the percentage of pairs having the copula family stated as a best fit. For approximately 64% of the pairs a Frank copula or Gaussian copula are the best fit. This means the dependence structure of the considered nodes do not show a high or lower tail dependency (see section A.0.1). For those pairs the assumption of a Gaussian copula is fair.

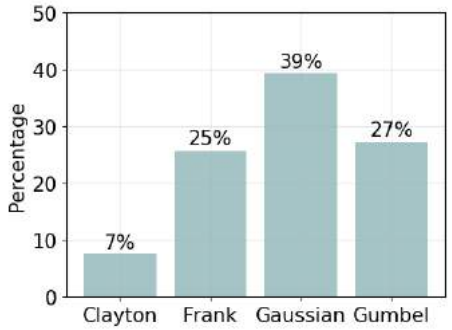



Figure 4.8: Cramer-von-Mises tests summary. Percentages of pairs having the copula family stated in the x-axis as best fit.

For the NPBN the Gaussian copula is used to build the joint distribution. Therefore, the CvM statistics for all pairs using the Gaussian copula are visualized (Figure 4.9). Pairs including variables 'Region','D_SD25', 'D_CDD' and 'W_P20' generally score high. Especially the 'Region' variable or 'W_P20' indicator. The three indicators take discrete values by definition

and the 'Region' variable is categorical (1 to 5), but are assumed to be continuous for application in Py_Banshee. This creates a very "binned" empirical copula which, which, when compared to any fitted copula, results in heightened CvM scores. Especially the 'Region' and 'W_P20' indicators, as they have a small range of 5 and 8 respectively. All pairs of the other variables show lower CvM scores. It should, however, be noted that the CvM statistics is a relative score for a pair and fitted copula. Therefore, differences in CvM scores for the Gaussian copula and their best fitted copula in case of a Gumbel or Clayton copula are presented (Figure 4.10).

The Gumbel or Clayton copula are suggested for pairs with stronger higher or lower tail dependencies. The differences vary, where pairs including the nodes 'Region' and 'W_P20' are most pronounced. It is likely these differences are larger due to the "binned" empirical copula. Most of existent pairs of dry indicators in figure 4.10 suggest higher tail dependencies. The same holds for the existent pairs of wet indicators. A mixture of better fit exists for the pairs between wet and dry indicators.

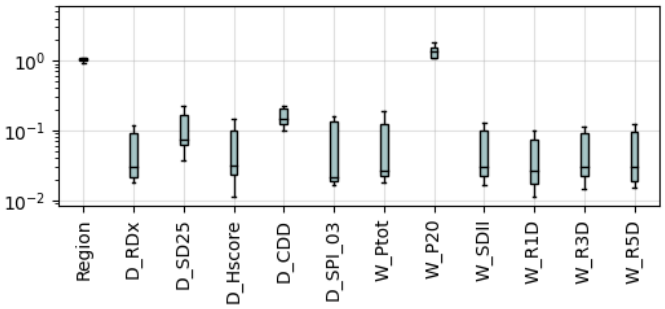



Figure 4.9: Boxplots of the Cramer-von-Mises scores for the Gaussian copula per indicator. High values are observed for pairs including the 'Region' node and 'W_P20' node. All other pair of nodes show a CvM score $< 0.23$.

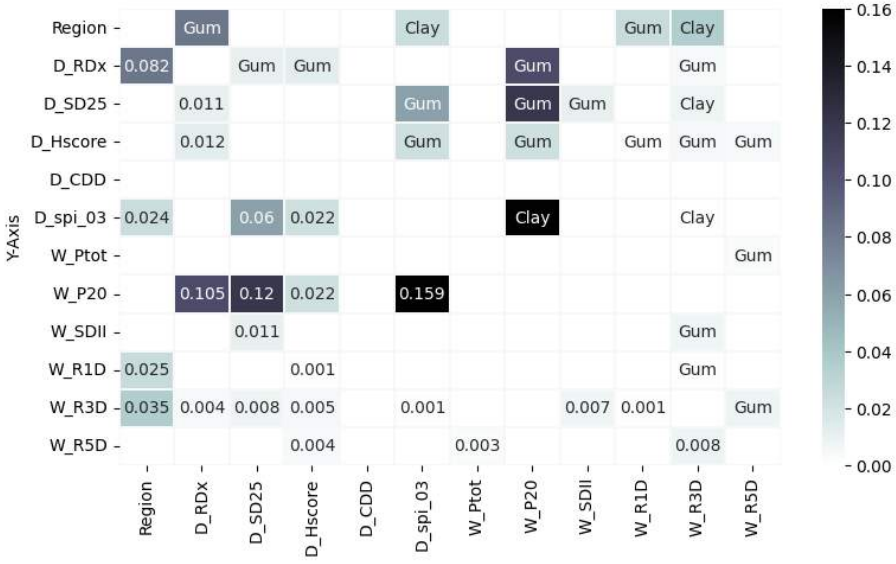



Figure 4.10: CvM scores comparison of the best fitted copula and Gaussian copula. The upper triangle shows the name of the best fitted copula, i.e. Gumbel (Gum) or Clayton (Clay). The lower triangle shows the CvM score of Gaussian copula minus the CvM score of the best fit copula. The larger the difference in CvM scores, the larger the number in the lower triangle and the higher the intensity of the color. Blank cells are present on the diagonal (non existent pairs) and for pairs where the best fit copula is either a Frank or Gaussian copula.

The *d-calibration score (ERC, NRC)* comparing the empirical rank correlation matrix with the empirical normal rank correlation matrix results in a score of 0.92 (2500 samples), which is within the [0.90, 0.93] confidence interval. For the restricted network the *d-calibration score (NRC,BNRC)* comparing the empirical normal rank correlation matrix with the Bayesian network rank correlation matrix results in a score of 0.80 (1200 samples), which is outside the [0.86, 0.90] confidence interval. This suggests that the joint normal copula is not valid for the particular structure. It should, however, be noted that this test can be rather severe on large data sets with a high number of variables (Hanea et al., 2015).

### 4.2.3. Dependencies of networks

The rank correlations of both SN and RN show strong dependencies between dry to dry and wet to wet indicators (Figure 4.11). Crossing dependencies from dry to wet are very limited. More details of the rank correlation matrices can be found in Appendix F.

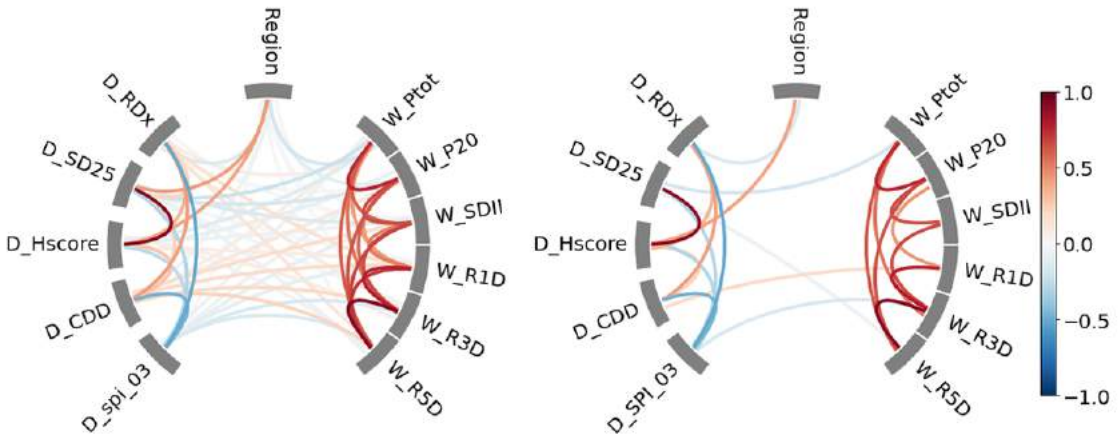



Figure 4.11: Saturated network (left) and restricted network (right) rank correlations visualized. The intensity of the color indicates the strength of the correlation, where red is a positive rank correlation value and blue is a negative rank correlation value.

### 4.2.4. Conditioning the networks

In this section both the SN and RN are conditioned on their dry condition to make inference of wet indicators. Forecasting expected wet conditions of the subsequent months of observed dry conditions in summer can be of value for planning for possible hazards leading from these conditions. Conditioning is a straightforward process, especially when using software like UNINET (UNINET Team, 2019). For example, the conditioned saturated network like in figure 4.12. When conditioning all dry indicators in the RN, indicators "D_SPI_03", "D_CDD" and "D_SD25" d-separates the other two dry indicators, leaving their conditioning to make no difference. The "Region" node is also d-separated and of no relevance in the RN.

Both networks their dry indicators are conditioned on the input data (Figure 4.13). Expected values for the wet indicators are retrieved by calculating the median of the inferred wet indicators from this conditioning. This conditioning shows very limited effect on wet indicators retrieved by stated approach. A well performing model would follow the calibration curve (diagonal line) in figure 4.13, where modelled values are equal to the original data. Not a single indicator of the RN or SN shows behaviour close to this calibration curve.

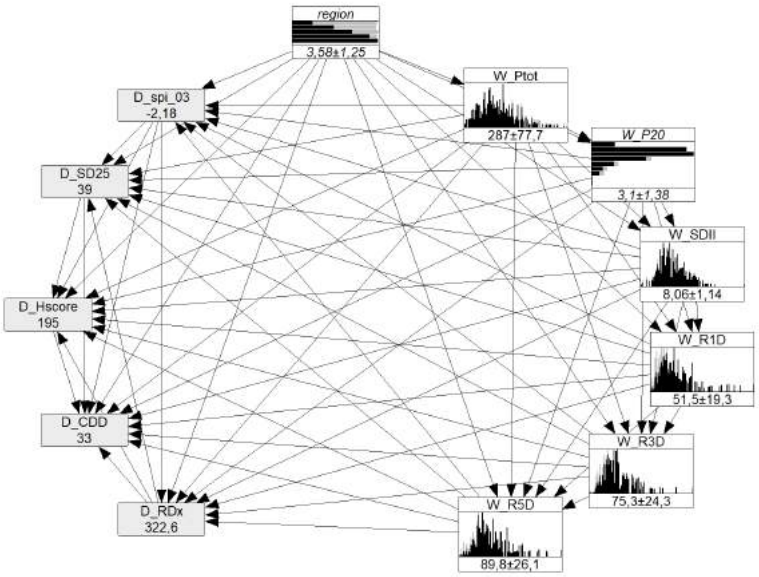



Figure 4.12: Saturated network conditioned on hypothetical extremely dry conditions in the Uninet software (UNINET Team, 2019). The inferred wet indicators only show minor changes in distributions based on the preconditions set.

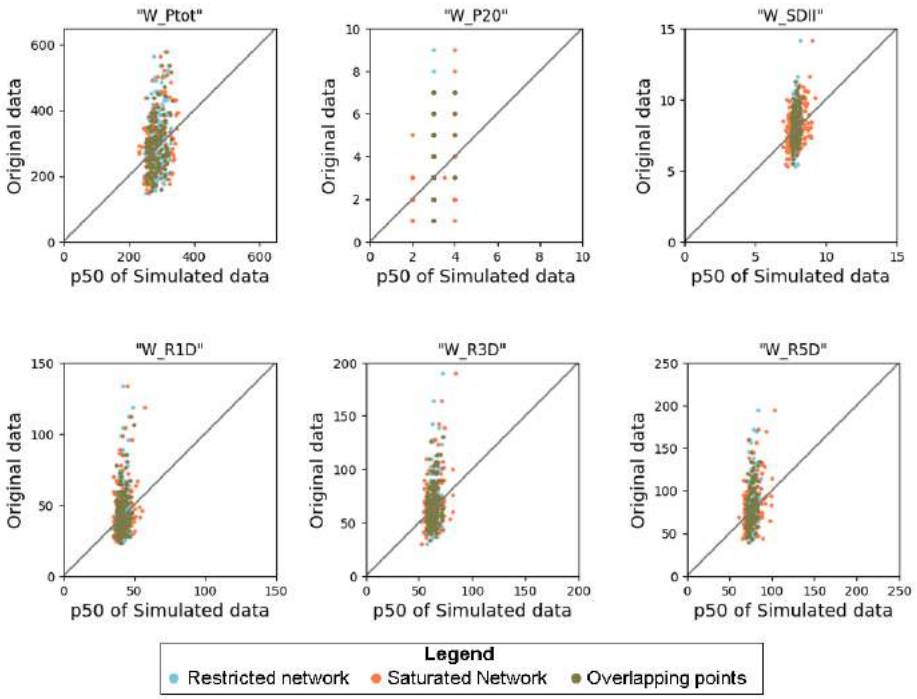



Figure 4.13: Median of inferred wet variables by conditioning the dry indicators from the input dataset versus corresponding values of the wet indicators from the input dataset. Restricted network in blue, Saturated network in orange and overlapping values in brown. A perfect model would follow the calibration curve (diagonal line). Both networks show substandard performance with stated approach.

To see if a quantile range would be more suitable to estimate wet indicator from the input dataset, each value of a wet indicator is tested to be within a quantile range of its inferred

distribution. The number of values observed within this range divided by the total number of values (i.e 280; length of the dataset) is visualized per increasing quantile range for each network (Figure 4.14). For increasing quantile range, the quantile range is expanding from the median in both directions. In other words, a quantile range of 2 is equal to the range between q49 and q51, and a quantile range of 4 is equal to the range between q48 and q52. This processes is repeated up to the quantile range of 98, representing the range between q01 and q99. For all indicators except "P20" and for both networks the number of data points increases near linearly by increasing quantile range. For both the RN and SN the "P20" indicator starts of at a number of approximately 30 values in the quantile range of 2. The somewhat step-wise increase observed for this indicator in the RN can be explained by the binned character of this value as observed in figure 4.13.

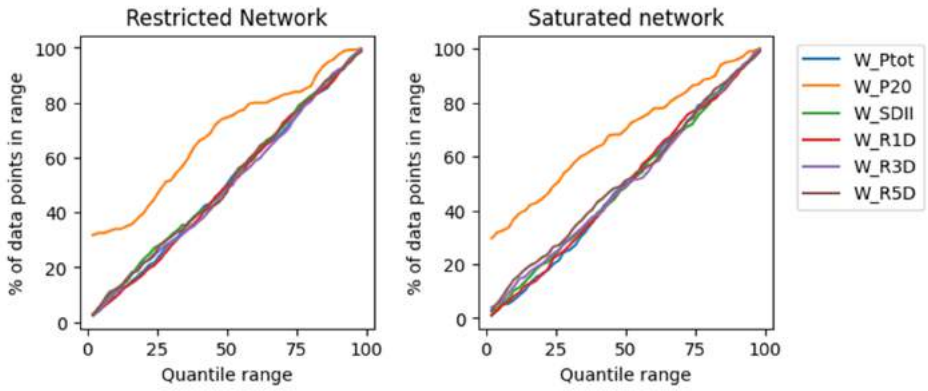



Figure 4.14: Percentage of wet indicator from input data within quantile range of inferred indicator against selected quantile range for the RN (left) and SN (right). For increasing quantile range, the quantile range is expanding from the median in both directions. All show a linear increasing number of data points with increasing quantile range of inferred wet indicators starting near zero. Except the P20 indicator, which shows a initial increased share in both networks.

Alternatively, both the SN and RN their dry indicators are conditioned on hypothetical dry conditions following table 4.4. With this conditioning of dry indicators, no corresponding wet indicator values are known. Figure 4.15 presents box plots of the values of each wet indicator conditioned on the dry indicators going from extremely mild summer conditions (left) to extremely dry conditions (right). Both the SN and the RN show a mild decrease in "W_Ptot" for mild to extreme dry conditions. In either network no effect is observed on "W_P20". For the "W_SDII" indicator a mild increase is observed both networks. The "W_R1D", "W_R3D" and "W_R5D" indicators all show a distinguishable increase from extremely mild to extremely dry summer conditions.

