UNIVERSITÀ DEGLI STUDI DI PALERMO

Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale
Area Ingegneria Idraulica e Ambientale

XXVI Ciclo del Dottorato di Ricerca in Ingegneria Civile e Ambientale
Tesi per il conseguimento del titolo

Objective regional frequency analysis and future change of extreme precipitation in Sicily, Italy

Angelo Forestieri

Tutor:
Prof. Leonardo V. Noto

Coordinatore del Dottorato:
Prof. Orazio Giuffré

Sede Amministrativa: Università degli Studi di Palermo
Settore Scientifico Disciplinare: Costruzioni idrauliche e marittime e idrologia (ICAR/02)

Palermo, Gennaio 2016

To my mother

Contents

List of Figures

List of Tables

Introduction

Extreme precipitation events have large impacts on society and are likely to continue to do so under global warming (Tebaldi et al., 2006). Indeed, extreme precipitation shows intensification in many regions and this is of key importance to society as a result of the large impact through flooding (Trenberth et al., 2003). For design and management decisions, particularly around hydraulic works, accurate estimates of precipitation magnitudes at different durations are needed.

Most of the typical hydrological applications need estimation of the physical quantity (e.g. precipitations or discharge) where measurements stations either are not available or the data provided result inadequate for the application to be implemented. Such a last case is often verified when the probability distributions adopted for the modelling of a physical process linked to measurements needs the knowledge of the parameters estimated using high-order statistics that can be obtained only if the sample is wide and purposeful.

Frequency analysis is related to an information problem. Indeed when the dataset of flood flows, rainfall, or low flows is available with a sufficiently long record, the frequency distribution for a site could be precisely determined, so long as change over time due to urbanization or natural processes did not alter the relationships of concern. Frequently the historically available rainfall series are unsuitable for this estimation process. This forces hydrologists to exploit the practical knowledge of the processes involved, and the adoption of efficient and robust statistical techniques, to develop the best estimates of hazard. Among these techniques, *Regional Freqeuncy analysis* is one of the most useful tool.

Extreme rainfall has been classically estimated using rain-gauge data at single point locations to fit extreme value distributions to annual maxima series to define extremes with given return periods on an annual basis and then interpolating these estimates bin the space (Prudhomme and Reed, 1999). More recently, regional frequency analysis has been more widely used (Hosking and Wallis, 2005, Lo Conti et al., 2007, Nguyen et al., 2002, Norbiato et al., 2007, Stedinger and Vogel, 1993).

Such an approach involves the regional pooling of standardized annual maxima for different duration to allow the estimation of long return-period rainfall events when individual records are too short to allow their reliable estimation (Fowler and Kilsby, 2003a, Hosking and Wallis, 2005). However, this methodology relies on the identification of homogeneous regions which share the same extreme rainfall characteristics: the assumption made is that the shape of the growth curve is the same for every site in the region, and only the magnitude of the events differs among different sites.

The *Regional Frequency Analysis* (RFA) has been shown to reduce the uncertainties in quantile estimation of extreme events (Hosking and Wallis, 1988) and plays an important role for several civil structure designs and non-structural problems involving natural hazards associated with extreme rainfall events. RFA uses data derived from a number of measuring sites belonging to a "region", defined as a group of sites, each of which could be modelled from the same frequency distribution. Therefore, the analysis involves the identification of such regions, testing whether the proposed regions are really homogeneous, and the choice of appropriate distributions to fit each region's data.

An important strength of the RFA approach is that the parameters of the probability model could be estimated with a wider historical observation dataset, using all the stations belonging to the same region. Previous works have demonstrated that RFA is able to provide more reliable estimates of extreme rainfall quantiles for different return even in the case when the dataset is very large (Hosking and Wallis, 1988). RFA is also able to resolve the problem of the evaluation of precipitation extremes at ungauged sites within the same region without interpolation processes involving quantiles at gauged sites.

Advantages and features of the regional analysis have been described by Hosking and Wallis (1988) and further exploited by Gabriele and Arnell (1991). These last authors have applied an hierarchical approach for the RFA to extreme rainfall data using different combinations of probability distributions and parameters evaluation methods. Through Montecarlo simulations the authors were able to retrieve the efficiency of the assessment achieved identifying regions where respectively the skewness coefficient and the coefficient of variation are assumed to be constant; furthermore the same authors have observed that in the regions where the skewness coefficient is not constant, it can be worthy to use the same hierarchical approach that can be considered an alternative to the identification of completely independent regions.

Estimating of the regional frequency distribution can be achieved by the L-moments statistics. Hosking and Wallis (2005) stated that L-moments are an alternative system of describing the shapes of probability distributions. Historically they arose as modifications of the probability weighted moments (PWMs) of Greenwood et al. (1979). RFA based on L-moments has been implemented in several studies of extreme precipitation (Fowler et al., 2005, Fowler and Kilsby, 2003a,b, Schaefer, 1990, Zwiers and Kharin, 1998).

In Italy, the diffusion of the RFA methodology for the streamflow and the extreme rainfall was used from the VAPI (VAlutazione delle Piene in Italia; Flood evaluation in Italy) realized by GNDCI (a national group for the hydrological risk assessment). In this project the RFA was applied for the extreme rainfall modelling of different regions: Calabria (Ferro and Porto, 1988), Sardinia (Cao et al., 1991), Puglia (Copertino et al., 1992), Sicily (Cannarozzo et al., 1995), Campania (Rossi and Villani, 1994) and Veneto (Villi et al., 2001) and other regions.

Approaches used to identify homogeneous regions (Alexander et al., 2006, Coelho et al., 2008) usually consider all extremes of a station before applying spatial clustering, in order to identify regions which display similar behaviour, (e.g. Maraun et al., 2008). A final tool, adopted by several authors (e.g. Dales and Reed, 1989, Neal and Phillips, 2009, Wigley et al., 1984) for RFA analysis, is to simplify the rainfall characteristics through the Principal Component Analysis (PCA) followed by a clustering analysis to exploit regional groups (e.g. Blenkinsop et al., 2008, Jones et al., 2014).

The aim of this work consists in the design and the development of a RFA procedure for the area of Sicily, Italy, based on the selection of suitable procedures considering the data availability and the meteoclimatic features of the area. In the previous works related to the same area (Cannarozzo et al., 1995, Lo Conti et al., 2007), the choice of the number and the extension of the homogeneous regions were made with hydrological criterions related principally on watersheds boundaries. In this work, an objective method will be adopted to achieve the identification of the homogeneous regions with the research of the best distribution able to represent the characteristics of the regions identified. A procedure aimed to assess the quantile accuracy has been carried out to confirm the goodness of fit of proposed method.

The response of precipitation extremes under climate change has also been studied in this thesis using an ensables of climate models provided from *Coordinated Regional Climate Downscaling Experiment* (CORDEX). The CORDEX

project essentially has the twofold purpose of providing a framework to evaluate and benchmark model performance (*model evaluation framework*), and design a set of experiments to produce climate projections for use in impact and adaptation studies (*climate projection framework*). The evaluation of extreme precipitations will be carried out for two different climate scenarios, related to the increase of the CO_2, assessing the impact of the climate change on the extreme rainfall.

The manuscript is divided into six chapter: a theoretical part which includes the state of the art, the description of the phenomena and the tools used for the analysis (three chapters); a description of the area of study and the dataset used(one chapter) and an experimental part in which the two applications are described and discussed (two chapters). A detailed description of each chapter is given below.

Chapter 1 provides a general description of the extreme event analysing the different type of meteorological phenomena, focusing the attention on Mediterranean area that shows unique characteristics which condition the extreme precipitation events. Chapter 2 describes the state of the art and general principles of the frequency analysis, essential for the examination in depth of the peculiarities of this kind of approach. Chapter 3 describes the regional frequency analysis and the statistical tool that will be integrated in the different steps of the RFA; this approach will provide a new objective approach to regionalization procedure in Sicily. In chapter 4 the area of study, the Sicily, will be described together with its climatic and morphologic characteristics, showing the dataset used in the case of study as well. Chapter 5 will present the case study about RFA with the evaluation of the performance of identification of the new regions obtained after fitting the appropriate probability of distribution. Finally, Chapter 6 will derive the possible effects of climate change on the extreme precipitation through the use of the regional climate models (CORDEX project) that provide an indication of the possible evolution of the climate until the end of the 21th century.

Chapter 1

Extreme events

Human society is particularly vulnerable to severe weather and climate events that may cause damage to property and infrastructure, injury, and even loss of life. Although generally rare at any particular location, such events cause a disproportionate amount of loss.

Extreme weather and climate events are a major source of risk for all human societies. There is a pressing need for more research on such events. Various societal changes, such as increased populations in coastal and urban areas and increasingly complex infrastructure, have made us potentially more vulnerable to such events than we were in the past. In addition, the properties of extreme weather and climate events are likely to change in the 21^{th} century owing to anthropogenic climate change. The definition, classification, and diagnosis of extreme events are far from simple. There is no universal unique definition of what is an extreme event. This chapter discusses these issues and presents a simple framework for understanding extreme events focusing on the extreme rainfall in the Mediterranean area.

1.1 Definition of extreme events

Extreme events are generally easy to recognize but difficult to define. This is due to several reasons. First, there is no unique definition for what is meant by the word "extreme": several definitions are in common use. Second, the concept of "extremeness" is relative and so strongly depends on context. Third, the words "severe", "rare", "extreme", and "high-impact" are often used interchangeably. Here are some definitions of these terms (Stephenson et al., 2008):

- *Severe events* are events that create large losses in measures such as number of lives, financial capital, or environmental quality (e.g., loss of species). The

severity can be measured by the expected long-term loss, which is known as the risk. Risk depends on the product of the probability of the event (i.e., hazard), the exposure to the hazards (e.g., how many people are exposed), and the vulnerability (i.e., how much damage ensues when someone is hit by the event). Severity is a function of not only the meteorological hazard but also the human state of affairs. The severity of the events has increased considerably in recent years, mainly owing to increased numbers of people settled in the zone with more exposure producing increased of it.

- *Rare events* are events that have a low probability of occurrence. Because of the rarity of these events, human societies (and other ecosystems) are often not well adapted to them and so suffer large amounts of damage when they occur. Hence, despite their rarity, the large vulnerability associated with such events can often lead to large mean losses (and hence they are a type of severe event).

- *Extreme events* are events that have extreme values of certain important meteorological variables. Damage is often caused by extreme values of certain meteorological variables, such as large amounts of precipitation (e.g., floods), high wind speeds (e.g., cyclones), high temperatures (e.g., heat waves), etc.. Extreme is generally defined as either taking maximum values or exceedance above pre-existing high thresholds. Such events are generally rare; for example, extreme rainfall depth exceeding the 100-year return value, which have a probability of only 0.01 of occurring in any particular year.

- *High-impact events* are severe events that can be either short-lived weather systems (e.g., severe storms) or longer-duration events such as blocking episodes that can lead to prolonged heat waves and droughts. The *World Meteorological Organization* (WMO) uses the phrase "high-impact weather" rather than "severe weather" to help people avoid confusing the term severe with only short-lived events such as individual storms.

In addition to this potential source of confusion, extreme events have a variety of different attributes and so cannot be completely described by a single number. The multidimensional nature of extreme events is often overlooked in rankings of the events based on only one of the attributes (e.g., the category numbers for hurricanes based solely on maximum surface wind speed). Extreme events have attributes such as:

- rate of occurrence (probability per unit time);
- magnitude (intensity);
- temporal duration and timing;
- spatial scale (footprint);
- multivariate dependencies.

The temporal duration of extreme events plays an important role in the exposure and hence total losses. Temporal duration also provides a useful way of classifying extreme events. The duration is implicit when one describes an event as a "climate" extreme event rather than a "weather" extreme event. The difference between weather and climate is a measure of time. Weather is the condition of the atmosphere over a short period of time, and climate is how the atmosphere "behaves" over relatively long periods of time. Weather is basically the way the atmosphere is behaving, mainly with respect to its effects upon life and human activities. The difference between weather and climate is that weather consists of the short-term (minutes to months) changes in the atmosphere. In most places, weather can change from minute-to-minute, hour-to-hour, day-to-day, and season-to-season.

Weather is chaotic, which means that even a microscopic disturbance can lead to large scale changes. The climate is the long-term average of the weather over a number of years. It is shaped by global forces that alter the energy balance in the atmosphere, such as changes in the sun, the tilt of the Earth's axis, the amount of sunlight the Earth reflects back into space and the concentration of greenhouse gasses in the air.

If we consider the example of a man with his dog leash, it is possible to link the weather with the irregular sporadic pattern of the dog. Though it is difficult to predict where the dog is going, we can know the range of his meandering because he is on a leash. Conversely, man's straight path is like the climate, which is broadly predictable by observing long-term changes in global forces (Figure 1.1). Climate, however, is the average of weather over time and space.

FIGURE 1.1: Weather vs Climate (source: natgeotv.com)

With regard to the space and time scale of weather events, a definition of terms that describe these characteristics, among different meteorological phenomena, is essential. Phenomena having short time scales also tend to have small spatial scales, and vice versa (Figure 1.2).

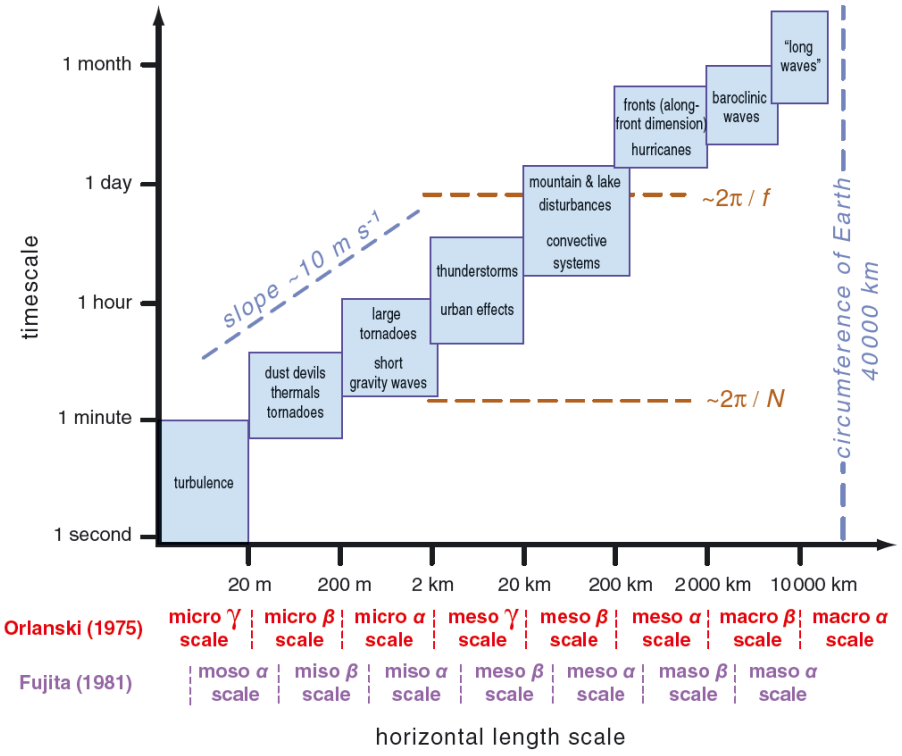


FIGURE 1.2: Scale definitions and the characteristic time and horizontal length scales of a variety of atmospheric processes.

Before defining what is meant by "mesoscale", it may be easiest first to define what is meant by the "synoptic" scale. The adjective "synoptic" is defined in the American Meteorological Society's *Glossary of Meteorology* as referring to meteorological data that are obtained simultaneously over a wide area in order to present a nearly instantaneous snap-shot of the state of the atmosphere. The early synoptic charts displayed the limited amount of data that could be collected routinely at the same times on a daily basis, and disturbances that could be resolved on these charts eventually were referred to as "synoptic-scale" disturbances. Thus, the term "synoptic", though not initially intended to define a scale, ultimately became a term used to describe the scale of large-scale weather systems, which were the only types of meteorological phenomena that could be resolved regularly by the coarse resolution observing platforms of the middle 20^{th} century.

The term "mesoscale" is believed to have been introduced by Ligda (1951), in order to describe phenomena smaller than the synoptic scale but larger than the "microscale", a term that was widely used at the time (and still is) in reference to phenomena having a scale of a few kilometres or less.

Today the mesoscale "officially" (according to the American Meteorological Society's *Glossary of Meteorology*) is defined as the 2–2000 km scale, with subclassifications of meso-α, meso-*beta*, and meso-*gamma* scales referring to horizontal scales of 200–2000 km, 20–200 km, and 2–20 km, respectively. In this classification scheme, the microscale is reserved for horizontal scales smaller than 2 km.

1.2 Extreme rainfall in the Mediterranean

Mediterranean Sea is located in a transitional zone where mid-latitude and tropical features variability are both relevant. Thus, the Mediterranean climate region evolves on the north to the Marine West Coast Climate (from 40° to subpolar regions) and on the south to the Subtropical Desert Climate (southward of 30° or 25°). Further, the Mediterranean climate is exposed to the *South Asian Monsoon* (SAM) in summer and the Siberian high pressure system in winter.

The southern part of the region is mostly under the influence of the descending branch of the Hadley cell, while the Northern part is more linked to the mid-latitude variability, characterized by the *North Atlantic Oscillation* (NAO) and other mid-latitude interconnections patterns. The NAO is the main cause of winter precipitation variability (and climate generally) in the North Atlantic region

covering central North America to Europe, extending into Northern Asia. The NAO is the seasaw of barometric pressure differences between the Azores subtropical high and the Icelandic polar low (Jones et al., 1997).

However, the climate variability patterns present a large amount of synoptic to mesoscale spatial variability, inter-seasonal and multi-decade to centennial time variability. An important consequence is that the analysis of the Mediterranean climate can be used to identify changes in the intensity and extension of global scale climate pattern like NAO, ENSO (*El Niño Southern Oscillation*) and the monsoons and their region of influence.

On the other hand, the large-scale atmospheric circulation exerts a strong influence on the cold season temperature and precipitation over the Mediterranean, though the strength of the relation varies with region. The largest amount of studies on the effect of the mid-latitude variability refers to the role of NAO which determines a large and robust signal on winter precipitation, which is anti-correlated with NAO over most of the western Mediterranean region (Xoplaki, 2002). However, in its Eastern part the advection of moisture from the Mediterranean itself produces a more complex situation, and eventually other large-scale patterns, like EA (*East Atlantic*), play an important role, and, in the central Mediterranean, the Scandinavian pattern has a strong influence (e.g. Xoplaki, 2002). This is superimposed with the effect of tropical variability, specifically with a reduction of cyclones in the Mediterranean area during *La Niña* events. Tropical variability events, like ENSO (Mariotti et al., 2002, Rodo et al., 2002), can be important in the parts where NAO influence is weaker (Rodó et al., 1997).

There are evidences that ENSO is significantly correlated with winter rainfall in the Eastern Mediterranean (Price et al., 1998, Yakir, 1996). However it is still open for debate, what could be the physical mechanisms for these links. In summer, when the advection of moisture from the Atlantic is weaker and the Hadley cell moves northward and attenuates, there are evidences of connections with the Asian and the African monsoons (stronger in the eastern part). The influence of NAO on the Mediterranean temperature is weaker than on precipitation and the observed correlation has been found to be non-linear and non-stationary (Pozo-Vázquez et al., 2001).

Mediterranean summer temperatures have no relation with the NAO, and they are not adequately linked to larger scale patterns. Rather, warm Mediterranean summers are connected with blocking conditions, subsidence, stability, a warm

lower troposphere and positive Mediterranean Sea surface temperature (Xoplaki et al., 2003).

The variability of large and locale scale meteorological factors that arise around Mediterranean Sea, are reflected on the variegate climatology that is observable on related lands. The western side of Africa coast is mostly classified as warm desert climate, while on the eastern side Algeria and Morocco coasts areas are classified as warm Mediterranean or cold semiarid climate as well as part of Spanish coast, some Italian, Greek and Turkish, costs and islands (then comprising Sicily). On the north temperate Mediterranean, temperate oceanic and warm oceanic/humid subtropical climate classes are attributed to French coasts, adjacent Spain and Italian lands and along the eastern Adriatic coast.

The Mediterranean region is characterized by the occurrence of short and very intense storms. These phenomena often cause landslides and floods with considerable loss of human life and enormous economic damage. These events (which can be classified as extreme events) are very rare when referring to specific sites. However their frequency is not negligible in a regional context. This seems to be more evident when analysing the annual maxima for rainfall and flow data series. In fact, there are some values much higher than the others and called "outliers" which do not seem to come from the same sample of data.

In the '90s the *Italian National Group for Defence from Hydrological disasters* (GNDCI) has developed a National Program for assessing floods in Italian watercourses, involving dozens of researchers in different operative units representing all Italian regions.

The latest developments in meteorology, for example, highlights how the extreme variability of rainfall is strongly influenced by the rainfall generating processes. These processes can be divided into three categories:

1. *frontal or cyclonic phenomena*, rain storm cell aggregates, which are the dominant structure in the Mediterranean weather (baroclinic cyclogenesis marked with orographic effects);

2. *isolated cells storms*, convective rainfall combined with evolving local phenomena, just like the typical summer storms;

3. *cyclonic vortex* (hurricane-like phenomena), rainfall generated by meteorological structures like meso-cyclonic vortices.

The first process produces ordinary events, while the other two processes generate extraordinary events. The "clustering" of high intensity rainfall areas embedded within rainfall areas of lower intensity can be analysed at many different scales (Willems, 2001). At the smallest scale, the individual rain cell forms the building-block of a spatial rainfall structure. The rain cells are embedded in a clustered way within "small mesoscale areas" (10^2–10^3 km^2). At larger scales, small mesoscale areas occur in a clustered way within lower intensity "large mesoscale areas", which in turn are embedded within some synoptic-scale lowest intensity rainfall field. The large mesoscale areas are observed as (sometimes elongated) bands and their spatial extent varies from 10^3 to 10^4 km^2.

The distribution of rainfall is strongly influenced by the orography (Tropeano and Furcolo, 2005; Cuomo and Guida, 2010) because, from a cinematic point of view, it promotes the channelling of atmospheric currents along preferential directions. On the other hand, from a thermodynamic point of view, it interacts with the overlying air, lifting the hot masses and transferring radiation through complex heating exchanges between the air and the soil.

The rainfall variability due to the orography and to the atmospheric circulation is difficult to be described at the different scale of interest. It proves to be even more difficult when only a few rainfall data sets are available (e.g. daily rainfall observations in few points of a region of interest). In the case of intense storms, especially for the convective storms, it may happen that the most violent rainfall might escape the observation. This condition is often caused by the low density of measurement points compared to the variability of the phenomenon and by the unavailability of a radar system.

1.2.1 Baroclinic cyclogenesis

The baroclinic cyclogenesis is the dominant meteorological phenomenon in the Mediterranean. The rainfall is associated with the movement of sub-horizontal baroclinic cyclones (e.g. Holton, 1992), when large masses of warm moist air move and meet cold air masses.

To the passage of cold fronts (Figure 1.3) can be associated with two types of facilities, as described by Bacchi (1996):

- Bands parallel to the line of the cold front, having a length of a few hundred kilometres in front of a width of a few kilometres. The sudden lifting of

warm air masses generate heavy rainfall, also the order of 100 mm/h, with a duration limited to a few minutes.

- Formation of bands of greater width (a few tens of kilometres), set back from the bands on either side of the face, accompanied by less intense rainfall (1-10 mm/h) duration of the order of the hour.

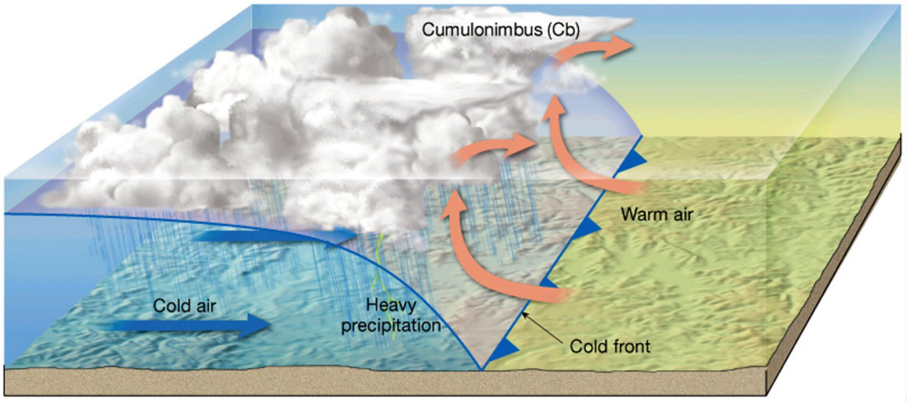


FIGURE 1.3: Cold front (source: Lutgens and Tarbuck, 2004)

Always Bacchi (1996), inspired by the work of Hobbs and Locatelli (1978) and Houze Jr et al. (1976) show that the warm fronts (Figure 1.4) are generally associated with stratiform precipitation by band structures, far between a few tens of kilometres and having extensions of the order of 50-100 km. The rainfall associated with this type of structures are not particularly intense (1-10 mm/h) and have durations of the order of the hour.

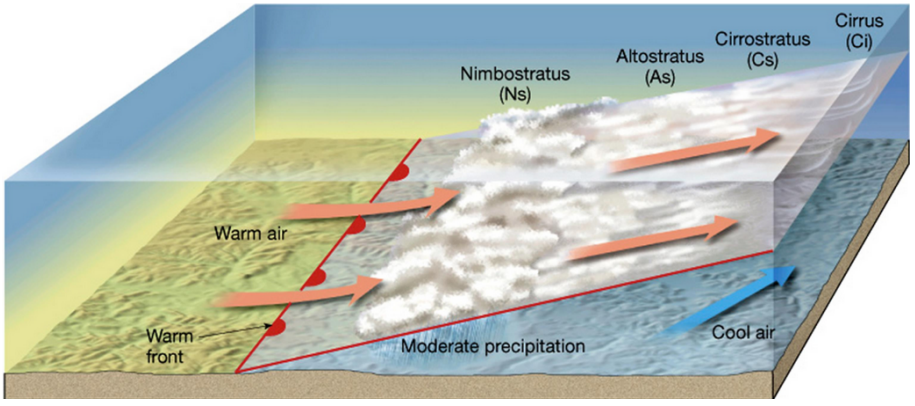


FIGURE 1.4: Warm front (source: Lutgens and Tarbuck, 2004)

The interaction of a large-scale baroclinic wave with an orographic obstacle like the Apennines is the cause of a smaller-scale, orographically induced, baroclinic lee cyclone that is generated to the lee of the obstacle (Reale and Atlas, 2001).

The precipitation associated with cyclonic and frontal storms are of the level of the "large mesoscale areas" that have an extensions of about 10^3–10^4 km^2 (Willems, 2001). Baroclinic structures, then, generally generate rainfall events with large extent, in which it's often possible to identify cells, or clusters of cells, with high intensity. Frontal events duration varies from several hours to several days. These events also move hundreds of kilometres accordingly to the atmospheric circulation and generate ordinary extreme events more frequent and less severe on average.

1.2.2 Isolated convective cells

In temperate zones, like the Mediterranean, it is common to observe, especially in the warm season, the formation of events with a low spatial scale, thunderstorms, which for this reason they are often referred to as summer storms. This type of event, which has convective genesis is denoted by one or more cumulus clouds (cumulonimbus) generated by convection of charged atmosphere of water vapour, made unstable by heating close to the ground. Rain cells and cell clusters most often appear in large and small mesoscale areas but they can also occur isolated outside these regions (e.g. in air mass thunderstorms) (Willems, 2001). Moisello (1999) showed that the life of a cell is about half an hour, but the storm can last for several hours as the cells succeed one behind the other.

The evolution of these cells, which is depicted schematically in Figure 1.5, can be conceptualized as occourring in three stage:

1. *cumulus stage*, during which the cells behave very much like the nonprecipitating convective clouds, although the updrafts may be somewhat stronger. The air is moving upward from the subcloud layer at all levels, and precipitations begin to form toward the end of this stage;

2. *mature stage* marks a dramatic transition in the dynamics of the convective cell. During this stage, precipitation is forming quite rapidly, and may reach the ground before the end of the stage. Large changes in the buoyancy distribution in the cloud result from the redistribution of condensed water mass, from the reevaporation of some of the precipitation in subsaturated environmental air, from mixing of cloudy air with its environment, and from melting of falling snow, graupel, or hail;

3. *dissipating stage* where the spreading cold air negates the ability of pressure perturbations associated with the buoyant cloud to draw up subcloud-layer

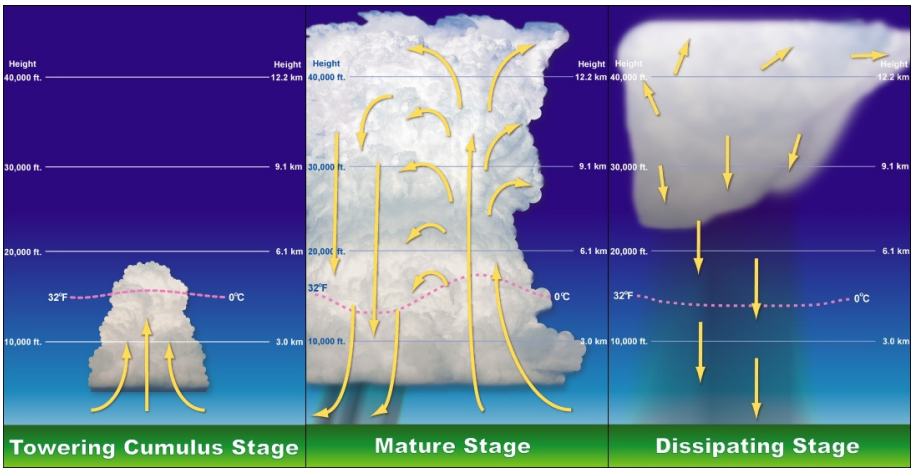


FIGURE 1.5: Isolated convective cells

air, because the latter has become decidedly stable and dense. The updraft collapses, and the downdraft decays on a time scale dictated by the fall speed of precipitation and the depth of the cloud. The remaining cloud at low levels evaporates through turbulent mixing with the surrounding, unsaturated air, while large concentrations of ice may remain in the high troposphere as an "anvil" cloud, produced by the outward-spreading updraft air. Here, the saturation vapour pressure is so small that the evaporation of even a small amount of condensed water quickly saturates the air.

The spatial extent of individual rain cells, in the order of 10–50 km^2 (Waymire et al., 1984), is indeed of the same order of magnitude as the spatial extent of most urban and small hydrographic catchments. Therefore the study of isolated cells requires a very dense rain gauge network or a radar system.

Convective precipitation systems are generally organized in isolated storm cells or clusters. This kind of phenomenon is associated with the convective movement of warm moist air masses towards the cold layers of the overlying atmosphere. It is also supported by contributions of energy and water vapour from limited areas, carried by a convergent flow of air masses. In the areas affected by convective rainfall the instantaneous and average intensity vary within a wide range, typically 10-100 mm/h within a few minutes. The spatial distribution of a convective rainfall is also heterogeneous and each cell usually lasts less than an hour. However before dissipating all the active cells, it may take several hours.

1.2.3 Tropical Like Cyclones

In the Mediterranean, vortices of great intensity and small scale were observed (Reale and Atlas, 2001), mainly concentrated in autumn (between August and November), which have characteristics of convective events more typical of tropical like cyclones (Figure 1.6) and they are named *Mediacane*, a *portmanteau* of the words Mediterranean and hurricane.

Mediterranean tropical cyclones are not considered to be formally classified tropical cyclones and their region of formation is not officially monitored by any agency. A majority of Mediterranean tropical cyclones form over two separate regions. The first, more conducive for development than the other, encompasses an area of the western Mediterranean bordered by the Balearic Islands, southern France, and the shorelines of the islands of Corsica and Sardinia. The second identified region of development, in the Ionian Sea between Sicily and Greece and stretching south to Libya, is less favourable for tropical cyclogenesis. An additional two regions, in the Aegean and Adriatic seas, produce fewer medicanes, while activity is minimal in the Levantine region.

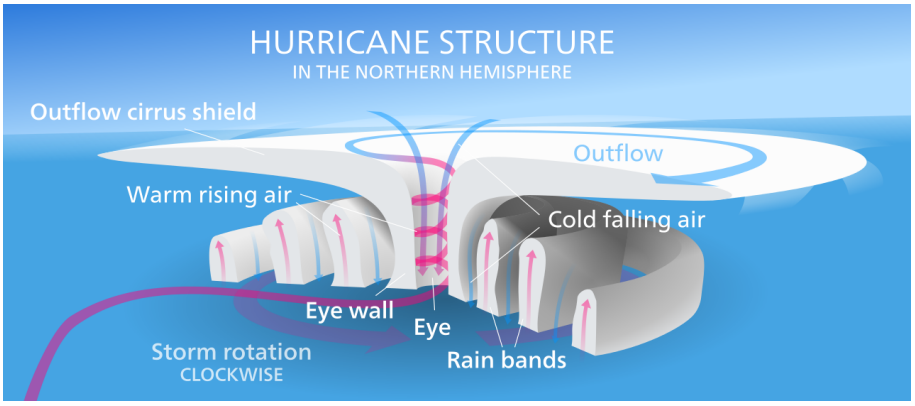


FIGURE 1.6: Diagram of a Northern hemisphere hurricane.

The geographical distribution of Mediterranean tropical cyclones is markedly different from that of other cyclones, with the formation of regular cyclones centring on the Pyrenees and Atlas mountain ranges, the Gulf of Genoa, and the island of Cyprus in the Ionian Sea. Although meteorological factors are most advantageous in the Adriatic and Aegean seas, the closed nature of the region's geography, bordered by land, allows little time for further evolution (Cavicchia et al., 2014).

The geography of mountain ranges bordering the Mediterranean are conducive for severe weather and thunderstorms, with the sloped nature of mountainous

regions permitting the development of convective activity (Homar et al., 2003). Although the geography of the Mediterranean region, as well as its dry air, which typically prevent the formation of tropical cyclones, under certain meteorological circumstances, these obstacles can be overcome (Emanuel, 2005).

The occurrence of tropical cyclones in the Mediterranean Sea is generally extremely rare, with an average of 1.57 forming annually and merely 99 recorded occurrences of tropical-like storms discovered between 1948 and 2011 in a modern study, with no definitive trend in activity in that period (Cavicchia et al., 2014). Few medicanes form during the summer season, though activity typically rises in autumn, peaks in January, and gradually decreases from February to May (Cavicchia et al., 2014). In the western Mediterranean region of development, approximately 0.75 such systems form each year, compared to 0.32 in the Ionian Sea region (Cavicchia et al., 2014).

Therefore, the Mediterranean is affected by meso-cyclonic vortices with very low frequency (not negligible though), similar to tropical cyclones. The rainfall due to the tropical like cyclones is characterized by an intensity much higher than that of the frontal phenomena. It generally occurs over a localized area with rapid advection over coastal areas. The events, to be considered "extraordinary", cause rainfall up to 30-50% of average annual rainfall, covering up to 100-1000 km^2 areas in less than 24 hours. The Figure 1.7 shows satellite imagery of a Mediterranean tropical-like cyclone.