Table 4.4: Values of the dry indicator dependent on the summer condition. These values are used to pre-condition the restricted and saturated networks to quantify its possible effect on wet indicators.

| Summer condition | RDx [mm] | SD25 [days] | Hscore [$^oC$] | CDD [days] | SPI-3 [-] |
|---|---|---|---|---|---|
| Extremely mild | 5.33 | 1 | 3.2 | 7 | 2.08 |
| Very mild | 67.2 | 4 | 17 | 9 | 1.30 |
| Median | 157.0 | 13 | 62.1 | 14 | 0 |
| Very dry | 268.6 | 30 | 135 | 24 | -1.28 |
| Extremely dry | 322.6 | 39 | 195 | 33 | -2.18 |

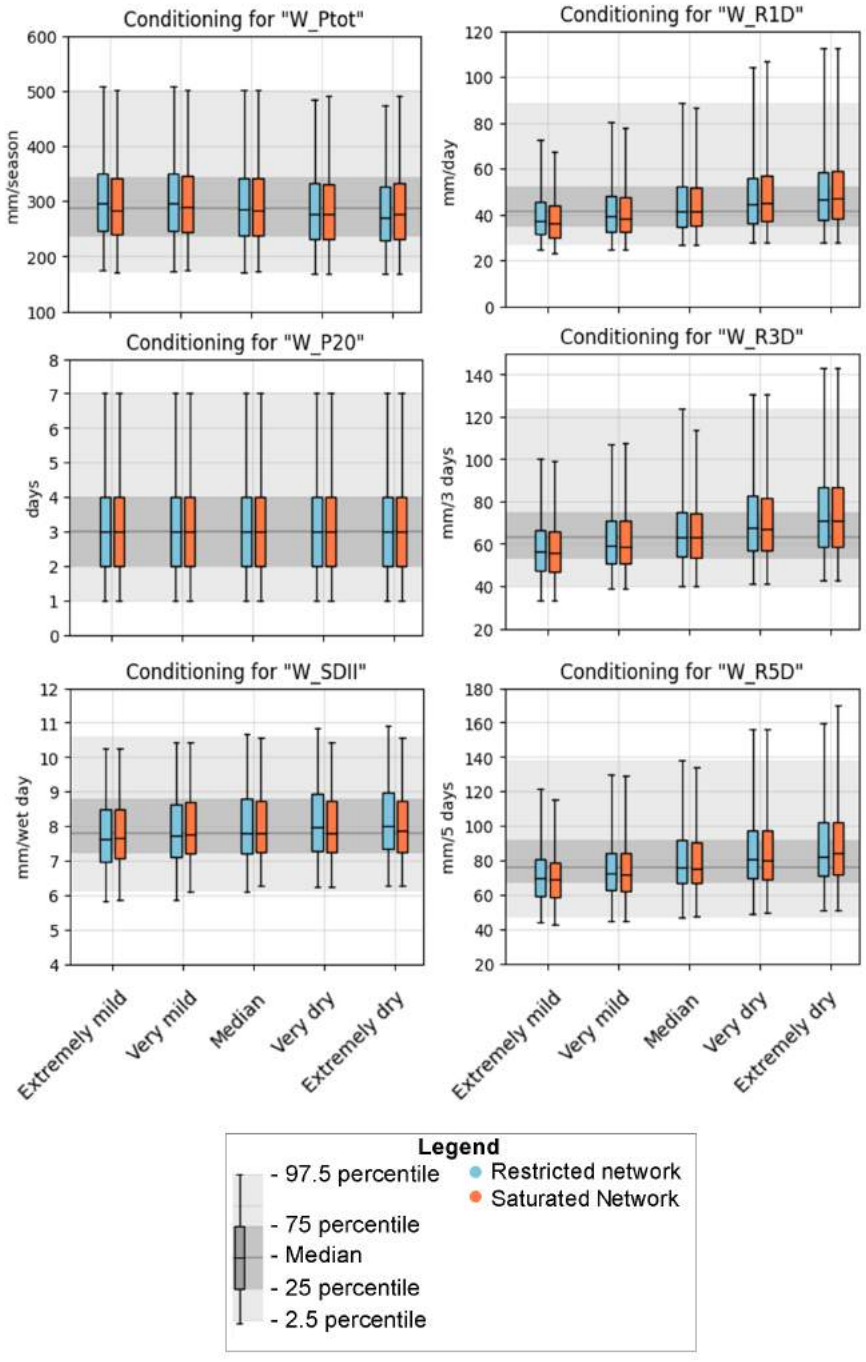



Figure 4.15: Summary of the conditioning performed on the restricted network (light blue box plots) and the saturated network (orange box plots). The original data is represented with corresponding quantiles (background shading) as a reference for the observed changes in distribution per wet indicator.

# 5

# Discussion and Recommendations

The Netherlands has an extensive amount of historical meteorological data allowing a systematic analysis of consecutive dry to wet (CDW) extremes over a period of fifty-six years. Research on the interaction between observed dry and wet events is still at its infancy. The aim of this thesis was therefore to improve knowledge on the interaction between dry and wet extremes with application to the Netherlands. First, settings were introduced to quantify the wet and dry extremes and to couple them spatial-temporally. The resulting dataset (or input data) was then applied two fold. Second, the Consecutive Events Graph (CEG) was introduced to quantify spatial and temporal changes in dry and wet indicators in consecutive seasons. Third, a fully probabilistic framework based on Non-parametric Bayesian Network was developed to model the dependence between dry and wet indicators. The results presented here are further discussed below. Possible processes that explain these results and sources that affect these results are highlighted.

## 5.1. Settings for the input data

The outcome of the CEG and the probabilistic framework are dependent on the chosen settings. These settings include a choice of indicators (to identify and quantify extremes), a choice of dry and wet periods (to link dry and wet temporally) and a choice of regions (to link dry and wet spatially). The settings used result from a straightforward approach. The chosen indicators are commonly used across literature or by the KNMI. The periods are selected based on the highest observed values or largest increases of the indicators. The dry period was required to always precede the wet period. The regions are selected such that wet and dry events were characterized on a regional scale. This approach allowed to identify and quantify consecutive dry to short-term wet events (e.g. cloudbursts) and consecutive dry to long-term wet events (e.g. sustained wet periods) on regional scales. Future observations can be used to update and thereby increase robustness of both applications. For example, year 2021 can now be added. Furthermore, decision-makers and other researchers can tailor the settings to identify and quantify other types of extremes, in different periods of interest and in other parts of the world. For example, a decision-maker may change the periods of interest based on their response time for a CDW event.

### Uncertainties and limitations

The defined summer period as dry period dot not always reflect the actual onset and ending of the dry period. Dry conditions can be persistent during the fall period as dry conditions in the Netherlands can stretch over more than three months (Daniëls, 2021). Short-term processes (i.e. in the order of days or weeks) that play a role in the interaction between dry conditions

and wet extremes may have an undesired effect on the current quantified wet extremes. For example, higher temperatures induce short-term convective precipitation extremes (Lochbihler et al., 2019). Dry conditions with accompanied higher temperatures persistent at the very end of the summer or beginning of fall could then lead to increased values of R1D, R3D, R5D, SDII and P20 quantified for the fall period.

Additionally, the current settings take little account of the areal magnitude of precipitation events. This translates itself into two issues. The first issue is the possible double quantification of short term precipitation events that either cover or propagate over multiple regions (as observed in the results). The second issue is that current wet indicators are quantified from a single precipitation station. This means that precipitation events of different size but with equal maximum precipitation quantities will be quantified as an equal extreme. A solution to this second issue would be an indicator that quantifies precipitation extremes over multiple neighbouring stations per period of choice. However, no such indicator was found across literature and introducing such indicator would be a tedious task. As an example, one could take the average of the "R1D" of the top five stations of a region. However, this would not guarantee the events took place on the same date nor would it guarantee stations to be neighbouring. And even if they are neighbouring, the area covered by these five stations would differ per set of stations. A more feasible solution to this would be the use of raster-data, but such data was not openly available. The inherent complexity to tackle this issue and time constraints for this study lead to disregarding this aspect to current extent.

## 5.2. Consecutive Events Graph

### Value of the CEG for identifying and quantifying consecutive events

The introduced CEG provides a framework to identify the occurrence and severity of CDW events. It synthesizes single events to a multi-event tool, which helps to move forward to a more comprehensive understanding of multi-hazard events. This study demonstrates that these events can be identified and quantified by introducing multiple climate indicators within the CEG. A straightforward tool that was not yet available across literature. The extremity of events is quantified by placing them in historical perspective. This way, decision-makers can use the CEG to get an understanding of the extremity of observed dry and wet events and various aspects of the dry events (e.g. temperature, precipitation) and wet events (e.g. long term and short term events). Other researchers may use the CEG for similar reasons whereby it can also be used as an exploratory tool for climate conditions of both extremes in the Netherlands.

### Uncertainties and limitations

As stated, the CEG visualizes extremes in historical perspective. This is done by transforming the indicators on a normal scale and presenting their values as a cumulative probability of occurrence. It should be noted that these statistical extremes may not result in (extreme) impacts. This can lead to misinterpretation of the extremities. Generally, the impact does not show a one-on-one relationship with the extremity of a climate or weather variable (Kumar et al., 2013). It is dependent on where and when the extreme event takes place or in terms of risk; the vulnerability and exposure also play a role. For consecutive events the quantification of impact can become increasingly complicated, whereas the preconditioning dry event may have different implications on the vulnerability of a system for the consecutive wet event to have impact. In other words, there is a multitude of combinations possible for the consecutive extremes where each combination can result in a different type of multi-hazard. A user of the CEG should draw up an inventory of the possible impacts of these combined events when using

it as a tool for e.g. risk management strategies or disaster preparedness. Such inventorization was not possible in this study due to time restrictions.

The CEG requires an extensive amount of historical (meteorological) data to place extremes in historical perspective. The study area in this research allows such analysis. However, in many cases the availability of similar data is limited, if not non-existent. A record length of at least 30 years is advised for a good estimated range of expected values for most indicators. If not available, the CEG might requires adjustment in settings or one should abstain from using the CEG under such conditions. This is required to prevent under-performance of the CEG.

## 5.3. Statistical model with NPBNs

### Value of statistical modeling with NPBNs

The statistical model used in this research provides quantitative insight on the interaction between CDW extremes. Application of Bayesian networks to natural hazards is still limited across literature (Tilloy et al., 2019). It is showed that the visual representation of NPBNs makes understanding of the interactions in CDW extremes more manageable. Additionally, making inference of wet indicators based on dry conditions and region is shown to be straightforward. This makes NPBNs well suited for research to interactions of natural hazards. The use case presented in this study will help to advance current modeling practices for the multi-hazard approach. As an application, the NPBNs can be used to forecast wet conditions in the fall period based on observed dry conditions in the summer period. It is showed that precise forecasting based on taking the 50th percentile is not suitable. Instead, the NPBNs should be used without further aggregation of the resulting distributions by inference. Forecasting of weather statistics for a forthcoming season (i.e. seasonal prediction) is extremely useful in helping to plan for possible hazards in fields such as health, agriculture or industry (Quesada et al., 2012).

### Interpretation of observed patterns by making inference

With the created inference of figure 4.15 we observed different patterns on the short-term and long-term wet events, which we will discuss separately. The interdependencies between these hydro-climatic variables can be observed across a large range of spatio-temporal scale as of coupling between Earth's atmosphere, land-surface and oceans (Leonard et al., 2014). However, extracting the exact physical processes behind the interaction of extremely dry and the consecutive extremely wet events can be rather difficult as of their complex nature.

Multiple short-term wet event indicators (R1D, R3D, R5D) and the SDII show a positive trend with increasing severity of dry preconditions in both NPBNs. The current framework does not discriminate for sustaining dry conditions in the early fall period. As stated in subsection 5.1 on the settings of the input data, it is possible that the dry conditions may persist at or near the beginning of the wet period and higher temperatures induce short-term convective precipitation extremes (Lochbihler et al., 2019). Daniëls (2021) shows a pairwise increase in duration and intensity of dry conditions based on a six months Standardized Precipitation Evapotranspiration Index. This could possibly cause the observed increase of indicators in both NPBNs. An additional explanation could be land-atmosphere coupling. Guillod et al. (2015), Petrova et al. (2018), Taylor et al. (2012) show that wet or dry soils can trigger precipitation events by positive or negative feedback mechanisms.

For the long-term wet events (Ptot) a small decreasing trend with increasing severity of dry preconditions in both NPBNs. The pairwise increase in duration and intensity presented by

Daniëls (2021) could be a possible explanation. Whereas dry conditions still persistent in the fall period can lead to decreased total amount of precipitation of that same period. Additionally, it is a possibility that large-scale dynamic processes have influence in the observed patterns. Common research directions of such processes are El Niño and La Niña Southern Oscillations, Atlantic Multidecadal Oscillation and changes in Arctic sea ice (He and Sheffield, 2020, Knight et al., 2006, Parry et al., 2013).

### Uncertainties and limitations

To learn the structure of the restricted network we assume the conditional probability densities to be linear Gaussians. The belief of this assumption to be inadequate is fair as multiple marginals show different behaviour than a Gaussian (Appendix C). The results from this method created a mixed picture where multiple logical and multiple illogical arcs or chains were created. For example, the "R1D" $>$ "R3D" $>$ "R5D" chain where it is likely that an extreme precipitation event in one day will be part of the largest multiple day precipitation event. On the other side, the arc from "Ptot" $>$ "P20" may be reversed as multiple 20 mm precipitation days will directly feed the total precipitation observed in a season instead of large total precipitation values to fall on 20 mm precipitation days. In general, the selection of the structure of a Bayesian network can be a challenging task. The number of structures grow super-exponentially with increasing number of variables used for the Bayesian network (Robinson-Garcia et al., 2020). Prior knowledge on relationships between dry and wet indicators in current framework is non-existent. This made learning the structure from the data with chosen method the best option available.

The "Region" node in both NPBNs is somewhat ambiguous. For the NPBNs the region is used as a variable in the network. This variable is categorical, but assumed to be continuous for application in PY_Banshee (Mendoza Lugo et al., 2021). Making the "Region" variable a node in the network was deemed necessary to have a lengthy number of data points per node, i.e. 280 instead of 56 per region. A longer dataset benefits the robustness of the network. For the RN the region node is D-separated when conditioning on all dry indicators. This made the node no problem for the RN within the application of this research. For the SN this is not the case. However, making variation in the order of values of this node showed very limited effects when making inference of wet indicators as performed in this thesis.

Non Parametric Bayesian Networks assumes the dependence structure to be a Gaussian copula (Kurowicka and Cooke, 2003). The adequacy of this assumption is evaluated statistically which shows that both the SN and RN are not able to model some important asymmetries observed in the data. This may lead to an over- or underestimation of inferred wet indicators when conditioning on dry indicators. To solve this issue a better dependence structure could be explored by using different copula families realized via e.g. the vine copula approach. This method was not considered as a complete theory of the vine copula approach conditionalization is non-existent to this date (Ragno et al., 2021).