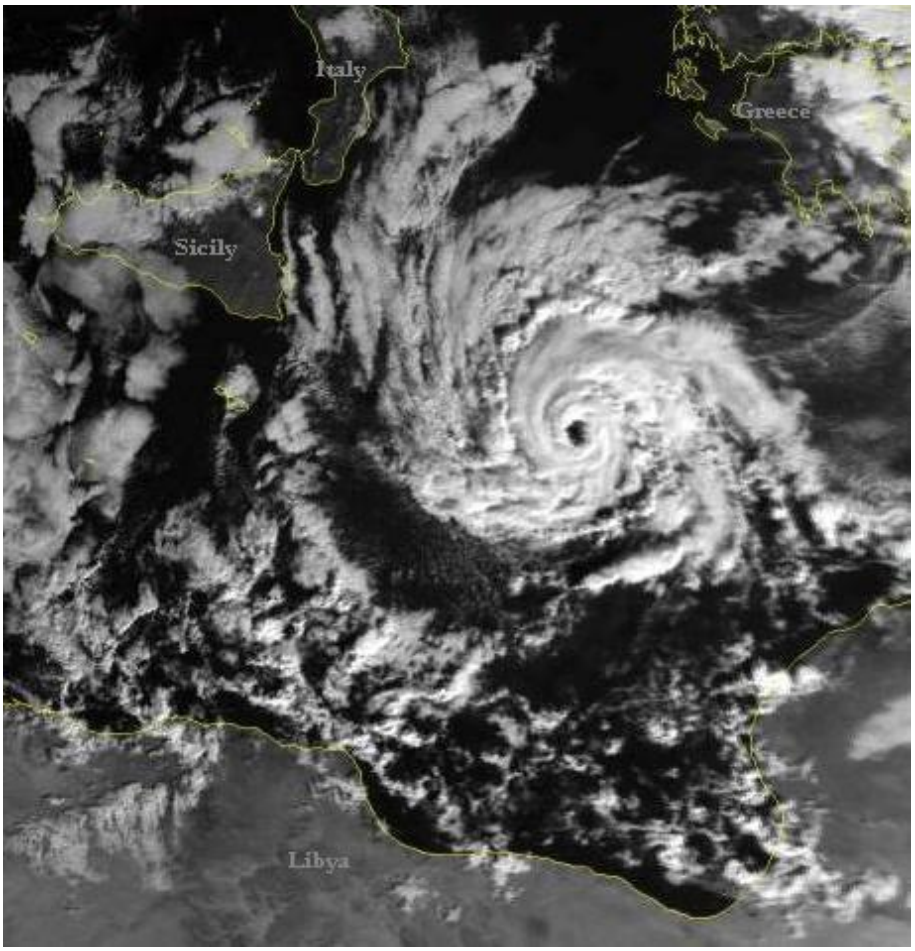


FIGURE 1.7: NOAA-7 visible satellite imagery of a Mediterranean tropical-like cyclone at 1236 UTC on 26 January 1982.

Chapter 2

Frequency analysis of extreme rainfall events

The purpose of frequency analysis in hydrology is to analyse past records of hydrologic variables so as to estimate future occurrence probabilities. The data used in the analysis must be evaluated in terms of the objectives, length of records available and completeness of records. It must also satisfy certain statistical criteria such as randomness, independence, homogeneity and stationarity. A frequency analysis can be performed using single-site data, regional data or both. It can also include historical information and reflect physical constraints.

Because hydrological phenomena are characterized by great variability, randomness and uncertainty, it should, therefore be recognized that statistical analysis of hydrological data will not always yield a true answer. The sources of uncertainty in frequency analysis include representativeness of the analytical approach, selection of the probability distribution and estimation of parameters.

In the case of extreme events, our major interest is not in what has occurred, but the likelihood that further extreme and damaging events will occur at some point in the future. A probabilistic approach is required to incorporate the effects of such phenomena into decisions. If the occurrences can be assumed to be independent in time the frequency analysis can be used to describe the likelihood of any one or a combination of events over the time horizon of a decision.

In this chapter, the extreme events will be described with relative the state of the art of frequency analysis, subdividing the works among *at-site* and *regional* frequency analysis. General probability concepts will be introduced, as the L-moments and the statistical criteria that the data series must meet. In the case

of the screening of data, the discordancy measure, derived using L-moments concepts, will be introduced. The discordancy measure provides an useful tool for the initial screening of the data and indicates sites where the data may merit close examination. Some probability distribution of extreme events will be described concurrently at the method aimed to the parameter estimation and the methods use to choose an appropriate frequency distribution will be discussed. Finally, at-site frequency analysis of rainfall will be described, while the regional frequency analysis will be deepen in the next chapter.

2.1 State of the art of rainfall frequency analysis

A frequency analysis can be performed for a site for which sufficient rainfall data are available. Similar to flood frequency analysis, rainfall frequency analysis is also based on annual maximum series or partial duration series (e.g. Wang, 1991, Wilks, 1993).

Probability distributions can be assessed for the mean rainfall intensities at the relevant time scales, but most important for urban drainage and flash flooding applications are the rainfall extremes. These extremes can be extracted from the full rainfall series using the classical approach of annual maxima (Coles et al., 2001) where the annual maximum within a (hydrological) year is concluded in the extreme value analysis. Traditionally, this approach has been used for analysing rainfall extremes (e.g. Alila, 1999, Schaefer, 1990, Wallis et al., 2007).

Another approach considers events above threshold level in the extreme value analysis. This approach, referred to as the *Partial Duration Series* (PDS) or *Peak-Over-Threshold* (POT) method, has been used for analysing extreme rainfall at fine temporal scales in for example Beguería and Vicente-Serrano (2006), Madsen et al. (2002), Wang (1991), Wilks (1993), Willems (2009), Willems et al. (2007). The pros and cons of the *Annual Maxima Series* (AMS) approach versus the POT method have been discussed, amongst others, by Madsen et al. (1997), Stedinger and Vogel (1993).

For example, Wilks (1993) have investigated the performance of 8 three-parameter probability distributions for representing annual extreme and partial duration precipitation data at stations in the northeastern and southeastern United States. Particular attention was paid to fidelity on the right tail, through use of a bootstrap procedure simulating extrapolation on the right tail beyond the data. The authors have found that the beta-k distribution best describes the extreme right

tail of annual extreme series, and the beta-P distribution is best for the partial duration data. The conventionally employed two-parameter Gumbel distribution was found to substantially underestimate probabilities associated with the larger precipitation amounts for both annual extreme and partial duration data. Fitting the distributions using left-censored data did not result in improved fits to the right tail. Arguments in favour of either of these techniques are contained in the literature (Stedinger and Vogel, 1993, Watt et al., 1989).

Procedure for statistical frequency analysis of a single set of data are well established. However in some case, the dataset of the different measuring sites have insufficient data. In this case, if the meteorological or environmental observation of the same variables at different measuring sites made on samples of similar products, and if the event frequency are similar for the different observed quantities, a more accurate conclusion can be reached by analysing all of the data samples together than by using only a single sample. In environmental application this approach is know as *regional frequency analysis*, because the data samples analysed are typically observations of the same variables at number of measuring sites within a suitably defined "region".

For example, Schaefer (1990) used the index flood methodology to conduct regional analyses of annual maximum precipitation data in Washington State. It has been shown that climatically homogeneous regions can be defined based on the mean annual precipitation. Further, it was found that the coefficients of variation and skewness of annual maximum rainfalls vary systematically with the mean annual precipitation. Hence, all sites within a homogeneous region could be characterized by a specific three-parameter probability distribution, such as the generalized extreme value, having fixed values of the coefficients of variation and skewness. However, the use of mean annual precipitation as an index variable may not be appropriate for other regions with different climatic or topographic conditions. For instance, the median of annual maximum rainfalls at a site was recommended as the index variable for regional estimation of extreme rainfalls in the United Kingdom (Faulkner, 1999).

In general, one of the main difficulties in the application of this technique is related to the definition of homogeneous regions. Various methods have been proposed for determining regional homogeneity, but there is no generally accepted procedure in practice (Mills, 1995, Nguyen et al., 2002).

Zwiers and Kharin (1998) used L-moments to study precipitation data produced by global climate models to compare current climate to modelled climate

under CO_2 doubling. Fowler and Kilsby (2003a,b) and Fowler et al. (2005) used L-moments to study extreme precipitation in the United Kingdom via RFA.

Approaches used to identify homogeneous regions (Alexander et al., 2006, Coelho et al., 2008) usually consider all extremes of a station before applying spatial clustering, in order to identify regions which display similar behaviour, (e.g. Maraun et al., 2008). A final tool, adopted by several authors (e.g. Dales and Reed, 1989, Neal and Phillips, 2009, Wigley et al., 1984) for RFA analysis, is to analyse the rainfall characteristics through the Principal Component Analysis (PCA) followed by a clustering analysis to identify regional groups (e.g. Blenkinsop et al., 2008, Jones et al., 2014).

In Italy the diffusion of the methodology for extreme rainfall was used from the VAPI (*VAlutazione delle Piene in Italia; Flood evaluation in Italy*) realized by GNDCI. In this project the RFA was applied for the extreme rainfall modelling of different regions: Calabria (Ferro and Porto, 1988), Sardinia (Cao et al., 1991), Puglia (Copertino et al., 1992), Sicily (Cannarozzo et al., 1995), Campania (Rossi and Villani, 1994) and Veneto (Villi et al., 2001) and other regions.

With regard to the works in Sicily, Cannarozzo et al. (1995) presented a regional frequency analysis for the Sicilian region which used the TCEV distribution and a hierarchical regionalization technique. The rainfall analysis, carried out with rainfall data recorded by Sicilian raingauges, showed a duration dependence of the TCEV parameters and allowed even for the determination of the rainfall depth-duration relationship at ungauged sites and at gauged sites with short records.

At the first level of the hierarchical regionalization procedure, Sicily is considered to be a homogeneous region with regard to the skewness and kurtosis coefficient and the *duration-dependent* hypothesis for the TCEV parameters Λ^* and θ^* was considered.

At the second level the region is divided into smaller areas, named homogeneous sub-regions, in which the λ_1 parameter is assumed constant. For each duration d, generating samples by a Monte Carlo technique, a Λ_1^* parameter (dependent on the duration) was used equal to the mean value of the λ_1 parameters of all the recording raingauges located in each sub-region; the relationships dependent on the duration d was evaluated for the λ_1 parameters.

The last step of the hierarchical procedure aims to determine a regional criterion to estimate μ_R even for ungauged sites or for sites with short records. Since the theoretical μ_R values were fixed equal to the empirical one, m_c, the third regionalization level was developed via m_c values calculated through a power-law

equation depending on the duration.

Lo Conti et al. (2007) carried out a regional frequency analysis procedure and applied to precipitation data in the island of Sicily, Italy. Annual maximum series (AMS) for rainfall durations of 1, 3, 6, 12 and 24 h from about 300 raingauges, for the period 1928-2004, were used for various statistical analyses. The identification of homogeneous regions was performed using only site characteristics (latitude, longitude, mean annual precipitation) and required a minimum of subjective relocation of stations in order to obtain relatively homogeneous regions evaluated through the Hosking and Wallis test (HW).

Two regional probability distributions were used in order to obtain depth-duration frequency curves: *generalized extreme value* (GEV) and *two component extreme value* (TCEV) distribution. The regional parameters of these distributions were estimated using the L-moment ratios approach. Different statistics were computed and compared with those resulting from at-site Monte Carlo simulation showing that the regional procedures improves the quantiles prediction.

Gabriele and Chiaravalloti (2013) provide a new approach based on meteo-climatic information in the regional daily rainfall frequency analysis. By starting with reanalysis meteorological data, the thermodynamic and dynamic atmospheric instability indexes *Convective Available Potential Energy* (CAPE), *vector Divergence* (QD), and the *Vertically Integrated Moisture Flux* (J) were used to search regional homogeneity in the absence of *a priori* rainfall data.

In the first step, the proposed methodology attempted to verify the possibility of identifying large areas that are related to climatic characteristics and contain prevalent high CAPE values or low QD values.

In these areas, which are prone to experience intense convective precipitation, the meteorological context, which is related to the annual maxima, is expected to be characterized by high CAPE values. In leeward areas, the direct exposure to deep fronts may characterize large areas in which annual maxima daily rainfall are generated by long precipitation with a high frequency of negative QD values.

The application of the *Hot Spot Analysis*, that is a statistical tool able to evaluate possible clustering of vector variables able to emphasize the tail of the spatial distribution of the frequency of CAPE and QD, confirmed the presence of homogeneous large areas named *macro-regions*; in these *macro-regions*, the characteristics of the CAPE and QD indexes were considered constant.

In the second step, the mean direction of the *Vertically Integrated Moisture Flux* (J) was investigated inside each macro-region by the use of kriging analysis. The

results showed that, especially for stratiform rainfall, flux is strongly influenced by orographic barriers and, determined by these barriers, there were many sub-areas that display a constant mean flux direction in the Italian Peninsula. In order to verify the proposed procedure, two regional analyses were performed; one was performed in the north-east in a stratiform macro-region (identified by high frequencies of low QD values). While a second was performed in the south Sicily in a convective macro-region (identified by high frequencies of high CAPE values). The results of the heterogeneity test clearly confirmed the rainfall homogeneity and verify the goodness fit, by the Z-statistic test, for the *Generalized Logistic distribution*. The proposed approach was evaluated for the two regions, which contained a well-defined climatic homogeneity of CAPE, QD and J, seemed capable of detecting rainfalls homogeneity without the use of direct observations. To reduce the number of indexes computation the authors used rainfalls to select the events to be analysed. Since there was no guarantee of the correspondence of the days of annual maximum of rainfalls and atmospheric indexes, more experiments were performed to find the representative meteorological events in the absence of a priori selection based on rainfall. The GLO distribution and the TCEV have shown better performance.

2.2 Probability concepts

In frequency analysis, a series is a convenient sequence of data, such as hourly, daily, seasonal or annual observations of a hydrological variable. If the record of these observations contains all the events that occurred within a given period, the series is called a complete duration series. Regardless of the method employed to obtain rainfall series, through observation or through stochastic generation, a regardless of the temporal scale, statistical analysis is required to quantify the probabilities of the rainfall intensities.

The AMS method considers only the maximum events within a year although other events in the year may exceed annual maxima of the other years. The use of the annual maximum series is very common in frequency analyses for two reasons. The first is for convenience, as most data are processed in such a way that the annual series is readily available. The second is that there is a simple theoretical basis for extrapolating the frequency of annual series data beyond the range of observation. A limitation of annual series data is that each year is represented by

only one event. The second highest event in a particular year may be higher than the highest in some other years, yet it would not be contained in the series.

The use of POT can address this issue because all peaks above the specified threshold are considered (WMO, 2008). The POT approach provides a more consistent definition of the extreme values be considering all events above a threshold. However, as opposed to the AMS approach that generally assures independents events, independence criteria have to be defined to ensure independence between extreme events in the POT series. In addition, the POT method includes selection of threshold level, which will introduce some sort of subjectivity in the extreme value analysis. Due to its simpler structure, the AMS-based method is more popular in practice.

The POT analysis, however, appears to be preferable for short records, or where return periods shorter than two years are of interest. The theory and application of the AMS approach have been well documented in hydrologic and engineering literature (Chow, 1964b, Stedinger and Vogel, 1993, Watt et al., 1989). As for any statistical analyses, both the quantity and quality of the data used are important.

Both the methods above described are directed to evaluate the quantile of the distribution, usually as function of the return period (T), with the aim to design and management of hydraulic works. The main steps of the quantiles evaluation are:

1. analysis of the dataset;
2. evaluation of statistic parameter;
3. population statistics and parameter estimation;
4. goodness-of-fit test of distribution.

All these steps will be discussed in the following sections providing also some information about the at-site frequency analysis, while the regional frequency analysis will described in the next chapter.

2.2.1 The rainfall data and hypothesis testing

Rainfall data used for frequency analysis are typically available in the form of annual maximum series, or are converted to this form using continuous records of hourly or daily rainfall data.

The precipitation data should be collected for a long period of time. A sufficiently long record of precipitation data provides a reliable basis for frequency determinations. It is known that a data sample of size n, in the absence of a priori distributional assumptions, can provide information only about exceedance probabilities greater than approximately $1/n$ (NRCC, 1998). It is a common rule of thumb to restrict extrapolation of at-site quantile estimates to return periods (years) of up to twice as long as the record length (Watt et al., 1989). Hence, long-term precipitation data are extremely valuable for determining statistically based rainfall estimates of reasonable reliability, especially for extreme rainfalls with high return periods, such as those greater than 100 years.

The quality of precipitation data may affect its usability and proper interpretation in flood frequency analysis studies. Precipitation measurements are subject to both random and systematic errors (Sevruk, 1985). The random error is due to irregularities of topography and microclimatical variations around the gauge site. Random errors are also caused by inadequate network density to account for the natural spatial variability of rainfall. The systematic error in point precipitation measurement is, however, believed to be the most important source of error. The largest systematic error component is considered to be the loss due to wind field deformation above the orifice of elevated precipitation gauges. Other sources of systematic error are wetting and evaporation losses of water that adheres to the funnel and measurement container, and rain splash.

The outliers could provide critical information for describing the upper tail of the rainfall distribution. Hence, high outliers are considered to be historical data if sufficient information is available to indicate that these outlying observations are not due to measurement errors. Regarding data inconsistency, there are many causes. Changes in gauging instruments or station environment may cause heterogeneity in precipitation time series.

The data series must meet certain statistical criteria such as randomness, independence, homogeneity and stationarity in order to guarantee for the results of a frequency analysis to be theoretically valid. These statistical criteria are listed in Table 2.1. A more detailed description of many of tests for the different criterion can be found in Helsel and Hirsch (1992).

Statistical tests can only indicate the significance of the observed test statistics and do not provide unequivocal findings. It is therefore important to clearly understand the interpretation of the results and to corroborate findings with physical evidence of the causes, such as land use changes. When data do not satisfy the

TABLE 2.1: Statistical tests and statistical criteria

Criterion	Explanation	Applicable statistical tests
Randomness	In a hydrologic context, randomness means essentially that the fluctuations of the variable arise from natural causes.	No suitable tests for hydrological series are available.
Independence	Independence implies that no observation in the data series has any influence on any following observations. Even if events in a series are random, they may not be independent. The elements of annual series of short-duration rainfall may, in practice, be assumed to be independent.	Anderson as described in Chow (1964a). Spearman rank order serial correlation coefficient as described in NERC (1975).
Homogeneity	Homogeneity means that all the elements of the data series originate from a single population. Elderton (1953) indicated that statistics are seldom obtained from strictly homogeneous material. When the variability of the hydrological phenomenon is too high, as in the case of extreme precipitation, non-homogeneity tends to be difficult to decipher (Miller, 1972), but non-homogeneity in yearly precipitation sums is easier to detect.	Terry (1952).
Stationarity	Stationarity means that, excluding random fluctuations, the data series is invariant with respect to time. Types of non-stationarity include trends, jumps and cycles. Trends may be caused by gradual changes in climatic conditions or in land use, such as urbanization. Cycles may be associated with long-term climatic oscillations.	Spearman rank correlation coefficient test for trend NERC (1975). Wald and Wolfowitz (1943) test for trend. No satisfactory method of testing is available for long-period cycles. Mann–Kendall test for trend Yue et al. (2002)

assumptions, then a transformation can often be employed so that the transformed observations would meet the criteria required for analysis. Caution is advised in interpolation and extrapolation when data do not meet the assumptions.

2.2.2 The *L-moments*

It is standard statistical practice to summarize a probability distribution or an observed data set by its moments (i.e. expected value, variance, skewness, etc.). In this method, while fitting a probability distribution to a sample, the parameters are estimated by equating the sample moments to those of the theoretical moments of the distribution. It is sometimes difficult to assess exactly what information about the shape of a distribution is conveyed by its moments of third and higher order; the numerical values of sample moments, particularly when the sample is small, can be very different from those of the probability distribution from which the sample was drawn; and the estimated parameters of distributions fitted by the method of moments are often markedly less accurate than those obtainable

by other estimation procedures such as the method of maximum likelihood. The alternative approach described here is based on quantities called *L-moments*.

Hosking and Wallis (1993) has defined *L-moments*, which are analogous to conventional moments, and can be expressed in terms of linear combinations of order statistics. Basically, *L-moments* are linear functions of probability weighted moments (*PWMs*) (Greenwood et al., 1979). *L-moments* have theoretical advantages over conventional moments of being able to characterize a wider range of distributions and, when estimated from a sample, of being more robust to the presence of outliers in the data. Experience also shows that, compared with conventional moments, *L-moments* are less subject to bias in estimation and approximate their asymptotic normal distribution more closely in finite samples. Parameter estimates obtained from *L-moments* are sometimes more accurate in small samples than even the maximum likelihood estimates. Lettenmaier et al. (1987) and Haktanir (1992) have obtained, in terms of statistical robustness, good performance with parameters estimated using the moments weighted into probability (PWM). They are in fact able to produce estimates of the quantiles with low distortion and small standard error observed even if the sample belongs to a population different from that assumed.

Just as the variance, or coefficient of skewness, of a random variable are functions of the moments E[X], E[X^2], and E[X^3], *L-moments* can be written as functions of *probability-weighted moments*, which can be defined as

$$\beta_r = E\{X[F(X)]^r\} \tag{2.1}$$

where *F(X)* is the *cdf* for X and r the order of the moment. Probability-weighted moments are the expectation of X times powers of *F(X)*. For $r = 0$, β_0 is the population mean μ_0.

The *L-moments* provide excellent performance in terms of robustness and accuracy of the estimation of the parameters especially in the context of regionalization, like the *PWM* from which they are derived. *L-moments* have been defined for a probability distribution, but in practice must often be estimated from a finite sample. Estimation is based on a sample size n, arranged in ascending order. Let $x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n}$ be the ordered sample. It is convenient to begin with an estimator of the probability weighted moment (β_r).

An unbiased estimator of β_r may be written in general as:

$$\beta_r = n^{-1} \sum_{j=1}^{r} \frac{(j-1)(j-2)...(j-r)}{(n-1)(n-2)...(n-r)} x_j \tag{2.2}$$

where x_j, for $j = 1, ..., n$ is the sample ordered ascending and n is the sample size.

The sample estimate of the *L-moments* can be done through the expressions:

$$\lambda_1 = \beta_0 \tag{2.3}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \tag{2.4}$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{2.5}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - b_0 \tag{2.6}$$

and in general

$$\lambda_r = \sum_{k=0}^{r-1} p^*_{r-1,k} \beta_k \tag{2.7}$$

where

$$p^*_{r,k} = \binom{r}{k} \binom{r+1}{k} \tag{2.8}$$

The merit of the *L-moments* compared to *PWM* is that they offer the advantage of an easier statistical interpretation being able to read as a measure of the shape parameters of the probability distributions of the theoretical and observed the sample. The *L-moments*, unlike the moments "classics", appear to be less subject to the presence of adverse sampling conditions, such as time series reduced or measurement errors (Kottegoda and Rosso, 1997).

In statistical, regional analyses are particularly suitable dimensionless ratios between the *L-moments* always proposed by Hosking and Wallis (1993).

The *L-moments ratio* are defined as dimensionless version of the *L-moments*; these are achieved by dividing the higher order *L-moments* by the scale measure λ_2. The *L-moments* ratios are defined:

$$\tau_r = \frac{\lambda_r}{\lambda_2} \qquad r = 3, 4, ... \tag{2.9}$$

L-moments ratios measure the shape of a distribution independently of its scale of measurement. The *L-CV* value was defined

$$L - CV = \frac{\lambda_2}{\lambda_1} \tag{2.10}$$

This quantity is analogous to the ordinary coefficient of variation, C_v. The *L-moments* are similar but the most convenient moments weighed in probability because more easily interpretable as a measure of shape of the distribution: in fact, λ_1 is the average, or a position measurement, λ_2 is a measure of scale, τ_3 and τ_4 (*L-skewness* and *L-kurtosis*) are, respectively, measures of skewness and kurtosis.

The estimator τ_r and τ are not unbiased, but their biases are very small in moderate or large samples. The quantities λ_1, λ_2, τ_3, and t_4 are statistical parameters useful to summarize the characteristics of a data sample and be used, as well as to judge which distributions are consistent for a given data sample (Hosking and Wallis, 1993), to estimate the parameters when fitting a distribution to a sample, by equating the *L-moments* of the population to those of the sample (Hosking and Wallis, 1993). In fact, many researches (Hosking and Wallis, 1988, Hosking et al., 1985, Lettenmaier and Potter, 1985, Lettenmaier et al., 1987, Potter and Lettenmaier, 1990, Wallis and Wood, 1985) have shown that the procedure of the flow rate based on the probability weighted moments, or *L-moments ratios*, produces estimates of quantile robust and accurate.

Zafirakou-Koulouris et al. (1998) mentioned that like ordinary product moments, L-moments summarize the characteristics or shapes of theoretical probability distributions and observed samples. Both moment types offer measures of distributional location (mean), scale (variance), skewness (shape), and kurtosis (flattening). The authors further mentioned that L-moments offer significant advantages over ordinary product moments, especially for environmental data sets, because of the following:

- *L-moment ratios*: *L-CV*, *L-skewness*, and *L-kurtosis* are almost undistorted, regardless of the probability distribution from which the observations arise (Hosking and Wallis, 1993);
- *L-moment ratios*: *L-CV*, *L-skewness*, and *L-kurtosis* may show fewer distortions than traditional statistical moments, especially for samples extremely asymmetrical;
- *L-moment ratios*: *L-CV*, and *L-skewness*, have no limits and depend on the size of the sample unlike what happens to the classic moments;
- the*L-moments* samples are linear combinations of observations, which are less sensitive to the observations in a sample larger compared to the classic moments products, that depend on the square or cube of the observations;

- *L-moment ratios* diagrams are particularly good at identifying the distributional properties of highly skewed data, whereas ordinary product moment diagrams are almost useless for this task (Peel et al., 2001, Vogel and Fennessey, 1993).

Hosking and Wallis (2005) suggested to use, in place of the ordinary moments, the L-moments because suitable to describe most distributions and more robust in estimating from samples inconsistent data in the presence of outliers resulting less subject to distortion in estimate.

2.2.3 Parameter estimation

Fitting a distribution to datasets provides a compact and smoothed representation of the frequency distribution revealed by the available data, and leads to a systematic procedure for extrapolation to frequencies beyond the range of the dataset. When hydro-climatic variable is well-described by some family of distributions, a task for the hydrologist is to estimate the parameters of the distribution so that required quantiles and expectations can be calculated with the "fitted" model. Appropriate choices for distribution functions can be based on examination of the data using probability plots and moment ratios, the physical origins of the data, and previous experience.

Several general approaches are available for estimating the parameters of a distribution. A simple approach is the *method of moments* which uses the available sample to compute an estimate $\hat{\Theta}$ of Θ so that the theoretical moments of the distribution of X exactly equal the corresponding sample moments. Alternatively, parameters can be estimated using the sample *L-moments*, corresponding to the *method of L-moments*. Still another method that has strong statistical motivation is the *method of maximum likelihood*. *Maximum likelihood estimators* (MLEs) have very good statistical properties in large samples, and experience has shown that they generally do well with records available in hydrology. However, often MLEs cannot be reduced to simple formulas, so estimates must be calculated using numerical methods. MLEs sometimes perform poorly when the distribution of the observations deviates in significant ways from the distribution being fit.

A different philosophy is embodied in Bayesian inference, which combines prior information and regional hydrologic information with the likelihood function for

available data. Advantages of the Bayesian approach are that it allows the explicit modelling of uncertainty in parameters, and provides a theoretically consistent framework for integrating systematic flow records with regional and other hydrologic information.

Occasionally *non-parametric methods* can be employed to estimate frequency relationships. These have the advantage that they do not assume that floods are drawn from a particular family of distributions. Modern non-parametric methods have not yet seen much use in practice and have rarely been used officially. However, curve-fitting procedures which employ plotting positions are non-parametric procedures often used in hydrology.

Issue of concern are the bias, variability, and accuracy of parameter estimation $\hat{\Theta}[X_1, ..., X_n]$, where this notation emphasizes that an estimator $\hat{\Theta}$ is a random variable whose value depends on observed sample values $[X_1, ..., X_n]$. Studies of estimators evaluate an estimator's bias, defined as

$$Bias[\hat{\Theta}] = E[\hat{\Theta}] - \Theta \tag{2.11}$$

and sample-to-sample variability, described by Var$[\hat{\Theta}]$. One wants estimators to be nearly unbiased so that on average they have nearly the correct value, and also to have relatively little variability. One measure of accuracy which combines bias and variability is the mean square error, defined as

$$MSE[\hat{\Theta}] = E[(\hat{\Theta} - \Theta)^2] = (Bias[\hat{\Theta}])^2 - Var[\hat{\Theta}] \tag{2.12}$$

An *unbiased* estimator (Bias $[\hat{\Theta}] = 0$) will have a mean square error equal to its variance. For a given sample size n, estimators with the smallest possible mean square errors are said to be *efficient*.

Bias and mean square error are statistically convenient criteria for evaluating estimators of a distribution's parameters or of quantiles. In particular situations, the hydrologists can also evaluate the expected probability and under or over-design, or use economic loss functions related to operation and design decisions.

2.2.4 Probability distribution of extreme events

Several probability distributions have been applied to describe the distribution of extreme rainfall intensities at a single site (e.g. Benjamin and Cornell, 1970, Chow, 1964a). Common distribution that have been applied to the analysis of

AMS include the Gumbel (Watt et al., 1989), Generalized Extreme Value (GEV) (NERC, 1975), Log-normal (Pilgrim et al., 1998), and Log-Pearson type 3 (Janusz, 1982, Pilgrim et al., 1998) distribution. Among these distributions the GEV and its special form, the Gumbel distribution, have received dominant applications in modelling the AMS of rainfall.

The Gumbel distribution was found, however, to underestimate the extreme precipitation amounts in several cases (Wilks, 1993). Studies using rainfall data from tropical and no-tropical climatic regions (Nguyen et al., 2002, Wilks, 1993, Zalina et al., 2002) suggest also that a three-parameter distribution can be provide sufficient flexibility to rappresent extreme precipitation data. In particular, the GEV distribution has been found to be the most convenient, since it requires a simpler method of parameter estimation and it is more suitable for regional estimation of extreme rainfall at-sites with limited or without data (Nguyen et al., 2002). When the return period associated with frequency-based rainfall estimates greatly exceed the length of record available, discrepancies between commonly used distribution tend to increase.

For POT extremes, following the extreme value theory of Pickands III (1975), the distribution's tail of POT extremes converges asymptotically to a Generalized Pareto Distribution (GPD). Many other distributions have also been successfully employed in hydrologic applications, including the five-parameter Wakeby distribution, the Boughton distribution, and the TCEV distribution (corresponding to a mixture of two Gumbel distributions).

This section provides a short description of some families of distributions commonly used in hydrology. These include the normal family, the GEV family, the *Generalized Pareto Distribution* (GPD), and the *Two Components Extreme Value* (TCEV).

2.2.4.1 Normal family

The normal (N), or Gaussian distribution is certainly the most popular distribution in statistics. It is also the basis of the lognormal (LN) and three-parameter lognormal (LN3) distributions which have seen many applications in hydrology. This section describes the basic properties of the three-parameter lognormal (LN3) distributions.

Three parameters Lognormal distribution. The three-parameter lognormal distribution is a skewed distribution that is useful for modelling continuous positive random variables with support set $[\xi, \infty)$ for some $\xi \geq 0$. Limpert et al. (2001) illustrated how the log-normal distributions are widespread through the science. The probability density function (pdf) of the three-parameter lognormal distribution is:

$$F(x) = \frac{1}{(x-\xi)\kappa\sqrt{2\pi}} exp\{-\frac{[ln(x-\xi)-\alpha]^2}{2\kappa^2} \tag{2.13}$$

where $x > \xi \geq 0$, $-\infty < \alpha < \infty$, $\sigma > 0$, and ξ is the threshold parameter or location parameter that defines the point where the support set of the distribution begins; α is the scale parameter that stretch or shrink the distribution and κ is the shape parameter that affects the shape of the distribution.

If X is a random variable that has a three-parameter log-normal probability distribution, then $Y = ln(X - \xi)$ has a normal distribution with mean α and variance 2κ. The two-parameter lognormal distribution is a special case of the three-parameter lognormal distribution when $\xi = 0$.

The shape parameters κ of the LN3 distribution in terms of L-moments, is a function ot τ_3 alone. No explicit solution is possible, but the following approximation has relative accuracy better than 2.5 x 10^{-6} for $|\tau_3| \leq 0.94$, corresponding to $|\kappa_3| \leq 0.94$:

$$\kappa \approx \tau_3 \frac{E_0 + E_1\tau_3^2 + E_2\tau_3^4 + E_3\tau_3^6}{1 + F_1\tau_3^2 + F_2\tau_3^4 + F_3\tau_3^6} \tag{2.14}$$

The coefficient used in the approximation are given in Table A.1. of the book by Hosking and Wallis (2005). The other parameters are then given by

$$\alpha = \frac{\lambda_2 \kappa e^{-\kappa^2/2}}{1 - 2\Phi(-k/\sqrt{2})} \tag{2.15}$$

$$\xi = \lambda_1 - \frac{\alpha}{\kappa}(1 - e^{\kappa^2/2}) \tag{2.16}$$

where Φ is the cumulative distribution function of the standard Normal distribution defined as:

$$\phi(x) = (2\pi)^{-1/2} exp(\frac{x^2}{2}), \Phi(x) = \int_{-\infty}^{x} \phi(x)dt \tag{2.17}$$

The quantiles of the LN3 distribution have no explicit analytical form.

2.2.4.2 GEV family

Many random variables in hydrology correspond to the maximum of several similar processes, such as the maximum rainfall or flood discharge in a year, or the lowest stream flow. The physical origin of such random variables suggests that their distribution is likely to be one of several *extreme value (EV) distributions* described by Gumbel. The cdf of the largest of n independent variates with common cdf $F(x)$ is simply $F(x)^n$. For large n and many choices for $F(x)$, $F(x)^n$ converges to one of three extreme value distributions, called types I, II, and III. Unfortunately, for many hydrologic variables this convergence is too slow for this argument alone to justify adoption of an extreme value distribution as a model of annual maxima and minima. In this section, the *generalized extreme value distribution* (GEV) and the *two components extreme value* (TCEV) are then introduced. It spans the three types of extreme value distributions for maxima popularized by Gumbel.

The Generalized Extreme Value Distribution. This is the general mathematical form which incorporates Gumbel's type I, II and III extreme value distribution for maxima. The GEV distribution's cdf can be written:

$$F(x) = exp\{-[1 - \frac{\kappa(x-\xi)}{\alpha}]^{1/\kappa}\} \tag{2.18}$$

The Gumbel distribution is obtained when $\kappa = 0$. For $|\kappa| < 0.3$, the general shape of GEV distribution is similar to the Gumbel distribution, though the right-hand tail is thicker for $\kappa < 0$ and thinner for $\kappa > 0$.

Here ξ is a location parameter, α is a scale parameter, and κ is the important shape parameter. For $\kappa > 0$ the distribution has a finite upper bound at $\xi + \alpha/\kappa$ and correspond to the EV type III distribution for maxima that are bounded above; for $\kappa < 0$ the distribution has a thicker right-hand tail and corresponds to the EV type II distribution for maxima from thick-tailed distributions like the generalized Pareto distribution with $\kappa < 0$.

The parameter of the GEV distribution in terms of L-moments are:

$$\kappa = 7.8590c + 2.9554c^2 \tag{2.19}$$

$$\alpha = \frac{\kappa\lambda_2}{\Gamma(1+\kappa)(1-2^{-k})} \tag{2.20}$$

$$\xi = \lambda_1 + \frac{\alpha}{\kappa[\Gamma(1+\kappa)-1]} \tag{2.21}$$

where

$$c = \frac{2\lambda_2}{\lambda_3 + 3\lambda_2} - \frac{ln(2)}{ln(3)} \tag{2.22}$$

The quantiles of the GEV distribution can be calculated from

$$x_p = \xi + \frac{\alpha}{\kappa}\{1 - [-ln(p)]^k\} \tag{2.23}$$

where p is the cumulative probability of interest.

Two Component Extreme Value distribution. The *Two Component Extreme Value* (TCEV) distribution, as proposed by Rossi et al. (1984), was used for the regional frequency analysis of maximum rainfall in different Italian regions (Rossi and Villani, 1994) and United Kingdom flood data (Arnell and Beran, 1987). The authors hypothesized that rainfall could be described by two probability distributions, an ordinary component and an extraordinary component. The distribution is characterized of two independently extreme value distributions *EV1*: the basic, which takes into account the usual values; and the outlying component, which takes into account the extreme values.

The TCEV distribution of annual maxima has the distribution function:

$$F(x) = \exp\left[-\lambda_1 e^{-x/\theta_1} - \lambda_2 e^{-x/\theta_2}\right] \tag{2.24}$$

or, in the standardised regional case

$$F(y) = \exp[-e^{-y} - \Lambda^* e^{-y/\Theta^*}] \tag{2.25}$$

where

$$\Theta^* = \frac{\Theta_2}{\Theta_1} \tag{2.26}$$

$$\Lambda^* = \frac{\Lambda_2}{(\Lambda_1^{1/\Theta^*})} \tag{2.27}$$

$$y = \frac{x - \theta_1^{'} \ln \lambda_1^{'}}{\theta_1^{'}} \tag{2.28}$$

$\theta_1^{'}$ and $\lambda_1^{'}$ are site parameters, while θ_1, θ_1, λ_1 and λ_2 are regional parameters.

This choice is justified when the maximum rainfalls are due to storms with different meteorological characteristics (Rossi and Villani, 1994).

The implementation of the TCEV distribution with the parameters obtained was achieved through the following expression:

$$F(x^{'}) = \exp\left[-\lambda_1(\exp\alpha)^{-x^{'}} - \Lambda^*\lambda_1^{1/\Theta^*}\left(\exp\frac{\alpha}{\Theta^*}\right)^{-x^{'}}\right] \tag{2.29}$$

wherein x' is the dimensionless variable obtained as $\frac{x}{\mu}$. Inside the equation there is the parameter $\alpha = \frac{\mu}{\theta_1}$ (Gabriele and Arnell, 1991) assessed with this expression:

$$\alpha = ln\lambda_1 + 0.5772 - \sum_{j=1} \infty \frac{(-1)^j \Lambda^* \Gamma(\frac{j}{\Theta^*})}{j!} \tag{2.30}$$

The valuation of the parameters Λ_* and Θ_* was achieved matching the expressions of theoretical and sample values of the L-skewness and L-kurtosis, through the following expressions:

$$\frac{\ln(9/8) + 2D_2 - D_1}{\ln 2 + D_1} - M_3 = 0 \tag{2.31}$$

$$\frac{\ln(2^{16}/3^{10}) + 5D_3 - 5D_2 - D_1}{\ln 2 + D_1} - M_4 = 0 \tag{2.32}$$

wherein

$$D_r = T_r - Tr - 1 \tag{2.33}$$

$$T_r = \sum_{j=1}^{\infty}(-1)^{j-1}\Lambda_j^*(r+1)^{j-1/\Theta^*}\Gamma(j/\Theta^*)/j! \tag{2.34}$$

where M_3 and M_4 are the sample values respectively of the L-skewness and L-kurtosis. The parameter λ_1 was evaluated with the similar expression matching with the theoretical value with the sample value L-CV, and solving with the parameters Λ^* and Θ^* before obtained.

$$\frac{\ln 2 + D_1}{\gamma + \ln \lambda_1 + T_0} - M_2 = 0 \tag{2.35}$$

where M_2 is the sample values of the L-CV.

The quantiles of the TCEV distribution have no explicit analytical form.

2.2.4.3 Generalised Pareto distribution

The Generalised Pareto Distribution (GPD) model has been used in hydrology to study floods (Ashkar and Ouarda, 1996) since it allows to model the excess over a threshold. The GPD arises as the limit distribution for the excess over a threshold and for sufficiently high thresholds, it tends asymptotic ally to the General Extreme Values (GEV) distribution.

The GPD has a cumulative distribution function

$$F(x) = \begin{cases} 1 - (1 - \frac{\kappa x}{\alpha})^{1/\kappa}, & \text{se } \kappa \neq 0 \\ 1 - exp(\frac{-x}{\alpha}), & \text{se } \kappa = 0 \end{cases} \tag{2.36}$$

The parameter of the GPD distribution in terms of L-moments are given by:

$$\kappa = \frac{1 - 3\tau_3}{1 + tau_3} \tag{2.37}$$

$$\alpha = (1 + \kappa)(2 + \kappa)\lambda_2 \tag{2.38}$$

$$\xi = \lambda_1 - (2 + \kappa)\lambda_2 \tag{2.39}$$

The quantiles of the GPD distribution can be calculated from

$$x(F) = \begin{cases} \xi + \alpha\{1 - (1 - F)^{\kappa}\}/\kappa, & \text{se } \kappa \neq 0 \\ \xi - \alpha \log(1 - F), & \text{se } \kappa = 0 \end{cases} \tag{2.40}$$

2.2.5 Goodness-of-fit test of distribution

The reliability of precipitation frequency estimates depends on how well the fitted model represents the parent distribution. Several rigorous statistical tests

are available and are useful in hydrology to determine whether it is reasonable to conclude that a given set of observations was drawn from a particular family of distributions (Stedinger and Vogel, 1993).

Several methods are available for testing the goodness fo fit of a distribution to data from single sample. These include quantile-quantile plot, chi-squared, Kolmogorov-Smirnov, and other general goodness-of-fit test and test based on moment or L-moment statistics.

The probability plot correlation test is a more effective test of whether a sample has been drawn from a postulated distribution (Chowdhury et al., 1991, Vogel and Fennessey, 1993). The Kolmogorov–Smirnov test provides bounds within which every observation on a probability plot should lie if the sample is actually drawn from the assumed distribution (Kottegoda and Rosso, 1997). Discussion of the development and interpretation of probability plots is provided by Stedinger and Vogel (1993) and Kottegoda and Rosso (1997).

The L–moments can be used to assess if a proposed distribution is consistent with a dataset (Chowdhury et al., 1991, Hosking, 1990). Using L-moments, it is natural to base test statistics on-site L-moments and the position they occupy on an L-moment ratio diagram Figure 2.1. Cong et al. (1993) have constructed statistics based on the scatter of the point for different sites on a L-moment ratio diagram about the L-skewness-L-kurtosis relation for different three-parameters distributions. Their aim was to choose the distribution that gives the best fit to the data in this sense.

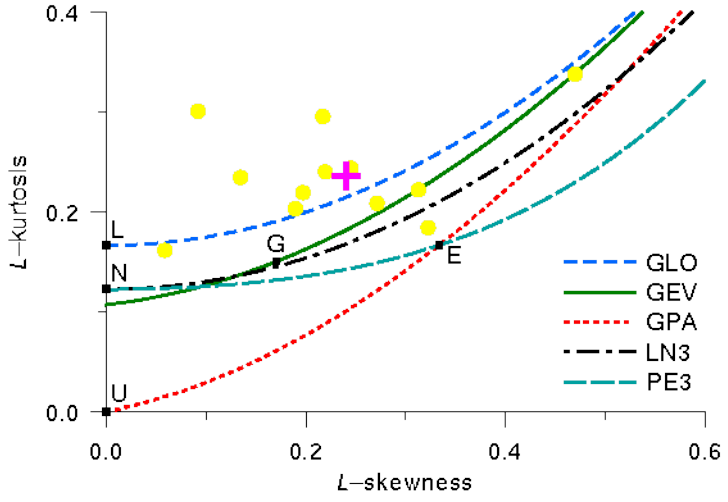


FIGURE 2.1: An L-moment ratio diagram. The yellow dots rappresent the value for each site, while the blue cross is the point that rappresent the mean value of the site. The curves show the candidate distributions.

The use of the best-fitting distribution for each data sample provides frequency estimates that are too sensitive to the sampling variations in the data and the period of record available. Current distribution selection procedures adopted by many countries are based on a combination of regionalization of some parameters and split-sample Monte-Carlo evaluations of different estimation methods to find distribution-estimation procedure combinations that give reliable quantile estimates (Stedinger and Vogel, 1993).

2.3 At-site frequency analysis of rainfall

A frequency analysis can be performed for a site for which sufficient rainfall data are available. Owing to its simpler structure, the AMS based method is more popular. The PDS or POT analysis, however, appears to be preferable for short records, or where return periods shorter than two years are of interest. The choice of an appropriate technique should depend on the purpose of the analysis and characteristics of the available data in terms of both quantity and quality. Improved reliability of the results can be generally achieved with the use of sophisticated and comprehensive analysis methods.

Virtually all hydrological estimates are subject to uncertainty. Therefore, it is often advisable to produce estimates using two or more independent methods and to perform a sensitivity analysis to gain information regarding the potential reliability of results.

Briefly, the steps below should be followed to determine the frequency distribution of annual maximum rainfall, for a fixed duration and for a given site:

1. obtaining a data sample and perform an assessment of data quality based on hydrological and statistical procedures;

2. selecting a candidate distribution model for the data and estimate the model parameters;

3. evaluating the adequacy of the assumed model in terms of its ability to represent the parent distribution from which the data were drawn.

The assessment of data quality is an important step in all statistical computations. The basic assumption in precipitation frequency analysis is that the data are independent and identically distributed as previously said. As mentioned above,

precipitation measurements could be subject to various sources of error, inconsistency and heterogeneity. Detailed examination and verification of the raw data are needed to identify invalid data in the record caused by instrument malfunction and/or human error. There is no general agreement as to which distribution or distributions should be used for precipitation frequency analysis. A practical method for selecting an appropriate distribution is by examining the data with the use of probability plots. Probability plots, which require the use of a plotting position formula, are an effective tool to display graphically the empirical frequency distribution of the data and to assess whether the fitted distribution appears consistent with the data.

There are several plotting position formulae available in practice among (see Nguyen et al. (1989)) which the Hazen, Weibull, and Cunnane formulas are the most popular. The differences between these three formulae are small for observations that are neither the largest nor the smallest; however, they can be significant for the largest three or four values in the data series. An alternative method for making a good choice among different distributions is based on the L–moment diagram (Stedinger and Vogel, 1993, Watt et al., 1989).

Common distributions that have been applied to the analysis of annual maximum series include the Gumbel (EV type I), generalized extreme value (GEV), log-normal (LN3), and log-Pearson type III distributions (LP3). Among these distributions, the GEV and its special form, the EVI, have received dominant applications in modelling the annual maximum rainfall series.

Even a long record may be a relatively small sample of a climatic regime. A better measure of the regime at a station may be given by a smoothed map, which includes information from nearby stations that can influence point data, and thus broadens the sample. The degree of smoothing should be consistent with the spacing of observation stations and the sampling error of the stations. Too little smoothing tends to confound sampling error with spurious regional variation. Rainfall frequency atlases have been produced by interpolation and smoothing of at-site frequency analysis results (Bonnin et al., 2006, Bureau, 1962).

Chapter 3

Rainfall regional frequency analysis

One of the aims of this thesis is to achieve an objective method for the identification of the homogeneous regions and the evaluation of the best frequency distribution. In order to obtain such a regionalization, some classical statistical tools were used in synergy with the regional frequency analysis.

In this chapter, the method of the regional rainfall frequency analysis will be described. For each step of the regional frequency analysis, the statistical tools used in the procedure will be described, analysing in details their characteristics.

3.1 Steps followed in a regional frequency analysis

Frequency analysis often is a problem in hydrology because sufficient information are seldom available at a site to adequately determine the frequency of rare events. When only 30 years of data are available to estimate the event exceeded with a chance of 1 in 100 (the 1 percent exceedance event), extrapolation is required.

The *National Research Council* (NRC) proposed three principles for modelling of the hydrological variables:

1. "substitute space for time";
2. introduction of more "structure" into models;

3. focus on extremes or tails as opposed to, or even to the exclusion of, central characteristics.

One substitutes space for time by using hydrologic information at different locations to compensate for short records at a single site. This is easier to do for rainfall which in regions without appreciable relief should have fairly uniform characteristics over large areas.

Regional frequency analysis (RFA), which involves data from many sites, has been shown to reduce the uncertainties in quantile estimation of extreme rainfall events (Hosking and Wallis, 1988). Regional frequency analysis involves the following basic step, which will be described in details in the next sections, that are:

1. screening of the data;
2. identification of homogeneous;
3. evaluation of homogeneous region obtained
4. choice of a frequency distribution;
5. accuracy of estimated quantile.

The RFA approach based on L-moments has been widely applied in hydrology since it was first introduced by Hosking and Wallis (1993). The method is based on an "index flood" or "index value" approach.

The *index value* method, introduced by Dalrymple (1960), is an efficient method of pooling a summary statistics from the different data sample. The assumptions made in this method are that the sites form a homogeneous region and that the frequency distribution of the different sites is identical except a site-specific scaling factor, called the *"scale factor"* (or *"index value"*).

The most practical application of the RFA, for rainfall, is aimed to the creation of the rainfall *Depth-Duration-Frequency* (DDF) curves. The DDF curves, which describe rainfall depth as a function of duration for given return periods, are essential for the design and verification of hydraulic infrastructures. For a homogeneous region, the DDF curve are given by:

$$H_{d,T} = \mu_R(d)h(T) \tag{3.1}$$

in which $H_{d,T}$ is the rainfall depth of fixed duration d and return period T, $h(T)$ represents the regional growth curve (function of return period T) as a dimensionless quantile common to every site or region, and $\mu_R(d)$ is the mean for fixed duration, that usually is modelled with a power-law of duration.

It is possible to assume that the scale factor is represented by central tendency the frequency of distribution and that can be estimated through the sample mean or median of the at-site value; consequently the mean of rescaled data is equal to 1 for each site, and so the regional average of these means is 1.

The index value procedure uses summary statistics of data at each site and combines them by averaging to form the regional estimates. If the summary statistics are the L-moments ratios of the at-site data, this method is called the *regional L-moments algorithm*.

The index-value method on a theoretical basis is not superior to the maximum-likelihood approach but is an efficient and intuitively method to summarize the information originated from different sites. The procedure requires a computational burden simpler than those of maximum-likelihood estimation, because the method does not involve the entire regional dataset (Hosking and Wallis, 2005).

The *L-moments* are effective to fit, to the data from the sites in a homogeneous region, a single frequency distribution (the regional frequency distribution) that describes the distribution of the observation at each site after scaling by at-site scaling factor. This distribution is scaled appropriately at each site in order to estimates quantiles of the at-site frequency distribution. In the hypothesis that analysis regards a group of n stations, the ratios between the *L-moments* calculated from the data of the generic station i are denoted as $t^{(i)}$ or *L-CV*, $t_3^{(i)}$ or *L-skewness*, $t_4^{(i)}$ or *L-kurtosis*. The regional average *L-moment ratios* are obtained through the weighted proportionally to the site's records length:

$$t^R = \frac{\sum_{i=1}^{N} n_i t^{(i)}}{\sum_{i=1}^{N} n_i} \tag{3.2}$$

$$t_r^R = \frac{\sum_{i=1}^{N} n_i t_r^{(i)}}{\sum_{i=1}^{N} n_i} \qquad r = 3, 4, ... \tag{3.3}$$

Homogeneity of the region is then defined in terms of constant second and higher order moments.

3.2 Screening of the data

As already discussed in the Section 2.2.1, the first essential stage of any statistical data analysis is to check that data are appropriate for the analysis. For frequency analysis, the data collected at a site must be a true representation of the quantity being measured and must all be drawn from the same frequency distribution. An initial screening of the data should aim to verify that these requirements are satisfied.

For the extreme rainfall, two kinds of error are particularly important and plausible.

First, data values may be incorrect. Incorrect recording or transcription of the data values is easily done and casts doubts on any subsequent frequency analysis of the data.

Second, the circumstances under which the data were collected may have changed over time. The measuring device may have been moved to different location or trends over time may have arisen from changes in the environmental of the measuring device. This means that the frequency distribution from which the data were sampled is not constant over time, and frequency analysis of the data will not be a valid basis for estimating the probability distribution of future measurement at the site.

Test for outlier and trend are well established in the statistical literature (e.g., Barnett and Lewis, 1994, Kendall, 1975). In the context of using L-moments, Hosking and Wallis (2005) have found that useful information can be obtained by comparing the sample L-moment rations for different sites. Incorrect data values, outliers, trends, and shift in the mean of a sample can all be reflected in the L-moments of the samples. A convenient amalgamation of the L-moment ratios into a single statistics, a measure of the discordancy between the L-moments ratios of site and average L-moment ratios group of similar sites, is described in the next section.

3.2.1 Discordance measure

The *discordance measure* (D) can be used for different aspects of the RFA. Usually it is applied before the identification of homogeneous regions to check whether errors are present in the data. At this stage, the data could be analysed carefully to find possible errors in the recording of transcription of data or for sources of unreality in the data. The discordance measure can also be used when

the homogeneous regions have been, at least tentatively, identified and if any site was discordant with the region as a whole, some sites could be removed from the regions, or a completely different assignment of sites to regions has to be tried.

The discordance measure in terms of the L-moments indicates the stations that are grossly discordant with the group as a whole. Regard the sample L-moments ratio of site as a point in three dimensional space. A group of sites will yield a cloud of such points. Flag as discordant any points that are far from the centre of the cloud. "Far" is interpreted in such a way as to allow for correlation between the sample L-moment ratio.

If $\mathbf{u}_i$ is the vectors containing the L-moments $[t^{(i)}, t_3^{(i)}, t_4^{(i)}]$ for *i-th* site (i = *1,...,* N), then the discordance measure for the same *i-th* site is defined as

$$D_i = \frac{1}{3} N(\mathbf{u}_i - \bar{\mathbf{u}})\mathbf{A}^{-1}(\mathbf{u}_i - \bar{\mathbf{u}}) \tag{3.4}$$

where

$$\bar{\mathbf{u}} = N^{-1} \sum_{i=1}^{N} \mathbf{u}_i \tag{3.5}$$

$$\mathbf{A} = (\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}}) \tag{3.6}$$

A site i is declared discordant if D_i is large. The definition of "large" depends on the number of sites group.

Wilks (1963) proposed an outlier measure that is equivalent to the largest of D_i. It is not easy to choose a value of D_i that can be used as a criterion for deciding whether at site is discordant. Hosking and Wallis (1993) initially suggested the criterion $D_i \geq 3$, but this is not satisfactory for small regions. It was proposed an algebraic bound to satisfy

$$D_{cr} = \frac{(N-1)}{3} \tag{3.7}$$

where N is the number of site in region.

Thus value of D_i larger than 3 can occur only in regions having 11 or more sites. To some extent, the criterion for the discordance should be increasing function of number of sites in the region. This is because large regions are more likely to contain sites with large value of D_i. However, Hosking and Wallis recommend that any site with $D_i > 3$ be regarded as discordant, as such sites have L-moments ratios that are markedly different from average for the other sites in the region (Hosking and Wallis, 1993). The sites that have regarded discordant will be moved in another region or deleted if they will result yet discordant.

3.3 Identification of homogeneous regions

Among all the stages in a RFA involving many sites, the identification of homogeneous region is usually the most difficult and requires the greatest amount of subjective judgement. For these reasons two statistical tools, the *principal components analysis* (PCA) and the k-means, are combined among them in order to obtain a solution as objective as possible. The PCA and the clustering techniques are used in many works regarding the hydrological variables.

The principal component analysis is here used to reduce the dimensionality and retained the most of the variation present in all of the original variables.

Instead, the k-means, that is a standard method of statistical multivariate analysis for dividing a dataset into groups, has been successfully used to form homogeneous regions for the RFA.

In the following sections the tools used in the identification of homogeneous regions will be described, providing, at the same time, a description of the them.

3.3.1 Principal Component Analysis

Possibly the most widely used multivariate statistical technique in the atmospheric sciences is *principal component analysis*, often denoted as PCA. The technique became popular for analysis of atmospheric data following the paper by Lorenz (1956), who called the technique *empirical orthogonal function* (EOF) analysis. Both names are commonly used, and refer to the same set of procedures. Sometimes the method is incorrectly referred to as factor analysis, which is a related but distinct multivariate statistical method.

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming it to a new set of variables, the *principal components* (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. PCA reduces a data set containing a large number of variables to a data set containing fewer (hopefully many fewer) new variables (Jolliffe, 2002).

These new variables are linear combinations of the original ones, and these linear combinations are chosen to represent the maximum possible fraction of the variability contained in the original data. That is, given multiple observations of a $(K\times 1)$ data vector $\boldsymbol{x}$, PCA finds $(M\times 1)$ vectors $\boldsymbol{u}$ whose elements are linear

combinations of the elements of the $\boldsymbol{x}$, which contain most of the information in the original collection of $\boldsymbol{x}$. PCA is most effective when this data compression can be achieved with $M \ll K$. This situation occurs when there are substantial correlations among the variables within $\boldsymbol{x}$, in which case $\boldsymbol{x}$ contains redundant information. The elements of these new vectors $\boldsymbol{u}$ are called the principal components (PCs).

Data for atmospheric and other geophysical fields generally exhibit many large correlations among the variables x_k, and a PCA results in a much more compact representation of their variations. Beyond mere data compression, however, a PCA can be a very useful tool for exploring large multivariate data sets, including those consisting of geophysical fields.

The PCA has the potential for yielding substantial insights into both the spatial and temporal variations exhibited by the field or fields being analysed, and new interpretations of the original data $\boldsymbol{x}$ can be suggested by the nature of the linear combinations that are most effective in compressing the data. Usually it is convenient to calculate the PCs as linear combinations of the anomalies $\boldsymbol{x}'=\boldsymbol{x}-\bar{\boldsymbol{x}}$. The first PC, u_1, is that linear combination of $\boldsymbol{x}'$ having the largest variance. The subsequent principal components u_m, $m = 2, 3, ..., M$, are the linear combinations having the largest possible variances, subject to the condition that they are uncorrelated with the principal components having lower indices. The result is that all the PCs are mutually uncorrelated.

The new variables or PCs that is, the elements u_m of $\boldsymbol{u}$ that will account successively for the maximum amount of the joint variability of $\boldsymbol{x}'$ (and therefore also of $\boldsymbol{x}$) are uniquely defined by the eigenvectors of the covariance matrix [S] of $\boldsymbol{x}$. In particular, the m^{th} principal component, u_m is obtained as the projection of the data vector $\boldsymbol{x}'$ onto the m^{th} eigenvector, $\boldsymbol{e}_m$,

$$u_m = \mathbf{e}_m^T \mathbf{x}' = \sum_{k=1}^{K} e_{km} x_k', \quad m = 1, ..., M \tag{3.8}$$

Notice that each of the M eigenvectors contains one element pertaining to each of the K variables, x_k. Similarly, each realization of the m^{th} principal component in Equation (3.8) is computed from a particular set of observations of the K variables x_k. That is, each of the M principal components is a sort of weighted average of the x_k values. Although the weights (the e_{km}) do not sum to 1, their squares do because of the scaling convention $\|e_m\| = 1$.

If the data sample consists of n observations (and therefore of n data vectors $\boldsymbol{x}$, or n rows in the data matrix [X]), there will be n values for each of the principal

components, or new variables, u_m. Each of these constitutes a single-number index of the resemblance between the eigenvector e_m and the corresponding individual data vector $\boldsymbol{x}$.

Geometrically, the first eigenvector, e_1, points in the direction (in the K-dimensional space of $\boldsymbol{x}$') in which the data vectors jointly exhibit the most variability. This first eigenvector is the one associated with the largest eigenvalue, λ_1. The second eigenvector e_2, associated with the second-largest eigenvalue λ_2, is constrained to be perpendicular to e_1, but subject to this constraint it will align in the direction in which the $\boldsymbol{x}$' vectors exhibit their next strongest variations. Subsequent eigenvectors e_m, m =3, 4, ..., M, are similarly numbered according to decreasing magnitudes of their associated eigenvalues, and in turn will be perpendicular to all the previous eigenvectors. Subject to this orthogonality constraint these eigenvectors will continue to locate directions in which the original data jointly exhibit maximum variability. Put another way, the eigenvectors define a new coordinate system in which to view the data. In particular, the orthogonal matrix [E] whose columns are the eigenvectors defines the rigid rotation

$$\mathbf{u} = [E]^T \mathbf{x}' \tag{3.9}$$

which is the simultaneous matrix-notation representation of $M = K$ linear combinations of the form of Equation (3.8) (i.e., here the matrix [E] is square, with K eigenvector columns).

This new coordinate system is oriented such that each consecutively numbered axis is aligned along the direction of the maximum joint variability of the data, consistent with that axis being orthogonal to the preceding ones. These axes will turn out to be different for different data sets, because they are extracted from the sample covariance matrix $[S_x]$ particular to a given data set. That is, they are orthogonal functions, but are defined empirically according to the particular data set at hand. This observation is the basis for the eigenvectors being known in this context as empirical orthogonal functions (EOFs).

It is a remarkable property of the principal components that they are uncorrelated. That is, the correlation matrix for the new variables u_m is simply [I]. This property implies that the covariances between pairs of the u_m are all zero, so that the corresponding covariance matrix is diagonal. Each PC represents a share of the total variation in $\boldsymbol{x}$ that is proportional to its eigenvalue,

$$R_m^2 = \frac{\lambda_m}{\sum_{k=1}^{K} \lambda_k} * 100\% = \frac{\lambda_m}{\sum_{k=1}^{K} s_{k,k}} * 100\% \tag{3.10}$$

Here R^2 is used in the same sense that is familiar from linear regression. The total variation exhibited by the original data is completely represented in (or accounted for by) the full set of K u_m, in the sense that the sum of the variances of the centred data $\boldsymbol{x}'$ $\sum_k s_{k,k}$, is equal to the sum of the variances $\sum_m \lambda_m$ of the principal component variables $\boldsymbol{x}$. Equation (3.10) expresses the transformation of a ($K \ll 1$) data vector $\boldsymbol{x}'$ to a vector $\boldsymbol{u}$ of PCs.

In principal component analysis, the number of components extracted is equal to the number of variables, necessitating to decide how many of these components are truly meaningful and worthy of being retained for rotation and interpretation. In general, the first few components will account for meaningful amounts of variance, and that the later components will tend to account for only trivial variance.

Mathematically, there are as many eigenvectors of [S] or [R] as there are elements of the data vector $\boldsymbol{x}$'. However, it is typical of atmospheric data that substantial covariances (or correlations) exist among the original K variables, and as a result there are few or no off-diagonal elements of [S] (or [R]) that are near zero. This situation implies that there is redundant information in $\boldsymbol{x}$, and that the first few eigenvectors of its dispersion matrix will locate directions in which the joint variability of the data is greater than the variability of any single element, x'_k, of $\boldsymbol{x}$. Similarly, the last few eigenvectors will point to directions in the K-dimensional space of $\boldsymbol{x}$' where the data jointly exhibit very little variation. To the extent that there is redundancy in the original data $\boldsymbol{x}$', it is possible to capture most of their variance by considering only the most important directions of their joint variations. That is, most of the information content of the data may be represented using some smaller number $M<K$ of the principal components u_m.

In effect, the original data set containing the K variables x_k is approximated by the smaller set of new variables u_m. If $M \ll K$, retaining only the first M of the principal components results in a much smaller data set. This data compression capability of PCA is often a primary motivation for its use. The truncated representation of the original data can be expressed mathematically by a truncated version of the analysis formula, Equation (3.9), in which the dimension of the truncated $\boldsymbol{u}$ is ($M \ll 1$), and [E] is the (nonsquare, $K \ll M$) matrix whose columns

consist only of the first K eigenvectors (i.e., those associated with the largest M eigenvalues) of [S].

There is no clear criterion that can be used to choose the number of principal components that are best retained in a given circumstance. The choice of the truncation level can be aided by one or more of the many available principal component selection rules, but it is ultimately a subjective choice that will depend in part on the data at hand and the purposes of the PCA.

Subjective criteria. Some approaches to truncating principal components are subjective, or nearly so. Perhaps the most basic criterion is to retain enough of the principal components to represent a sufficient fraction of the variances of the original $\boldsymbol{x}$. That is, enough principal components are retained for the total amount of variability represented to be larger than some critical value. Of course the difficulty comes in determining how large the fraction of the critic value must be in order to be considered sufficient. Ultimately this will be a subjective choice, informed by the analysts knowledge of the data at hand and the uses to which they will be put.

Another essentially subjective approach to principal component truncation is based on the shape of the graph of the eigenvalues λ_m in decreasing order as a function of their index $m=1,\ldots,K$, known as the eigenvalue spectrum. Plotting the eigenvalue spectrum with a linear vertical scale produces what is known as the scree plot or screen test (Figure 3.1) (Cattell, 1966).

In this method the eigenvalues associated with each component are plotted and will be looking for a "break" between the components with relatively large eigenvalues and those with small eigenvalues. The component numbers are listed on the horizontal axis, while eigenvalues are listed on the vertical axis. The components that appear before the break are assumed to be meaningful and are retained for rotation; those appearing after the break are assumed to be unimportant and are not retained. Sometimes a scree plot will display several large breaks. When this is the case, the last big break before the eigenvalues begin to level off should look for. Only the components that appear before this last large break should be retained.

The scree test can be expected to provide reasonably accurate results, provided the sample is large (over 200) and most of the variable communality are large (Stevens, 2012). However, this criterion has its own weaknesses as well, most notably the ambiguity that is often displayed by scree plots under typical research

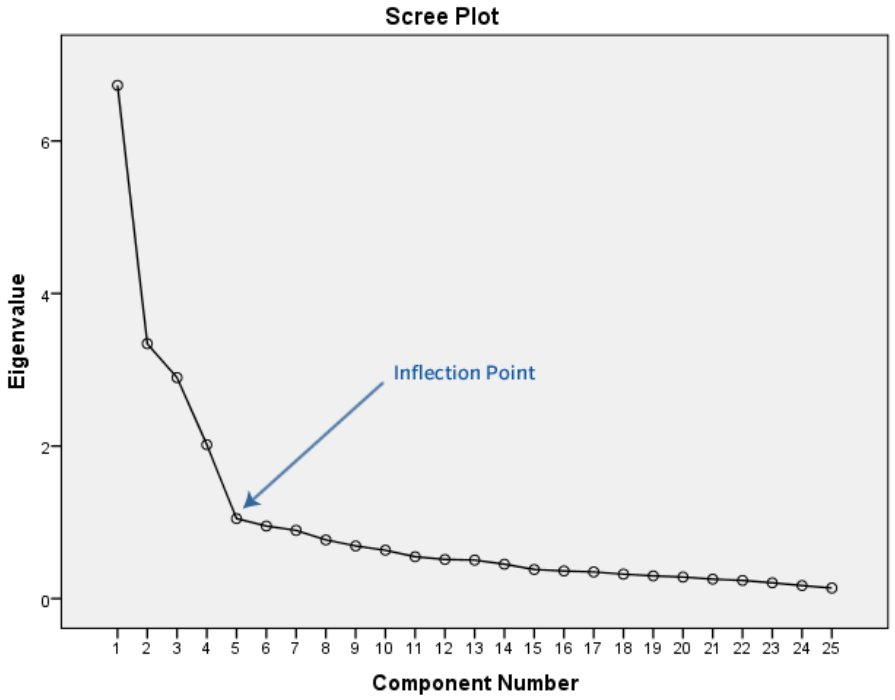


FIGURE 3.1: Example scree plot

conditions, very often, it is difficult to determine exactly where in the scree plot a break exists, or even if a break exists at all.

Cumulative percentage of total variation. Perhaps the most obvious criterion for choosing m is to select a (cumulative) percentage of total variation which one desires that the selected PCs contribute, say 80% or 90%. The required number of PCs is then the smallest value of m for which this chosen percentage is exceeded. It remains to define what is meant by "percentage of variation accounted for by the first m PCs", but this poses no real problem. Principal components are successively chosen to have the largest possible variance, and the variance of the k^{th} PC is λ_m. Furthermore, $\sum_{m=1}^{M} \lambda_m = \sum_{k=1}^{K} s_{kk}$, that is the sum of the variances of the PCs is equal to the sum of the variances of the elements of $\boldsymbol{x}$. The obvious definition of "percentage of variation accounted for by the first m PCs" is therefore

$$t_m = 100 \cdot \frac{\sum_{m=1}^{M} \lambda_m}{\sum_{k=1}^{K} s_{k,k}} = 100 \cdot \frac{\sum_{m=1}^{M} \lambda_m}{\sum_{m=1}^{K} \lambda_m} \tag{3.11}$$

which reduces to

$$t_m = \frac{100}{K} \sum_{m=1}^{M} \lambda_m \tag{3.12}$$

in the case of a correlation matrix.

Choosing a cut-off t^* somewhere between 70% and 90% and retaining m PCs, where m is the smallest integer for which $t_m > t^*$, provides a rule which, in practice, preserves in the first m PCs most of the information in $\boldsymbol{x}$. The best value for t^* will generally become smaller as K increases, or as n, the number of observations, increases. Although a sensible cutoff is very often in the range 70% to 90%, it can sometimes be higher or lower depending on the practical details of a particular data set. For example, a value greater than 90% will be appropriate when one or two PCs represent very dominant and rather obvious sources of variation. Here the less obvious structures beyond these could be of interest, and to find them a cutoff higher than 90% may be necessary. Conversely, when K is very large choosing m corresponding to 70% may give an impractically large value of m for further analyses. In such cases the threshold should be set somewhat lower.

Size of the last retained eigenvalue. Another class of principal-component selection rules involves focusing on how small an "important" eigenvalue can be. This set of selection rules can be summarized by the criterion

$$Retain\ \lambda_m\ if \lambda_m > \frac{T}{K} \sum_{k=1}^{K} s_{k,k} \tag{3.13}$$

where $s_{k,k}$ is the sample variance of the k^{th} element of $\boldsymbol{x}$, and T is a threshold parameter.

A simple application of this idea, known as *Kaiser's rule*, involves comparing each eigenvalue (and therefore the variance described by its principal component) to the amount of the joint variance reflected in the average eigenvalue. Principal components whose eigenvalues are above this threshold are retained. That is, *Kaiser's rule* uses Equation (3.13) with the threshold parameter *T=1*. Jolliffe (2002, 1972) has argued that *Kaiser's rule* is too strict (i.e., typically seems to discard too many principal components). He suggests that the alternative *T=0.7* often will provide a roughly correct threshold, which allows for the effects of sampling variations. A third alternative in this class of truncation rules is to use the broken stick model, so called because it is based on the expected length of the m^{th}

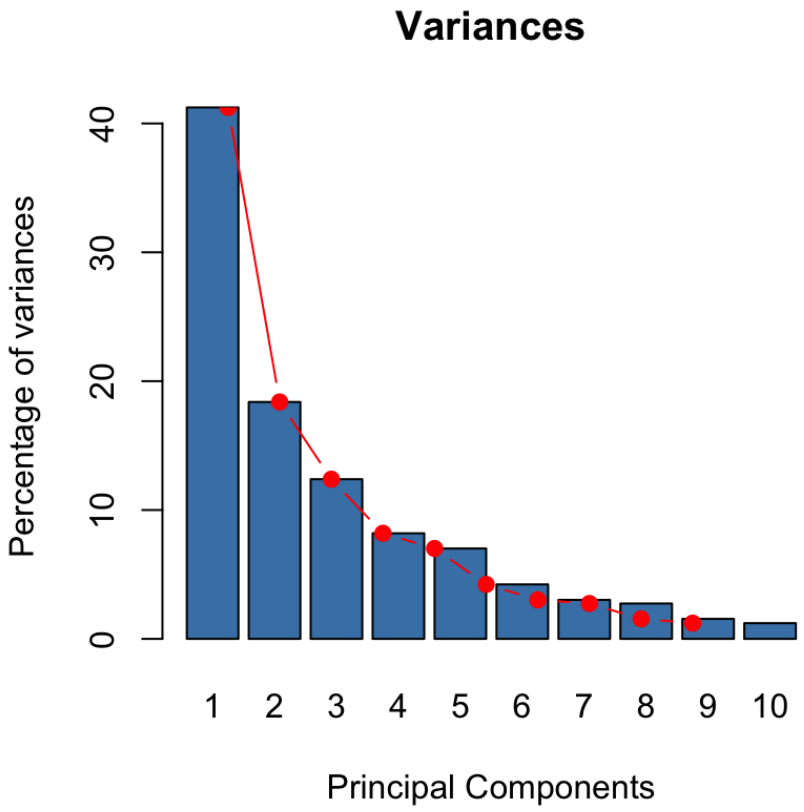


FIGURE 3.2: Example cumulative percentage of total variation

longest piece of a randomly broken unit line segment. According to this criterion, the threshold parameter in Equation (3.13) is taken to be

$$T(m) = \frac{1}{K} \sum_{j=m}^{K} \frac{1}{j} \tag{3.14}$$

This rule yields a different threshold for each candidate truncation level that is, $T = T_m$, so that the truncation is made at the smallest m for which Equation (3.13) is not satisfied, according to the threshold in Equation (3.14).