## 5.4. Recommendations

Future research directions that either build upon or that could depart from this research are:

- An alternative approach on selection of the dry and consecutive wet period. The dry and wet period are defined by summer and fall respectively. However, we know that dry conditions can find their ending earlier or later in than the current set date. Defining

the exact ending of drought is a difficult objective. A potential solution to this is the Threshold Level Method (Heudorfer and Stahl, 2017)

- Instead of the alternative approach above, one could apply the methods of this research to a region where the dry season and rain season are more distinguishable. There are multiple parts of the world where such strong seasonality exist. For example, in SE-Asia multiple countries experience a strong rain season. No such strong division exist in the Netherlands. This also lead to the difficulty in defining the end of the dry season and start of the wet season. When selecting a different area in the world, one has to bear in mind the required historical meteorological data for the analysis.
- Introduce a wet indicator that includes the spatial extent of wet events. Current wet indicators do not account for this. Precipitation events recorded over multiple stations affect a larger area and should have a higher weight. Introducing a wet indicator that covers this aspect will allow better characterization of wet events and thus CDW events.
- To include additional indicators from other parts of the hydrological cycle. The current research is limited to meteorological variables only, but flood hazards originate from multiple other sources such as rivers or seas. Including indicators from these types of sources will create more comprehensive picture of possible consecutive extremely dry to extremely wet events.
- Couple the methods with climate models. The current study sets consecutive extremely dry and wet events in historical perspective. However, severity and frequency of both dry and wet extremes are expected to raise due to a changing climate (Alfieri et al., 2015, IPCC, 2012, Visser-Quinn et al., 2019). It would be interesting to include a future perspective by full coupling climate models, as the past may not be the best guide to the future.
- To work towards a unique indicator for dry and wet extremes. Dry conditions show to correspond with high values of all or most dry indicators. Wet conditions show to correspond with one or more high values of wet indicators. For example, one could condition the input dataset based on (arbitrary) thresholds for e.g. the sum of squared dry indicators and for one or more wet indicators. A suitable thresholds can be defined based on impacts (leading to the next point).
- An impact-driven approach, as promoted by (Leonard et al., 2014), (Lloyd-Hughes and Saunders, 2002) and (Zscheischler et al., 2019). The consecutive events analysis performed in this study is based on its drivers rather than their impact. The definition of impact is stakeholder specific. Starting from an hydraulic engineer perspective on flood hazards one could built out to include other stakeholders like ecologist, urban planners or emergency services.

As an additional remark, the multi-hazards approach shows increasing interest in recent years from all corners of the earth. In the course of this thesis, additional research was published on topics related to the multi-hazard approach (Rashid and Wahl, 2021, Sanuy et al., 2021, Ward et al., 2021, Zhang et al., 2021a, Zscheischler et al., 2021). A large variety of terms are introduced to describe the different types of multi-hazards causing confusion and redundancy. To overcome future fragmentation across literature it is of great importance to increase communications between those who are active in this field of research.

# 6 Conclusion

This research explored the concept of consecutive dry and wet extremes in the Netherlands. The introduced CEG shows potential to identify and quantify these CDW extremes. Region-to-region and year-to-year comparison is possible to quantify changes between the considered years or regions. Generally, all dry indicators show values in higher percentiles for all regions in years a drought was present. For several regions within these dry years, higher percentiles of one or more wet indicators is observed. The combination of these indicators in higher percentiles define the occurrence of a CDW extreme. Future research on the application of this CEG is recommended, e.g. by changing settings or by adapting to an impact approach.

Quantitative insight on the interaction between the dry and wet indicators is achieved by making use of NPBNs. The NPBNs disclosed limited interdependencies across dry and wet indicators. Conditioning the NPBNs to infer wet variables of the fall period based on dry conditions in the summer period and region is shown to be straightforward. Using the NPBNs for precise forecasting of expected wet conditions by calculating the median of inferred wet indicators based on observed dry conditions in the summer season is deemed unsuitable due to its low precision. Additionally, preconditioning of hypothetical dry conditions going from very mild to extremely dry revealed a number of trends in wet indicators when inferred. These are positive trends in short term precipitation event indicators (R1D, R3D and R5D) and the Simple precipitation Intensity Index (SDII). For the long term precipitation indicator (Ptot) a mild negative trend is observed for increasing dry conditions. Extracting the physical processes behind the interaction of extremely dry and the consecutive extremely wet events can be rather difficult as of their complex nature.

Extreme dry events, extreme wet events and consecutive occurrences of these events are inevitable. Furthermore, it is expected that these phenomena will occur more frequently and become more severe due to a changing climate (Alfieri et al., 2015, IPCC, 2012, Visser-Quinn et al., 2019). To this day research on CDW events is still in its infancy. Findings from this thesis will help to smooth the path towards better understanding of the identification, quantification and interaction of CDW events or multi-hazard events in general. From here on, we will be one step closer to the future practice in modeling climate extremes (Leonard et al., 2014).

# Bibliography


Adikari, Y. and Yoshitani, J. (2009). Global Trends in Water-Related Disasters : an insight for policymakers The United Nations World Water Development Report 3 Water in a Changing World. *Unesco*, page 28.

AghaKouchak, A. (2018). How do natural hazards cascade to cause disasters? *Springer Nature Limited*, (561):458 – 460.

Aghakouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., and Sadegh, M. (2020). Climate Extremes and Compound Hazards in a Warming World. *Annual Review of Earth and Planetary Sciences*, 48:519–548.

Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., and Salmerón, A. (2011). Bayesian networks in environmental modelling. *Environmental Modelling and Software*, 26(12):1376–1388.

Alfieri, L., Burek, P., Feyen, L., and Forzieri, G. (2015). Global warming increases the frequency of river floods in Europe. *Hydrology and Earth System Sciences*, 19(5):2247–2260.

Beersma, J., Hakvoort, H., Jilderda, R., Overeem, A., and Versteeg, R. (2019). Neerslagstatistiek en reeksen voor het waterbeheer 2019. page 205.

Beersma, J. J. and Buishand, T. A. (2007). Drought in the Netherlands - Regional frequency analysis versus time series simulation. *Journal of Hydrology*, 347(3-4):332–346.

Beleidstafel Droogte (2019). Nederland beter weerbaar tegen droogte - Eindrapportage Beleidstafel Droogte. pages 1–77.

Brunner, M. I., Slater, L., Tallaksen, L. M., and Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *Wiley Interdisciplinary Reviews: Water*, 8(3):1–32.

Buras, A., Rammig, A., and S. Zang, C. (2020). Quantifying impacts of the 2018 drought on European ecosystems in comparison to 2003. *Biogeosciences*, 17(6):1655–1672.

Chadwick, A., Glasson, J., and Therivel, R. (2005). *Introduction to Environmental Impact Assessment: Principles and Procedures, Process, Practice and Prospects*. Routledge, 2 edition.

Cherubini, U. and Luciano, E. (2002). Bivariate option pricing with copulas. *Applied Mathematical Finance*, 9(2):69–85.

Czado, C. (2019). *Analyzing dependent data with vine copulas: A practical guide with R*, volume 222.

Dai, L., Wörner, R., and van Rijswick, H. F. (2018). Rainproof cities in the Netherlands: approaches in Dutch water governance to climate-adaptive urban planning. *International Journal of Water Resources Development*, 34(4):652–674.

Daniëls, E. (2021). Recente droogtes in historisch perspectief.

De Bruin, H. A. and Stricker, J. N. (2000). Evaporation d'une pelouse dont le sol est bien alimenté en eau. *Hydrological Sciences Journal*, 45(3):391–406.

de Ruiter, M. C., Couasnon, A., van den Homberg, M. J., Daniell, J. E., Gill, J. C., and Ward, P. J. (2020). Why We Can No Longer Ignore Consecutive Disasters. *Earth's Future*, 8(3).

Di Baldassarre, G., Martinez, F., Kalantari, Z., and Viglione, A. (2017). Drought and flood in the Anthropocene: Feedback mechanisms in reservoir operation. *Earth System Dynamics*, 8(1):225–233.

Didde, R. (2021). *Nederland Droogteland*. Uitgeverij Lias B.V.

Dracup, J. A., Lee, K. S., and Paulson, E. G. (1980). On the definition of droughts. *Water Resources Research*, 16(2):297–302.

Eden, J. M., Kew, S. F., Bellprat, O., Lenderink, G., Manola, I., Omrani, H., and van Oldenborgh, G. J. (2018). Extreme precipitation in the Netherlands: An event attribution case study. *Weather and Climate Extremes*, 21(November 2017):90–101.

Elidan, G. (2013). Copulas in Machine Learning. pages 39–60.

Embrechts, P., Lindskog, F., and Mcneil, A. (2003). Modelling Dependence with Copulas and Applications to Risk Management. *Handbook of Heavy Tailed Distributions in Finance*, pages 329–384.

European Union (2007). Directive 2007/60/EC of the European Counil and European Parliment of 23 October 2007 on the assessment and management of flood risks. *Official Journal of the European Union*, (2455):27–34.

Federal Emergency Management Agency (1999). Types of Flooding.

Genest, C., Favre, A. C., Béliveau, J., and Jacques, C. (2007). Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research*, 43(9):1–12.

Gilissen, H. K. (2014). The integration of the adaptation approach into EU and dutch legislation on flood risk management. *Journal of Water Law*, 24(3-4):157–165.

Gill, J. C. and Malamud, B. D. (2014). Reviewing and visualizing the interactions of natural hazards. *Reviews of Geophysics*, 52(4):680–722.

Glantz, M. H. and Katz, R. W. (1977). When is a drought a drought? *Nature*, 267(5608):192–193.

Guenang, G. M. and Mkankam Kamga, F. (2014). Computation of the standardized precipitation index (SPI) and its use to assess drought occurrences in Cameroon over recent decades. *Journal of Applied Meteorology and Climatology*, 53(10):2310–2324.

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., and Seneviratne, S. I. (2015). Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nature Communications*, 6.

Güneralp, B., Güneralp, I., and Liu, Y. (2015). Changing global patterns of urban exposure to flood and drought hazards. *Global Environmental Change*, 31:217–225.

Guttman, N. B. (1999). Accepting the standardized precipitation index: A calculation algorithm. *Journal of the American Water Resources Association*, 35(2):311–322.

Hanea, A., Morales Napoles, O., and Ababei, D. (2015). Non-parametric Bayesian networks: Improving theory and reviewing applications. *Reliability Engineering and System Safety*, 144:265–284.

Hanea, A. M. (2008). *Algorithms for non - parametric bayesian belief nets*. PhD thesis, TU Delft.

Hanea, A. M., Kurowicka, D., and Cooke, R. M. (2006). Hybrid method for quantifying and analyzing bayesian belief nets. *Quality and Reliability Engineering International*, 22(6):709–729.

Hanwei, Y. and Shixuan, Z. (2012). 2011 . 36(5).

Hao, Z., Singh, V. P., and Hao, F. (2018). Compound extremes in hydroclimatology: A review. *Water (Switzerland)*, 10(6):16–21.

Hazarika, S., Biswas, A., and Shen, H. W. (2018). Uncertainty Visualization Using Copula-Based Analysis in Mixed Distribution Models. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):934–943.

He, X. and Sheffield, J. (2020). Lagged Compound Occurrence of Droughts and Pluvials Globally Over the Past Seven Decades. *Geophysical Research Letters*, 47(14).

Hegger, D. L., Driessen, P. P., Dieperink, C., Wiering, M., Raadgever, G. T., and van Rijswick, H. F. (2014). Assessing stability and dynamics in flood risk governance: An empirically illustrated research approach. *Water Resources Management*, 28(12):4127–4142.

Hekman, A., Läkamp, R., van der Kooij, S., van de Velde, I., and van Hussen, K. (2019). Economische schade door droogte in 2018 | Rapport | Rijksoverheid.nl. pages 28–31.

Herrera-Estrada, J. E., Satoh, Y., and Sheffield, J. (2017). Spatiotemporal dynamics of global drought. *Geophysical Research Letters*, 44(5):2254–2263.

Heudorfer, B. and Stahl, K. (2017). Comparison of different threshold level methods for drought propagation analysis in Germany. *Hydrology Research*, 48(5):1311–1326.

IPCC (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation*, volume 3.

IPCC (2021). Annex VII: Glossary. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, (August):73.

IWR (2021). Droogteseizoen 2020. Technical report, WMCN-LCW.

Jahn, M. (2015). Economics of extreme weather events: Terminology and regional impact models. *Weather and Climate Extremes*, 10:29–39.

Jeuken, A., van Beek, E., van Duinen, R., Bocalon, A., Delsman, J., van der Zee, S., Stuyfand, P., Zuurbier, K., Creusen, R., Appelman, W., Katschnig, D., Rozema, J., Kwakkel, J., Thissen, W., Veraart, J., Tolk, L., and de Vries, A. (2012). Balancing supply and demand of fresh water under increasing drought and salinisation in the Netherlands. Technical report.

Ji, Z., Li, N., and Wu, X. (2018). Threshold determination and hazard evaluation of the disaster about drought/flood sudden alternation in Huaihe River basin, China. *Theoretical and Applied Climatology*, 133(3-4):1279–1289.

Jonkman, S. N. (2005). Global perspectives on loss of human life caused by floods. *Natural Hazards*, 34(2):151–175.

Jordi Huirne (2020). 2020 bij 8 droogste jaren sinds 1906.

Kappes, M. S., Keiler, M., von Elverfeldt, K., and Glade, T. (2012). Challenges of analyzing multi-hazard risk: A review. *Natural Hazards*, 64(2):1925–1958.

Knight, J. R., Folland, C. K., and Scaife, A. A. (2006). Climate impacts of the Atlantic multidecadal oscillation. *Geophysical Research Letters*, 33(17):2–5.

KNMI (1976). Meteorologische aspekten van de droge zomer van 1976. (mm):1976–1978.

KNMI (2000). *Handboek Waarnemingen*, volume 1.

KNMI (2018). Quality enhancement of quantitative precipitation estimates.

KNMI (2021a). Daggegevens van het weer Nederland.

KNMI (2021b). Daily homogenised precipitation stations.

KNMI (2021c). Extreme neerslag.

KNMI (2021d). KNMI Klimaatsignaal'21. page 75.

KNMI (2022). Archive months and season overview.

Koninklijk Nederlands Metereologisch Instituut (2015). KNMI'14 Klimaatscenario's voor Nederland; Leidraad voor professionals in klimaatadaptatie. *Herziene uitgave 2015*, page 34.

Kramer, N., Mens, M., Beersema, J., and Kielen, N. (2019). Hoe extreem was de droogte van 2018? *H2O-Online*, page 7.

Kreibich, H., Blauhut, V., Aerts, J. C., Bouwer, L. M., Van Lanen, H. A., Mejia, A., Mens, M., and Van Loon, A. F. (2019). How to improve attribution of changes in drought and flood impacts. *Hydrological Sciences Journal*, 64(1):1–18.

Krysanova, V., Buiteveld, H., Haase, D., Hattermann, F. F., van Niekerk, K., Roest, K., Martínez-Santos, P., and Schlüter, M. (2008). Practices and lessons learned in coping with climatic hazards at the river-basin scale: Floods and droughts. *Ecology and Society*, 13(2).

Kuksina, L. and Golosov, V. (2020). Flash floods: Formation, study and distribution. *E3S Web of Conferences*, 163(January).

Kumar, A., Chen, M., Hoerling, M., and Eischeid, J. (2013). Do extreme climate events require extreme forcings? *Geophysical Research Letters*, 40(13):3440–3445.

Kurowicka, D. and Cooke, R. (2003). Distribution - Free Continuous Bayesian Belief Nets.

Leelaruban, N. and Padmanabhan, G. (2017). Drought occurrences and their characteristics across selected spatial scales in the contiguous United States. *Geosciences (Switzerland)*, 7(3).

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., Mcinnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdisciplinary Reviews: Climate Change*, 5(1):113–128.

Lloyd-Hughes, B. (2012). A spatio-temporal structure-based approach to drought characterisation. *International Journal of Climatology*, 32(3):406–418.

Lloyd-Hughes, B. (2014). The impracticality of a universal drought definition. *Theoretical and Applied Climatology*, 117(3-4):607–611.

Lloyd-Hughes, B. and Saunders, M. A. (2002). A drought climatology for Europe. *International Journal of Climatology*, 22(13):1571–1592.

Lochbihler, K., Lenderink, G., and Siebesma, A. P. (2019). Response of Extreme Precipitating Cell Structures to Atmospheric Warming. *Journal of Geophysical Research: Atmospheres*, 124(13):6904–6918.

Madakumbura, G. D., Kim, H., Utsumi, N., Shiogama, H., Fischer, E. M., Seland, ., Scinocca, J. F., Mitchell, D. M., Hirabayashi, Y., and Oki, T. (2019). Event-to-event intensification of the hydrologic cycle from 1.5 °C to a 2 °C warmer world. *Scientific Reports*, 9(1).

Makkink, G. (1957). Testing the Penman formula by means of lysimeters. *Journal of the Institution of Water Engineers*, 11(3):277–288.

Manning, C., Widmann, M., Bevacqua, E., Van Loon, A. F., Maraun, D., and Vrac, M. (2018). Soil moisture drought in Europe: A compound event of precipitation and potential evapotranspiration on multiple time scales. *Journal of Hydrometeorology*, 19(8):1255–1271.

Marco Scutari (2010). Learning Bayesian Networks wirth bnlearn R package. *Journal of Statistical Software*, 35(3).

May, F. (2007). Cascading Disaster Models in Postburn Flash Flood. *International Journal of Disaster Risk Reduction*, 12:350–356.

McKee, T., Doesken, N. J., and Kleist, J. (1993). The relationship of drought frequency and duration to time scales.

McNeil, A. J., Rüdiger, F., and Embrechts, P. (2015). *Quantitative risk management : concepts, techniques and tools*. Number July. Princeton University Press.

Mendoza Lugo, M. A., Koot, P., and Morales Napoles, O. (2021). Py_BANSHEE.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391(1-2):202–216.

Moftakhari, H. R., AghaKouchak, A., Sanders, B. F., Allaire, M., and Matthew, R. A. (2018). What Is Nuisance Flooding? Defining and Monitoring an Emerging Challenge. *Water Resources Research*, 54(7):4218–4227.

Montz, B. E. and Gruntfest, E. (2002). Flash flood mitigation: Recommendations for research and applications. *Environmental Hazards*, 4(1):15–22.

Morales-Nápoles, O., Delgado-Hernández, D. J., De-León-Escobedo, D., and Arteaga-Arcos, J. C. (2014a). A continuous Bayesian network for earth dams' risk assessment: Methodology and quantification. *Structure and Infrastructure Engineering*, 10(5):589–603.

Morales-Nápoles, O., Hanea, A. M., and Worm, D. T. (2014b). Experimental results about the assessments of conditional rank correlations by experts: Example with air pollution estimates. *Safety, Reliability and Risk Analysis: Beyond the Horizon - Proceedings of the European Safety and Reliability Conference, ESREL 2013*, (2009):1359–1366.

Munich Re (2007). Natural catastrophes 2006 – Analyses, assessments, positions. *Rapport final*, pages 1–54.

Niazi, M. H. K., Morales Nápoles, O., and Van Wesenbeeck, B. K. (2021). Probabilistic characterization of the vegetated hydrodynamic system using non-parametric bayesian networks. *Water (Switzerland)*, 13(4).

Nicklin, H., Leicher, A. M., Dieperink, C., and Van Leeuwen, K. (2019). Understanding the costs of inaction-An assessment of pluvial flood damages in two European cities. *Water (Switzerland)*, 11(4):1–18.

Nixon, S. (2015). *EU overview of methodologies used in preparation of Flood Hazard and Flood Risk Maps*.

Palmer, W. (1965). Meterological Drought. *US Department of Commerce, Weather Bureau*, (45):58.

Paprotny, D., Morales-Nápoles, O., Worm, D. T., and Ragno, E. (2020). BANSHEE–A MATLAB toolbox for Non-Parametric Bayesian Networks. *SoftwareX*, 12:100588.

Parry, S., Marsh, T., and Kendon, M. (2013). 2012: From drought to floods in England and Wales. *Weather*, 68(10):268–274.

Pearl, J. (1986). A constraint-propagation approach to probabilistic reasoning. *Machine Intelligence and Pattern Recognition*, 4(C):357–369.

Pearl, J. (1991). A statistical semantics for causation. *UCLA Department of Statistics Papers*, (September):4.

Pescaroli, G. and Alexander, D. (2015). A definition of cascading disasters and cascading effects : Going beyond the " toppling dominos " metaphor. *GRF Davos Planet@Risk*, 3(1):58–67.

Pescaroli, G. and Alexander, D. (2018). Understanding Compound, Interconnected, Interacting, and Cascading Risks: A Holistic Framework. *Risk Analysis*, 38(11):2245–2257.

Petrova, I. Y., Miralles, D. G., van Heerwaarden, C. C., and Wouters, H. (2018). Relation between convective rainfall properties and antecedent soil moisture heterogeneity conditions in North Africa. *Remote Sensing*, 10(6).

Philip, S. Y., Kew, S. F., Van Der Wiel, K., Wanders, N., Jan Van Oldenborgh, G., and Philip, S. Y. (2020). Regional differentiation in climate change induced drought trends in the Netherlands. *Environmental Research Letters*, 15(9).

Prokić, M., Savić, S., and Pavić, D. (2019). Pluvial flooding in Urban Areas Across the European Continent. *Geographica Pannonica*, 23(4):216–232.

Quesada, B., Vautard, R., Yiou, P., Hirschi, M., and Seneviratne, S. I. (2012). Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change*, 2(10):736–741.

Quiring, S. M. (2009). Developing objective operational definitions for monitoring drought. *Journal of Applied Meteorology and Climatology*, 48(6):1217–1229.

R Core Team (2019). R: A Language and Environment for Statistical Computing.

Ragno, E., Hrachowitz, M., and Morales-Nápoles, O. (2021). Applying Non-Parametric Bayesian Network to estimate monthly maximum river discharge: potential and challenges. *Hydrology and Earth System Sciences Discussions*, (May):1–25.

Rashid, M. M. and Wahl, T. (2021). *New paradigm in hydrologic risk: Importance of integrating consecutive events of dry and wet extremes for improved infrastructure resilience*. PhD thesis, university of Central Florida.

Raymond, C., Horton, R. M., Zscheischler, J., Martius, O., AghaKouchak, A., Balch, J., Bowen, S. G., Camargo, S. J., Hess, J., Kornhuber, K., Oppenheimer, M., Ruane, A. C., Wahl, T., and White, K. (2020). Understanding and managing connected extreme events. *Nature Climate Change*, 10(7):611–621.

Rijkswaterstaat and Unie van Waterschappen (2019). Watermanagement in Nederland. Technical report.

RIVM (1988). *Zorgen voor morgen; nationale milieuverkenning 1985-2010*.

Robinson-Garcia, N., Costas, R., Sugimoto, C. R., Larivière, V., and Nane, G. F. (2020). Task specialization across research careers. *eLife*, 9:1–23.

Rosenzweig, B. R., McPhillips, L., Chang, H., Cheng, C., Welty, C., Matsler, M., Iwaniec, D., and Davidson, C. I. (2018). Pluvial flood risk and opportunities for resilience. *Wiley Interdisciplinary Reviews: Water*, 5(6):1–18.

Runhaar, H. A., Uittenbroek, C. J., van Rijswick, H. F., Mees, H. L., Driessen, P. P., and Gilissen, H. K. (2016). Prepared for climate change? A method for the ex-ante assessment of formal responsibilities for climate adaptation in specific sectors. *Regional Environmental Change*, 16(5):1389–1400.

Sanuy, M., Rigo, T., Jiménez, J. A., and Llasat, M. C. (2021). Classifying compound coastal storm and heavy rainfall events in the north-western Spanish Mediterranean. *Hydrology and Earth System Sciences*, 25(6):3759–3781.

Satoh, Y., Shiogama, H., Hanasaki, N., Pokhrel, Y., Boulange, J. E. S., Burek, P., Gosling, S. N., Grillakis, M., Koutroulis, A., Müller Schmied, H., Thiery, W., and Yokohata, T. (2021). A quantitative evaluation of the issue of drought definition: A source of disagreement in future drought assessments. *Environmental Research Letters*, 16(10).

Schneider, W. (2021). On the distribution of the sum of dependent standard normally distributed random variables using copulas. pages 1–10.

Sheffield, J. and Wood, E. F. (2011). *Drought : Past Problems and Future Scenarios*. Taylor & Francis Group, London, UNITED KINGDOM.

Sheng, W., Tian, H., Ding, X. J., Xie, W. S., and Yin, T. (2009). Climate Characteristics of Precipitation and Phenomenon of Drought-flood Abrupt Alternation during Main Flood Season in Huaihe River Basin. *Chinese Journal of Agrometeorology*, 30(1):31–34.

Simon Wang, S. Y., Yoon, J. H., Becker, E., and Gillies, R. (2017). California from drought to deluge.

Sklar, M. (1959). Fonctions de r´epartition 'a n dimensions et leurs marges. *Universit´e Paris 8*.

Sluijter, R., Plieger, M., Oldenborgh, G. J. v., Beersma, J., and de Vries, H. (2018). Een analyse op basis van het potentiële neerslagtekort. *Knmi*.

Smakhtin, V. U. and Schipper, E. L. F. (2008). Droughts: The impact of semantics and perceptions. *Water Policy*, 10(2):131–143.

Smit, M. (2021). Onderzoek naar gevolgen van extreme droogte voor de Nederlandse watersector.

Smith, M. S. (2021). Implicit Copulas: An Overview.

Spekkers, M., Rözer, V., Thieken, A., ten Veldhuis, M.-C., and Kreibich, H. (2017). A comparative survey of the impacts of extreme rainfall in two international case studies. *Natural Hazards and Earth System Sciences Discssions*, (July 2012):1–38.

STOWA (2019). Nieuwe normering van waterveiligheid. page 9.

Swain, D. L., Langenbrunner, B., Neelin, J. D., and Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, 8(5):427–433.

Szewrański, S., Chrusćiński, J., Kazak, J., Swíader, M., Tokarczyk-Dorociak, K., and Zmuda, R. (2018). Pluvial Flood Risk Assessment Tool (PFRA) for rainwater management and adaptation to climate change in newly urbanised areas. *Water (Switzerland)*, 10(4).

Tallaksen, L. M. and Lanen, H. A. J. v. (2004). *Hydrological drought. Processes and estimation methods for streamflow and groundwater*.

Tank, A. M. K., Zwiers, F. W., and Xuebin Zhang (2011). Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation. *Impact of Weather and Climate Extremes*, (72):1–269.

Task Force Fact-finding hoogwater 2021 (2021). Hoogwater 2021. Technical report.

Taylor, C. M., De Jeu, R. A., Guichard, F., Harris, P. P., and Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, 489(7416):423–426.

Tilloy, A., Malamud, B. D., Winter, H., and Joly-Laugel, A. (2019). A review of quantification methodologies for multi-hazard interrelationships. *Earth-Science Reviews*, 196(May):102881.

UNDRR (2020). Hazard Definition & classification review: Technical Report. pages 1–88.

UNINET Team (2019). UNINET: An uncertainty analysis program.

United Nations (2015). Sendai Framework for Disaster Risk Reduction 2015 - 2030. *UNISDR*.

van Haren, R., van Oldenborgh, G. J., Lenderink, G., Collins, M., and Hazeleger, W. (2013). SST and circulation trend biases cause an underestimation of European precipitation trends. *Climate Dynamics*, 40(1-2):1–20.

van Loon, A. F. (2015). Hydrological drought explained. 2(August).

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Van Lanen, H. A. (2016). Drought in the Anthropocene. *Nature Geoscience*, 9(2):89–91.

Van Loon, A. F. and Van Lanen, H. A. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, 16(7):1915–1946.

Van Ootegem, L., Verhofstadt, E., Van Herck, K., and Creten, T. (2015). Multivariate pluvial flood damage models. *Environmental Impact Assessment Review*, 54:91–100.

Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.

Vicente-Serrano, S. M. (2006). Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula. *Water Resources Management*, 20(1):37–60.

Vinet, F. (2011). Floods : Volume 1 - Risk Knowledge. *Elsevier*, 1(2017):1–5.

Visser-Quinn, A., Beevers, L., Collet, L., Formetta, G., Smith, K., Wanders, N., Thober, S., Pan, M., and Kumar, R. (2019). Spatio-temporal analysis of compound hydro-hazard extremes across the UK. *Advances in Water Resources*, 130(November 2018):77–90.

Vojinovic, Z. and Huang, J. (2015). *Differentiating between Different Types of Floods and Their Impacts*.

Wahl, T., Jain, S., Bender, J., Meyers, S. D., and Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5(12):1093–1097.

Wanders, N., Lanen, H. A. J. v., and Loon, A. F. v. (2010). Indicators for Drought Characterization on a Global Scale (Technical Report No . 24). (24):1–93.

Ward, P. J., Daniell, J., Duncan, M., Dunne, A., Hananel, C., Hochrainer-Stigler, S., Tijssen, A., Torresan, S., Ciurean, R., Gill, J. C., Sillmann, J., Couasnon, A., Koks, E., Padrón-Fumero, N., Tatman, S., Tronstad Lund, M., Aerts, C., Alabaster, A., Bulder, B., Campillo Torres, C., Critto, A., Hernández Martín, R., Machado, M., Mysiak, J., Orth, R., Palomino, I., Petrescu, E.-C., Reichstein, M., Tiggeloven, T., van Loon, A. F., Vuong Pham, H., and de Ruiter, M. C. (2021). Invited perspectives: A research agenda towards Disaster Risk Management pathways in multi-risk assessment. (November):1–15.

Ward, P. J., de Ruiter, M. C., Mård, J., Schröter, K., Van Loon, A., Veldkamp, T., von Uexkull, N., Wanders, N., AghaKouchak, A., Arnbjerg-Nielsen, K., Capewell, L., Carmen Llasat, M., Day, R., Dewals, B., Di Baldassarre, G., Huning, L. S., Kreibich, H., Mazzoleni, M., Savelli, E., Teutschbein, C., van den Berg, H., van der Heijden, A., Vincken, J. M., Waterloo, M. J., and Wens, M. (2020). The need to integrate flood and drought disaster risk reduction strategies. *Water Security*, 11(2020).

Weber, P., Medina-Oliva, G., Simon, C., and Iung, B. (2012). Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, 25(4):671–682.

Werick, W. J. and Whipple, W. (1994). Managing Water for Drought. *IWR Report 94-NDS-8*, page 210.

WHO (2021). Health topics - drought and floods.

Wilhite, D. A. (2000). *Drought : a global assessment. Vol. II.* Routledge, London SE - 304 blz. ; .. cm.

Wilhite, D. A. and Glantz, M. H. (1985). Understanding the drought phenomenon: The role of definitions. *Planning for Drought: Toward A Reduction of Societal Vulnerability*, 10:3:111–120.

WMO (1969). Manual for Depth-Area-Duration Analysis of Storm Precipitation. *World Meteorological Organization (WMO)*, 1969(237):131.

WMO (1987). Standardized Precipitation Index User Guide. *Journal of Applied Bacteriology*, 63(3):197–200.

Wolters, D., Homan, C., and Bessembinder, J. (2011). Ruimtelijke klimatologische verschillen binnen Nederland.

Yevjevich, V. (1969). An objective approach to definitions and investigations of continental hydrologic droughts. *Journal of Hydrology*, 7(3):353.

Zhang, W., Luo, M., Gao, S., Chen, W., Hari, V., and Khouakhi, A. (2021a). Compound Hydrometeorological Extremes: Drivers, Mechanisms and Methods. *Frontiers in Earth Science*, 9(October):1–20.

Zhang, W., Murakami, H., Khouakhi, A., and Luo, M. (2021b). *Editorial: Compound Climate Extremes in the Present and Future Climates: Machine Learning, Statistical Methods and Dynamical Modelling*, volume 9.

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., and Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change*, 2(6):851–870.

Zhao, Y., Weng, Z., Chen, H., and Yang, J. (2020). Analysis of the evolution of drought, flood, and drought-flood abrupt alternation events under climate change using the daily SWAP index. *Water (Switzerland)*, 12(7).

Zhou, H., Liu, Y., and Liu, Y. (2019). An Approach to Tracking Meteorological Drought Migration. *Water Resources Research*, 55(4):3266–3284.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., Maraun, D., Ramos, A. M., Ridder, N. N., Thiery, W., and Vignotto, E. (2020). A typology of compound weather and climate events. *Nature Reviews Earth and Environment*, 1(7):333–347.