This criterion has a number of positive features that have contributed to its popularity. Perhaps the most important reason for its widespread use is its simplicity, it do not need any subjective decisions, but merely retain components with eigenvalues greater than one. On the positive side, it has been shown that this criterion very often results in retaining the correct number of components, particularly when a small to moderate number of variables are being analysed and the variance in each of the original variables is explained by the extracted factors are high. Stevens (2012) reviews studies that have investigated the accuracy of

the eigenvalue-one criterion, and recommends its use when less than 30 variables are being analysed and communality are greater than 0.70, or when the analysis is based on over 250 observations and the mean communality is greater than or equal to 0.60.

There are a number of problems associated with the eigenvalue-one criterion, however. It can lead to retaining the wrong number of components under circumstances that are often encountered in research (e.g., when many variables are analysed, when communality are small). Also, the mindless application of this criterion can lead to retaining a certain number of components when the actual difference in the eigenvalues of successive components is only trivial. There is a possibility than more a component has a value near to 1.00, (e.g. component 2 displays an eigenvalue of 1.001 and component 3 displays an eigenvalue of 0.99), then component with value greater than 1.00 will be retained but component with value lower 1.00 will not retained; this may mislead you into believing that the third component was meaningless when, in fact, it accounted for almost exactly the same amount of variance as the second component. In short, the eigenvalue-one criterion can be helpful when used judiciously, but the thoughtless application of this approach can lead to serious errors of interpretation.

3.3.2 Cluster analysis

Cluster analysis divides data into groups (clusters) that are meaningful, useful, or both. If meaningful groups are the goal, then the clusters should capture the natural structure of the data. In some cases, however, cluster analysis is only a useful starting point for other purposes, such as data summarization.

Whether for understanding or utility, cluster analysis has long played an important role in a wide variety of fields: psychology and other social sciences, biology, statistics, pattern recognition, information retrieval, machine learning, and data mining. There have been many applications of cluster analysis to practical problems. In this section an overview on the partition methods will be provide with particular regard to the k-means method.

Cluster analysis groups data objects based only on information found in the data that describes the objects and their relationships. The goal is that the objects within a group be similar (or related) to one another and different from (or unrelated to) the objects in other groups. The greater similarity (or homogeneity) within a group and the greater the difference between groups, the better or more

distinct the clustering. In many applications, the notion of a cluster is not well defined.

Cluster analysis is related to other techniques that are used to divide data objects into groups. For instance, clustering can be regarded as a form of classification in that it creates a labelling of objects with class (cluster) labels. However, it derives these labels only from the data. In contrast with the *"supervised classification"*; i.e., new, unlabelled objects are assigned to a class label using a model developed from objects with known class labels. For this reason, cluster analysis is sometimes referred to as *"unsupervised classification"*. When the term classification is used without any qualification within data mining, it typically refers to *"supervised classification"*. Also, while the terms segmentation and partitioning are sometimes used as synonyms for clustering, these terms are frequently used for approaches outside the traditional bounds of cluster analysis.

For example, the term partitioning is often used in connection with techniques that divide graphs into sub-graphs and that are not strongly connected to clustering. Segmentation often refers to the division of data into groups using simple techniques; e.g., an image can be split into segments based only on pixel intensity and colour. Nonetheless, some work in graph partitioning and in image and market segmentation is related to cluster analysis.

3.3.2.1 Clustering methods

There are many clustering algorithms in the literature. It is difficult to provide a crisp categorization of clustering methods because these categories may overlap so that a method may have features from several categories. The main reason for having many clustering methods is the fact that the notion of "cluster" is not precisely defined Estivill-Castro and Yang (2000). Consequently many clustering methods have been developed, each of which uses a different induction principle.

Fraley and Raftery (1998) suggest dividing the clustering methods into two main groups: hierarchical and partitioning methods. Han and Carlin (2001) suggest categorizing the methods into additional three main categories: density-based methods, model-based clustering and grid-based methods. An alternative categorization based on the induction principle of the various clustering methods is presented in (Estivill-Castro and Yang, 2000).

These methods are briefly summarized in Table 3.1. Some clustering algorithms integrate the ideas of several clustering methods, so that it is sometimes difficult to classify a given algorithm as uniquely belonging to only one clustering

method category. Furthermore, some applications may have clustering criteria that require the integration of several clustering techniques.

TABLE 3.1: Overview of clustering methods discussed in this section. Note that some algorithms may be combined in various methods.

Method	**General characteristic**
Partitioning methods	– Find mutually exclusive clusters of spherical shape – Distance-based – May use mean or medoid (etc.) to represent cluster centre – Effective for small- to medium-size data sets
Hierarchical methods	– Clustering is a hierarchical decomposition – Cannot correct erroneous merges or splits – May incorporate other techniques like microclustering or consider object "linkages"
Density-based methods	– Can find arbitrarily shaped clusters – Clusters are dense regions of objects in space that are separated by low-density regions – Cluster density: Each point must have a minimum number of points within its "neighbourhood" – May filter out outliers
Grid-based methods	– Use a multi-resolution grid data structure – Fast processing time (typically independent of the number of data objects, yet dependent on grid size)

Partitioning methods. Given a set of n objects, a partitioning method constructs k partitions of the data, where each partition represents a cluster and $k \leq n$. That is, it divides the data into k groups such that each group must contain at least one object. In other words, partitioning methods conduct one-level partitioning on data sets. The basic partitioning methods typically adopt exclusive cluster separation. That is, each object must belong to exactly one group. This requirement may be relaxed, for example, in fuzzy partitioning techniques. Most partitioning methods are distance-based. Given k, the number of partitions to construct, a partitioning method creates an initial partitioning. It then uses an iterative relocation technique that attempts to improve the partitioning by moving objects from one group to another. The general criterion of a good partitioning

is that objects in the same cluster are "close" or related to each other, whereas objects in different clusters are "far apart" or very different. There are various kinds of other criteria for judging the quality of partitions.

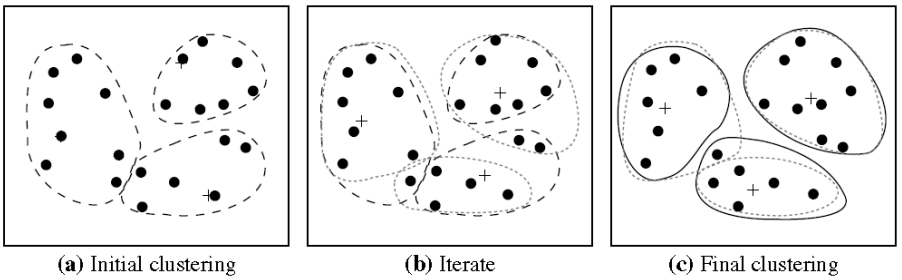


FIGURE 3.3: Clustering of a set of objects using the k-means method; for (b) update cluster centers and reassign objects accordingly (the mean of each cluster is marked by a C. Source (Han et al., 2011)

Traditional partitioning methods can be extended for subspace clustering, rather than searching the full data space. This is useful when there are many attributes and the data are sparse. Achieving global optimality in partitioning-based clustering is often computationally prohibitive, potentially requiring an exhaustive enumeration of all the possible partitions. Instead, most applications adopt popular heuristic methods, such as greedy approaches like the *k-means* and the *k-medoids* algorithms, which progressively improve the clustering quality and approach a local optimum. These heuristic clustering methods work well for finding spherical-shaped clusters in small- to medium-size databases. To find clusters with complex shapes and for very large data sets, partitioning-based methods need to be extended (Han et al., 2011).

Hierarchical methods. A hierarchical method creates a hierarchical decomposition of the given set of data objects. A hierarchical method can be classified as being either agglomerative or divisive, based on how the hierarchical decomposition is formed. The *agglomerative approach*, also called the *bottom-up approach*, starts with each object forming a separate group. It successively merges the objects or groups close to one another, until all the groups are merged into one (the topmost level of the hierarchy), or a termination condition holds.

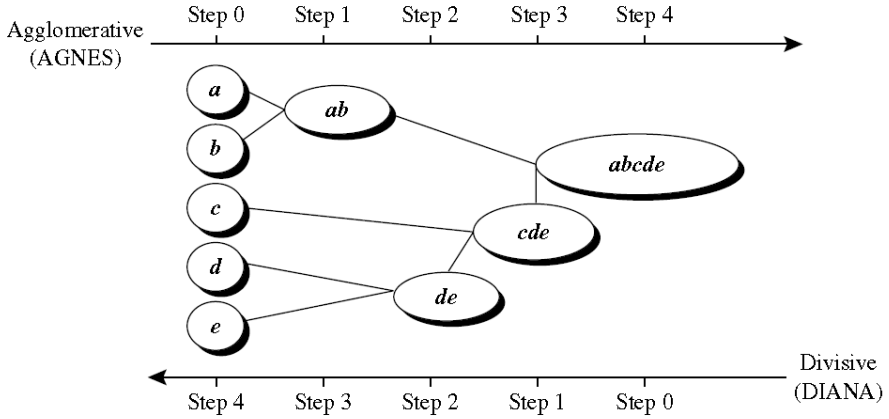


FIGURE 3.4: Agglomerative and divisive hierarchical clustering on data objects *[a,b,c,d,e]*. Source (Han et al., 2011)

.

The *divisive approach*, also called the *top-down approach*, starts with all the objects in the same cluster. In each successive iteration, a cluster is split into smaller clusters, until eventually each object is in one cluster, or a termination condition holds. Hierarchical clustering methods can be *distance-based* or *density-and continuity based*. Various extensions of hierarchical methods consider clustering in subspaces as well.

Hierarchical methods suffer from the fact that once a step (merge or split) is done, it can never be undone. This rigidity is useful in that it leads to smaller computation costs by not having to worry about a combinatorial number of different choices. Such techniques cannot correct erroneous decisions; however, methods for improving the quality of hierarchical clustering have been proposed (Han et al., 2011).

Density-based methods. Most partitioning methods cluster objects based on the distance between objects. Such methods can find only spherical-shaped clusters and encounter difficulty in discovering clusters of arbitrary shapes. Other clustering methods have been developed based on the notion of density. Their general idea is to continue growing a given cluster as long as the density (number of objects or data points) in the "neighbourhood" exceeds some threshold. For example, for each data point within a given cluster, the neighbourhood of a given radius has to contain at least a minimum number of points. Such a method can be used to filter out noise or outliers and discover clusters of arbitrary shape. Density-based methods can divide a set of objects into multiple exclusive clusters, or a hierarchy of clusters. Typically, density-based methods consider exclusive

clusters only, and do not consider fuzzy clusters. Moreover, density-based methods can be extended from full space to subspace clustering (Han et al., 2011).

Grid-based methods. Grid-based methods quantize the object space into a finite number of cells that form a grid structure. All the clustering operations are performed on the grid structure (i.e., on the quantized space). The main advantage of this approach is its fast processing time, which is typically independent of the number of data objects and dependent only on the number of cells in each dimension in the quantized space. Using grids is often an efficient approach to many spatial data mining problems, including clustering. Therefore, grid-based methods can be integrated with other clustering methods such as density-based methods and hierarchical methods. Some clustering algorithms integrate the ideas of several clustering methods, so that it is sometimes difficult to classify a given algorithm as uniquely belonging to only one clustering method category. Furthermore, some applications may have clustering criteria that require the integration of several clustering techniques (Han et al., 2011).

3.3.2.2 Dissimilarity of numeric data

In this section, the distance measures, that are commonly used for computing the dissimilarity of objects analysed by numeric attributes, will be described. These measures include the *Euclidean* and *Manhattan distance*.

In some cases, the data are normalized before applying distance calculations. This involves transforming the data to fall within a smaller or common range, such as *[-1, 1]* or *[0.0, 1.0]*.

In general, expressing an attribute in smaller units will lead to a larger range for that attribute, and thus tend to give such attributes greater effect or "weight". Normalizing the data attempts to give all attributes an equal weight. It may or may not be useful in a particular application.

The most popular distance measure is *Euclidean distance* (i.e., straight line or "as the crow flies"). If $i = (x_{i1}, x_{i2}, ..., x_{ip})$ and $j = (x_{j1}, x_{j2}, ..., x_{jp})$ are two objects described by p numeric attributes. The Euclidean distance between objects i and j is defined as

$$d(i,j) = \sqrt{(x_{i1} + x_{j1})^2 + (x_{i2} + x_{j2})^2 + ... + (x_{p1} + x_{p1})^2} \tag{3.15}$$

Another well-known measure is the *Manhattan (or city block) distance*, named so because it is the distance in blocks between any two points in a city (such as 2 blocks down and 3 blocks over for a total of 5 blocks). It is defined as

$$d(i,j) = |x_{i1} + x_{j1}| + |x_{i2} + x_{j2}| + ... + |x_{p1} + x_{p1}| \tag{3.16}$$

Both the Euclidean and the Manhattan distance satisfy the following mathematical properties:

- *non-negativity*. Distance is a non-negative number, $d(i,j) \geq 0$;
- *identity of indiscernible*. The distance of an object to itself is 0, $d(i,j) = 0$;
- *symmetry*. Distance is a symmetric function, $d(i,j) = d(j,i)$;
- *triangle inequality*. Going directly from object i to object j in space is no more than making a detour over any other object k, $d(i,j) \leq d(i,k)+d(k,j)$.

3.3.3 The *k*-means algorithm

The simplest and most fundamental version of cluster analysis is partitioning, which organizes the objects of a set into several exclusive groups or clusters. To keep the problem specification concise, the number of clusters can be assumed given as background knowledge. This parameter is the starting point for partitioning methods. In this section, the k-means that is the most well-known and commonly used partitioning methods will be described.

Formally, given a data set, D, of n objects, and k, the number of clusters to form, a partitioning algorithm organizes the objects into k partitions ($k \leq n$), where each partition represents a cluster. The clusters are formed to optimize an objective partitioning criterion, such as a dissimilarity function based on distance, so that the objects within a cluster are "similar" to one another and "dissimilar" to objects in other clusters in terms of the data set attributes.

Suppose a dataset, D, contains n objects in Euclidean space. Partitioning methods distribute the objects in D into k clusters, C_1, ..., C_k. An objective function is used to assess the partitioning quality so that objects within a cluster are similar to one another but dissimilar to objects in other clusters. This is the objective function aims for high intra-cluster similarity and low inter-cluster similarity.

A centroid-based partitioning technique uses the centroid of a cluster, C_j, to represent that cluster. Conceptually, the centroid of a cluster is its centred point. The centroid can be defined in various ways such as by the mean or medoid of the objects (or points) assigned to the cluster. The medoid can be defined as the object of a cluster whose average dissimilarity to all the objects in the cluster is minimal. The dissimilarity between an object $p \in C_i$ and c_i, the representative of the cluster, is measured by *dist(p,c$_i$)*, where *dist()* is the Euclidean distance between two points p and c_i.

The k-means algorithm defines the centroid of a cluster as the mean value of the points within the cluster.

First, it randomly selects k of the objects in D, each of which initially represents a cluster mean or centroid. For each of the remaining objects, an object is assigned to the cluster to which it is the most similar, based on the Euclidean distance between the object and the cluster mean. The k-means algorithm then iteratively improves the within-cluster variation. For each cluster, it computes the new mean using the objects assigned to the cluster in the previous iteration. All the objects are then reassigned using the updated means as the new cluster centres. The iterations continue until the assignment is stable, that is, the clusters formed in the current round are the same as those formed in the previous round. The k-means procedure is summarized in Figure 3.5.

The k-means method is not guaranteed to converge to the global optimum and often terminates at a local optimum. The results may depend on the initial random selection of cluster centres. In order to obtain good results in practice, it is common to run the k-means algorithm multiple times with different initial cluster centres.

There are several variants of the k-means method. These can differ in the selection of the initial k-means, the calculation of dissimilarity, and the strategies

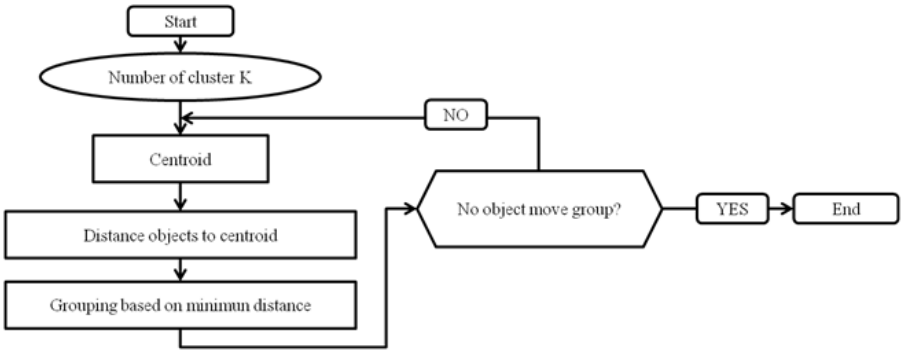


FIGURE 3.5: The k-means partitioning algorithm

for calculating cluster means. The k-means method can be applied only when the mean of a set of objects is defined. This may not be the case in some applications such as when data with nominal attributes are involved.

The necessity for users to specify k, the number of clusters, in advance can be seen as a disadvantage. There have been studies on how to overcome this difficulty; however, the k value can be evaluated by providing an approximate range of k values, and subsequently using an analytical technique, the best value will be determined k by comparing the clustering results obtained for the different k values (Han et al., 2011).

Two of possible evaluation measures are described, with different theoretical backgrounds and demands. These methods, of course, not represent an exhaustive list of clustering evaluations, but tries to give an impression of the possible methods which are concerned with clustering evaluation.

Sum-of-Squared-Error Criterion. Summing over the squared distances between the clustering objects and their cluster representatives (i.e. the respective cluster centroids) is a standard cost function. The evaluation defines a measure for the homogeneity of the clustering results with respect to the object description data. The *sum-of-squared-error* (E) originally refers to Euclidean distance, but is applicable to further distance measures. The definition is given in Equation (3.17), where the quality of cluster C_i can be measured by the within-cluster variation, which is the sum of squared error between all objects in C_i and the centroid c_i , defined as

$$E = \sum_{i=1}^{k} \sum_{p \in C_d} dist(p, c_i)^2 \tag{3.17}$$

where E is the sum of the squared error for all objects in the data set; p is the point in space representing a given object; and c_i is the centroid of cluster C_i (both p and c_i are multidimensional).

In other words, for each object in each cluster, the distance from the object to its cluster centred is squared, and the distances are summed. This objective function tries to make the resulting k clusters as compact and as separate as possible.

Silhouette Value. Rousseeuw (1987) presented the *silhouette plot* as a means for clustering evaluation. With this method, each cluster is represented by a silhouette

displaying which objects lie well within the cluster and which objects are marginal to the cluster.

The silhouettes are useful when the proximities are on a ratio scale (as in the case of Euclidean distances) and when one is seeking compact and clearly separated clusters. Indeed, the definition makes use of average proximities as in the case of group average linkage, which is known to work best in a situation with roughly spherical clusters.

In order to construct silhouettes, only two things need: the partition obtained (by the application of some clustering technique) and the collection of all proximities between objects. For each object i, a certain value $s(i)$ will be introduced, and then these numbers will be combined into a plot.

Let us define first the numbers $s(i)$ in the case of dissimilarities. Take any object i in the data set, and denote by A the cluster to which it has been assigned. When cluster A contains other objects apart from i, then it is possible compute $a(i)$, the average dissimilarity of i to all other objects of A. In Figure 3.6, this is the average length of all lines within A.

Let us consider now any cluster C which is different from A, and compute $d(i,C)$, the average dissimilarity of i to all objects of C. In Figure 3.6, this is the

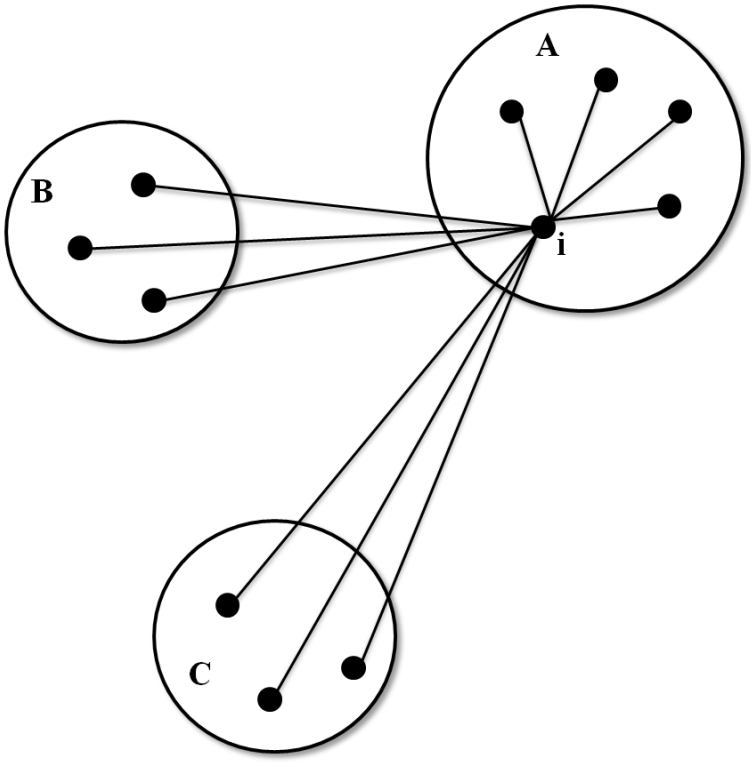


FIGURE 3.6: An illustration of the elements involved in the computation of $s(i)$, where the object i belongs to cluster A.

average length of all lines going from i to C. After computing *d(i,C)* for all clusters $C \neq A$, the smallest value of those numbers will be denote by *b(i) = minimum d(i, C)*.

The cluster B for which this minimum is attained (that is, *d(i,B) = b(i)*) it is called the neighbour of object i. This is like the second-best choice for object i: if it could not be accommodated into cluster A, the cluster B would be the best choice. In Figure 3.6, cluster B indeed appears to be *"closest"* (on the average) to object i, when A itself is discarded. Therefore, it is very useful to know the neighbour of each object in the dataset. Note that the construction of *b(i)* depends on the availability of other clusters apart from A, so it is necessary to assume that the number of clusters k is more than one.

The number *s(i)* is obtained by combining *a(i)* and *b(i)* as follows:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases} \tag{3.18}$$

It is even possible to write this in one formula:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \tag{3.19}$$

When cluster A contains only a single object it is unclear how *a(i)* should be defined, and then the value will be set *s(i)* equal to zero. This choice is of course arbitrary, but a value of zero appears to be most neutral. Indeed, from the above definition it is evident that $-1 \leq s(i) \leq 1$ for each object i.

When *s(i)* is at its largest (that is, *s(i)* close to 1) this implies that the *"within"* dissimilarity *a(i)* is much smaller than the smallest *"between"* dissimilarity *b(i)*. Therefore, it can say that i is *"well-clustered"*, as there appears to be little doubt that i has been assigned to a very appropriate cluster: the second-best choice *(B)* is not nearly as close as the actual choice *(A)*.

A different situation occurs when *s(i)* is about zero. Then *a(i)* and *b(i)* are approximately equal, and hence it is not clear at all whether i should have been assigned to either A or B. Object i lies equally far away from both, so it can be considered as an *"intermediate case"*.

The worst situation takes place when *s(i)* is close to -1. Then *a(i)* is much larger than *b(i)*, so i lies on the average much closer to B than to A. Therefore it

would have seemed much more natural to assign object i to cluster B, so it can almost conclude that this object has been *"misclassified"*.

In order to obtain an overview, an example of the silhouettes of the different clusters are shown in Figure 3.7. In this way the entire clustering can be displayed by means of a single plot, which enables us to distinguish "clear-cut" clusters from "weak" ones.

Figure 3.7 shows the silhouettes for the clustering into k = 3 clusters of the twelve stations. The grey bar shows the silhouette value for each station while on the right are shown the number of the station into cluster and the mean value of the silhouette, respectively. In the bottom the total average value of the silhouette is shown. In this case, all the station have a positive value obtaining a good performance.

To conclude, *s(i)* measures how well object i matches the clustering at hand (i.e., how well it has been classified). In the special case where there are only two clusters ($k = 2$), it possible note that shifting object i from one cluster to the other will convert *s(i)* to *-s(i)*.

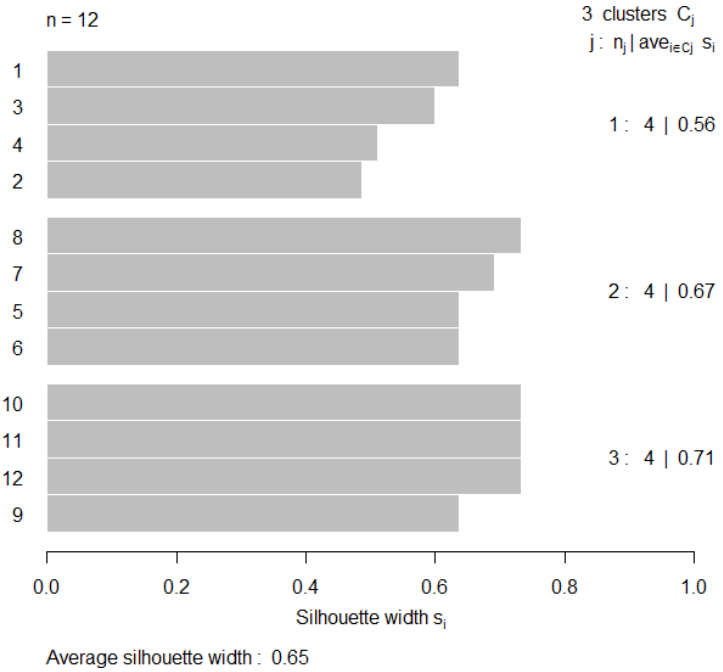


FIGURE 3.7: An illustration of the elements involved in the computation of *s(i)*, where the object i belongs to cluster A.

3.4 Evaluation of homogeneous regions

The assessment of regional homogeneity is a critical point in regional frequency analysis. The procedures for the analysis of a single set of data are well-established, but often observations of the same variable at different measuring sites are available, and more accurate conclusions can be reached by analysing many data samples together. This constitutes the basis for regional frequency analysis (Hosking and Wallis, 2005). Critical points of the regional approach to frequency analysis are in the choice of the method to group the data samples together, and in the assessment of the plausibility of the obtained groupings. This involves testing whether the proposed regions may be considered homogeneous or not. The hypothesis of homogeneity implies that frequency distributions for different sites are the same, except for a site-specific scale factor.

Commonly used homogeneity tests, based on L-moments ratios, are considered here in a comparison with two rank tests that do not rely on particular assumptions regarding the parent distribution. The performance of these tests is assessed in a series of Monte Carlo simulation experiments. In particular, the power and type I error of each test are determined for different scale and shape parameters of the regional parent distributions. The tests are also evaluated by varying the number of sites belonging to the region, the series length, the type of the parent distributions and the degree of heterogeneity.

In the next section, three different test will be described. The first two test, the regional homogeneity statistic, HW, (Hosking and Wallis, 1993) and the Anderson-Darling rank test, AD, (Stedinger et al. 1993), evaluate the extreme rainfall regions obtained assessing the homogeneity. The last test, the discordance measure for each station, D (Hosking and Wallis, 2005), examines possibly anomalous behaviour of individual stations.

3.4.1 Hosking and Wallis homogeneity test

Hosking and Wallis (1993, 2005) proposed a statistical test for assessing the homogeneity of a group of basins at three different levels by focusing on three measures of dispersion for different orders of the sample L-moment ratios.

In a homogeneous region all site have the same population L-moments ratios. Thus, a natural question to ask is if the between-site dispersion of the sample L-moments ratios, for the group of sites under consideration, is larger than would be expected of a homogeneous region. The latter is estimated through repeated

simulations of homogeneous regions with samples drawn from a four parameter kappa distribution. The four-parameter distribution for the simulation is capable of representing many of the distribution occurring in the environmental sciences.

The weighted standard deviation of the at-site sample L-CVs, is calculated as:

$$V = \left\{ \frac{\sum_{i=1}^{N} n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^{N} n_i} \right\}^{1/2} \tag{3.20}$$

After that the parameters of a four-parameters kappa distribution have been fitted to the regional averaged L-moment ratios t^R, t_3^R andt_4^R, a large number N_{sim} of realizations of sets of k samples are generated. The *i-th* site sample in each set has a kappa distribution as its parent and record length equal to n_i. For each simulated homogeneous set, the statistic V is calculated, obtaining N_{sim} values. From the simulations, the mean μ_V and standard deviation σ_V of the N_{sim} were determined. An heterogeneity measure, which is called here HW, is finally found as

$$HW = \frac{V - \mu_V}{\sigma_V} \tag{3.21}$$

HW can be approximated by a normal distributed variable with zero mean and unit variance: following Hosking and Wallis (2005), the region under analysis can therefore be regarded as "acceptably homogeneous" if $HW<1$, "possibly heterogeneous" if $1\leq HW<2$, and "definitely heterogeneous" if $HW\geq 2$.

Hosking and Wallis (2005) suggested that these limits should be treated as useful guidelines. Even if the HW statistic is constructed like a significance test, significance levels obtained from such a test would, in fact, be accurate only under special assumptions: to have independent data both serially and between sites, and the true regional distribution being four-parameters kappa.

3.4.2 The bootstrap Anderson-Darling test

A test that does not make any assumption on the parent distribution is the *Anderson-Darling* (AD) rank test (Scholz and Stephens, 1987). The AD test is the generalization of the classical Anderson-Darling goodness of fit test (e.g., D'Agostino, 1986), and it is used to test the hypothesis that k independent samples belong to the same population without specifying their common distribution function.

The test is based on the comparison between local and regional empirical distribution functions. The empirical distribution function, or sample distribution function, is defined by

$$F(x_j) = \frac{j}{\eta} \tag{3.22}$$

where η is the size of the sample and $x_{(j)}$ are the ordered statistics, i.e. the observations arranged in ascending order $x_{(j)} \leq x < x_{(j+1)}$.

If the empirical distribution function of the *i-th* sample (local) is denoted by $\hat{F}_i(x)$, and that of the pooled sample of all $N = n_1 + ... + n_k$ observations (regional) by $H_N(x)$, the k-sample Anderson-Darling test statistic is then defined as

$$\theta_{AD} = \sum_{i=1}^{k} n_i \int_{allx} \frac{[\hat{F}_i(x) - H_N(x)]^2}{H_N(x)[1 - H_N(x)]} \, dH_N(x) \tag{3.23}$$

If the pooled ordered sample is $Z_1 < ... < Z_N$, the computational formula to evaluate Equation (3.23) is:

$$\theta_{AD} = \frac{1}{N} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{N-1} \frac{(NM_{ij} - jn_i)^2}{j(N-j)} \tag{3.24}$$

where M_{ij} is the number of observations in the *i-th* sample that are not greater than Z_j. The homogeneity test can be carried out by comparing the obtained θ_{AD} value to the tabulated percentage points reported by Scholz and Stephens (1987) for different significance levels (Viglione et al., 2007).

The statistic θ_{AD} depends on the sample values only through their ranks. This guarantees that the test statistic remains unchanged when the samples undergo monotonic transformations, an important stability property not possessed by *HW* heterogeneity measures. However, problems arise in applying this test in a common index value procedure.

In fact, the index value procedure corresponds to dividing each site sample by a different value, thus modifying the ranks in the pooled sample. In particular, this has the effect of making the local empirical distribution functions much more similar to the other, providing an impression of homogeneity even when the samples are highly heterogeneous. The effect is analogous to that encountered when applying goodness-of-fit tests to distributions whose parameters are estimated from the same sample used for the test (e.g., D'Agostino, 1986, Laio, 2004). In both cases, the percentage points for the test should be opportunely redetermined. This can be

done with a non-parametric bootstrap approach (Viglione et al., 2007), presenting the following steps:

1. building up the pooled sample S of the observed non-dimensional data, the sample with replacement from S and generate k artificial local samples, of size $n_1, ..., n_k$;

2. dividing each sample for its index value, and calculate $\theta_{AD}^{(1)}$.

3. repeating the procedure for N_{sim} times and obtain a sample of $\theta_{AD}^{(1)}, j = 1, ..., N_{sim}$ values, whose empirical distribution function can be used as an approximation of $G_{H_0}(\theta_{AD})$, the distribution of θ_{AD} under the null hypothesis of homogeneity.

The acceptance limits for the test, corresponding to any significance level α, are then easily determined as the quantiles of $G_{H_0}(\theta_{AD})$ corresponding to a probability $(1 - \alpha)$. The test obtained with the above procedure will be called the bootstrap Anderson-Darling test, hereafter referred to as AD.

3.5 Choice and accuracy of the distribution

Analytical goodness-to-fit criteria are useful for gaining an appreciation for weather the lack of fit is likely to be due to sample-to-sample variability, or whether a particular departure of the data from a model is statistically significant. In the most case several distribution will provide statistically acceptable fits to available data so that goodness-of-fit tests are unable to identify the "true" or "best" distribution to use. Such tests are valuable when they can demonstrate that some distribution appear inconsistent with the data.

Unfortunately, the true distribution is probably too complex to be of practical use. Still, L-moments skewness-kurtosis and CV-skewness diagram are good for investigating what simple family of distribution are consistent with available data sets for a region (Stedinger and Vogel, 1993, Watt et al., 1989). Standard goodness-of-fit statistics, such as probability plot correlation have also been used to see how well a member of each family of distribution can fit a sample. Unfortunately, the goodness-of-fit statistics likely do not identify the actual family of the sample, rather, the test will choose the most flexible families that generally achieves the best fitting of the data. Regional L-moments diagram focuses on the character of

sample statistics which describe the "parent" distribution for available samples, rather than goodness-of-fit.

In the next section, the approach proposed by Hosking and Wallis (2005), that works directly with the regional average L-moment statistics, will be illustrated.

Calculation of goodness-of-fit measures involves the estimate of sampling variability of L-moment ratios in a region whose record lengths and average L-moment ratios match those of the data. If the region is homogeneous and data at different sites are statistically independent, then if one of the distributions is the true distribution for the region, its goodness-of-fit measure should have approximately a standard normal distribution.

If one supposes that the region has N sites with record length n_i, indicating the sample of L-moments ratios for each site and regional with t^i, t_3^i and t_4^i and t^R, t_3^R and t_4^R respectively, for each m-th simulated region, the regional average L-skewness $t_3^{[m]}$ and L-kurtosis $t_4^{[m]}$ can be calculated for the selected distribution and the kappa distribution, the latter used as reference distribution, by the regional average L-moments. The bias and standard deviation of the $t_4^{[m]}$ are:

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} \lim (t_3^{[m]} - t_4^R) \tag{3.25}$$

$$\sigma_4 = \left\{ (N_{sim} - 1)^{-1} \left[\sum_{m=1}^{N_{sim}} \lim (t_3^{[m]} - t_4^R)^2 - N_{sim} B_4^2 \right] \right\}^{1/2} \tag{3.26}$$

For each distribution, the goodness-of-fit measure is:

$$Z^{DIST} = (t_4^{DIST} - t_4^R + B_4)/\sigma_4) \tag{3.27}$$

The value of Z^{DIST} would be sufficiently close to zero to provide the best fit, a reasonable criterion is given to $|Z^{DIST}| \leq 1.64$, to considerer the distribution accepted.

3.6 Accuracy of estimated quantile

In the traditional statistics, the assessment of accuracy of the estimated quantile is achieved by the construction of confidence intervals for estimated parameters and quantiles. In the regional frequency analysis through the regional L-moments algorithm it is possible to constructs confidence intervals for estimation in homogeneous regions, at least as a large-sample approximation when sample L-moments

may be taken to be normally distributed. The confidence intervals are of limited utility in practice, because rarely the model used is the "correct" and in the case of the regional L-moments algorithm all the assumptions of the index-flood must be satisfy.

A reliable assessment of the accuracy should should considerer the possibility of heterogeneity in the region, misspecification of frequency distribution, and statistical dependence between observation at different sites.

The procedure, described by Hosking and Wallis (2005), for simulation of the regional L-moments algorithm (omitting the intersite dependence) that involves Monte Carlo simulation. The steps which characterises the algorithm, needed to evaluate the accuracy of estimated quantiles, can be summarised as following:

1. specifying N and for each of N sites its record length n_i and L-moments of its frequency distribution.

2. Calculating the parameters of the at-site frequency distribution given their l-moment ratios.

3. For each of M repetitions of the simulation procedure, carting out the following steps:

   (a) generating sample data for each site. This simply requires the generation of a random sample of size n_i from the frequency distribution for site i, $i = 1,\ldots, n$.

   (b) Applying the regional L-moments algorithm to the sample of regional data. This involved the following steps:

      i. Calculating at-site L-moment ratios and regional average L-moment ratios;

      ii. Fitting the chosen distribution;

      iii. Calculating estimates of the regional growth curve and at-site quantiles.

   (c) Calculating the relative error of the estimated regional growth curve and at-site quantiles, and accumulate the sum needed to calculate overall accuracy measures.

4. Calculating overall measures of the accuracy of the estimated quantiles and regional growth curve.

The Monte Carlo simulation is a reasonable approach to estimate the accuracy of the estimates quantiles. The simulation should considerer the particular characteristic of the data from which the estimates are calculated.

The L-moments ratio at the individual sites should be chosen to yield a region whose heterogeneity is consistent with the regional measures calculated from the data. Some preliminary simulation should may be needed to establish the variation at-site of the L-moments ratio to yield the observed value of the heterogeneity measure.

An important point is that between-site variation in population of L-moments ratio for the simulated region should be less of the sample L-moments ratio of the actual data, because sampling variability causes sample L-moment ratios to be much more widely scattered than the corresponding population L-moment ratio. The use in the simulation the observed sample L-moments ratios as the population L-moments ratio of the simulated region, because this would yield a simulated region that has much more heterogeneity that actual data (Hosking and Wallis, 2005).

In the simulation procedure, quantile estimates are calculates for various non-exceedance probabilities. At the m^{th} repetition, let the site-i quantile estimate for non exceedance probability F be $\hat{h}_i^m(F)$. The relative error of this estimate is $(\hat{h}_i^m(F) - \hat{h}_i(F))/\hat{h}_i(F)$. This quantity can be squared and averaged over all M repetitions to approximate the relative RMSE of the estimators. The relative BIAS and relative RMSE are approximated, for large M, by

$$BIAS_i(F) = M^{-1} \sum_{m=1}^{M} (\hat{h}_i^m(F) - \hat{h}_i(F))/\hat{h}_i(F) \tag{3.28}$$

and

$$RMSE_i(F) = \left[M^{-1} \sum_{m=1}^{M} \left\{ (\hat{h}_i^m(F) - \hat{h}_i(F))/\hat{h}_i(F) \right\}^2 \right]^{1/2} \tag{3.29}$$

A summary of the accuracy of estimated quantiles overall of the sites in the region is given by the regional average relative RMSE of the estimates quantile,

$$BIAS^R(F) = N^{-1} \sum_{i=1}^{N} B_i(F) \tag{3.30}$$

and the regional average relative RMSE of the estimated dimensionless quantiles,

$$RMSE^R(F) = N^{-1} \sum_{i=1}^{N} R_i(F) \tag{3.31}$$

Chapter 4

Area of study

In this chapter, the area of study will be described together with its characteristics meteo-climatic and geographical. In the next sections, some more specific information about Sicily and the dataset used in the application of different method shown in the next chapters will be described.

4.1 Climatology of Sicily

Sicily is the largest island in the Mediterranean Sea situated in the South of Italy with an area about 25.000 km^2 and its location is from 36° to 38° north and 12° to 15° east Figure 4.1.

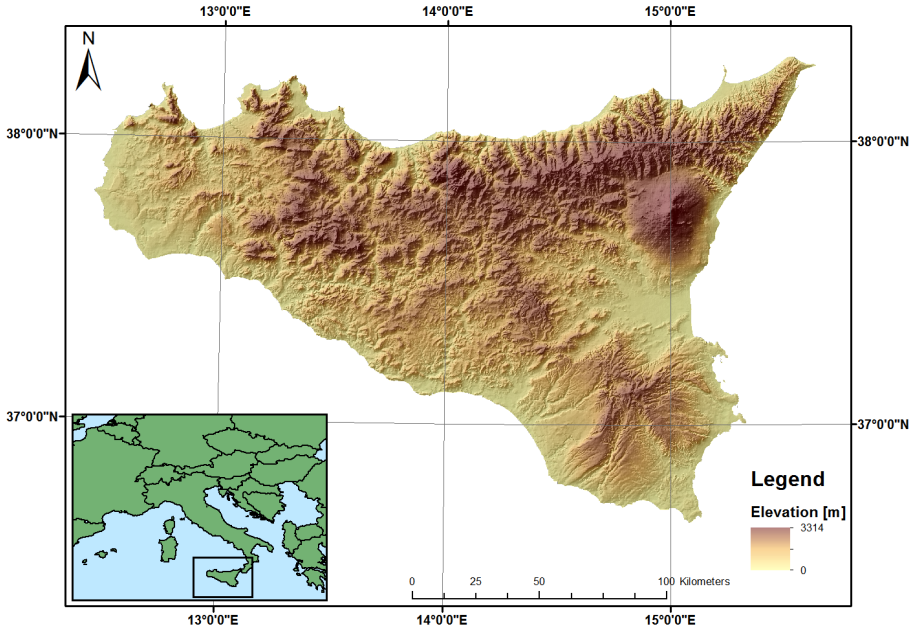


FIGURE 4.1: Study area location with digital elevation model

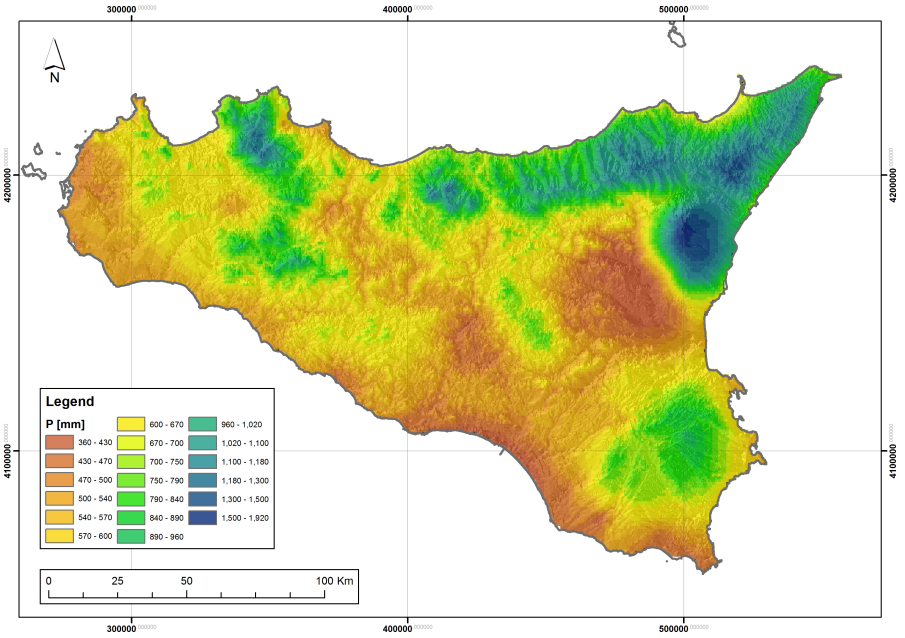


FIGURE 4.2: Map of mean annual precipitation (Di Piazza et al., 2011)

The mean annual precipitation over Sicily (Figure 4.2) is about 715 mm and it is possible to identify a summer and winter season, with rainfall concentrated in the winter period while July-August months are usually rainless.

Considerable spatial variability of precipitation is observed, ranging from an average of 1500 mm in the Northern-Eastern part to an average of 400 mm in the South-Eastern part (Di Piazza et al., 2011).

The mean annual temperature (Figure 4.3) ranges from 11° C to 20° C (Drago et al., 2002). The lowest temperature are registered in the mountainous while the highest temperature along the coasts. In particular, the south coast results the driest and the warmest (Piazza et al., 2015).

Even though the morphology of Sicily is very complex, it is useful in terms of meteorological features, to consider three subareas related to the three main sides of the island: the northern side, the south-western side and the eastern side (Drago et al., 2002).

The Northern Side includes the whole northern part of Sicily where is the mountain range considered as part of the Apennines Mountains. The rainfall regime is characterized by a rainy season (autumn-winter) and a dry season (spring-summer). The precipitation is more frequent in winter. Indeed the number of rain days (total daily amount > 1 mm) is greater than 70 and the precipitation is typically orographically-induced, with precipitation often of longer duration and rarely of extreme intensity.

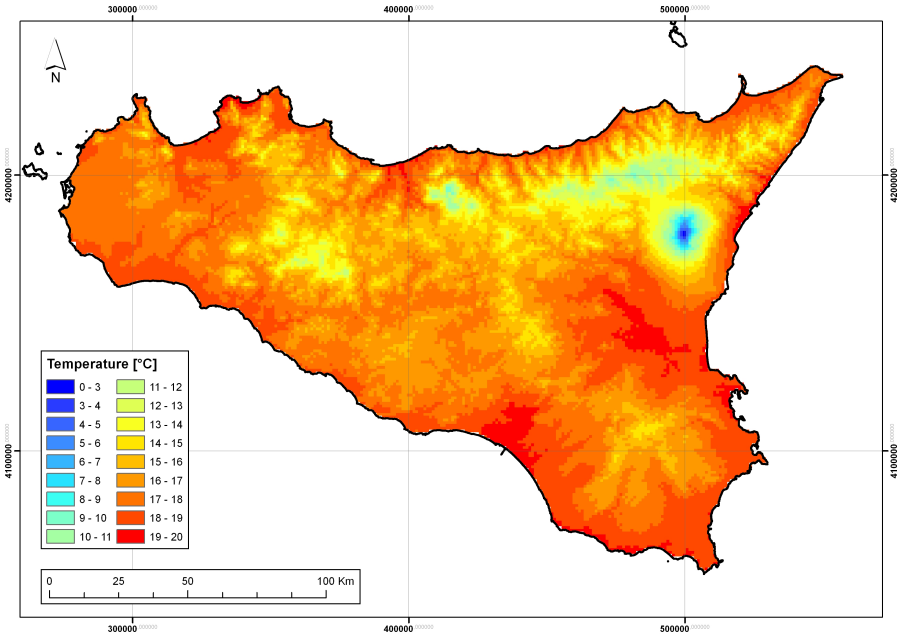


FIGURE 4.3: Map of mean annual temperature (Piazza et al., 2015)

The eastern side includes whole east part of the island and is characterized by the presence of the Etna volcano (3329 m high). In this zone the rainfall is greatest in the winter season. Precipitation is infrequent compared to the northern zone except for the zone near to the Etna volcano and the number of rain days does not exceed 60. The regime is typically oriental with greater events coming from the east. Precipitations are often brief and sometimes extreme. This behaviour is due to low pressure areas that bring rain from Africa produced because of the elevated thermal contrast.

The Southern Side includes the whole of the coastline of the Mediterranean Sea, the Sicily canal and the central zone. As well as for the entire region, the rainy season is winter. The number of rain-days is lower than the northern zone (<60 days). The rainfall regime is conditioned from airflows mainly coming from the south-west.

4.2 Rainfall dataset

In order to develop the RFA, the extreme rainfall data published by the *Servizio Osservatorio delle Acque* (the former *Hydrologic service*) have been used. The rainfall data used in this research are relative to the annual maxima rainfalls with duration (AMR_d) equal to 1, 3, 6, 12 and 24 h and the annual maxima daily (AMD). There is a total of 314 stations for AMR_d data used for the analysis

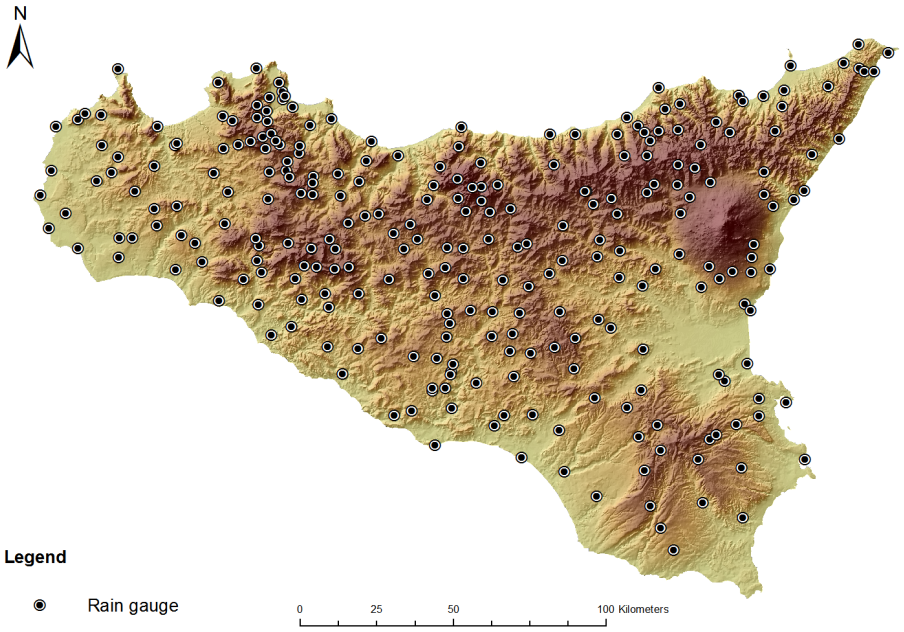


FIGURE 4.4: Rain-gauges distribution

spanning the period 1928-2010 (Figure 4.4), while for AMD there are 382 stations spanning the period 1928-2009.

In Figure 4.5 it is possible to see as some stations have many temporal gaps due to the presence of some non-operative period. The density plot shows that there is a time window with a greater presence of stations operating (years between the years 1972 and 2003). This period, characterized by the maximum number of stations operating, was selected for the next analysis.

This choice was made to achieve the maximum overlap of the rainfall events among different stations. The use of different ranges of operation could provide analysis of rainfall events registered just from few stations. The dataset comprises 124 stations, each with a minimum record length of 20 years (Figure 4.6).

The Table 4.1 shows the stations, with coordinates, elevation and activity period, used in the analysis that will be described in the next chapter.

TABLE 4.1: Hydrologic service network rain-gauges list with location and working period

ID	LOCATION	North ED50 [m]	East ED50 [m]	Elevation [m]	S-year	E-year
20	S. SABA	4237314	543535	22	1973	2002
40	MONFORTE S.GIORGIO	4223520	533660	321	1972	2002
60	MILAZZO	4230437	521320	0	1959	2010
70	CASTROREALE	4216925	518572	395	1930	2007
80	BARCELLONA	4222320	519185	49	1974	2003
130	MONTALBANO ELICONA	4208511	501538	876	1932	2011
170	FICARRA	4217863	485273	437	1973	2004
180	CAPO D'ORLANDO	4223229	478228	100	1959	2008
220	TORTORICI	4209363	484671	471	1930	2011

Table 4.1: continued on next page

Table 4.1: continued from previous page

ID	LOCATION	North ED50 [m]	East ED50 [m]	Elevation [m]	S-year	E-year
280	ALCARA LI FUSI	4208618	473580	396	1971	2003
310	S. FRATELLO	4207875	464802	655	1965	2011
323	S. STEFANO DI CAMASTRA	4207974	442803	60	1971	2004
380	GERACI SICULO	4191461	425673	991	1973	2004
430	CASTELBUONO	4198604	420013	397	1953	2011
460	ISNELLO	4203887	412808	729	1953	2010
520	SCILLATO	4191270	404563	390	1972	2008
550	ALIA	4182001	386755	713	1972	2008
600	VICARI	4187837	374201	596	1929	2004
640	SAMBUCHI	4192472	380281	550	1960	2008
660	CIMINNA	4194981	373384	473	1928	2009
670	MONUMENTALE	4205600	384292	10	1939	2008
700	TURDIEPI	4204596	354362	658	1928	2003
720	FICUZZA	4194006	357434	691	1939	2001
780	PIOPPO	4213382	347017	378	1929	2007
790	ALTOFONTE	4212098	350419	392	1933	2004
840	S. MARTINO DELLE SCALE	4217362	346879	547	1938	2011
850	PARTINICO	4213798	335792	174	1957	2010
870	PALERMO (Ist. Zootecnico)	4219903	351044	124	1958	2005
880	PALERMO (Oss.Astronomico)	4219404	355462	41	1953	2010
910	PALERMO (Ist.Castelnuovo)	4224919	354160	48	1940	2008
920	PALERMO (Piazza Verdi)	4220406	356009	23	1928	2005
930	ISOLA DELLE FEMMINE	4229524	346760	0	1960	1998
9806	CALATAFIMI	4197667	313283	306	1931	2008
1000	SPECCHIA C.C.	4204258	296153	177	1971	2001
1010	LENTINA (Contrada)	4214350	296057	100	1948	2003
1020	S. ANDREA BONAGIA	4214703	290759	64	1956	2003
1030	TRAPANI	4210451	281182	0	1928	2009
1040	CASTELLAMMARE DEL GOLFO	4210401	314425	68	1970	2009
1050	CAPO S.VITO	4229238	301460	1	1970	2010
1060	FASTAIA	4200568	301579	228	1961	2009
1070	DIGA RUBINO	4195352	299314	185	1972	2006
1080	BORGO FAZIO	4192617	294514	183	1953	2008
1100	BIRGI NUOVO	4196090	279791	20	1953	2010
1110	CIAVOLO (Contrada)	4182184	284316	83	1968	2009
1120	MARSALA	4187992	276125	8	1942	2010
1140	MAZARA DEL VALLO	4170625	288266	17	1942	2006
1160	DIGA TRINITA'	4174134	301889	75	1972	2010
1180	PARTANNA	4178077	314256	429	1928	2008
1190	CASTELVETRANO	4174003	306087	201	1928	2007
1210	PIANA DEGLI ALBANESI	4207075	348899	734	1928	2002
1220	CASA DINGOLI	4208205	351545	725	1932	2004
1230	S. CRISTINA GELA	4205713	353081	670	1928	2006
1240	DIGA MAGANOCE	4203364	349734	625	1928	2007
1270	CORLEONE	4186778	350533	564	1952	2008
1310	GIBELLINA	4184438	320714	420	1969	2005
1330	MONTEVAGO	4174938	322230	375	1969	2004
1370	DIGA ARANCIO	4166248	328972	269	1955	2009
1390	CALTABELLOTTA	4160464	342463	768	1958	2010
1400	SCIACCA (Scuola Agraria)	4153499	334511	111	1929	2008
1420	PIANO DEL LEONE	4170590	364733	845	1931	2011
1430	PALAZZO ADRIANO	4172402	357101	613	1932	2010
1440	CHIUSA SCLAFANI	4171528	347504	620	1928	2004
1480	BIVONA	4164898	362292	515	1964	2010
1490	LERCARA FRIDDI	4178883	376710	669	1958	2009
1520	DIGA FANACO (Platani)	4170403	372444	678	1955	2005
1540	CASTRONOVO DI SICILIA	4170817	364798	875	1933	1999
1600	S. CATERINA VILLARMOSA	4160714	414358	606	1955	2010
1630	MARIANOPOLI	4162540	402994	520	1933	2006
1680	MUSSOMELI	4160541	390072	786	1960	2008
1740	PIETRANERA (Azienda)	4155963	369048	160	1982	2005
1750	CIANCIANA	4153963	361566	377	1973	2004
1760	CATTOLICA ERACLEA	4145127	358205	185	1971	2009
1810	AGRIGENTO (Ispett. Agrario)	4129672	374931	182	1928	2009
1840	CANICATTI'	4135377	398198	460	1972	2009
1850	PALMA DI MONTECHIARO	4116128	391730	145	1969	2005
1860	PETRALIA SOTTANA	4186213	420281	950	1931	2008
1900	ALIMENA	4173714	422570	769	1974	2008
1920	RESUTTANO	4170791	414489	591	1972	2004
1960	ENNA	4158307	435811	877	1930	2009

Table 4.1: continued on next page

Table 4.1: continued from previous page

ID	LOCATION	North ED50 [m]	East ED50 [m]	Elevation [m]	S-year	E-year
2020	CALTANISSETTA	4150518	416707	618	1928	2009
2040	PIETRAPERZIA	4142058	423781	478	1956	2010
2240	GELA	4102365	433501	13	1928	2009
2250	PIAZZA ARMERINA	4138324	444269	677	1928	2008
2260	DIGA DISSUERI	4116398	437034	174	1935	2010
2290	MONTEROSSO ALMO	4104685	478933	624	1958	2002
2300	CHIARAMONTE GULFI	4098259	473650	615	1974	2010
2310	DIGA RAGOLETO	4109137	471789	325	1977	2010
2320	VIZZINI	4112996	477763	562	1957	1997
2350	VITTORIA	4089771	457849	165	1928	2002
2370	RAGUSA	4086500	475456	514	1931	2009
2380	MODICA	4079425	478831	375	1929	2009
2440	CASTELLUCCIO	4087577	492655	197	1933	2010
2450	NOTO	4082655	505748	77	1940	2002
2470	PALAZZOLO ACREIDE	4101831	491141	677	1928	2010
2490	SORTINO	4113242	503643	465	1976	2003
2550	AUGUSTA	4120485	519911	2	1951	2009
2580	LENTINI (Città)	4127327	499904	52	1939	2010
2590	LENTINI (Bonifica)	4133115	507231	3	1928	2003
2660	MANIACI	4191500	484278	745	1953	2005
2690	CESARO'	4188865	474475	1108	1948	2008
2750	TROINA	4181917	464776	1001	1933	1996
2790	BRONTE	4182182	485339	781	1928	2009
2800	NICOSIA	4178089	446907	706	1928	2009
2870	RAGALNA	4164810	494769	702	1969	2004
2900	ADRANO	4168791	484989	549	1948	1998
2910	PATERNO'	4158035	492108	245	1928	2004
2940	LEONFORTE	4166668	446794	647	1948	2004
2960	VALGUARNERA	4150071	445902	590	1928	2005
2970	TORRICCHIA	4161108	465352	238	1969	2002
2980	CATENANUOVA	4158420	472896	175	1975	2002
3000	RADDUSA	4147520	458685	325	1928	2002
3050	CALTAGIRONE	4121762	457304	486	1930	2008
3060	MINEO	4124444	472594	500	1928	2010
3070	RAMACCA	4137784	473085	263	1931	2000
3110	ZAFFERANA ETNEA	4172087	509384	583	1929	2009
3120	LINGUAGLOSSA	4188191	512771	525	1968	2010
3150	ACIREALE	4163912	514640	183	1929	2010
3160	CATANIA (Ist. Agrario)	4152468	506400	68	1977	2010
3170	CATANIA (G.C.OO.MM.)	4150371	508351	0	1928	2002
3290	ALCANTARA	4186523	522413	38	1949	2005
3310	TAORMINA	4189603	525753	247	1929	2009
3350	CAMARO (Caserma Forestale)	4229511	543857	544	1929	2005
3380	MESSINA (Ist. Geofisico)	4228394	548574	42	1950	2011
3400	GANZIRRI	4234573	553279	0	1953	2006
3410	ALI' TERME	4206499	537177	13	1959	2002

4.3 Variables used for the RFA

In order to identify the homogeneous regions, the regional frequency analysis needed to select the variables to be used to a such aim. This step is very important to obtain a regional frequency analysis with good performance in terms of homogeneous regions.

The available data for the formation of the regions are quantities often calculated from the variable of interest through simple or advanced statistical operation, denoted as *at-site statistics* (e.g. quantile), and other site descriptors called *site characteristics*. In environmental applications, the *site characteristics* would typically include the geographical location at site, its elevation, and other physical

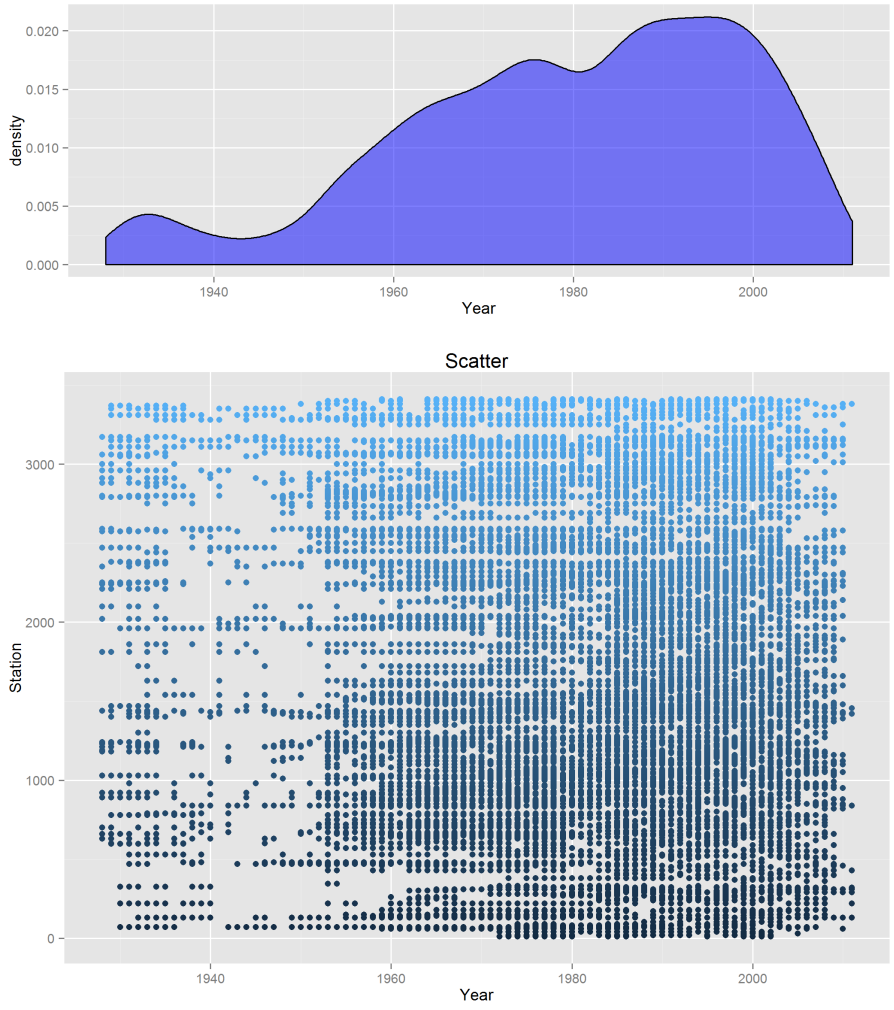


FIGURE 4.5: Consistency plot (bottom) and density plot (top) for the value of AMR_d for all the stations spanning the period 1928-2010

properties associated with the site while the *at-site statistics* would include the quantile calculated or percentile value of the variable analysed (Hosking and Wallis, 2005).

In principle, *site characteristics* are quantities that are known even before any data are measured at a site. However, it is reasonable to include among the *site characteristics* some quantities that are estimated from data measured at site, provided that these measurements are not too highly correlated with the variable values themselves (Hosking and Wallis, 2005).

The choice of the variables has been conducted with the necessity of describing the behaviour of the extreme events, and consequently, increasing the comprehension related to the high variability of events over the Sicily.

In this study, among all the possible variables that can be used in the RFA, following previous studies a particular set was selected. This variables set, listed

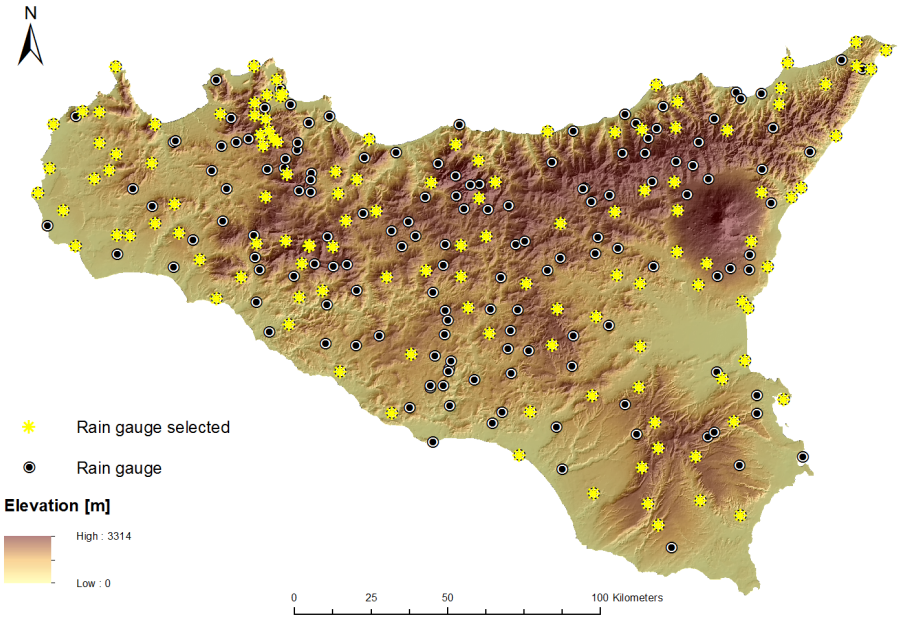


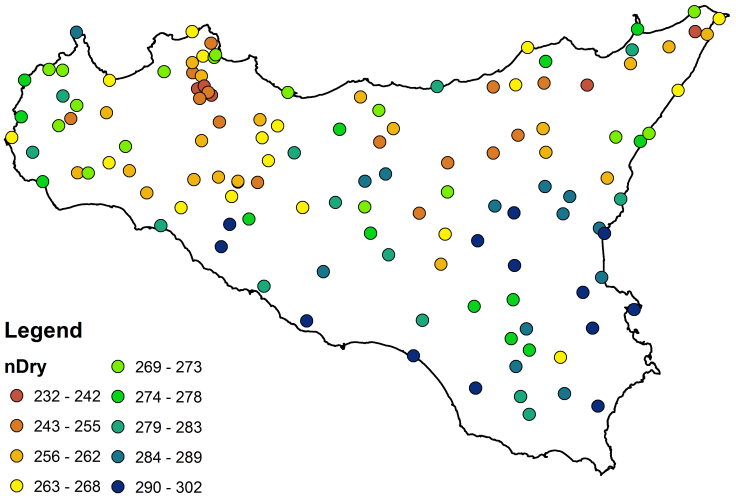
FIGURE 4.6: Station selected spanning the period 1972-2003

in Table 4.2, contains the following variables: the raingauge elevation Z_i; the annual maxima rainfalls with duration, $AMR_{d,i}$; $nDry_i$ that depicts the mean annual number of dry day; Rs_i/Rw_i that is the ratio between summer rainfall (i.e. the cumulative rainfall recorded between April-September) and the winter rainfall (i.e. cumulative rainfall recorded between October-May). These variables have been selected because provides an idea as the rainfall are subdivided during the year, highlighting possible seasonal characteristics of seasonality.

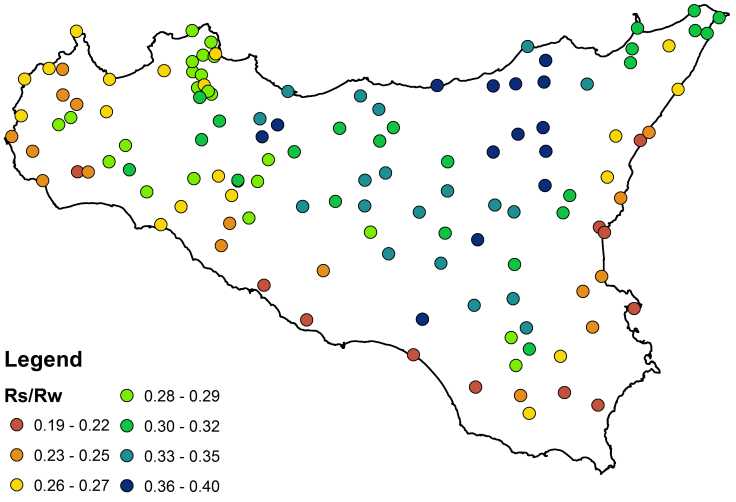
The Figure 4.7(a) and Figure 4.7(b) show the value of the $nDry_i$ and Rs_i/Rw_i, respectively, for each stations. The south coast and the south-east coast show the highest number of dry day and the lowest value of Rs_i/Rw_i. The opposite

TABLE 4.2: Variables used in rainfall region development

Variable	Description
Z_i	Station elevation
$A\bar{M}R_{d,i}$	Annual maxima rainfall for fixed duration
$nDry_i$	Number of days < 1 mm
Rs_i/Rw_i	Ratio between summer rainfall (April - September) and winter rainfall (October - May)
$\bar{\theta}_i$	Mean date of events, represents a measure of the average time of occurrence of rainfall events (i = 1, 3, 6, 12 and 24 hours).
$\bar{r}_i$	Seasonality vector provides a dimensionless measure of the spread of the data (i = 1, 3, 6, 12 and 24 hours).



(a) *nDry*



(b) *Rs/Rw*

FIGURE 4.7: Maps of number of dry day (nDry) and ratio among summer rainfall and winter rainfall (Rs/Rw)

behaviour is shown on the stations in the north-east coast, where the precipitations are distributed during all the year.

Reed (1994) have suggested that measures of the seasonality and timing can be used as variables to define similar regions, as shown in the *Flood estimation handbook* (Reed et al., 1999). In order to describe the seasonality, two indices, named θ and $\bar{r}$, have been used; they respectively represent a measure of the average time of occurrence of rainfall events and the measure of the time spread of the data.

These variables that can be assimilated to *site characteristics* able to describe

the seasonality of hydrological time series are the *directional statistics* (Fisher, 1993, Magilligan and Graber, 1996, Mardia, 1975).

Following Bayliss and Jones (1993), Burn (1997) and Cunderlik and Ouarda (2009), the *Julian date* of occurrence of an annual hydrologic event i can be transformed to a directional statistic:

$$\theta_i = (julianedata)\frac{2\pi}{365} - \frac{2\pi}{3} \tag{4.1}$$

where θ_i is the angular value (in radians).

Since the majority of the events occur over or between October and March, an angular value equal to $2\pi/3$ is subtracted from the transformation of the Julian date in order to remove the ostensible difference of the values because of the change of the year.