Zscheischler, J., Sillmann, J., and Alexander, L. (2021). Introduction to the special issue: Compound weather and climate events. *Weather and Climate Extremes*, (August):0–2.

Zscheischler, J., Van Den Hurk, B., Ward, P. J., and Westra, S. (2019). *Multivariate extremes and compound events*. Elsevier Inc.

Zscheischler, J., Westra, S., Hurk, B. J. J. M. V. D., Seneviratne, S. I., Ward, P. J., Pitman, A., Aghakouchak, A., Bresch, D. N., and Leonard, M. (2018). Future climate risk from compound events. 8(June).

Zurich Insurance (2020). Four common types of flood explained. *July 22, 2020*, pages 1–5.

# A

# Introduction to copulas

### A.0.1. What is a Copula?

Copula is derived from the word Copulare, which means "to join" or "to couple" in Latin and its Mathematical concept was first introduced in 1959, known as Sklar's theorem (Sklar, 1959). Copulas are an elegant and flexible way to model multivariate dependencies, as the study of the margins and the study of the dependence structure are separated. In other words, the marginal behavior of each component of a random vector is described by the margins, whereas the the copula describes the dependence structure among the components. For multivariate data from the same distribution (e.g. a multivariate Gaussian distribution) this might not be as relevant, but it becomes advantageous when modelling multivariate data with a combination of different distributions (e.g. Student, Gamma and Beta). As of this flexibility, copulas are extremely popular in the fields of financial modeling, risk management and more recently machine learning problems (Cherubini and Luciano, 2002, Elidan, 2013, Embrechts et al., 2003).

As an example we consider a set of $n$ random variables $\{X_1, \dots, X_n\}$ and their joint distribution function $F(X_1, \dots, X_n)$ which can be separated into two parts. The first part is that each random variable (rv) can be fully described by its marginal, or its cumulative distribution function (cdf) $F_i(x) := P(X_i \leq x)$. Here, $x$ is a realization of $X_i$ and $P$ is a probability. This marginal cdf has two interesting properties:

> Property 1: If $U \sim U[0,1]$ and $F_i$ is a cdf, then the inverse cdf (i.e. $F_i^{-1}$) of $U_i$ corresponds to the rv $X_i$.
>
> $$P(F_i^{-1}(U_i) \leq x) = F_i(x_i) \tag{A.1}$$
>
> Property 2: If a real-valued rv $X_i$ has a cdf $F_i$ then:
>
> $$F_i(X_i) \sim U_i[0,1] \tag{A.2}$$

The first property, also known as the inverse cdf transform or quantile transform, is a commonly known way to simulate data from a uniform distribution with a cdf of choice. The second property is also known as the probability transform, or probability integral transformation. Following from the set of marginals $\{F_1, \dots, F_n\}$, we move to the second element, which is the copula. The second element its decomposition relies on Sklar's theorem (Sklar, 1959) which states that for every set of random variables $\{X_1, \dots, X_n\}$ with joint distribution function $F_X$ marginals $\{F_{X_1}, \dots, F_{X_n}\}$ there exist an N-dimensional "copula function" $C : [0,1]^n \to [0,1]$ such

that

$$F(x_1, \dots, x_n) = C(F_{X_1}(x_1), \dots, F_{X_n}(x_n)) \tag{A.3}$$

This can be inverted to compute the copula from multivariate distributions:

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \tag{A.4}$$

If equation A.4 is used to extract a copula, the copula is considered implicit. This family of copulas is versatile and shares an auxiliary representation that allows estimations in high dimensions to be tractable (Smith, 2021). The most commonly used implicit copula is the Gaussian copula. This copula is constructed from a multivariate Gaussian or normal distribution which is based on a $n \times n$ correlation matrix $\Omega$ and correlation coefficient $\rho$ that captures the dependency relationship between the rvs. The two dimensional Gaussian copula is defined as

$$C_\rho^{Gauss}(u_1, u_2) = \Phi_\Omega\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho\right) \tag{A.5}$$

Where $x_1 := \Phi^{-1}(u_1)$ and $x_2 := \Phi^{-1}(u_2)$ are two rvs, $\Phi_\Omega(\cdot)$ denotes the cdf of the bivariate normal distribution and $\Phi^{-1}(\cdot)$ as an inverse of the standard normal function.

The copula function is per definition a cdf. These graphs are relatively hard to interpret as they are monotonically increasing (Figure A.1 and A.2). As of this reason, copulas are commonly visualised using their density function, which can be derived from the copula cdf:

$$c(u_1, \dots, u_n) = \frac{\delta^n C(u_1, \dots, u_n)}{\delta u_1, \dots, \delta u_n} \tag{A.6}$$

Although not all copulas are differentiable. This means some copulas do not have density functions. The two dimensional Gaussian copula is sufficiently differentiable and is denoted as

$$c_\rho^{Gauss}(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} exp\left[-\frac{\rho^2(x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)}\right] \tag{A.7}$$

Additional to implicit copulas there are explicit copulas. This type of copulas have a closed form and unlike to the Gaussian copula, their use is more demanding for uncertainty analysis as of the predefined objective of their usage (Embrechts et al., 2003, Hazarika et al., 2018). Examples of implicit copulas are the Frank, Clayton and Gumbel copulas. The two dimensional copulas of these families are denoted as

$$C_\theta^{Frank}(u_1, u_2) = -\frac{1}{\theta} ln\left[1 + \frac{(exp(-\theta u_1) - 1) \cdot (exp(-\theta u_2) - 1)}{exp(-\theta) - 1}\right], \quad \theta \in \mathbb{R}\backslash\{0\} \tag{A.8}$$

$$C_\theta^{Clayton}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}, \quad \theta \in [-1, \infty)\backslash\{0\} \tag{A.9}$$

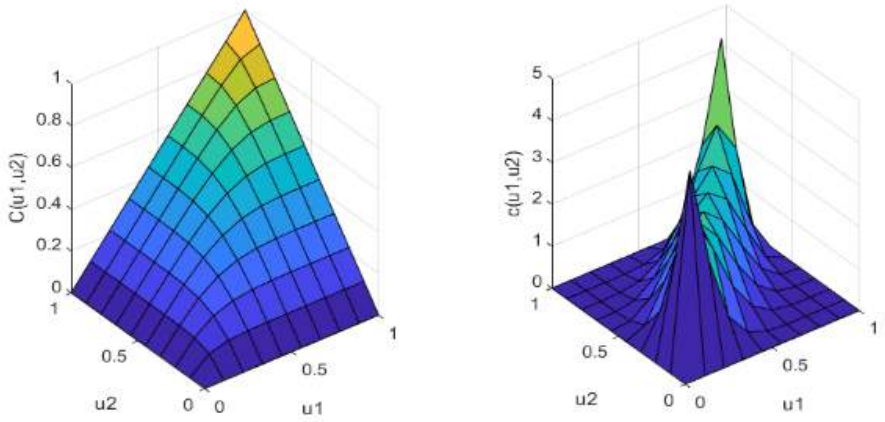



Figure A.1: The cumulative density function (left) and density function (right) of a two dimensional Gaussian copula with $\rho = 0.9$. Source: Schneider (2021) (Schneider, 2021)

$$C_\theta^{Gumbel}(u_1, u_2) = exp\left[-((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{\frac{1}{\theta}}\right], \quad \theta \in [1, \infty) \tag{A.10}$$

With according density functions

$$c_\theta^{Frank}(u_1, u_2) = \frac{-\theta g(1)(1 + g(u_1 + u_2))}{(g(u_1)g(u_2) + g(1))^2} \tag{A.11}$$

Where $g(x)$ is defined as

$$g(x) = e^{-\theta x} - 1$$

$$c_\theta^{Clayton}(u_1, u_2) = (1 + \theta)(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1+2\theta}{\theta}}(u_1 u_2)^{-(\theta+1)} \tag{A.12}$$

$$c_\theta^{Gumbel}(u_1, u_2) = \frac{C^{Gumbel}(u_1, u_2)}{u_1 u_2} \frac{((-\ln u_1)^\theta + (-\ln u_1)^\theta)^{\frac{2}{\theta}-2}}{(\ln u_2 \ln u_2)^{1-\theta}} \times \left(1 + (\theta - 1)((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{-\frac{1}{\theta}}\right) \tag{A.13}$$

There are two dependency measures that are commonly used for expressing the dependency between two variables. These are Spearmanns rho ($\rho$) and Kendall's tau ($\tau$). Both can be expressed in terms of the copula (Czado, 2019):

$$\rho = 12 \int_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3 \qquad \text{and} \qquad \tau = 4 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1$$

The relationship between Spearmanns rho and Kendall's tau for the bivariate normal distribution be denoted as (Czado, 2019)

$$\tau = \frac{2}{\pi}\arcsin(\rho) \tag{A.14}$$

As denoted before, equation A.7 of the Gaussian copula uses the parameter $\rho$ as dependency measure. As can be observed in equations A.8, A.9 and A.10 the copula parameter $\theta$ is introduced. This parameter can be derived from Kendall's tau with the following equation per copula family (McNeil et al., 2015)

$$\begin{aligned} c_\theta^{Frank} &: \tau = 1 - 4\theta^{-1}(1 - D_1(\theta)) \\ \text{Where} \quad D_1(\theta) &= \theta^{-1}\int_0^\theta \frac{t}{e^t - 1}dt \end{aligned} \tag{A.15}$$

$$c_\theta^{Clayton} : \tau = \frac{\theta}{\theta + 2} \tag{A.16}$$

$$c_\theta^{Gumbel} : \tau = 1 - \frac{1}{\theta} \tag{A.17}$$

In line with the example of Schneider (Schneider, 2021) in figure A.1, the corresponding values for $\theta$ are determined with the help of equations A.14, A.15,A.16 and A.17 (Figure A.2).

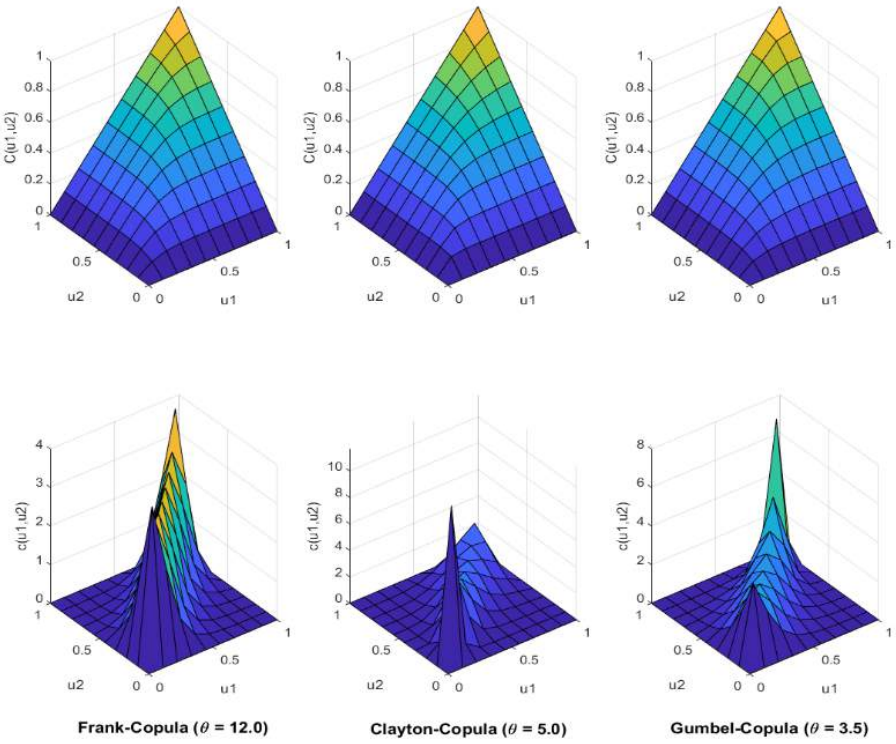



Figure A.2: The cumulative density function (top row) and density function (bottom row) for the Frank, Clayton and Gumbel copulas with corresponding $\theta$-parameters where $\rho = 0.9$. Different tail dependencies exist per copula. This may not be observed in the cumulative density functions, but are better visible in the density functions. In opposition to the Gaussian copula the Frank copula is lighter in the tails. The Clayton copula a larger lower tail and smaller higher tail dependency. The Gumbel copula has a similar lower tail and larger higher tail dependency. Source: Schneider (2021) (Schneider, 2021)

# B

# Temporal extent of relevant variables per KNMI-A station

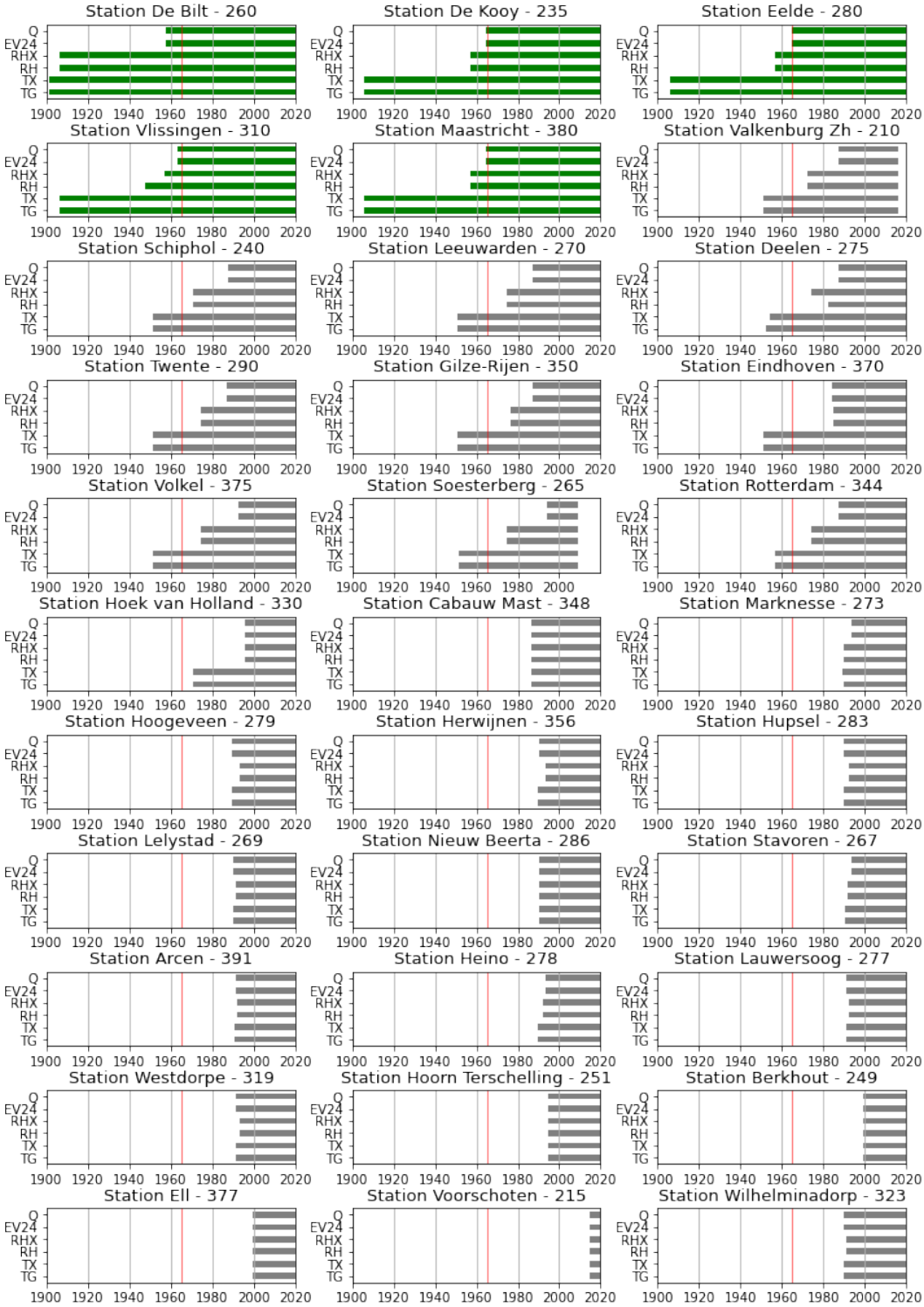



Figure B.1: Temporal extent of the variables: daily evapotranspiration (EV24), maximum daily precipitation intensity (RHX), daily precipitation (RH), maximum daily temperature (TX) and average daily temperature (TG) per KNMI-A station. The red vertical line indicates the year 1965, whereat the green barred figures have a larger temporal extent for all variables than this year. These are the five selected stations for the drought indicators.