Burn (1997) suggested for a sample of n events, the values of θ_i can be aggregated in order to estimate the $\bar{x}$ and $\bar{y}$ coordinates of the mean date of event occurrence:

$$\bar{x} = \frac{1}{n}\sum_{i}^{n}\cos\theta \tag{4.2}$$

$$\bar{y} = \frac{1}{n}\sum_{i}^{n}\sin\theta \tag{4.3}$$

The mean direction of the event dates is then obtained from:

$$\bar{\theta} = \begin{cases} tan^{-1}(\frac{\bar{y}}{\bar{x}}), & \bar{x} \geq 0, \bar{y} \geq 0 \\ tan^{-1}(\frac{\bar{y}}{\bar{x}}) + \pi & \bar{x} < 0 \\ tan^{-1}(\frac{\bar{y}}{\bar{x}}) + 2\pi & \bar{x} \geq 0, \bar{y} \geq 0 \end{cases} \tag{4.4}$$

A measure of the variability of the n rainfall occurrences about the mean date can also be determined. The variability measure is obtained by defining the mean resultant as:

$$\bar{r} = \sqrt{\bar{x}^2 + \bar{y}^2} \tag{4.5}$$

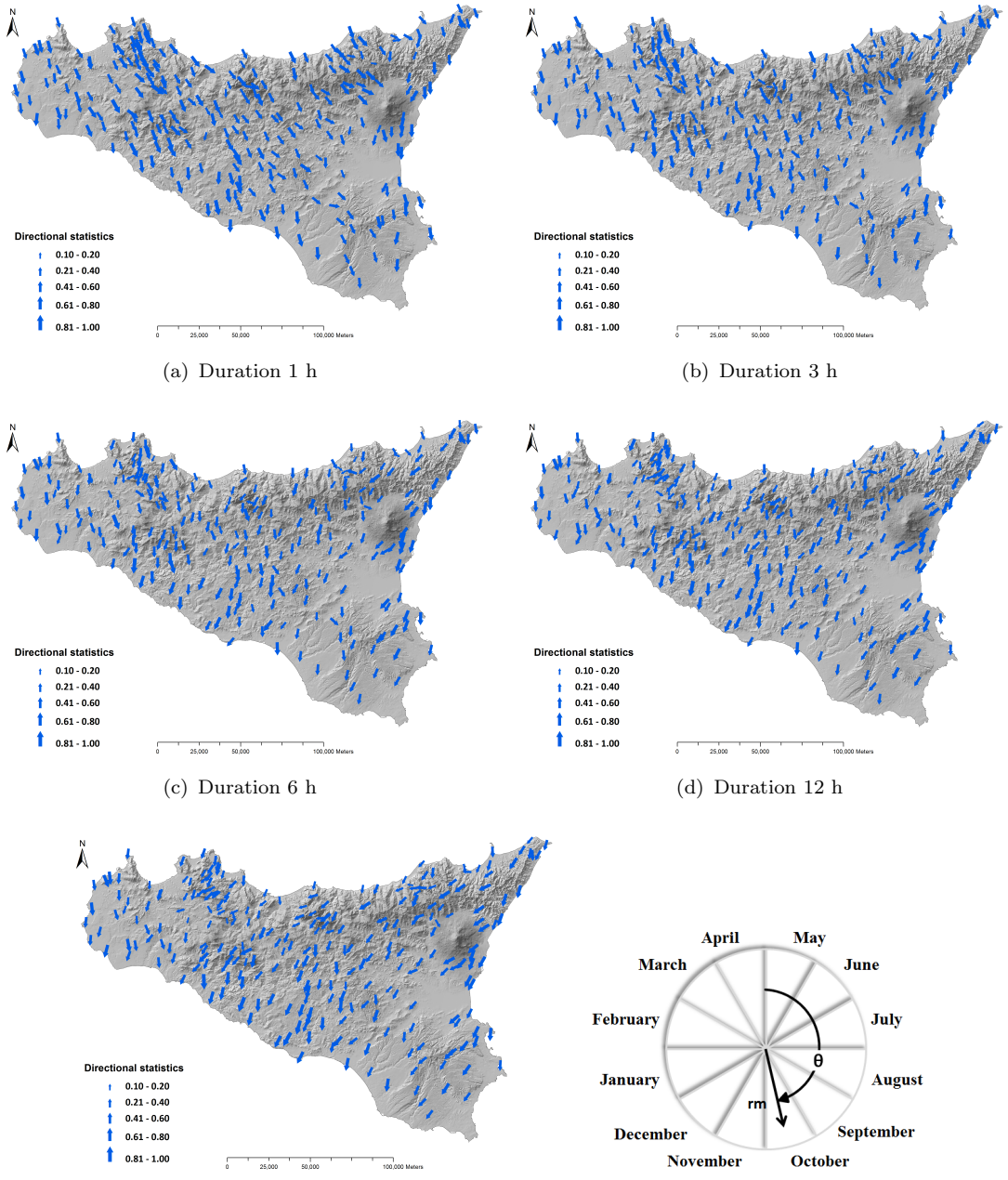
where $\bar{r}$ provides a dimensionless measure of the spread of the data, ranging from zero (high variability) to one (low variability).

The use of these variables, through the similarities of the timing and of the seasonality of the extreme events, can give an indication of the meteorological characteristics of the stations.

The maps (Figure 4.8) rappresents the two indices θ and $\bar{r}$ by a vector, wherein the angle indicates the timing and the thickness indicates the spread of the event for each stations. Observing these vectors for the different durations, it is possible to deduce how the events occurrance tends to shift from the autumn season to winter season.

The mean value of the directional statistics, shown in Figure 4.8 for each durations, were averaged and plotted in polar plots shown in Figure 4.9, where the colour of the slice represents the percent of the event that occur in each month.

The polar plots describes more clearly how the timing of the rainfall events moving from the October/November to December/January with the increasing of the duration. This result could explain the type of the event, in fact, the convective events usually are linked to the short precipitations that manifest themselves at the beginning of the autumn season, while the stratiform events usually associated to the long duration occur fully in the winter season.



(a) Duration 1 h

(b) Duration 3 h

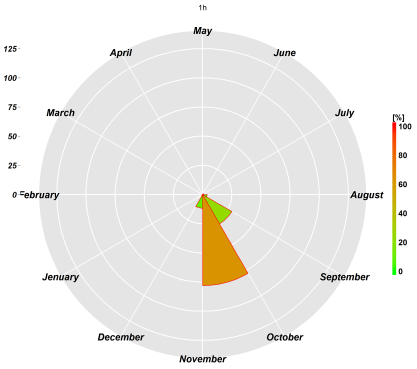
(c) Duration 6 h

(d) Duration 12 h

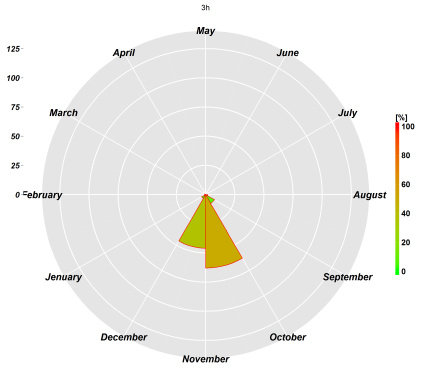
(e) Duration 24 h

(f) Legend

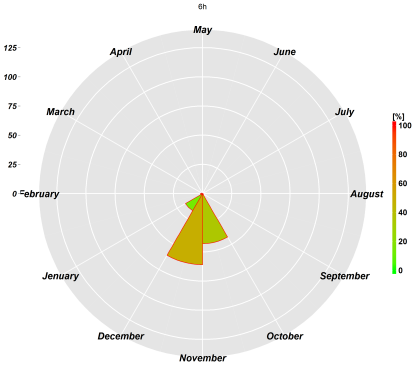
FIGURE 4.8: Maps of the mean direction and variability measure. The size of the arrow provides the value of variability measure while the direction of the arrow provides the timing of the event, as it is shown in the legend in bottom right.



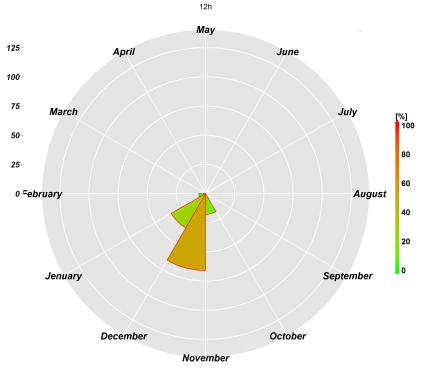
(a) Duration 1 h



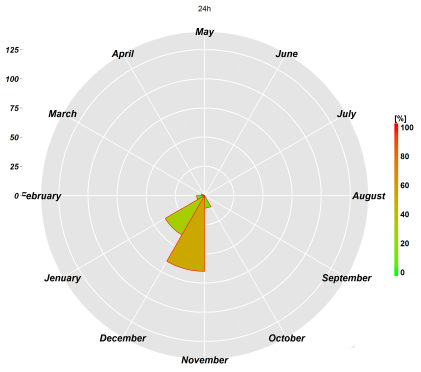
(b) Duration 3 h



(c) Duration 6 h



(d) Duration 12 h



(e) Duration 24 h

FIGURE 4.9: Plot polar for the directional statistics for each duration. The shaded value shows the $\bar{r}$ value in percent.

Chapter 5

Case study - Regional frequency analysis in Sicily

The aim of this work consists in the design and the development of a RFA procedure for the area of Sicily, Italy, based on the selection of suitable procedures considering the data availability and the meteoclimatic features of the area. In the previous works related to the same area (Cannarozzo et al., 1995, Lo Conti et al., 2007), the choice of the number of the homogeneous regions and their extension were made with geographical criterion related principally on watersheds boundaries.

In these works, the regionalization was based on an initial hypothesis related to geographical or hydrological characteristics. For the identification of homogeneous regions such hypothesis was subsequently verified. The new approach tries to obtain a regionalization starting from the meteo-climatic characteristics identified and defined in the previous chapter, considering not only those linked to the precipitation depth but even variables able to describe the seasonality and the timing of the events. In order to obtain a new regionalization, a combination of statistical tools was used to support the "traditional" regional frequency analysis.

For these reasons in this work, an objective method has been adopted to achieve the identification of the homogeneous regions researching, at same time, the best distribution able to represent the characteristics of the regions identified. An assessment of the accuracy has been carried out to confirm the goodness of fit as well.

5.1 Methodology

The methodology developed in this chapther follows the *index flood method* introduced by Dalrymple (1960), and previously described.

In this application, the mean of the frequency of distribution has been assumed as the scale factor and has been estimated through the sample mean of the at-site values; consequently the mean of rescaled data is 1 for each site, and so the regional average of these means is 1. The flow diagram, that describes the main steps of the methodology, is shown in Figure 5.1.

The selected variables were used as input in the Principal Components Analysis (PCA), whose primary purpose is to reduce the number of variables taken into account in the process of identification of homogeneous regions. As previously mentioned, PCA is particularly useful when a data reduction procedure that makes

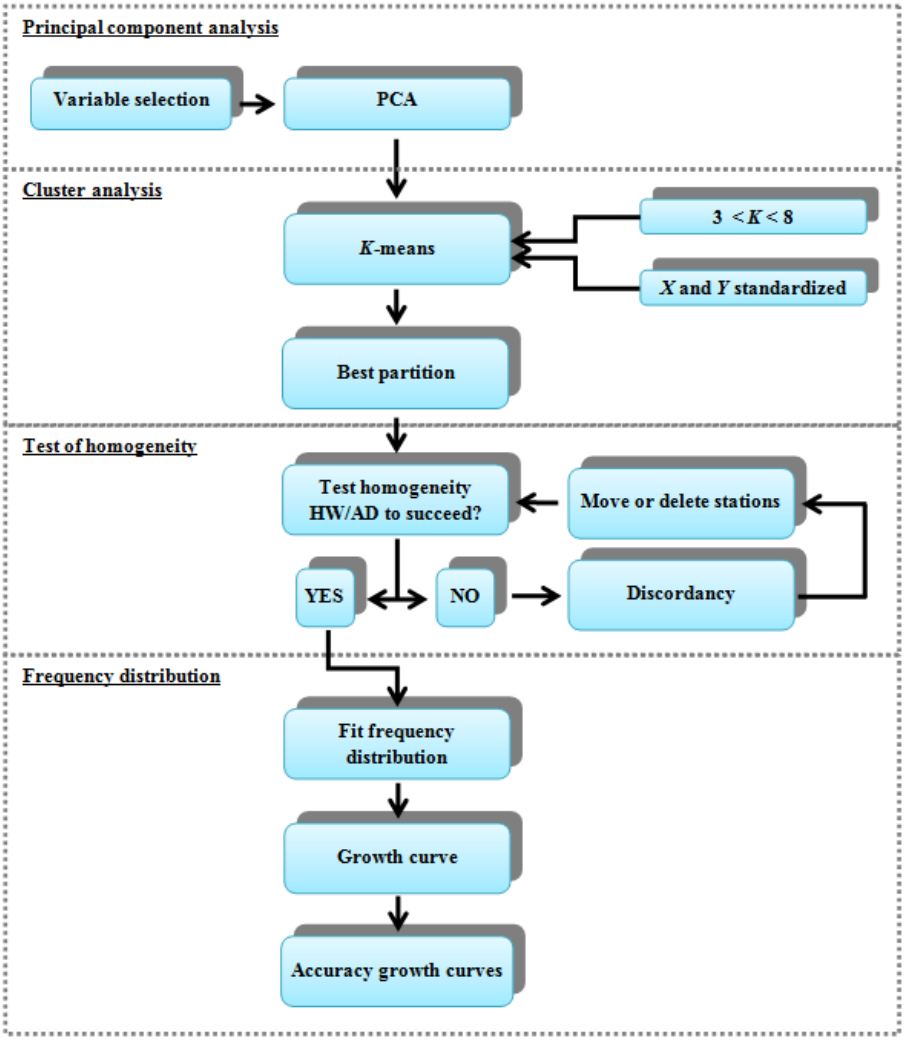


FIGURE 5.1: The flow diagram, that describes the main steps in the methodology

no assumptions is needed with regard to an underlying causal structure that is responsible for co-variation in the data. The new variables obtained after PCA, the i.e. principal components (PCs), have been provided as input for the cluster analysis process.

The cluster analysis (CA) has been carried out using, in addition the PCs selected from the PCA analysis, the normalized latitude and longitude of each station to support the grouping of spatially continuous regions. Operationally, the CA has been performed by means of the k-means clustering method (Hartigan and Wong, 1979).

Once a set of physically plausible regions have been defined, it is necessary to assess their degree of homogeneity that tests whether two or more homogeneous regions are sufficiently similar that they should be combined into a single region.

In this study three different tests were chosen, i.e., the discordance measure for each station D (Hosking and Wallis, 1993) to examine possibly anomalous behaviour of individual stations; the HW (Hosking and Wallis, 1993), and the Anderson-Darling rank test AD (Scholz and Stephens, 1987) to assess homogeneity of the extreme rainfall regions obtained.

After having identified the homogeneous regions, the next step is to test whether a given distribution fits the data acceptably. A related aim is the selection, from a number of candidate distributions, the distribution that gives the best fit to the data, through the goodness-of-fit procedure based on the computation of Z^{DIST} (Hosking and Wallis, 2005) described in section 3.5.

There are many families of distributions that might be candidates for being fitted by a regional data set. Their suitability as candidates can be evaluated by considering their ability to reproduce features of the data that are of particular importance in modelling. The goodness-of-fit is evaluated for each homogeneous region considering five different probability distributions: *generalized logistic* (GLO), *generalized extreme-value* (GEV), *lognormal* (LN3), *Pearson type III* (PE3) and *Generalized Pareto Distribution* (GPA).

For each of the homogeneous regions identified, the growth curves, for the distribution selected, were obtained setting $F(x') = 1 - 1/T$ and evaluating the quantiles function x' for the different return periods T [i.e. $x^{'}(T)$].

In order to evaluate the reliability of the results relative to quantile evaluation, the assessment of the magnitude of uncertainty has been carried out. In this work, the algorithm implemented by Hosking and Wallis (2005) has been used to evaluate the accuracy of the quantiles estimated.

5.2 Extreme rainfall regions

5.2.1 Principal component analysis

The results of a PCA are usually discussed in terms of component scores o principal components (PCs), sometimes called *factor scores*, that rappresent the transformed variable values corresponding to a particular data point, while the *loadings* rappresent the weight by which each standardized original variable should be multiplied to get the component score.

Some of the rainfall variables used for the identification of the homogeneous regions depend on the duration (Table 4.2) and for this reason, five groups of principal components have been obtained; consequently, five different identification of homogeneous regions have been obtained.

The residual percent variance method was used to select the principal components obtained from the analysis. The method is based on the idea that the residual variance should reach a steady state when the factors begin to account for random errors (Valle et al., 1999). In this study, it was decided to retain the principal components which accounted for at least 5% of the total variance in the input dataset, analysing also the values obtained for the different durations; then only the five principal components, that explain about 95% of data variance, were selected (Table 5.1).

Analysing the PCs obtained from the different durations (Table 5.1), the components have shown more or less the same proportional contribution to the variance, with exception for the duration of 1 h that has shown a slightly different behaviour. For this reason it has been considered the possibility to derive a unique

TABLE 5.1: Percentage variance explained for every duration for the different PCA. In bold, the values lower than threshold selected (5%) and parenthetical the cumulative of the explained variance

PC	1h	3h	6h	12h	24h
1	42.40 (42.40)	37.62 (37.62)	37.75 (37.75)	38.54 (38.54)	40.64 (40.64)
2	18.93 (61.33)	18.87 (56.49)	21.03 (58.78)	23.03 (61.57)	22.37 (63.01)
3	13.86 (75.19)	17.78 (74.27)	16.90 (75.68)	15.00 (76.57)	14.10 (77.11)
4	11.01 (86.19)	11.95 (86.22)	11.15 (86.83)	10.69 (87.27)	10.99 (88.10)
5	7.71 (93.91)	8.36 (94.58)	8.29 (95.12)	7.82 (95.08)	7.10 (95.21)
6	6.09 (100.00)	5.42 (100.00)	**4.88** (100.00)	**4.92** (100.00)	**4.79** (100.00)

group of regions able to explain and resume the characteristics relative to the different five duration data.

Analysing the loadings values in Table 5.2, where the bold type indicates most significant contributing variables, it is possible to suppose a connection among the PCs and the original variables. For example, since the most important information of the first PCs is provided from the Z and R_s/R_w, it is possible to suppose that the first PCs is linked with the morphology and R_s/R_w. Similar conclusion can be made for other PCs.

Observing loadings in Table 5.2, it was decided that the PCs for the different durations could be combined through an averaging operation.

In order to confirm the results provided by the value of loadings (Table 5.2), the Figure 5.2 shows the comparison among the interpolation of the score of the PCs by *Universal Kriging* and the observed values in each station, where the colours were inverted, respect to maps, to underline the correlation among them. The Figure 5.2(a) highlights the responses to orography, the mountains chains emphasized with the black coincide with the clear values inferred from the stations. The high value of precipitation (Figure 5.2(b)) and the higher value of the dry day show a link with the seasonality, in particular, in the east side where the rainfall events are usually short and intense (Figure 5.2(c)) are clearly correlated, respectively, to the second and third PCs.

5.2.2 *k*-means clustering method

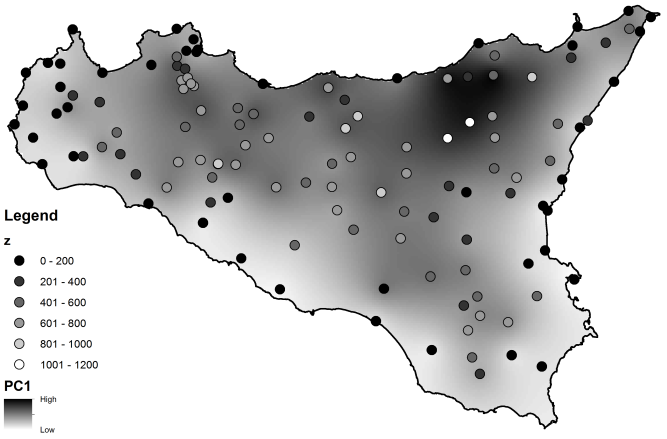
The cluster analysis has been performed by means of the k-means clustering method (Hartigan and Wong, 1979) using as input the variables obtained from the PCA analysis together with the normalized latitude and longitude of each station to support, as previously said, the grouping of contiguous regions.

In this study, the possible range of k value have been selected considering the minimum and a maximum number of region used in previous works realized with regard of Sicily (Cannarozzo et al., 1995, Gabriele and Chiaravalloti, 2013, Lo Conti et al., 2007).

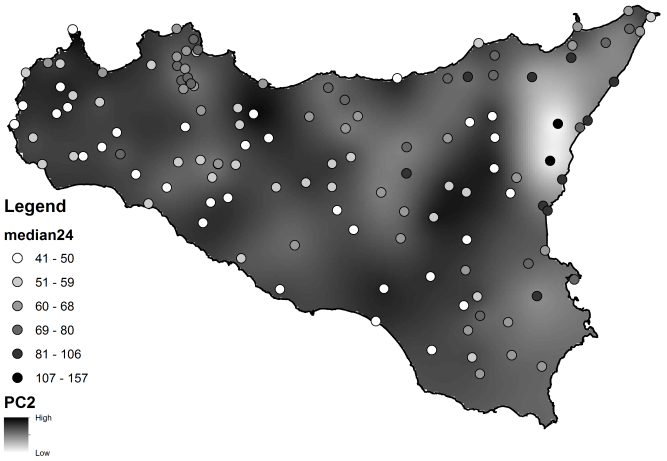
The silhouette method (Rousseeuw, 1987) has been used to evaluate the best number of regions in the range of k from 3 to 8. The entire clustering is displayed by combining the silhouettes into a single plot, and allowing for an appreciation of the relative quality of the regions and an overview of the data configuration (Figure 5.3). The average silhouette width provides an evaluation of clustering validity, and might be used to select an "appropriate" number of regions.

TABLE 5.2: Loadings of each variable within the first five principal component. Bold type indicates most significant contributing variables.

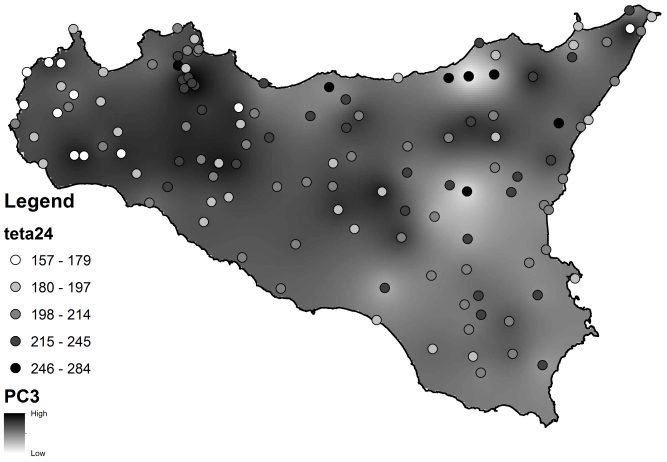
	Duration 1h				
Variable	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**
Z	**0.45**	-0.33	0.27	-0.36	**0.62**
AMR_d	-0.32	-0.36	**-0.74**	-0.43	0.18
θ	-0.42	0.02	**0.56**	**-0.57**	-0.04
$\bar{r}$	-0.34	**-0.56**	0.20	**0.60**	0.34
$nDry$	-0.34	**0.66**	-0.08	0.10	**0.66**
Rs/Rw	**0.53**	0.14	-0.15	0.05	0.19
	Duration 3h				
Z	**-0.48**	-0.38	0.14	0.41	-0.44
AMR_d	0.31	**-0.57**	0.23	**-0.65**	-0.31
θ	0.20	**-0.62**	**-0.56**	0.36	0.09
$\bar{r}$	0.41	0.05	**0.58**	**0.46**	-0.33
$nDry$	0.40	0.37	-0.51	0.02	**-0.62**
Rs/Rw	**-0.56**	0.08	-0.13	-0.25	**-0.46**
	Duration 6h				
Z	**0.48**	-0.20	0.35	**-0.57**	-0.16
AMR_d	-0.17	**-0.74**	0.09	0.42	-0.4
θ	0.20	**-0.57**	**-0.59**	-0.30	0.38
$\bar{r}$	-0.48	-0.16	0.41	**-0.53**	-0.10
$nDry$	-0.42	0.18	**-0.57**	-0.35	**-0.49**
Rs/Rw	**0.55**	0.16	-0.14	0.00	**-0.60**
	Duration 12h				
Z	**0.48**	-0.22	0.38	**-0.51**	0.18
AMR_d	-0.04	**-0.76**	-0.11	0.24	**-0.55**
θ	0.36	**-0.40**	**-0.62**	-0.08	0.50
$\bar{r}$	**-0.43**	**-0.40**	0.37	-0.48	0.07
$nDry$	**-0.43**	0.14	**-0.56**	**-0.58**	-0.16
Rs/Rw	0.52	0.21	-0.10	-0.33	**-0.62**
	Duration 24h				
Z	**0.47**	-0.15	0.45	**-0.54**	-0.05
AMR_d	0.07	**-0.77**	-0.09	0.26	0.56
θ	0.40	**-0.40**	**-0.51**	-0.08	**-0.62**
$\bar{r}$	-0.43	-0.39	0.35	**-0.48**	-0.11
$nDry$	-0.43	0.02	**-0.60**	-0.53	0.17
Rs/Rw	**0.50**	0.26	-0.22	-0.35	**0.51**



(a) PC1 describes the morphology

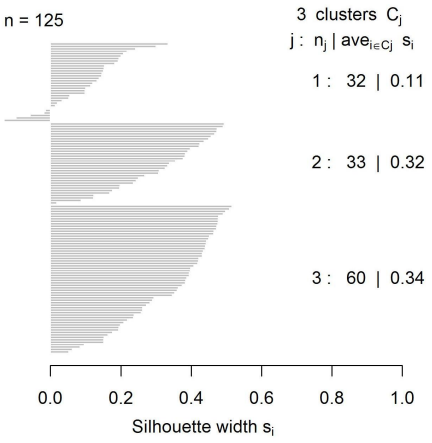


(b) PC2 describes the annual maximum

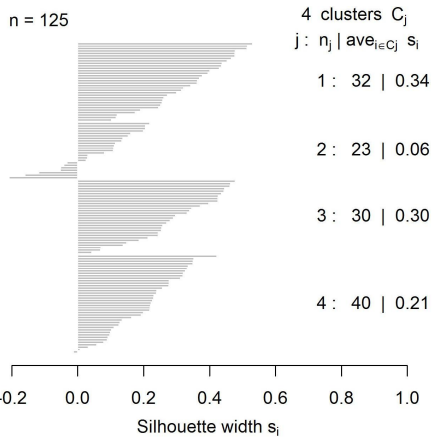


(c) PC3 describes the seasonality

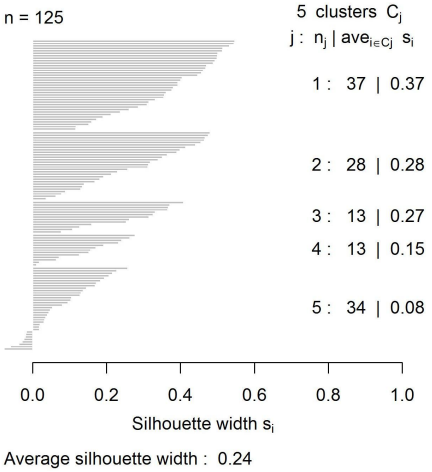
FIGURE 5.2: Some measures of extreme rainfall and the scores of principal components derived from the variables showed in Table 4.2. The dots are the values recorded of the stations which show a relationship with the score of the PCA.



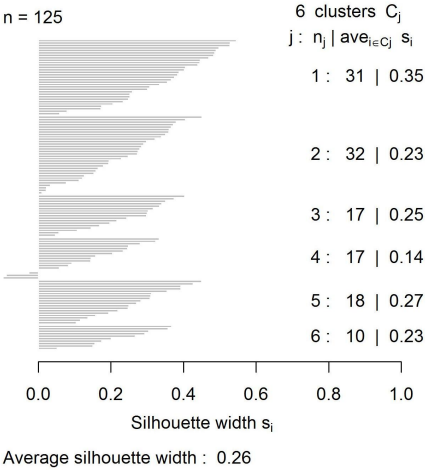
(a) 3-region



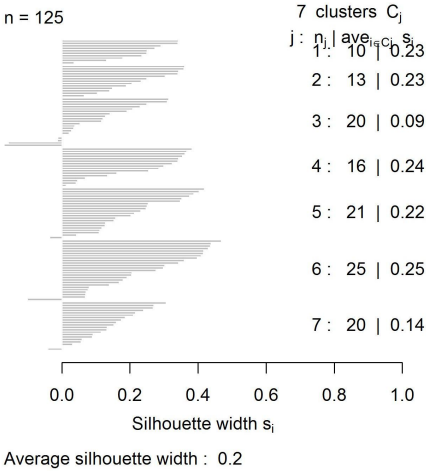
(b) 4-region



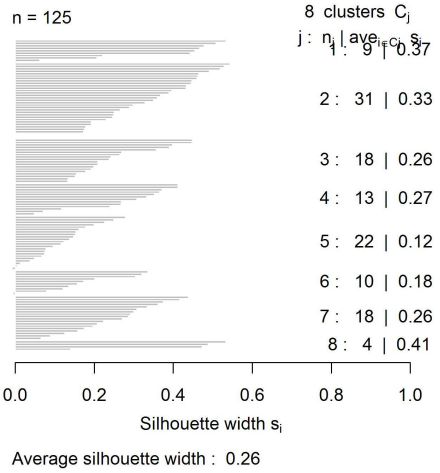
(c) 5-region



(d) 6-region



(e) 7-region



(f) 8-region

FIGURE 5.3: Silhouette value for different number of region

The values in Table 5.3 list min, max and average values of the silhouette. Analysing the values in the table, the most robust and optimal solution has been obtained for *k=6* wherein the silhouette value is lower than other k value. The silhouette value for *k=8* has the same mean value than *k=6* , but there is a cluster with only 4 stations, and for this reason this possible solution has not been taken into account.

The regions obtained from the cluster analysis are shown in Figure 5.4 while the characteristics of each region are summarized in Table 5.4. The Figure 5.4 has been obtained considering the limit of the watershed. Only two regions (4 and 5) result significantly more numerous than the others while the region 3 includes only ten stations. The low number of stations inside that present in a region could cause increase of the uncertainty, that could result higher than regions with a high number of stations.

In Table 5.4, it is possible to observe as the mean annual precipitation is not strongly related to the mean elevation; indeed the region 3, that shows the highest values of AMR_d, does not have the highest elevation, while opposite observation can be made for the region 5.

5.2.3 Test of regional homogeneity

The two different tests of homogeneity were applied at the regions obtained through the cluster analysis.

As previously said, Viglione et al. (2007) have shown that the Hosking and Wallis heterogeneity measure HW_1 (only based on L-CV) is preferable when skewness is low, while the bootstrap Anderson-Darling test should be used for more skewed regions. The authors suggested that the L-moment space can be divided into two regions: if the L-skewness coefficient for the region under analysis is lower than

TABLE 5.3: Max, min and mean Silhouette value for different numbers of region

N. of region	Max s(i)	Min s(i)	Mean s(i)	Mean n. stations	Max n. stations	Min n. stations
3	0.34	0.11	0.28	42	60	32
4	0.34	0.06	0.24	31	23	40
5	0.37	0.08	0.24	25	13	37
6	0.35	0.14	0.26	20	10	32
7	0.25	0.09	0.20	18	10	25
8	0.41	0.12	0.26	16	4	31

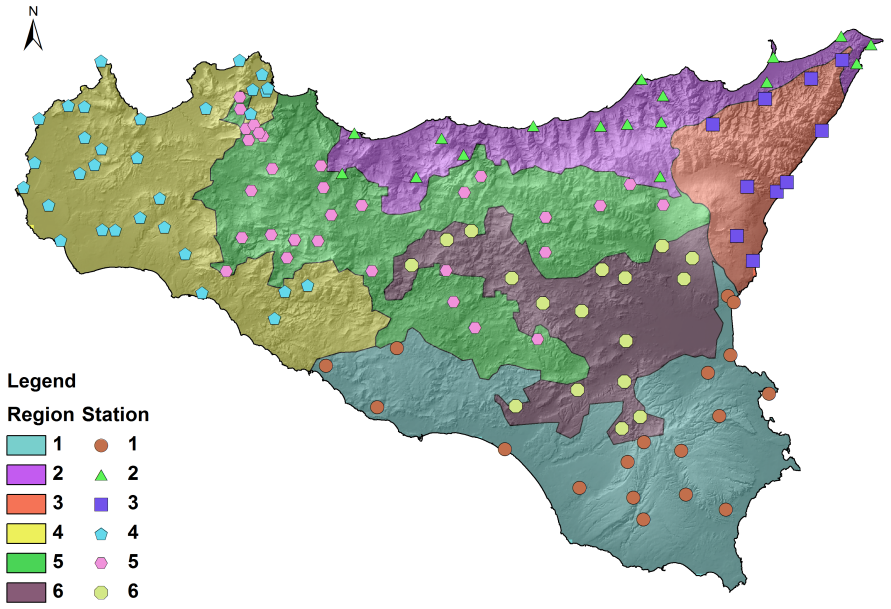


FIGURE 5.4: Homogeneous regions obtained with the k-means

TABLE 5.4: Characteristics of the homogeneous regions

Region	Stations	$A\bar{M}R_d$ (mm)					Mean annual rainfall (mm)	Mean elevation
		1h	3h	6h	12h	24h		
1	18	28	37	46	56	67	586	257
2	17	24	34	42	52	63	746	297
3	10	33	48	65	80	99	911	373
4	31	25	33	39	46	55	573	156
5	32	23	31	39	48	58	714	697
6	16	23	31	38	46	55	509	473
All	124	26	36	45	55	66	673	376

0.23, they propose to use the Hosking and Wallis heterogeneity measure HW_1; otherwise if L-skewness > 0.23, the bootstrap Anderson-Darling test is preferable. In the Table 5.5, the results of two tests applied are illustrated (white background AD where L-skewness >0.23 and grey background HW_1 where L-skewness <0.23).

When the test of homogeneity indicates that the regions cannot be accepted, the characteristics of sites that have showed marked differences have been carefully examined. If a site results discordant, it was evaluated to reassign it to other regions, but often there was no physical reason why the atypical sites should be different from the rest of the region.

For example, an extreme convective meteorological event could affect only a few stations in a region; this could yield discordant values for only a few events,

TABLE 5.5: Results of the test of homogeneity; Anderson-Darling test are indicated with white background (upper limit=0.95), while Hosking and Wallis are indicated with grey background (upper limit=1.00). In bold the values that exceed the upper limit.

	Test of homogeneity					
	1	**2**	**3**	**4**	**5**	**6**
1 h	0.66	**1.01**	-0.01	0.17	0.13	-1.58
3 h	0.97	0.13	0.41	0.97	0.06	-0.52
6 h	0.81	0.15	0.00	0.24	0.74	**1.55**
12 h	0.10	0.26	0.35	0.62	0.68	**1.66**
24 h	0.04	0.07	0.45	-0.73	0.87	0.76

and in this case, these sites could be to treat anyway into a homogeneous region (Hosking and Wallis, 2005).

The region 6 resulted, for two duration, "possibly heterogeneous", but it has no station with discordant value evaluated by the discordance measure. In other cases, some regions presented stations with discordant values but during the test of homogeneity, the region as a whole resulted homogeneous. In this case, the stations have not been deleted and then used during the subsequent analysis.

Since it is difficult in practice to ensure that sites used in a particular application of RFA constitute a region that is exactly homogeneous, the analysis of the performance is necessary to evaluate if the region can be accepted even when it is possibly heterogeneous, as the region 6.

5.2.4 Choice of frequency distribution

In a RFA analysis, a single frequency distribution is fitted to data coming from several sites. In general, the region will be slightly heterogeneous, and there will be no single "true" distributions that apply to each region. The aim is therefore not to identify a "true" distribution but to find a unique regional distribution that will yield accurate quantiles estimates for each site.

Calculation of heterogeneity and goodness-of-fit measures involve the estimate of sampling variability of L-moment ratios in a homogeneous region whose record lengths and average L-moment ratios match those of the data. In the hypothesis that the region is homogeneous and data at different sites are statistically independent, if one of the distributions correspond to the "true" distribution, its goodness-of-fit measure should have approximately a standard normal distribution.

In this application, the sampling variability is estimated by Montecarlo simulation using 1000 replications for each single distribution as explained in the Section 3.5.

The goodness-of-fit measure was made for each region and for different durations (1h, 3h, 6h, 12h and 24h). The Z^{DIST} statistics for the five candidate distributions are shown in Table 5.6, indicating that the LN3 and the GEV provided acceptably close fits to the regional average L-moments because Z^{DIST} resulted lower in absolute value than 1.645 for most cases. For the PE3, the test was passed only for some durations and regions, then this reason, this distribution was not used. The GPA has provided the worst performance, as previously said, this distribution is more conform to model the POT approach.

In order to compare this work with the other previous works carried out for Sicily (Cannarozzo et al., 1995, Lo Conti et al., 2007), in addition to the two chosen distributions, the TCEV distribution have been introduced to permit a possible comparison with previous works.

TABLE 5.6: Simulations results for the goodness of fit measure Z^{DIST}. In bold, the values with $|Z| < 1.645$. The grey background shows the best distributions.