# C

# Histograms of region and indicators

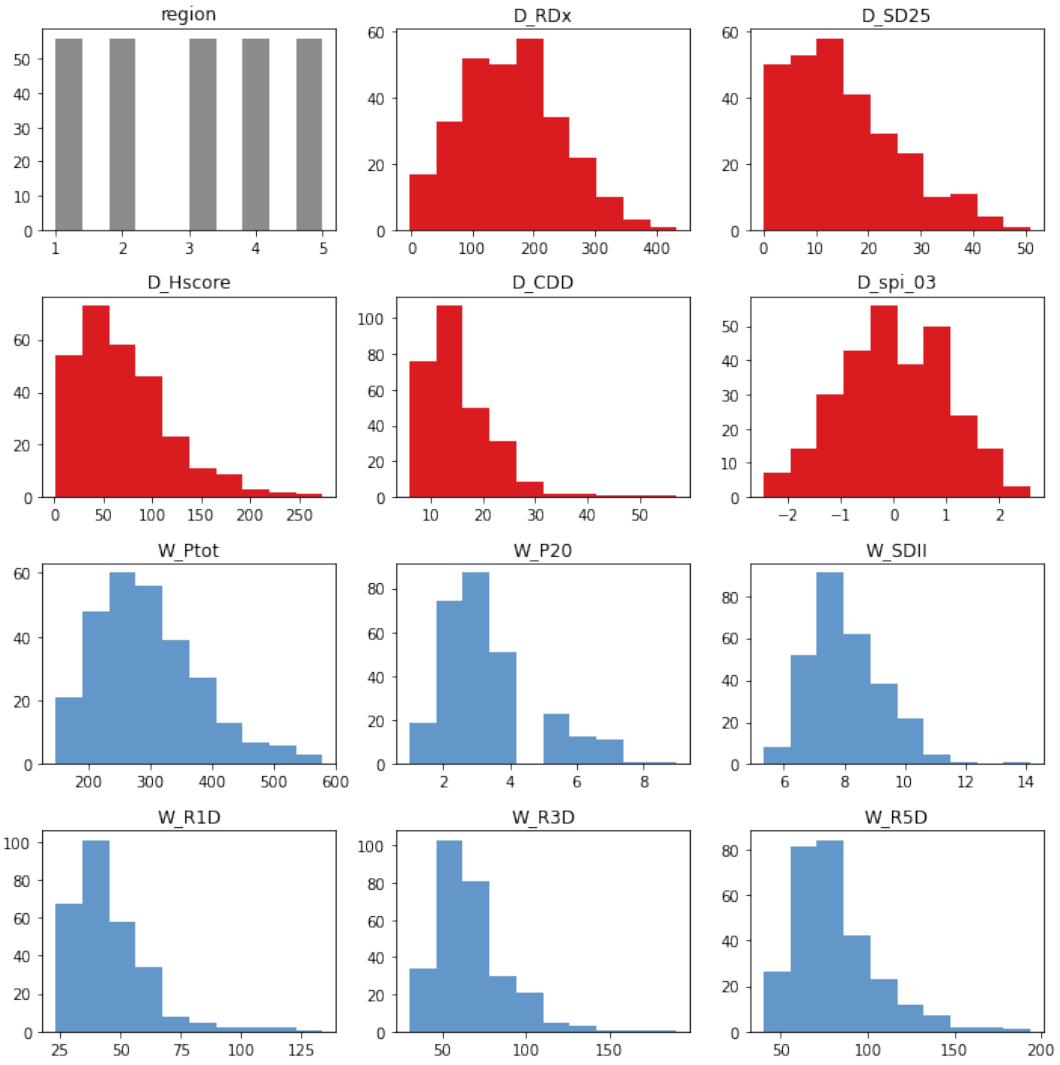



Figure C.1: Histograms of region (gray), dry inidicators (red) and wet indicators (blue)

# D

# Evapotranspiration of region 3

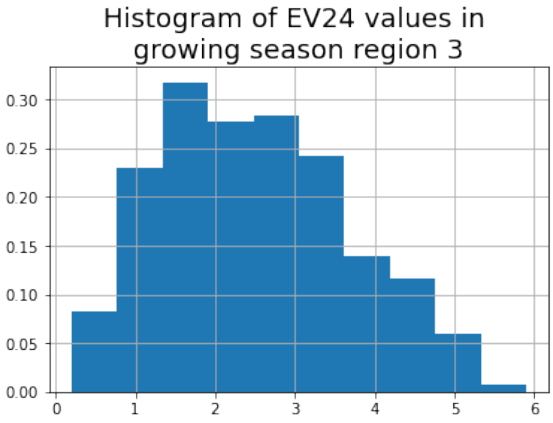



Figure D.1: Histogram of evapotranspiration (EV24) values [mm] in growing season (April - September) for region 3.

# E

# CvM statistics


Figure E.1: Cramer-von-Mises test scores per pair of nodes. Gaussian is colored different to the other families, as this is the family used in NPBNs.

# Correlation matrices of RN and SN

Table F.1: Correlation matrix of the saturated network (SN). The intensity of the color indicates the strength of the correlation, where red is a positive rank correlation value and blue is a negative rank correlation value.

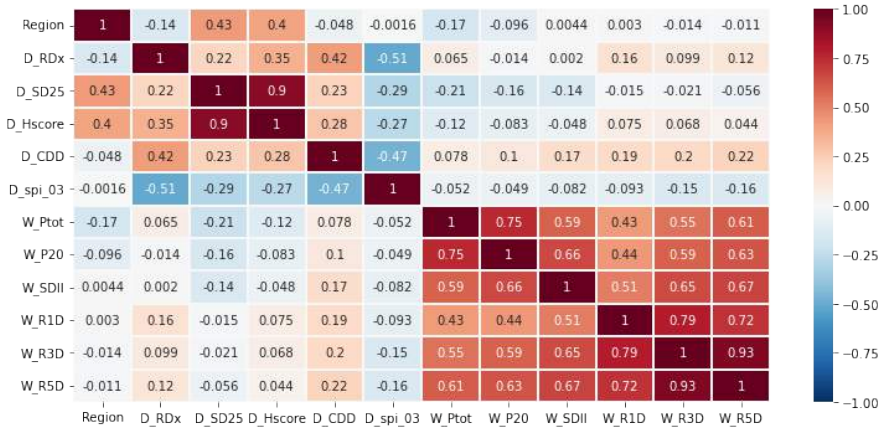
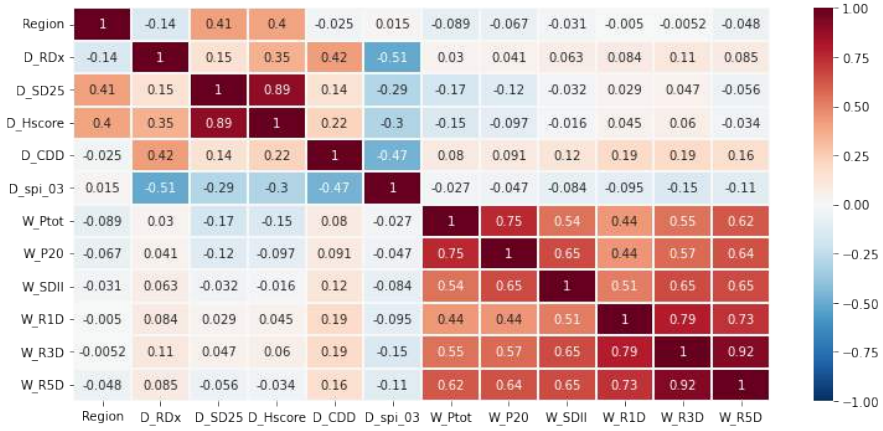
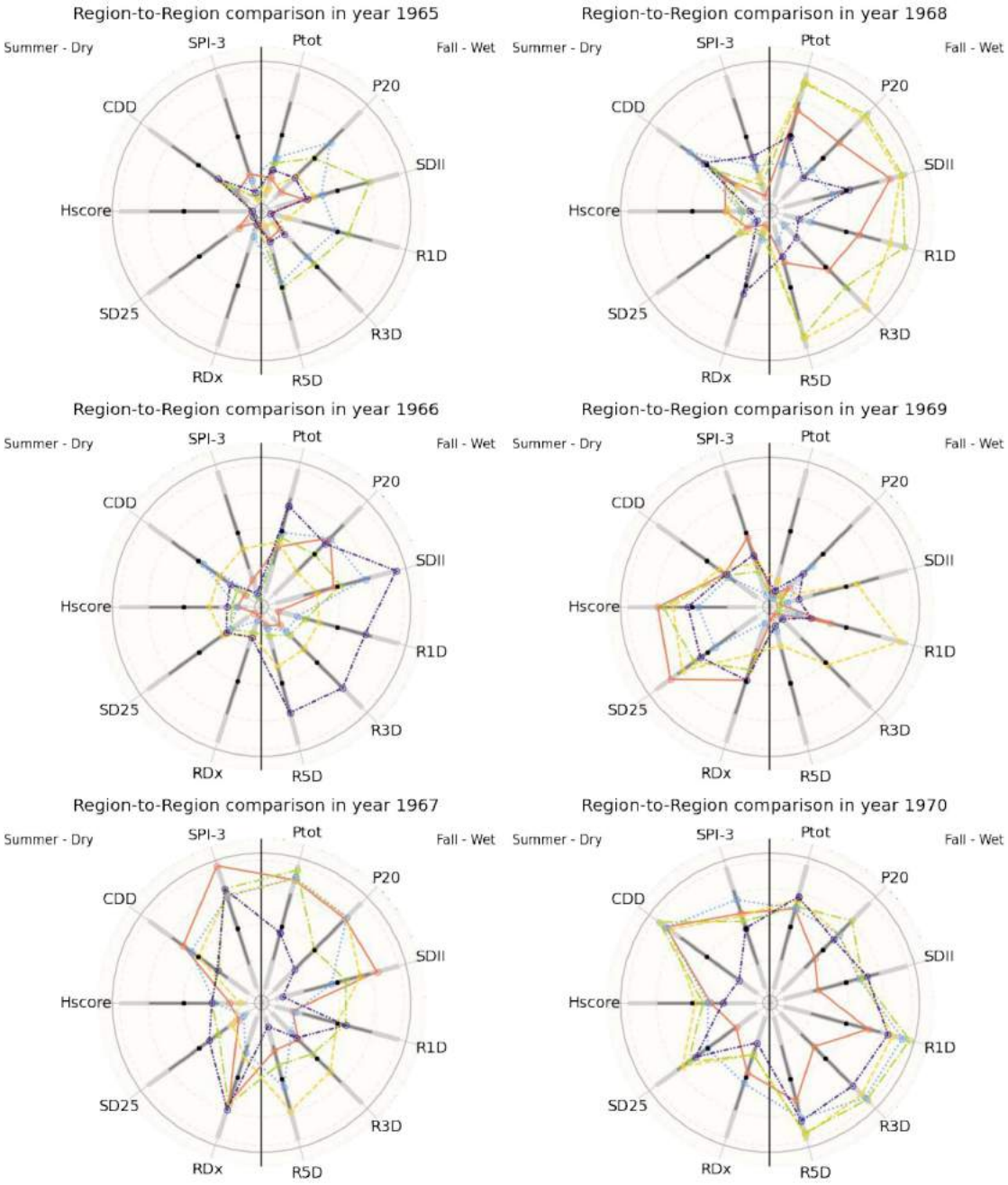


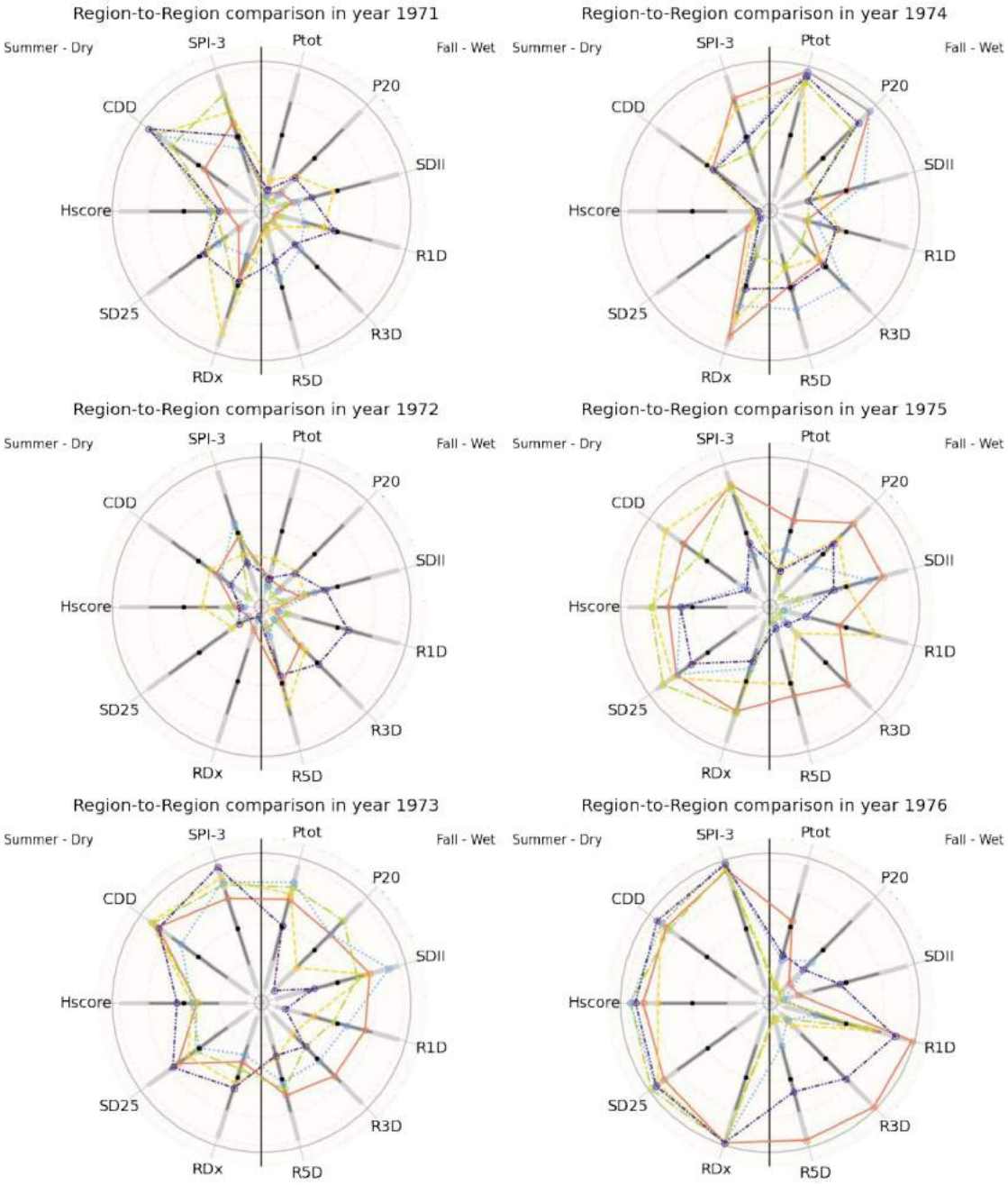

Table F.2: Correlation matrix of the restricted network (RN). The intensity of the color indicates the strength of the correlation, where red is a positive rank correlation value and blue is a negative rank correlation value.

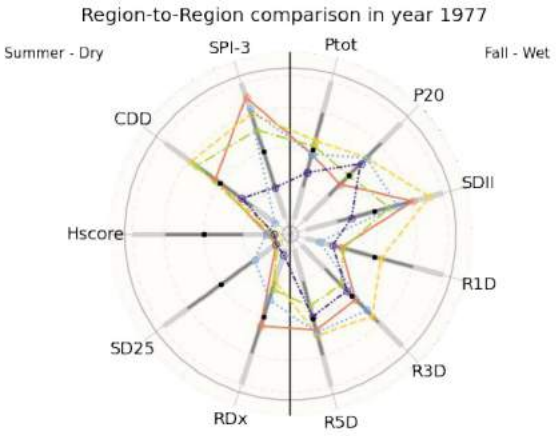


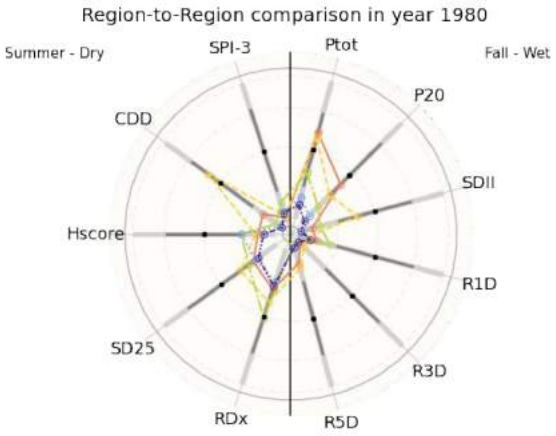
# G

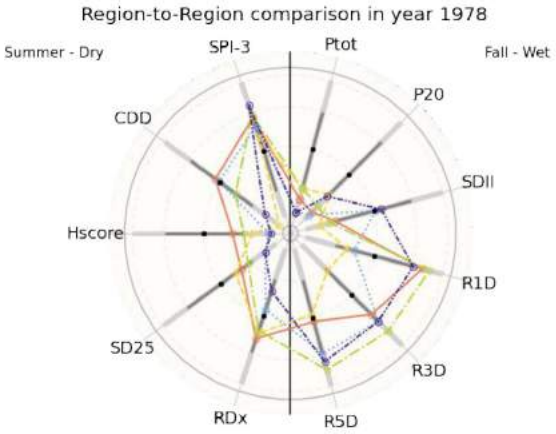
# R2R for all years

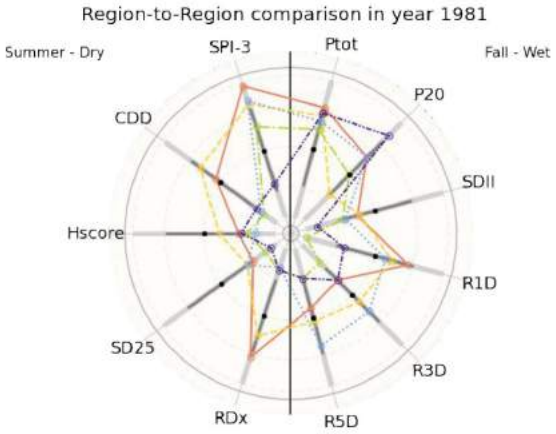

Region-to-Region comparison in year 1965
Region-to-Region comparison in year 1968
Region-to-Region comparison in year 1966
Region-to-Region comparison in year 1969
Region-to-Region comparison in year 1967
Region-to-Region comparison in year 1970
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

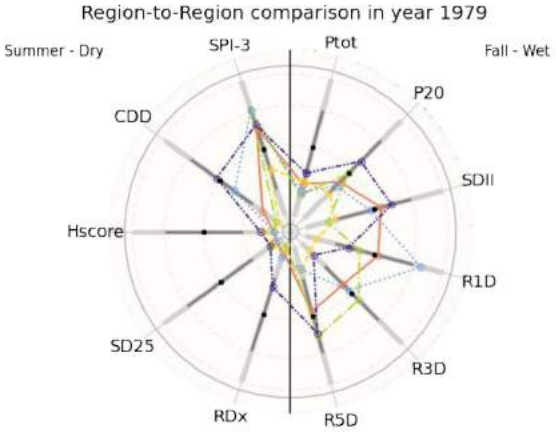

Region-to-Region comparison in year 1971
Region-to-Region comparison in year 1974
Region-to-Region comparison in year 1972
Region-to-Region comparison in year 1975
Region-to-Region comparison in year 1973
Region-to-Region comparison in year 1976
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

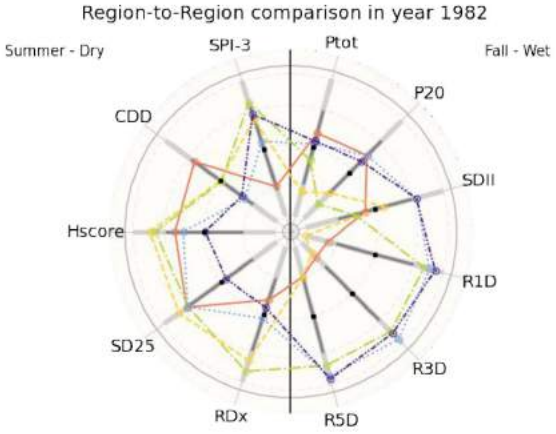

Region-to-Region comparison in year 1977
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

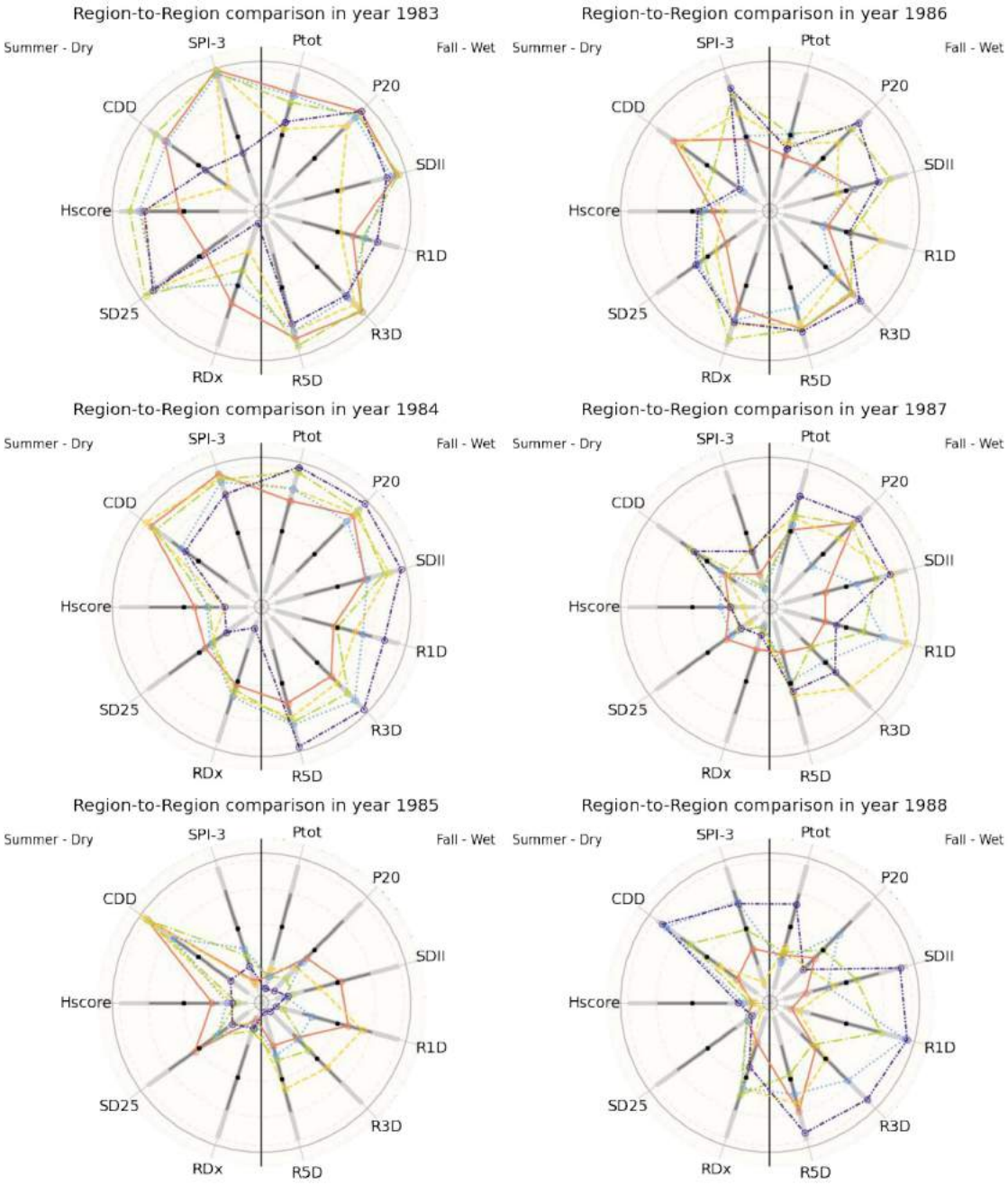

Region-to-Region comparison in year 1980
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

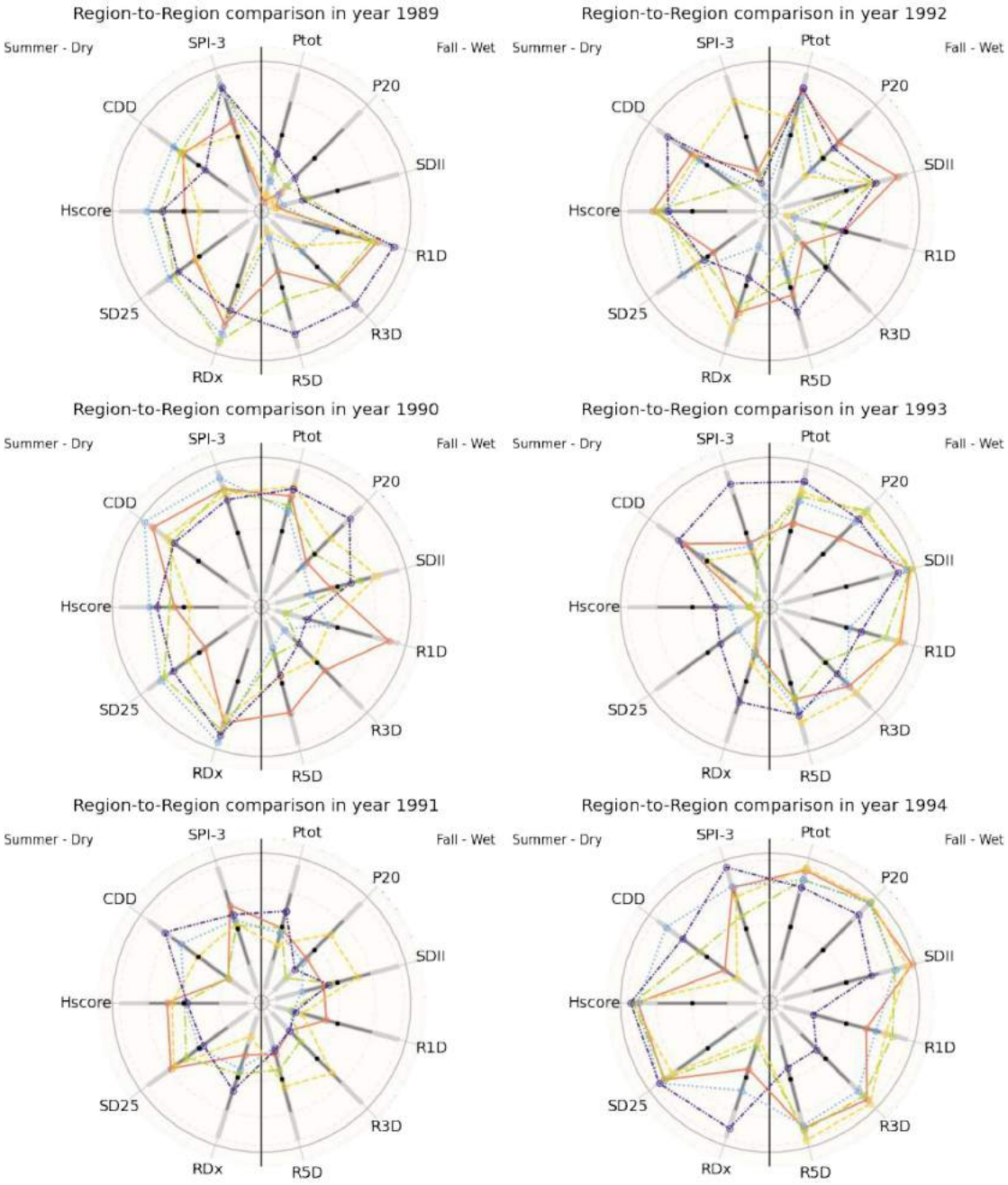

Region-to-Region comparison in year 1978
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

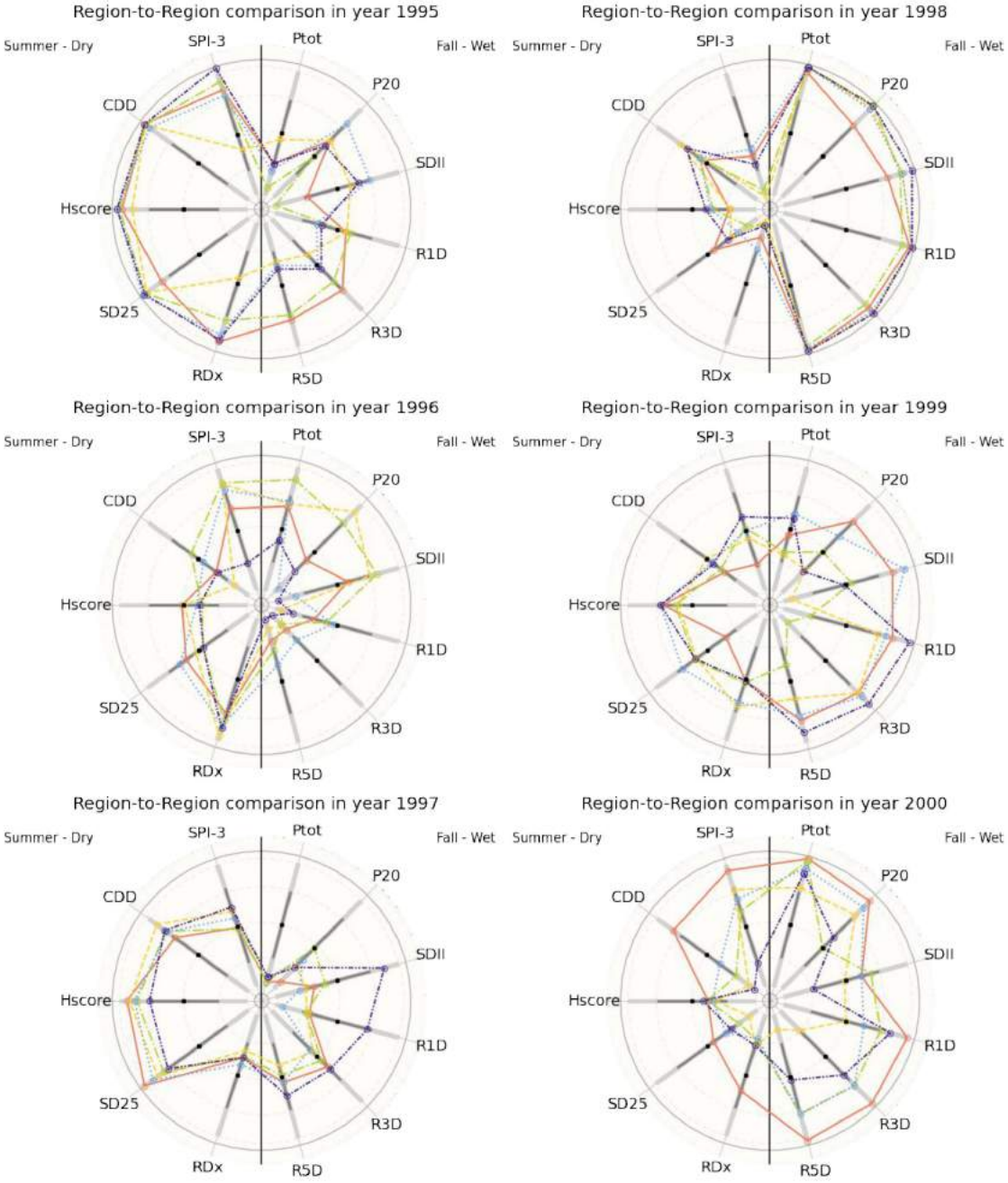

Region-to-Region comparison in year 1981
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