	Region 1					**Region 4**				
D	**GLO**	**GEV**	**LN3**	**PE3**	**GPA**	**GLO**	**GEV**	**LN3**	**PE3**	**GPA**
1h	3.66	**1.55**	**1.19**	**0.35**	-3.23	3.79	**1.11**	**0.56**	**-0.64**	-5.04
3h	2.75	**1.09**	**0.50**	**-0.62**	-2.90	3.27	**0.87**	**0.23**	**-1.05**	-4.75
6h	1.76	**0.32**	**-0.35**	**-1.58**	-3.29	3.05	**0.68**	**0.06**	**-1.20**	-4.87
12h	2.59	**1.30**	**0.52**	**-0.87**	-2.06	2.50	**-0.27**	**-0.68**	-1.68	-6.47
24h	1.87	**0.62**	**-0.22**	-1.70	-2.71	3.14	**0.27**	**-0.06**	**-0.98**	-6.07
	Region 2					**Region 5**				
1h	3.81	2.08	**1.50**	**0.39**	-2.08	3.13	**0.77**	**0.12**	**-1.16**	-4.78
3h	**1.06**	**-0.03**	**-0.77**	-2.06	-2.94	1.54	**-0.27**	**-1.18**	-2.82	-4.86
6h	**0.94**	**-0.37**	**-0.98**	-2.08	-3.63	0.55	**-1.27**	-2.04	-3.45	-5.76
12h	2.71	**1.20**	**0.58**	**-0.57**	-2.54	2.40	**0.50**	**-0.35**	-1.90	-4.24
24h	2.01	**0.68**	**0.00**	**-1.22**	-2.71	2.65	**0.84**	**-0.18**	-1.99	-3.84
	Region 3					**Region 6**				
1h	1.66	**0.37**	**-0.05**	**-0.84**	-2.73	2.11	**0.32**	**-0.04**	**-0.83**	-3.77
3h	**1.10**	**0.22**	**-0.41**	**-1.51**	-2.16	1.60	**-0.33**	**-0.58**	**-1.23**	-4.61
6h	**1.01**	**0.07**	**-0.50**	**-1.52**	-2.38	**0.77**	**-1.09**	**-1.30**	**-1.89**	-5.22
12h	**1.59**	**0.73**	**0.12**	**-0.95**	**-1.59**	2.28	**0.57**	**0.19**	**-0.62**	-3.38
24h	1.70	**0.76**	**0.16**	**-0.91**	-1.72	2.14	**0.86**	**0.20**	**-1.00**	-2.38

5.2.5 Parameters of distribution and hierarchical levels

After that the frequency distributions have been chosen, these have been fitted through data from the sites of each homogeneous region for each duration after scaling data by the at-site scaling factor. The maxima values for each station were normalized by the median of annual maximum values.

The distributions have been fitted using the method of the L-moments (Hosking and Wallis, 2005); in particular parameters have been estimated by equating the L-moments of the distribution to the sample L-moments calculated from the data.

An analysis of the dependence with the duration have been achieved, to confirm whether it was necessary to take into account a relation between the L-moments values and the duration. The values of the regionally weighted L-moments (L-kurtosis, L-skewness and L-CV) were calculated for the different durations (Equation 3.3) and these values were plotted in Figure 5.5 as a function of the duration for each region.

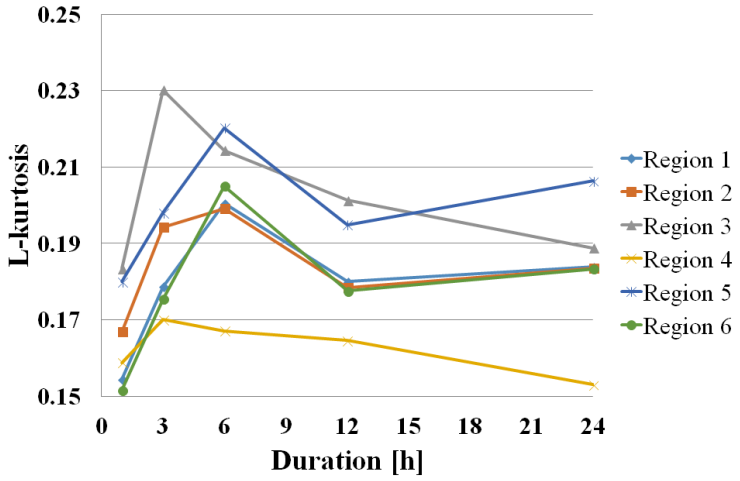
The Figure 5.5 shows the different behaviour of the L-moments of hourly rainfall between the duration with respect to the other durations, as it has been shown in the loadings value of the PCs. This behaviour could be due to the errors in the recorded data.

The L-skewness (Figure 5.5(b)) and L-CV (Figure 5.5(c)) for the region 4 and 6 have shown a possible dependence with the duration; in particular, for the region 6 the reason could be due to the "possibly heterogeneous" behaviour shown during the test of homogeneity.

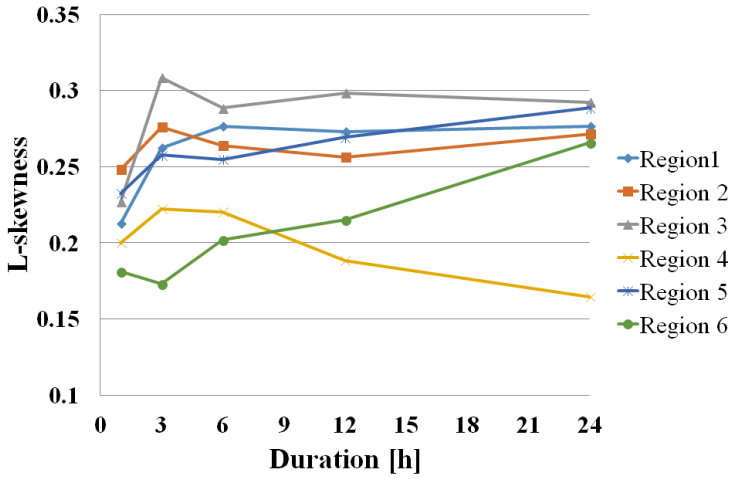
An analysis of the standard deviation of the L-moments values has been carried out comparing cases in which 1 h data is present and absent. The standard deviation of the L-moments with hourly data resulted almost always higher than that relative to L-moments values without the one-hour duration. This result supports the decision do not consider the shortest duration (i.e. 1 hour).

The parameters of the distributions have been evaluated through *hierarchical procedure* considering the mean value for each L-moment not including the values for the duration of 1h.

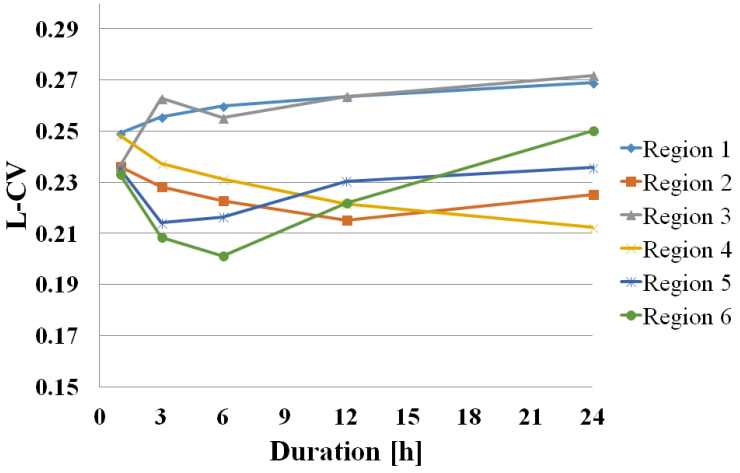
Fiorentino et al. (1987) and Gabriele and Arnell (1991) proposed a procedure that involved a hierarchy of regions. Relatively large regions were defined, where the shape parameters (related to the L-skewness for the LN3 and GEV) are assumed to be constant, and these regions are subdivided into smaller subregions over which the dispersion parameter is assumed to be constant.



(a) L-kurtosis



(b) L-skewness



(c) L-CV

FIGURE 5.5: Variation of the L-kurtosis (a), the L-skewness (b) for the different regions and durations and the L-CV (c) for the different regions and durations.

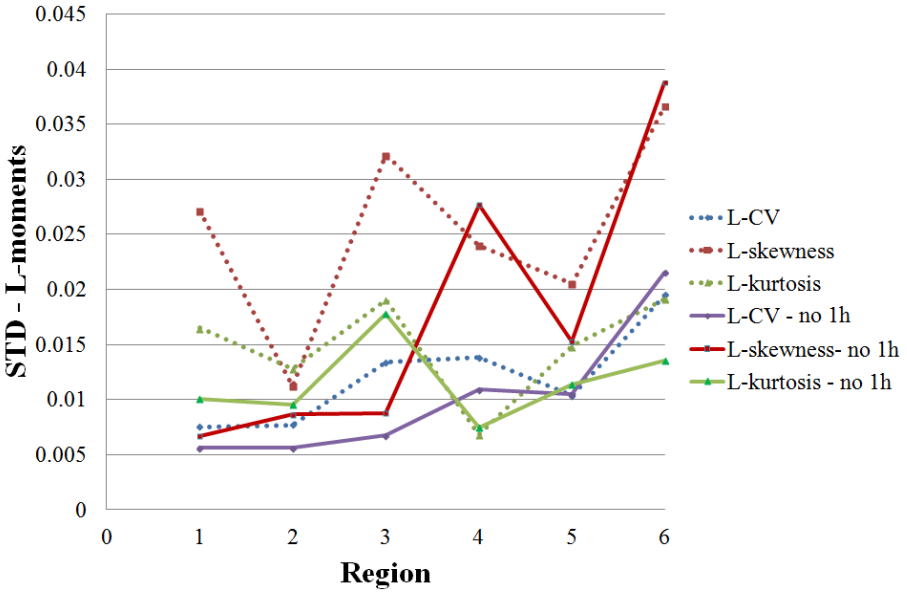


FIGURE 5.6: Standard deviation for the L-moments with and without for duration of 1h.

Lo Conti et al. (2007) carried out the regional analysis on three hierarchical levels estimating first the parameters that depend on higher order moment for the station inside a wide homogeneous region and progressively that depends on of the lower order moment, being these calculated for smallest homogeneous regions.

The hierarchical approach has been used in this work for the estimation of parameters for the selected distribution.

The first regional level considered is the entire Sicily, evaluated through the Hosking and Wallis test value of HW_3. The measure based on HW_3 should be an appropriated tool for assessing the heterogeneity of proposed regions when using the hierarchical approach (Hosking and Wallis, 2005). The value of HW_3 suggests that this region can be considered homogeneous respect to the value of L-skewness and L-kurtosis and then homogeneous for the first level of the regional analysis. This result coincides with that obtained from the two previous works.

Parameters of distributions depending on the values of L-kurtosis and L-skewness were assessed with the data of the entire Sicily, without considering the dependence on the duration and excluding the L-moments relative to hourly data.

For the second regional level of the analysis, the subdivision in six sub-regions described in the previous paragraph has been used to estimate the parameters of the distributions. In this case the parameters of the distributions linked to L-CV have been evaluated considering the homogeneity in every region; also in this case,

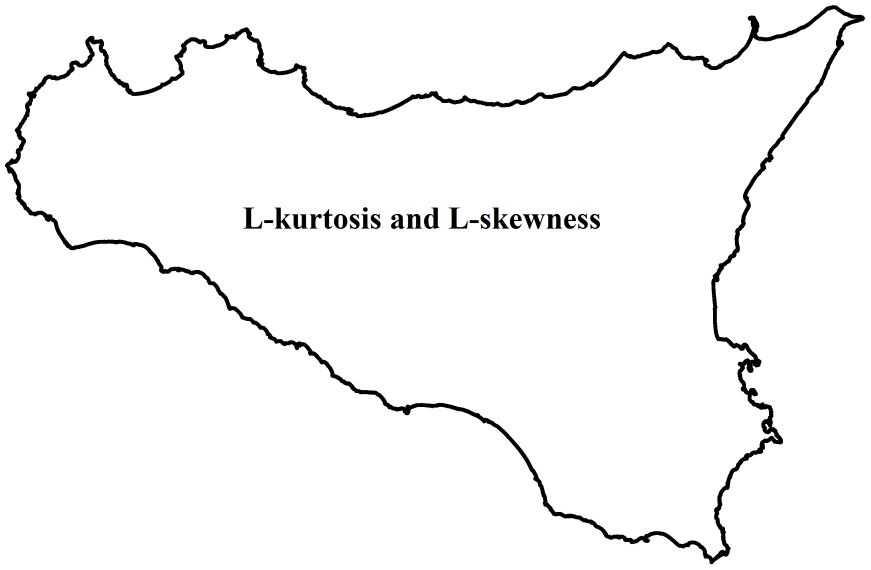


FIGURE 5.7: First level of regionalization with L-kurtosis and L-skewness homogeneous for all over Sicily.

the mean values of L-CV were calculated by excluding the hourly data and not considering the dependence on the duration.

In Table 5.7, 5.8, 5.9 the parameters of the distributions are listed. The parameters of the first level of regionalization are constant over Sicily; in the case of the TCEV are constant the first two parameters related to the L-kurtosis and L-skewness.

The Figure 5.9 shows the growth curves for each distribution and for each region; these curves are practically similar for return period lower than 100 years, while for greater return period, the distributions provide different values. In particular the GEV distribution provides the highest quantiles while the TCEV gives the lowest quantiles.

TABLE 5.7: Parameters of LN3

	Region					
	1	**2**	**3**	**4**	**5**	**6**
κ			-0.477			
ξ	0.893	0.909	0.892	0.907	0.909	0.910
α	0.425	0.362	0.428	0.367	0.360	0.357

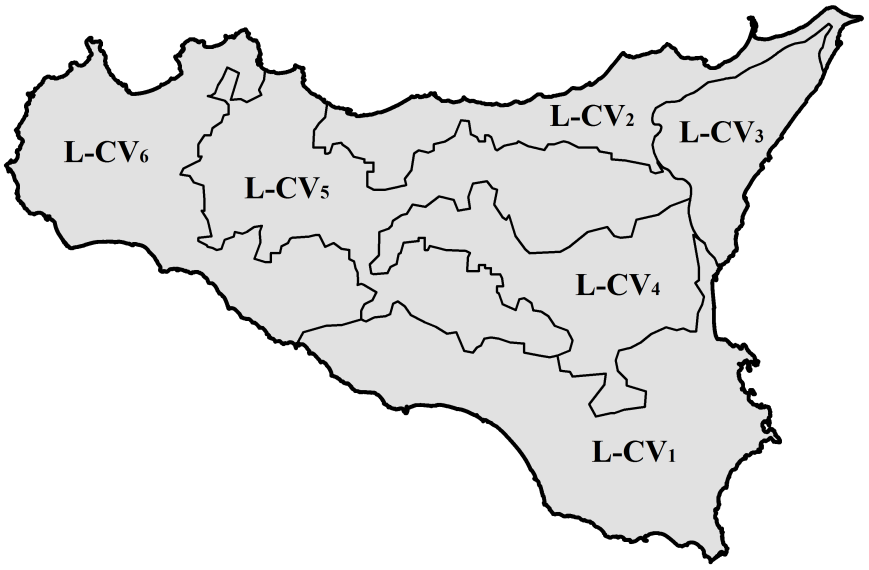


FIGURE 5.8: Second level of regionalization with L-CV homogeneous for each region.

A comparison with previous work of that Lo Conti et al. (2007) shows that the shape parameter for the GEV distribution resulted $\kappa = -0.1323$, while in this work is resulted equal to $\kappa = -0.091$. Indeed, in the case of the TCEV distribution, the parameters Λ^* and Θ^* result different from those estimated by Lo Conti et al. (2007) parameters, that were equal to 0.71 and 2.24 respectively.

The last step of the hierarchical procedure aims to determine a regional criterion to estimate $\mu_{i,d}$ for ungauged sites or for sites with short records.

The practical application of the results obtained from the RFA, requires the construction of the DDF (Depth Duration Frequency) curve, that will be obtained by multiplying the value of $h(T)$ (i.e. the growth curves), by average mean value of precipitation relative to different duration, $\mu_{i,d}$, for each station.

TABLE 5.8: Parameters of GEV

	Region					
	1	**2**	**3**	**4**	**5**	**6**
κ			-0.091			
ξ	0.765	0.800	0.764	0.797	0.801	0.803
α	0.347	0.296	0.349	0.300	0.294	0.292

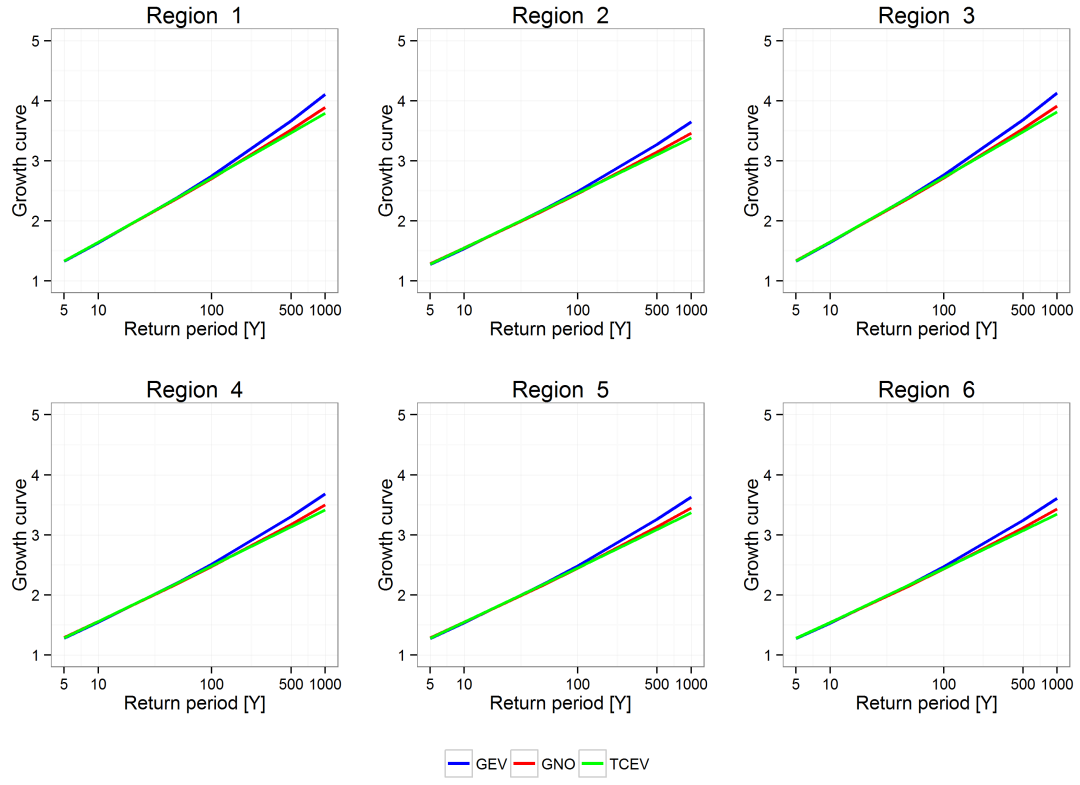


FIGURE 5.9: Value of growth curves for different return period

For the estimation of DDF in sites with measuring station, the values of $\mu_{i,d}$ can be set equal to the sample estimates m_d obtained from recorded data. In general, for each recording raingauge, m_d can be obtained as a function of duration d according to the power law relationship:

$$m_d = a \cdot d^n \tag{5.1}$$

in which a and n are site specific parameters.

TABLE 5.9: Parameters of TCEV

	Region					
Region	**1**	**2**	**3**	**4**	**5**	**6**
Λ^*	1.05					
Θ^*	1.98					
λ_1	9.122	18.974	8.865	17.691	19.494	20.305
α	4.214	4.946	4.185	4.876	4.973	5.014

It is possible to proceed to the spatialization of the values of a and n in order to have their value also in ungauged sites.

Previous works i Sicily have shown that the elevation can not be employed as a unique predictor variable of the variation for the parameter a, while the index n has shown a good correlation with elevation. Other parameters, especially the longitude, latitude, the concavity index and the distance from the sea, seem to have a significant influence on the variability of heavy rains. Moreover, the authors have obtained a estimate of the variable n better than parameter a. Therefore, the maps of values were obtained through a spatial interpolation by *Universal Kriging* such maps are shown in Figure 5.10.

5.2.6 Assessments of the accuracy

The comparison of the growth curve for the different distributions, shown in Figure 5.9, does not show substantial dissimilarity, especially for the lower return periods. The accuracy procedure applied in this work was described in the Section 3.6.

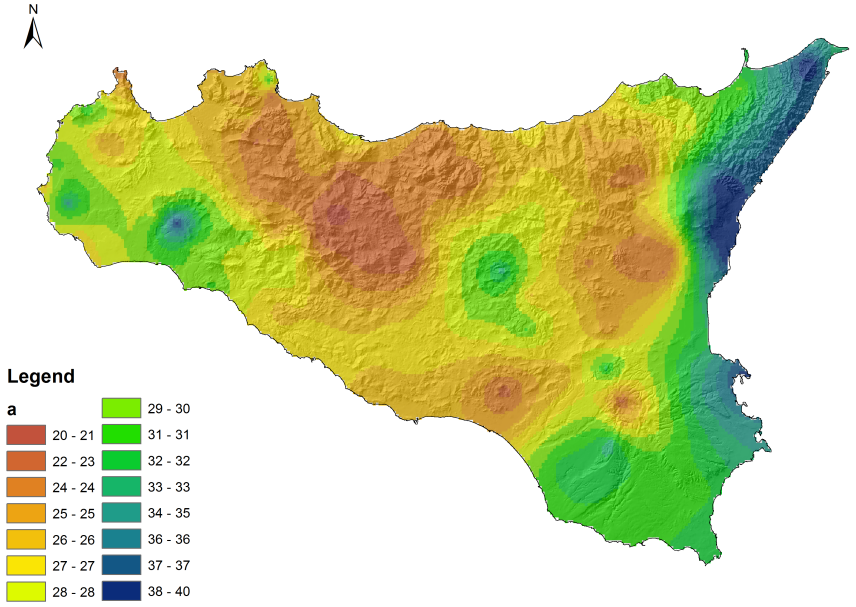
The probability distributions at the sites, fitted through L-moments method, are used to extract a random sample from these frequency distributions considering the effective record lengths. The regional frequency distribution specified is then fitted to the simulated data. The estimated dimensionless quantiles with the growth curves, obtained from random data are compared, and accuracy measures are calculated for the estimators. A number of 10,000 realizations (M) for each region have been carried out and the regional L-moment algorithm was used to fit the three different distributions to the data generated at each realization.

For each region, the accuracy was assessed using the relative BIAS and relative RMSE values for the different probability, considering each probability distribution.

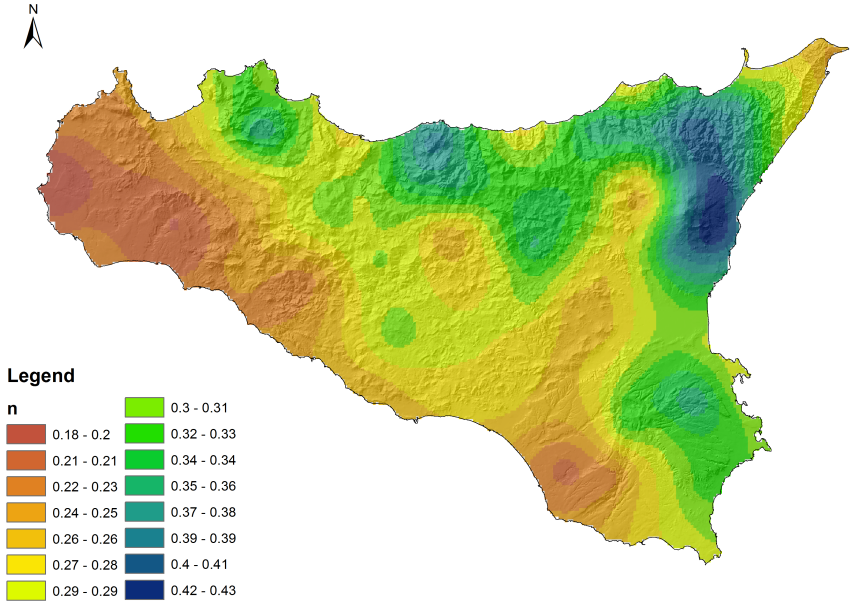
The *regional average relative bias*, $BIAS^R(F)$, calculated by Equation (3.30), measures the tendency of the simulated dimensionless quantiles to be uniformly higher or lower than the whole region.

The *regional average relative RMSE*, $RMSE^R(F)$, calculated by Equation (3.31), measures the overall deviation of the simulated dimensionless quantiles from the true quantiles.

The values of the relative BIAS (Figure 5.11) resulted negative for the LN3 and GEV distributions and near zero or positive for the TCEV distribution that shows the lowest absolute values.



(a) Parameter a



(b) Parameter n

FIGURE 5.10: Map of values of the parameters a and n

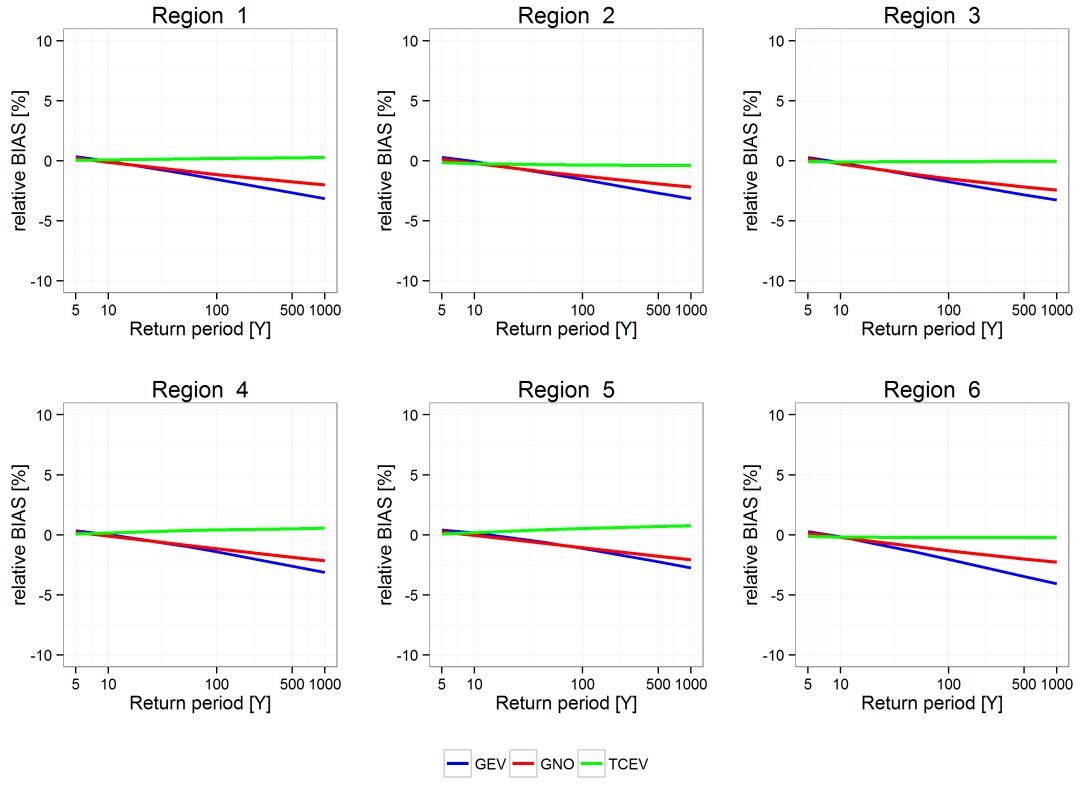


FIGURE 5.11: The relative BIAS (%) value for the different distribution

The LN3 and GEV distributions have produced relative RMSE values higher than TCEV (Figure 5.12). For high return periods, TCEV provides the best result; the reason of this behaviour could be due to the structure of the distribution given by the union of two EV1.

For high return periods, the difference among the TCEV with the other distributions is clear with the exception for the region 4 and 5, where the values are similar. The region 3 has shown the worst performance, probably due to the small number of stations inside the region. The region 6, despite being identified as "possibly heterogeneous", does not show RMSE and BIAS values significantly greater than those relative to the other regions.

As previously said, the values of the L-moment ratio for the hourly data have not been considered during the process of the evaluation of the parameters of the regional distributions. In order to verify the choice to exclude the L-moment for hourly duration, a sample of 1000 L-moments ratio values were extracted randomly from the hourly values, and by the Monte Carlo technique for each distribution dimensionless quantile values were generated.

Figure 5.13, 5.14 and 5.15 show, for each distribution, that the hourly CDF of

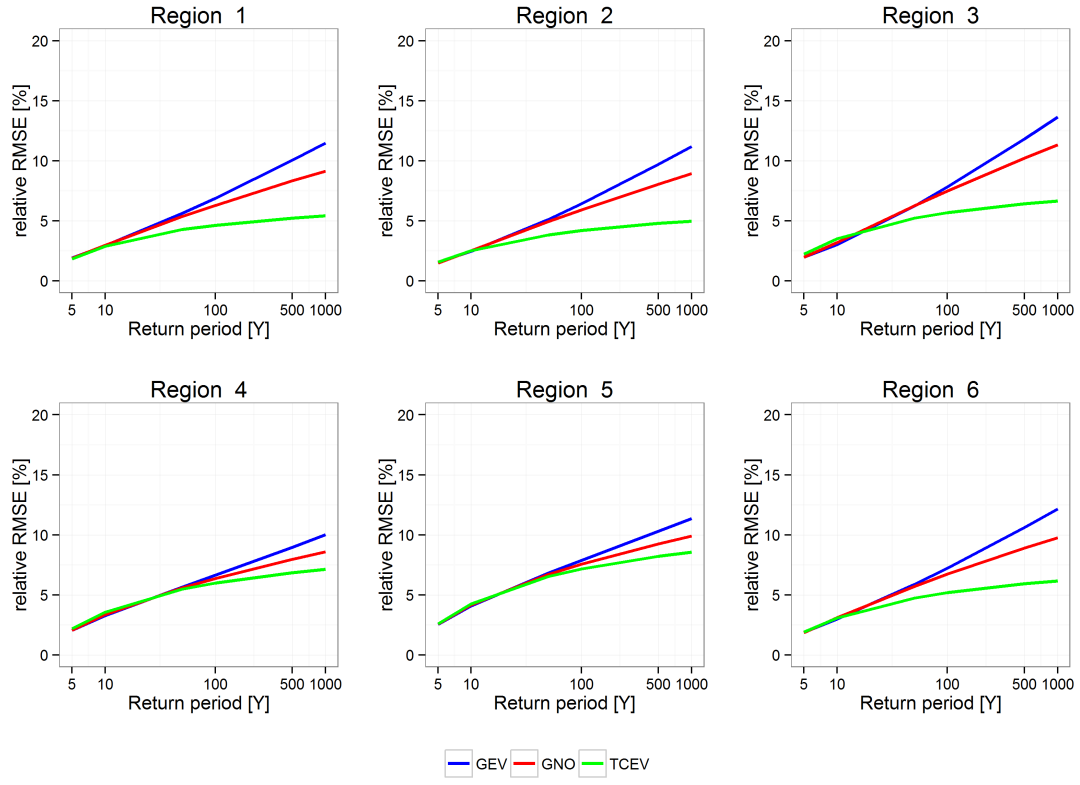


FIGURE 5.12: The relative RMSE (%) value for the different distribution

generated dimensionless quantiles follows the trend of the regional CDF beforehand evaluated without considering the hourly data.

The CDF of the quantiles for the LN3 (Figure 5.13) and TCEV (Figure 5.15) distributions show a limited dispersion around the regional CDF, while the CDF for the GEV (Figure 5.14) shows a scattering higher than other distributions, providing generally the worst performance. Even for the shortest duration, the LN3 and the TCEV achieved the best performance.

The results obtained confirm that the use of the regional growth curve obtained, disregarding the value for 1h, is reliable also for the lowest duration and consequently the accuracy of estimate uncertainty can be comparable with those regional curves obtained neglecting the hourly data.

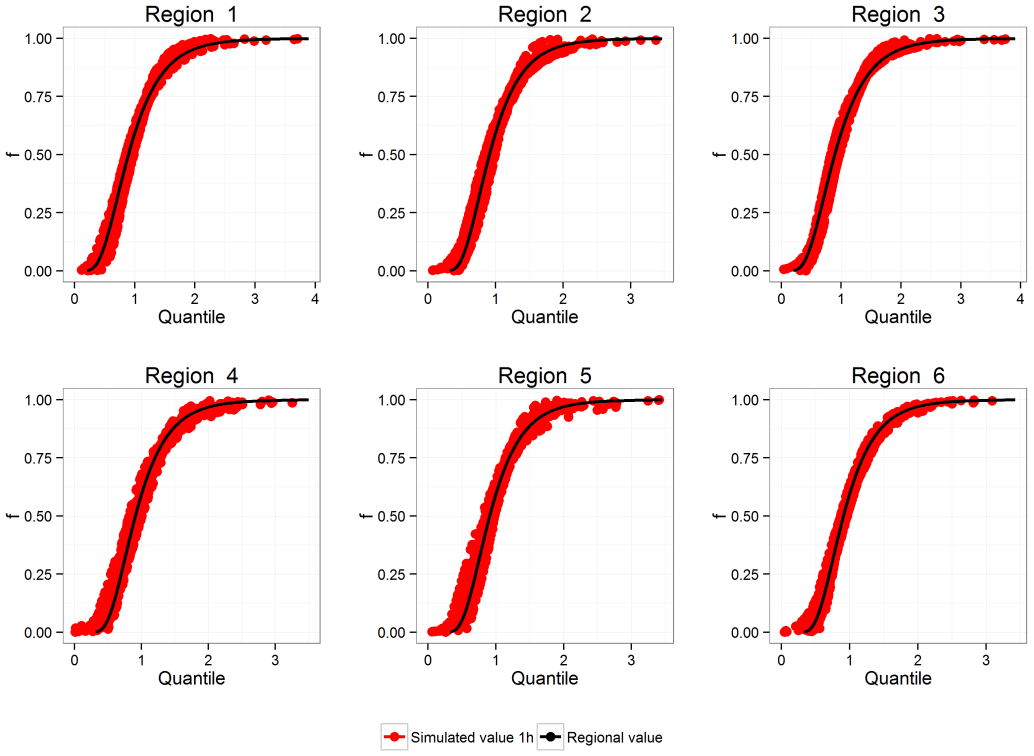


FIGURE 5.13: The hourly CDF of the dimensionless quantiles for LN3 distribution

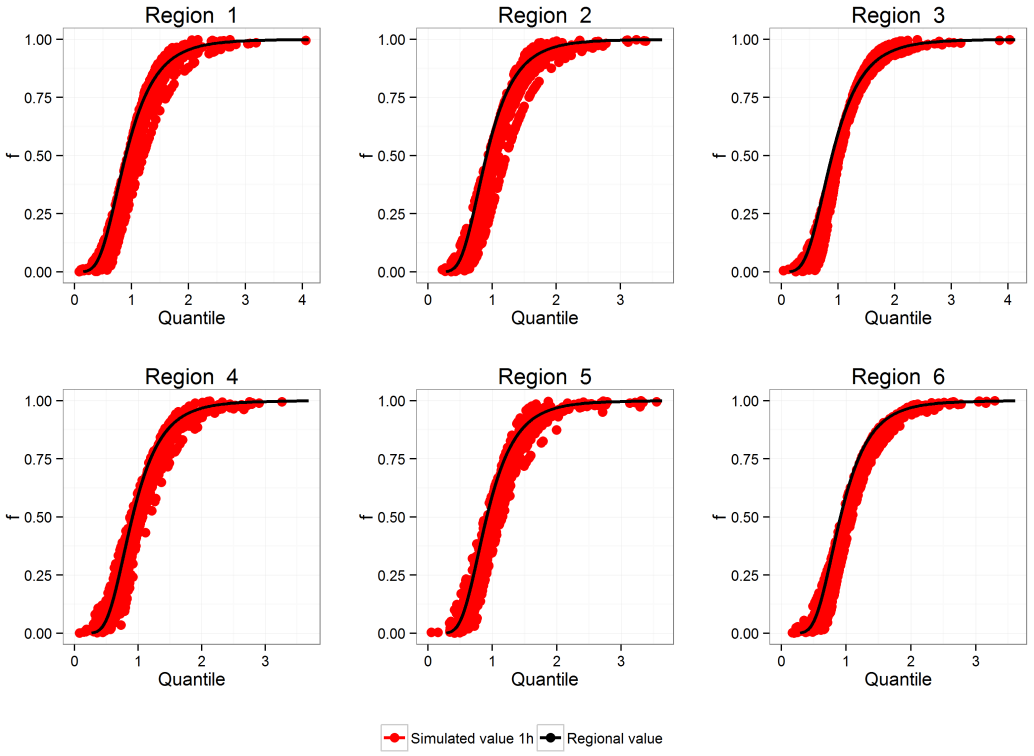


FIGURE 5.14: The hourly CDF of the dimensionless quantiles for GEV distribution

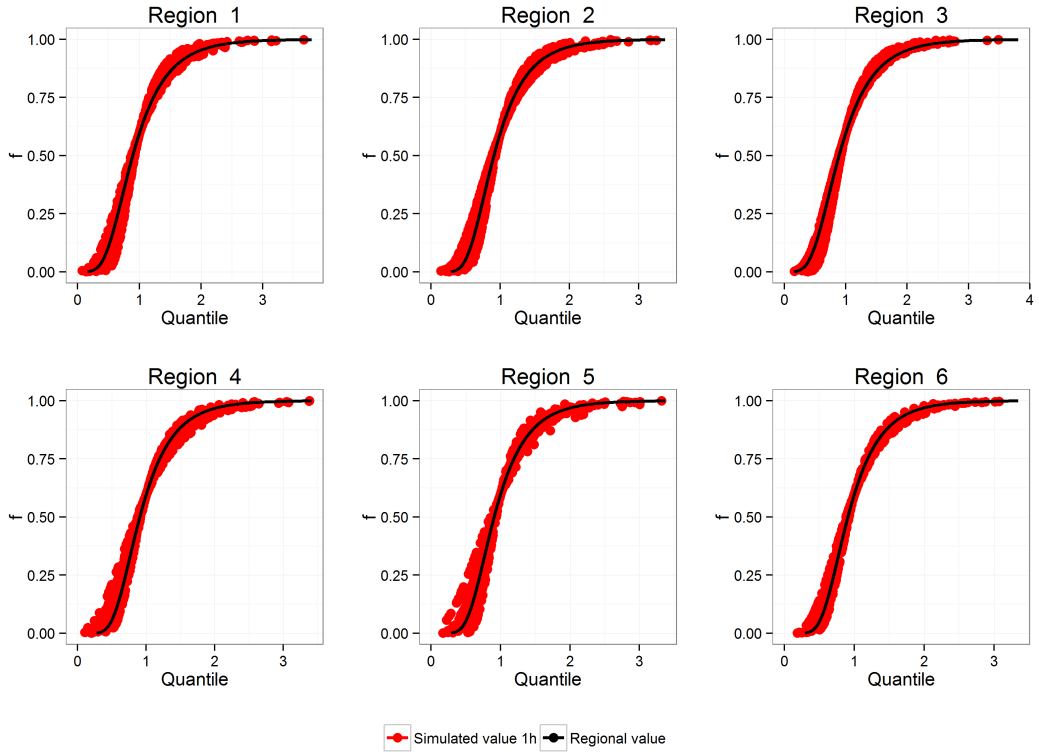


FIGURE 5.15: The hourly CDF of the dimensionless quantiles for TCEV distribution

Chapter 6

Climate change on the extreme precipitation

Global Climate Models (GCMs) are the most appropriate and powerful tool for understanding the behaviour of the global climate system over the coming centuries. They typically have grid-cells with dimensions 150-300 km or greater, so are only able to provide climate change information on large spatial scales. Thus, although GCMs can provide useful information about possible future changes in atmospheric circulation at the regional (e.g. continental) scale, they do not provide the detail required for regional and national assessments. This is particularly true for heterogeneous regions, where sub-GCM grid-scale variations in topography, vegetation, soils, and coastlines can strongly affect climate. In addition, extreme events, such as heavy precipitation, are often not captured or their intensity is unrealistically low at coarse resolutions.