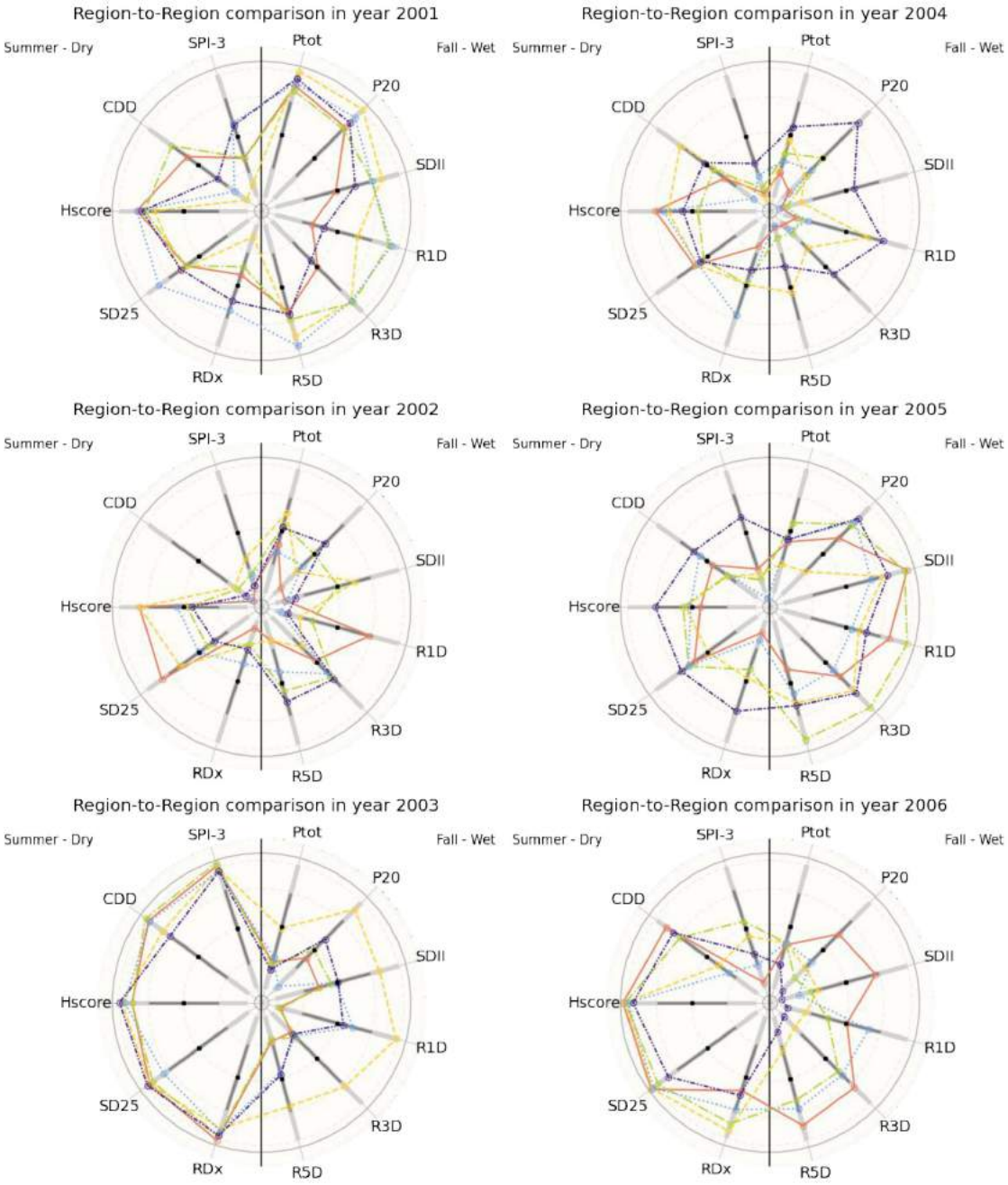

Region-to-Region comparison in year 1979
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

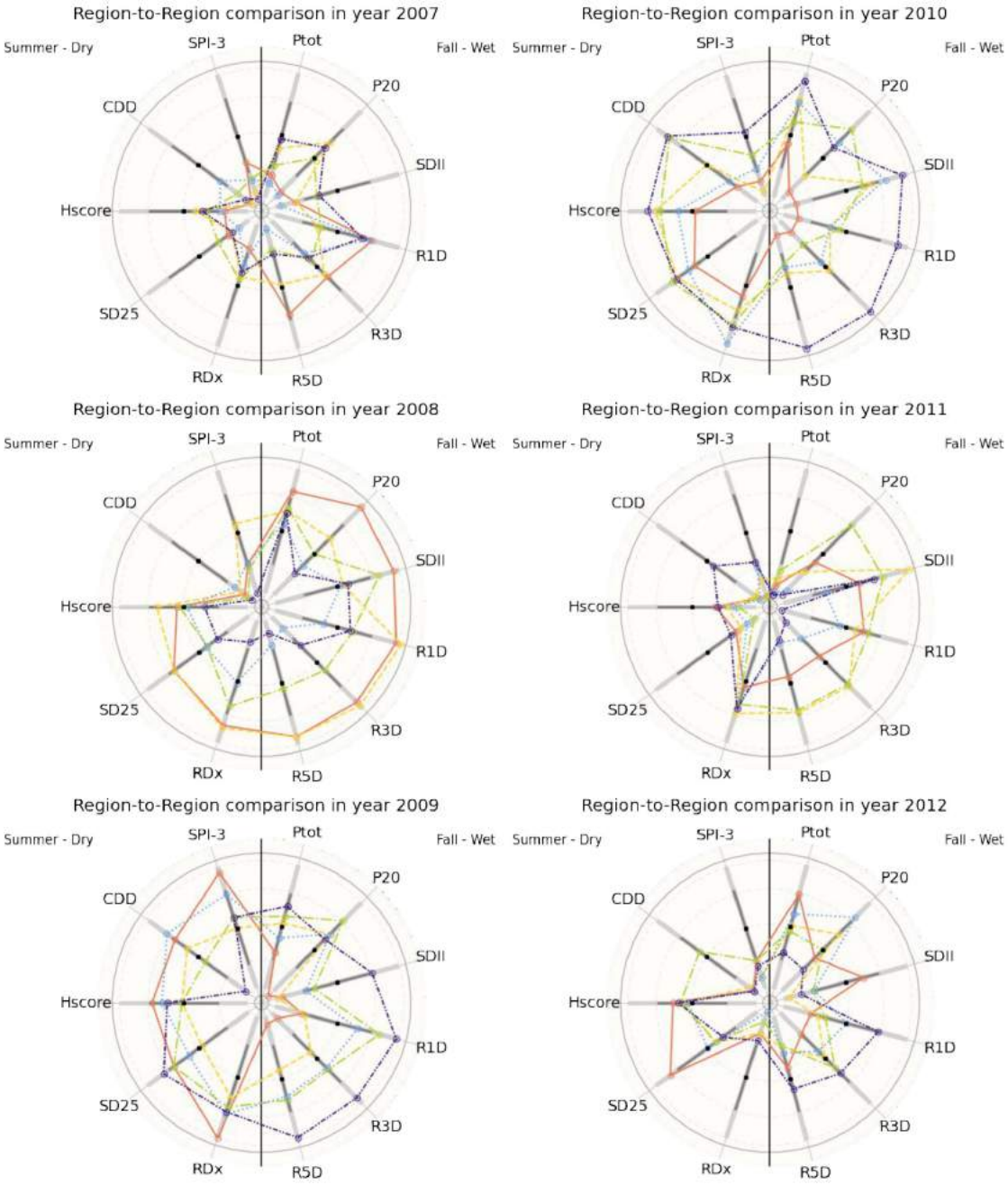

Region-to-Region comparison in year 1982
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

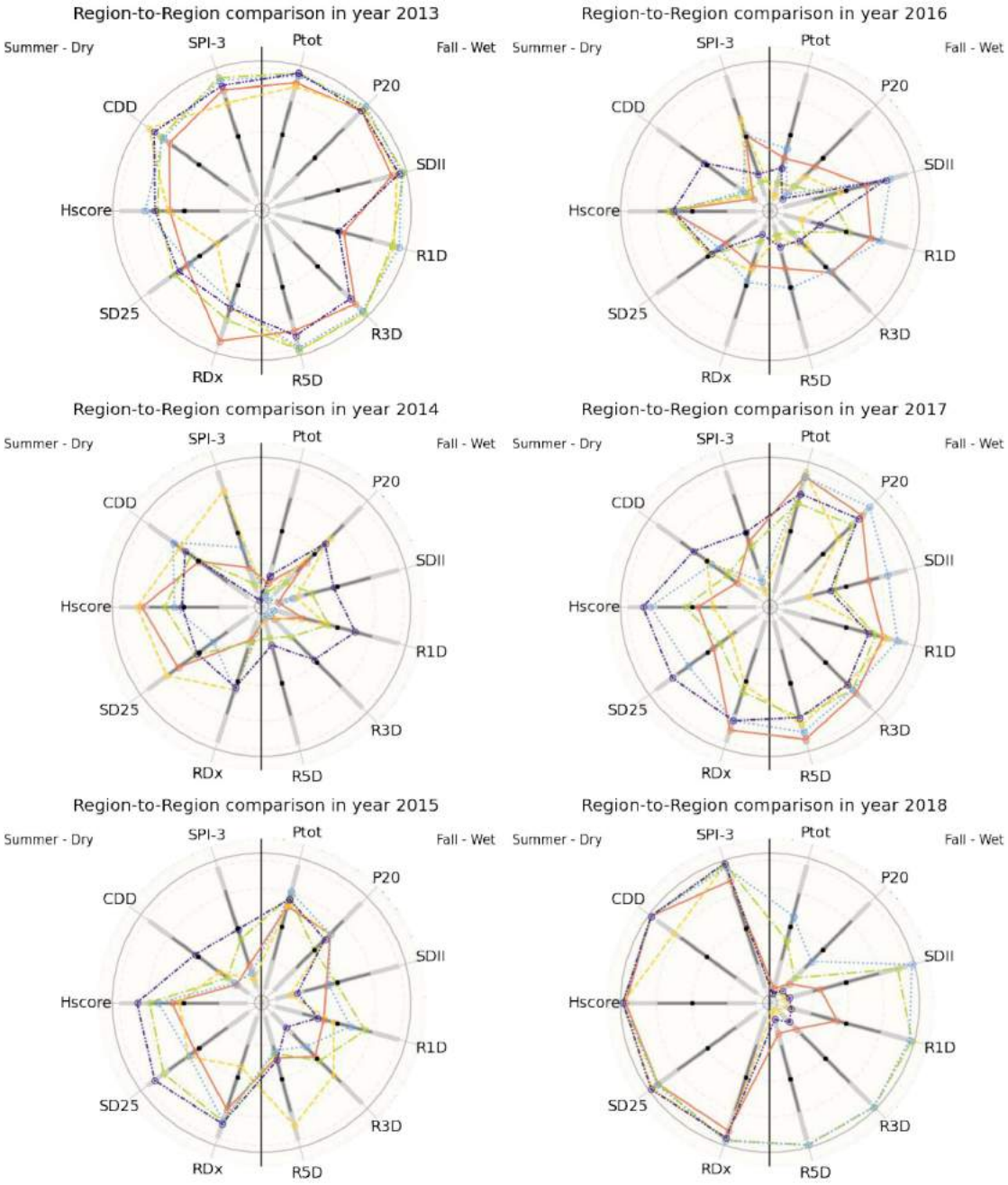

Region-to-Region comparison in year 1983
Region-to-Region comparison in year 1986
Region-to-Region comparison in year 1984
Region-to-Region comparison in year 1987
Region-to-Region comparison in year 1985
Region-to-Region comparison in year 1988
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD

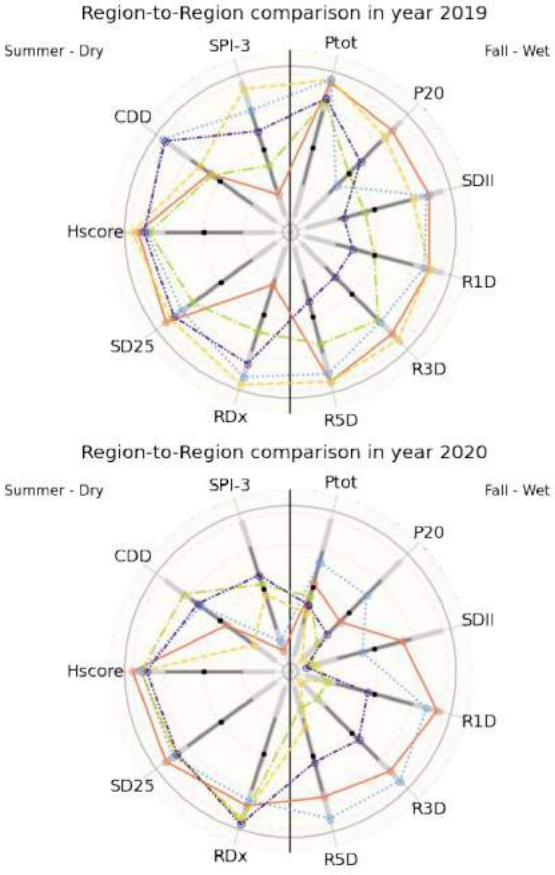

Region-to-Region comparison in year 1989
Region-to-Region comparison in year 1992
Region-to-Region comparison in year 1990
Region-to-Region comparison in year 1993
Region-to-Region comparison in year 1991
Region-to-Region comparison in year 1994
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD


Region-to-Region comparison in year 1995
Region-to-Region comparison in year 1998
Region-to-Region comparison in year 1996
Region-to-Region comparison in year 1999
Region-to-Region comparison in year 1997
Region-to-Region comparison in year 2000
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD


Region-to-Region comparison in year 2001
Region-to-Region comparison in year 2004
Region-to-Region comparison in year 2002
Region-to-Region comparison in year 2005
Region-to-Region comparison in year 2003
Region-to-Region comparison in year 2006
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD


Region-to-Region comparison in year 2007
Region-to-Region comparison in year 2010
Region-to-Region comparison in year 2008
Region-to-Region comparison in year 2011
Region-to-Region comparison in year 2009
Region-to-Region comparison in year 2012
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD


Region-to-Region comparison in year 2013
Region-to-Region comparison in year 2016
Region-to-Region comparison in year 2014
Region-to-Region comparison in year 2017
Region-to-Region comparison in year 2015
Region-to-Region comparison in year 2018
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD


Region-to-Region comparison in year 2019
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD
Region-to-Region comparison in year 2020
Summer - Dry
Fall - Wet
SPI-3
Ptot
P20
SDII
R1D
R3D
R5D
RDx
SD25
Hscore
CDD