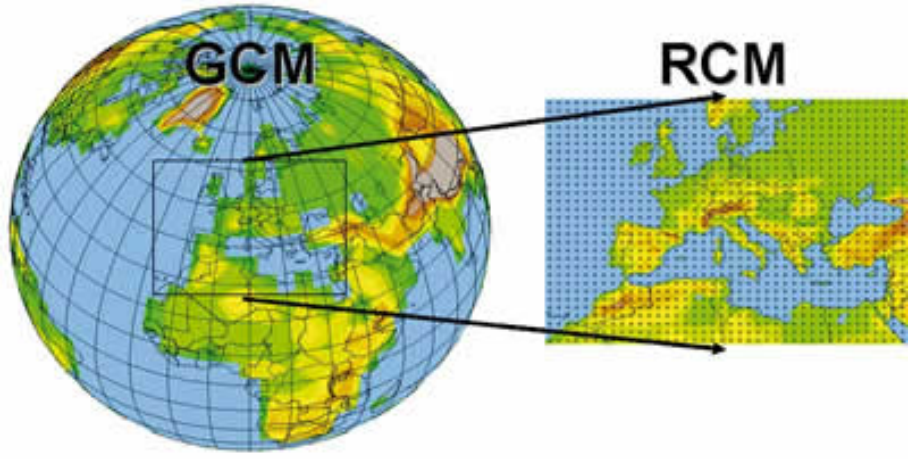


FIGURE 6.1: GCM and RCM

Climate change information at higher resolution can be derived by nesting *Regional Climate Models* (RCMs), which cover a limited area at a higher resolution (50 km or finer) than GCMs (Figure 6.1). This methodology is widely used to provide realistic spatial and temporal detail on how the climate may change locally. However, such future climate projections, at the local scale, are not necessarily reliable because uncertainty in the representation of processes in the GCMs and RCMs often leads to a range of projected changes, the likelihood of which often cannot be established (Déqué et al., 2007).

In this chapter, the possible effects of the climate change on the extreme precipitation will be studied through the use of the regional climate models that provide an indication of the possible climate evolution until the end of the 21th century. A procedure of bias correction will be implemented and used to correct daily RCMs data. Then a temporal downscaling method will be carried out to obtain a sub-daily values of precipitation. Finally, the final results will be discussed and analysis of the future trend of rainfall quantiles for different duration will be achieved.

6.1 A short history of climate models

Climate models have a lot in common with models that are used to predict the weather. In fact, they both have the same roots. Numerical modelling of the atmosphere had been envision as early as the early 20th Century.

In 1904, the Norwegian meteorologist Vilhelm Bjerknes first proposed the possibility of the numerical prediction of weather if the initial state and the physical laws were known accurately. Then, the English scientist Lewis Fry Richardson made a weather prediction using equations describing the physics of the atmosphere that he calculated by hand.

In 1922, Richardson explained his forecast in his book Weather Prediction by Numerical Process. Unfortunately, his forecasts were horribly incorrect, because the observations that he used for the initial conditions were not very reliable. Also, the equations he used were too complex, allowing atmospheric waves of all kinds including sound waves. These high frequency waves grew to be very large.

Later in the 1930s, Carl Gustav Rossby discovered this fatal mistake and reconfigured the equations to filter out these high frequency waves (Washington and Parkinson, 2005).

With the development of modern computers in the late 1940s, the idea of direct numerical modelling of the atmosphere could be revisited. At Princeton University's *Institute for Advanced Studies*, John von Neumann supervised the construction of one of these early computers, and he realized the potential of using it for weather forecasting. He subsequently established a team of scientist led by Jule Charney to develop a numerical weather prediction model. This team of scientists used Rossby's simplified equations (Washington and Parkinson, 2005). By this time, there were better data. What was missing from the set of observational data at Richardson's time were data from above the surface. By the 1940s, there were regular upper-air soundings made over land (Weart, 2011).

However, the first models had to be two-dimensional and regional for weather prediction purposes for the rudimentary computers of the time. Norman Phillips at the University of Chicago took a step towards global climate modelling Inspired by his sink experiments of features that resembled weather in a rotating pan of water, he developed a two-layer model on a cylinder instead of a sphere that produced features that resembled a jet stream and weather systems (Weart, 2011).

Encouraged by Phillips's results, Joseph Smagorinsky at the *U.S. Weather Bureau*, the predecessor to the *National Weather Service*, established a team to develop a *general circulation model* (GCM), a global three-dimensional model of the atmosphere. A key member of this team was Syukuro "Suki" Manabe. Smagorinsky and Manabe developed a nine-layer model that was the first to include physical processes as well as moisture fluxes from a global damp surface (Weart, 2011). This group grew to become the *Geophysical Fluid Dynamics Laboratory* now housed at Princeton University.

Another group developing a GCM at about the same time was Yale Mintz's group at the University of *California-Los Angeles* (UCLA). Mintz recruited Akio Arakawa to help in the development of numerical schemes for a GCM. One of those schemes was a staggered vertical grid to resolve complications that develop when calculating all quantities at the same grid points. Together, Mintz and Arakawa developed a two-layer model with separate land and ocean surfaces (Weart, 2011).

With the advancement of computers, the GCMs became increasingly more complex with the inclusion of more processes and even a return to the original equations that Richardson used. Over the years, separate models began to be developed for the oceans, land surface, and sea ice that were eventually coupled to the atmospheric model for more accurate simulations of the whole Earth system (or as close to the whole system as possible). Also, more and more groups started

developing their own GCMs first in the U.S. and then in other countries, but many of the later models are really offshoots from earlier models.

6.2 Global climate model (GCM)

Global Climate Models (GCMs) have evolved from the *Atmospheric General Circulation Models* (AGCMs) widely used for daily weather prediction. GCMs have been used for a range of applications, including investigating interactions between processes of the climate system, simulating evolution of the climate system, and providing projections of future climate states under scenarios that might alter the evolution of the climate system. The most widely recognized application is the projection of future climate states under various scenarios of increasing atmospheric carbon dioxide (CO_2).

At the core of a GCM is an *Atmospheric General Circulation Model* (AGCM) that dynamically simulates the circulation of the atmosphere, including the many processes that regulate energy transport and exchange by and within the atmospheric flow. The basic atmospheric flow is represented by fundamental equations that link the mass distribution and the wind field. These equations are represented on a spherically spatial grid field that has many levels representing the depth of the atmosphere.

The flow equations are modified by the representation of processes that occur on a scale below that of the grid-including such processes as turbulence, latent heat of condensation in cloud formation, and dynamic heating as solar and infrared radiation interact with atmospheric gases, aerosols, and clouds.

The oceans are at least as important as the atmosphere for the transport of energy. For that reason, the GCM also includes an *Ocean General Circulation Model* (OGCM) that simulates the circulation of the oceans. The OGCM is vital for climate simulations because the oceans represent a dynamic thermal reservoir that, through energy exchange with the atmosphere, dominates the evolution of the climate system. The specification of the processes that regulate heat, moisture, and momentum exchanges between the ocean and atmosphere is crucial to the integrity of a GCM.

Land surface, and how soil moisture and vegetation type regulate heat, moisture, and momentum with the atmosphere, plays a lesser but nevertheless important role in the simulation of climate. Soil moisture and vegetation respond to local precipitation and affect the exchange of heat, moisture, and momentum with the

atmosphere over time. The soil moisture and vegetation (and their regulation of land-atmosphere exchange processes) respond to the climate on the shorter time-scale of weather systems but, due to the varying accumulation of soil moisture, the influence of land surface on climate is on seasonal and interannual time-scales.

Surface ice sheets also have an important role in the evolution of the climate system. Their formation and expansion represent a lowering of the total energy of the climate system as a whole because latent heat is lost as water changes from the liquid to solid phase. Likewise, contraction of surface ice sheets represents an increase in the total energy of the climate system.

A primary function of the climate system is to transport energy from the tropics to higher latitudes; globally, there is an equilibrium between solar radiation absorption and infrared radiation loss to space. Of course, with such a complex system there is rarely perfect balance. At times, especially during the cycle of seasons, Earth is accumulating radiation energy and warming, whereas at other times it is losing energy and cooling. But the rate of radiation loss varies with temperature and acts as a natural thermostat: when Earth warms, the infrared radiation loss to space increases such that it exceeds the solar input and warming ceases; when Earth cools, the infrared radiation loss to space decreases such that solar radiation exceeds the infrared radiation loss and cooling ceases.

In general, clouds and their interaction with the climate system are difficult to model. Clouds are an outcome of vertical motion and saturation, but the feedback to the circulation through radiation processes is sensitive. Although cloud fields tend to be regulated by the larger scale circulation, the processes leading to cloud formation and dissipation are operating on scales very much smaller than that of the computation grid, with individual clouds often occupying only a small part of a grid. Thus it is necessary for models to specify the climate interaction of a multitude of differing clouds across a grid space by a single process.

6.3 Regional Climate Models (RCMs)

As previously said, a key limitation of GCMs is the fairly coarse horizontal resolution. For the practical planning of local issues such as water resources or flood defences, countries require information on a much more local scale than GCMs are able to provide. Regional models provide one solution to this problem.

Regional Climate Models (RCMs) work by increasing the resolution of the GCM in a small, limited area of interest. An RCM might cover an area the size of

western Europe, or southern Africa - typically 5000km x 5000km. The full GCM determines the very large scale effects of changing greenhouse gas concentrations and volcanic eruptions on global climate. The climate calculated by the GCM is used as input at the edges of the RCM for factors such as temperature and wind. RCMs can then resolve the local impacts given small scale information about orography (land height) and land use, giving weather and climate information at resolutions as fine as 50 or 25km.

In regions where the land surface is flat for thousands of kilometres, and there is no ocean anywhere near, the coarse resolution of a GCM may be enough to accurately simulate weather changes. However, most land areas have mountains, coastlines and changing vegetation characteristics on much smaller scales, and RCMs can represent the effects of these on the weather much better than GCMs.

6.3.1 Nested regional climate modelling

The nested regional climate modelling technique consists of using initial conditions, time-dependent lateral meteorological conditions and surface boundary conditions to drive high-resolution RCMs. The driving data is derived from GCMs (or analyses of observations) and can include *Greenhouse gas* (GHG) and aerosol forcing. A variation of this technique is to also force the large-scale component of the RCM solution throughout the entire domain (e.g., Cocke and LaRow (2000), KIDA et al. (1991), von Storch et al. (2000))

To date, this technique has been used only in one-way mode, i.e., with no feedback from the RCM simulation to the driving GCM. The basic strategy is, thus, to use the global model to simulate the response of the global circulation to large-scale forcing and the RCM to (a) account for sub-GCM grid scale forcing (e.g., complex topographical features and land cover inhomogeneity) in a physically-based way; and (b) enhance the simulation of atmospheric circulations and climatic variables at fine spatial scales.

The nested regional modelling technique essentially originated from numerical weather prediction, and the use of RCMs for climate application was pioneered by Dickinson et al. (1989) and Giorgi (1990). RCMs are now used in a wide range of climate applications, from palaeoclimatology (Hostetler, 1994, Hostetler et al., 1993) to anthropogenic climate change studies. They can provide high resolution (up to 10 to 20 km or less) and multi-decadal simulations and are capable of describing climate feedback mechanisms acting at the regional scale.

A number of widely used limited area modelling systems have been adapted to, or developed for, climate application. More recently, RCMs have begun to couple atmospheric models with other climate process models, such as hydrology, ocean, sea-ice, chemistry/aerosol and land-biosphere models.

Two main theoretical limitations of this technique are the effects of systematic errors in the driving fields provided by global models; and lack of two-way interactions between regional and global climate.

Practically, for a given application, consideration needs to be given to the choice of physics parametrizations, model domain size and resolution, technique for assimilation of large-scale meteorological conditions, and internal variability due to non-linear dynamics not associated with the boundary forcing (e.g., Giorgi and Mearns (1991, 1999), Ji and Vernekar (1997). Depending on the domain size and resolution, RCM simulations can be computationally demanding, which has limited the length of many experiments to date.

Finally, GCM fields are not routinely stored at high temporal frequency (6-hourly or higher), as required for RCM boundary conditions, and thus careful coordination between global and regional modellers is needed in order to perform RCM experiments.

6.4 CORDEX project

The nested regional climate modelling is the technique used in the *Coordinated Regional Climate Downscaling Experiment* (CORDEX) project which essentially has the twofold purpose of providing a framework to evaluate and benchmark model performance (*model evaluation framework*), and design a set of experiments to produce climate projections for use in impact and adaptation studies (*climate projection framework*).

The choice of common *regional climate downscaling* (RCD) domains is a prerequisite for the development of the model evaluation and climate projection frameworks. The goal of *CORDEX* is to provide a framework accessible to a broad scientific community with maximum use of results. *CORDEX* domains therefore encompass the majority of land areas of the world.

Figure 6.2 shows a first selection of common domain, where these should be interpreted as interior analysis domains, e.g. not including the lateral relaxation zone in RCMs. This selection is based partly on physical considerations (i.e. inclusion of processes important for different regions), partly on considerations

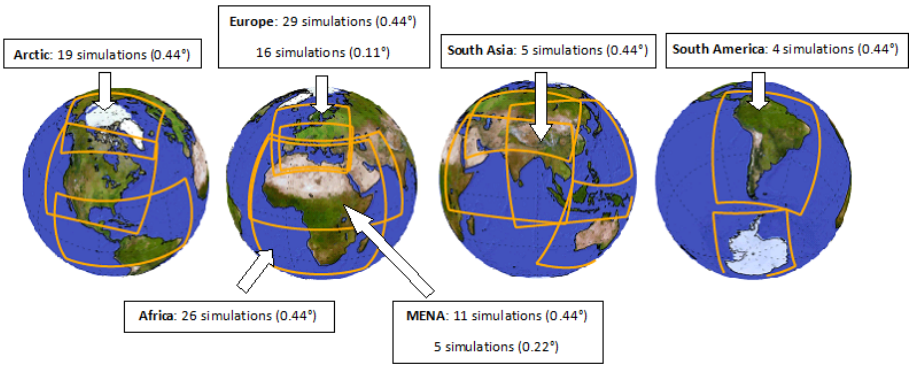


FIGURE 6.2: CORDEX simulations

of resources needed for the simulations and partly on the availability of ongoing programmes.

In order to allow wide participation the broader community decided to make the standard horizontal resolution for the first phase *CORDEX* simulations to be $\sim$ 50 km (or 0.5 degrees). Today, many groups are running RCMs with considerably higher resolution than this (up to $\sim$ 10 km) and they are encouraged to explore the benefits of increased RCM resolution within the *CORDEX* framework. Nevertheless, it was felt that a standard resolution, allowing contribution by many groups, would increase the sense of community ownership of the *CORDEX* project, while also increasing the size of any ensuing RCM scenario set for analysis and comparison purposes.

6.4.1 Model evaluation framework

The climate projection framework within CORDEX is based on the set of new global model simulations planned in support of the *IPCC Fifth Assessment Report* (referred to as CMIP5). This set of simulations includes a large number of experiments, ranging from new greenhouse-gas scenario simulations for the 21^{st} century, decade prediction experiments, experiments including the carbon cycle and experiments aimed at investigating individual feedback mechanisms (Taylor et al., 2009).

For its initial activities, CORDEX will focus on the scenario simulations. Different from the scenario runs employed in the fourth IPCC assessment cycle, which were based on the SRES GHG emission scenarios (IPCC, 2000), this next generation of scenario simulations is based on so-called *reference concentration pathways* (RCPs), i.e. prescribed greenhouse-gas concentration pathways throughout the 21^{st} century, corresponding to different radiative forcing stabilization levels by the

year 2100. Four RCPs have been selected, with stabilization levels at 2.9, 4.5, 8.5 and 11.2 W/m^2 (referred to as RCP2.9, RCP4.5, RCP8.5 and RCP11.2, respectively). Within CMIP5, the highest-priority global model simulations have been selected to be the RCP4.5 and RCP8.5, roughly corresponding to the IPCC SRES emission scenarios B1 and A1B, respectively. The same scenarios are therefore also planned to be the highest priority CORDEX simulations.

Ideally, all regional model simulations should span the period 1951-2100 in order to include a recent historical period, plus the entire 21^{st} century. For many groups, however, it may prove computationally too demanding to run CORDEX simulations for this entire time span. The 1951-2100 period has thus been divided into five 30-year time slices and participating groups are requested to simulate time slices in the following order of priority 1981-2010, 2041-2070, 2011-2040, 2071-2100, 1951-1980.

The first of these (1981-2010) represents the reference period for model evaluation and for the calculation of climate changes. The second priority time slice, covering a future time period, was selected as a compromise between the needs of the impact community in terms of future time horizon and the requirement to obtain a robust change signal. It is requested that all participating groups at a minimum perform these two time slices to have a reasonable set of simulations for analysis and intercomparison. In the initial phase of CORDEX, it is planned to simulate one realization for each RCP scenario selected, using driving data from multiple global models. In this way, CORDEX will explore the model configuration uncertainty but not the internal variability one. As mentioned above, this should not represent a major drawback, since previous experience has shown that the former is a much more important source of uncertainty when looking at long temporal scales.

6.4.2 EURO-CORDEX

As part of the global CORDEX framework, the EURO-CORDEX initiative provides regional climate projections for Europe at 50 km (EUR-44) and 12.5 km (EUR-11) resolution, thereby complementing coarser resolution data sets of former activities like, e.g., PRUDENCE and ENSEMBLES. The regional simulations are downscaling the new CMIP5 global climate projections (Taylor et al., 2012) and the new representative concentration pathways (RCPs) (Moss et al., 2010, Van Vuuren et al., 2011). Twenty-six modelling groups contributing 11 different

regional climate models, partly in different model configurations, actively support EURO-CORDEX.

In its initial phase, EURO-CORDEX mainly focussed on model evaluation in present-day climate (e.g., Vautard et al., 2013). So far more than 30 evaluation simulations have been conducted. Further activities include the coordinated analysis of future climate simulations, the joint analysis of dynamical and empirical–statistical methods and the design of suitable bias correction techniques to tailor EURO-CORDEX data for direct application in climate impact research. Particular emphasis is put on the construction of a simulation matrix that covers uncertainty in emission scenarios, the driving global climate model and the downscaling method in the best affordable manner.

Similarity to general CORDEX program, the EURO-CORDEX simulations consider the global climate simulations from the CMIP5 long-term experiments up to the year 2100. They are based on greenhouse gas emission scenarios (Moss et al., 2010, Nakicenovic and Swart, 2000, Van Vuuren et al., 2008).

6.4.3 Med-CORDEX

Med-CORDEX initiative has been proposed by the Mediterranean climate research community as a follow-up of previous and existing initiatives. Med-CORDEX takes advantage of new very high-resolution Regional Climate Models (RCM, up to 10 km) and of new fully coupled *Regional Climate System Models* (RCSMs), coupling the various components of the regional climate.

Med-CORDEX is a unique framework where research community will make use of both regional atmospheric, land surface, river and oceanic climate models and coupled regional climate system models for increasing the reliability of past and future regional climate information and understanding the processes that are responsible for the Mediterranean climate variability and trends.

The areas surrounding the Mediterranean basin have quite a unique character that results both from their complex morphology and socio-economic conditions. It is indeed surrounded by various and complex topography channelling regional winds (Mistral, Tramontane, Bora, Etesian, and Sirocco) than defined local climates and from which numerous rivers feed the Mediterranean sea. Many small-size islands limit the low-level air flow and its coastline is particularly complex. Strong land-sea contrast, land-atmosphere feedback, intense air-sea coupling and aerosol-radiation interaction are also among the regional characteristics to take into account when dealing the Mediterranean climate modelling.

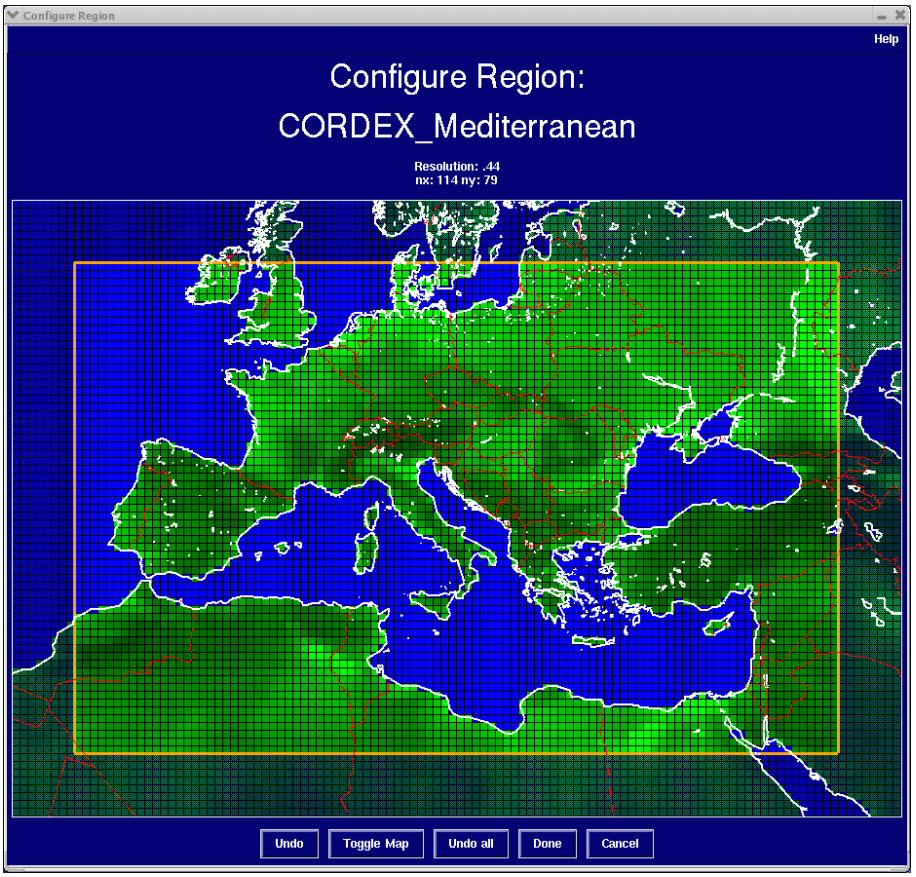


FIGURE 6.3: MED-CORDEX domain

In addition, the region features an enclosed sea with a very active regional thermohaline circulation. It is connected to the Atlantic Ocean only by Gibraltar strait and surrounded by very urbanized littorals. The Mediterranean region is consequently a good case study for climate regionalization and was naturally chosen as a CORDEX sub-domain (MED) leading to the Med-CORDEX initiative endorsed by Med-CLIVAR and HyMeX. The MED domain is defined in the CORDEX domain document as shown in the Figure 6.3.

6.5 Climate models

The EURO-CORDEX domain covers all countries in the European Union, but it does not map perfectly to the Europe region defined for the IPCC Fifth Assessment Report. For the eastern part of Turkey, unfortunately, no regional model projections are available.

A summary of the grid configuration and differences in the parametrization schemes for the participating regional models (ALADIN5.1: Colin et al., 2010, Hermann et al., 2011, CCLM: Rockel et al., 2008, HIRHAM: Christensen et al., 1998,

RACMO2: van Meijgaard et al., 2012, RCA4: Kupiainen et al., 2011, Samuelsson et al., 2011 REMO: Jacob et al., 2012, WRF Version 3.3.1: Skamarock et al., 2008) are described following and is given in the Table 6.4.

The RCP scenarios, the driving GCMs and the driven RCMs as well as the simulation length are listed in the Table 6.1. Seven different RCMs and five different GCMs have been supplied from the CORDEX project. Two of the RCMs were driven by five different GCMs; five GCM-RCM chains did simulate both RCP scenarios. They all provide data at least until the mid of the century. Eight RCP4.5 simulations and nine RCP8.5 simulations had reached the end of the century.

TABLE 6.1: Overview of the global and regional climate models for RCP4.5 and RCP8.5

GCM	GCM Member	RCM	Scenarios	Time
MPI-ESM-LR	r1i1p1	CCLM	RCP4.5/8.5	until 2100
CNRM-CM5-LR	r1i1p1	CCLM	RCP4.5	until 2100
EC-EARTH	r12i1p1	CCLM	RCP4.5	until 2100
HadGEM2-ES	r1i1p1	CCLM	RCP4.5	until 2100
CNRM-CM5-LR	r8i1p1	ALADIN V5.2	RCP4.5/8.5	until 2100
MPI-ESM-LR	r1i1p1	REMO	RCP4.5/8.5	until 2100
IPSL-CM5A-MR	r1i1p1	WRF331	RCP4.5	until 2100
EC-EARTH	r1i1p1	RACMO2	RCP4.5/8.5	until 2100
EC-EARTH	r12i1p1	RCA4	RCP8.5	until 2100
CNRM-CM5-LR	r1i1p1	RCA4	RCP8.5	until 2100
HadGEM2-ES	r1i1p1	RCA4	RCP8.5	until 2100
MPI-ESM-LR	r1i1p1	RCA4	RCP8.5	until 2100
IPSL-CM5A-MR	r1i1p1	RCA4	RCP8.5	until 2100
EC-EARTH	r3i1p1	DMI-HIRHAM	RCP4.5/8.5	until 2050

ALADIN-Climate. ALADIN-Climate can actually be considered as a version of ARPEGE-Climate since they share the same computer code.

The global model used, *Action de Recherche Petite Echelle Grande Echelle/ Integrated Forecasting System* (ARPEGE/IFS), is a spectral model developed for operational numerical weather forecast by *Météo-France* in collaboration with *European Centre for Medium-range Weather Forecast* (ECMWF). Its climate version has been developed in the 90s (Déqué et al., 1994). ARPEGE climate version has been used as the atmosphere part of the Météo-France earth modelling system

	REMO	RCA4	CCLM	RACMO2	WRF	ALADIN5.1	HIRHAM
Institution	CSC	SMHI	CLMCOM[1)]	KNMI	IPSL-INERIS	Météo-France	DMI
Grid resolution	0.11 º x 0.11º	0.11 º x 0.11º	0.11 º x 0.11º	0.11 º x 0.11º	0.11 º x 0.11º	0.11 º x 0.11º [2)]	0.11 º x 0.11º
Grid (lat*lon)	433 x 433	438 x 456	450 x 438	444 x 456	442 x 454	432 x 288	452 x 432
Rotation	lon -162° lat 39.25	lon -162° lat 39.25°	lon -162° lat 39.25°	lon -162° lat 39.25°	CORDEX specifications	0°E, 90°S	lon -162° lat 39.25°
Vertical levels	27	40	40	40	32	31	31
Boundary layer scheme	Louis 1979	Cuxart et al 2000	Louis 1979	Lenderink and Holtslag 2004; Siebesma et al. 2007	YSU, Hong et al. 2006	Ricard and Royer 1993	Louis 1979
Number of points (sponge zone)	8	10	12	8 (16)		8	
Convection	Mass flux Tiedtke 1989; Nordeng 1994 for CAPE closure; Pfeifer 2006	Kain and Fritsch 1990, 1993; Kain 2004; Jones and Sanchez 2002	Tiedtke 1989	Tiedtke 1989; Nordeng 1994; Neggers et al 2009	Grell and Devenyi 2002	Mass flux Bougeault 1985	Tiedtke 1989
Microphysics	Lohmann and Roeckner 1996	Rasch and Kristjánsson 1998	Doms et al. 2007; Baldauf and Schulz 2004	Tiedtke 1993; Tompkins et al, 2007; ECMWF-IFS 2007; Neggers 2009	Hong et al. 2004	Ricard and Royer 1993	Lohmann and Roeckner 1996
Radiation	Morcrette 1986; Giorgetta and Wild 1995	Savijärvi 1990; Sass et al. 1994	Ritter and Geleyn 1992	Fouquart and Bonnel 1980; Mlawer et al.1997	RRTMG, Lacono et al, 2008	Morcrette 1990	Morcrette et al 1986; Giorgetta and Wild 1995
Land surface scheme	Hagemann 2002; Rechid et al. 2009	Samuelsson et al. 2006	TERRA-ML Doms et al. 2007	Van den Hurk et al 2000; Balsamo et al. 2009	NOAH	Douville et al 2000	Hagemann 2002
Soil thermal layers	5	5	10	4	4	4	5
Soil moisture layers	1	3	8	4	4	2	1
Main references	Jacob et al. 2012	Samuelsson et al. 2011; Kupiainen et al. 2011	Rockel et al. 2008; http://www.cosmo-model.or	Meijgaard van et al. 2012	Version 3.3.1, Skamarock et al. 2008	Colin et al. 2010; Herrmann et al. 2011	Christensen et al. 1998

FIGURE 6.4: Summary of grid configurations and parametrizations for EURO-CORDEX models

(atmosphere, ocean, land-surface and sea-ice) for IPCC (2007). The full system is described by Salas-Mélia et al. (2005).

ALADIN-Climate is a bi-spectral RCM with a semi-implicit semi-lagrangian advection scheme. Its configuration includes a 11-point wide bi-periodization zone in addition to the more classical 8 point relaxation zone. This so-called extension zone allows the computation of the fast-Fourier transforms for the spectral-to-grid point space computation. More details can be found in Farda et al. (2010). In this version, the planetary boundary layer turbulence physics is based on Louis (1979) and the interpolation of the wind speed from the first layer of the model (about 30 m) to the 10 m height follows GELEYN (1988).

The version 5 used in the framework of the regional Med-CORDEX exercise and is close to the ARPEGE-Climate version used for the next CMIP5 exercise. ALADIN-Climate version 4 was used for the European ENSEMBLES project in which it was inter-compared with the state-of-the art of the European RCMs at 50 and 25 km (Boberg et al., 2010, Christensen et al., 2008, Sanchez-Gomez et al., 2009).

Different configurations of this model can be used. They are obtained by varying the spatial resolution, the size and position of the domain, and the dataset used for the large scale forcing. More complex and unusual options are also available as the use of spectral nudging and interactive airsea coupling techniques.

COSMO-CLM. The *Climate Limited-area Modelling-Community* (CLM) is an open international network of scientists, who accepted the CLM-Community agreement in all of its details. Members of the CLM-Community was applied and developed the COSMO-CLM or CCLM, which is the COSMO model in CLimate Mode.

The first version of the COSMO-CLM (named CLM) was developed by colleagues from HZG, PIK and BTU Cottbus on the basis of the Local Model (LM) version 3.1 (now COSMO model), originally developed by the German Meteorological Service.

The COSMO model is the nonhydrostatic operational weather prediction model applied and further developed by the national weather services joined in the *COnsortium for SMall scale MOdeling* (COSMO).

In 2005 CLM became the regional Community-Model for the German climate research. This model version has been applied on time scales up to centuries and spatial resolutions between 10 and 50 km in different regions of the world.

In 2007/08 both CLM and LM developments were unified to get a model version for regional climate modelling (COSMO-CLM) and operational weather forecast (COSMO). From this time, the idea of a uniform model version for weather and climate became a guiding principle of the model development.

The unified model version COSMO 5 for weather, climate, and environmental research has been released recently. An intensive evaluation on a climatological time scale is just being prepared.

HIRHAM. HIRHAM is a regional atmospheric climate model (RCM) based on a subset of the HIRLAM (Unden et al., 2002) and ECHAM models (Roeckner et al., 2003), combining the dynamics of the former model with the physical parametrization schemes of the latter.

The HIRLAM model - *High Resolution Limited Area Model* - is a numerical short-range weather forecasting system developed by the international HIRLAM Programme (http://hirlam.org) and is used for routine weather forecasting at a number of meteorological institutes, i.e. DMI (Denmark), FMI (Finland), IMS (Iceland), KNMI (The Netherlands), met.no (Norway), INM (Spain), and SMHI (Sweden).

The ECHAM global climate model (GCM) is a general atmospheric circulation model developed at the *Max Planck Institute of Meteorology* (MPI) in collaboration with external partners. The original HIRHAM model was collaboration between DMI, the *Royal Netherlands Meteorological Institute* (KNMI) and MPI.

The latest running version of the HIRHAM model system, version 4, was released in 1996 (Christensen et al., 1996). An increasing demand for the use of state-of-the-art physical parametrization schemes and higher model resolution, however, recently made it crucial that a major upgrade be performed. Initial work on this upgrade began in 2005 and commenced throughout 2006.

HIRHAM5 is based on release 7.0 of the HIRLAM model and release 5.2.02 of the ECHAM model.

Regional Atmospheric Climate Model (RACMO2). In the 1990s the KNMI developed in cooperation with the Danish Meteorological Institute the research model RACMO based on the *High Resolution Limited Area Model* (HIRLAM) numerical weather prediction model.

In 1993 UU/IMAU started to modify the model such that it better represented the extreme conditions over glacier surfaces. This first version of RACMO,

RACMO1, combined the dynamical core of the HIRLAM model with ECHAM4 physics. The polar modified version of RACMO1 was mainly applied to the Antarctic Ice Sheet.

The second version, RACMO2, combines the dynamical core of the HIRLAM model with the *European Centre for Medium-range Weather Forecasts* (ECMWF) *Integrated Forecast System* (ISF) physics. RACMO versions 2.0 and 2.1 included HIRLAM version 5.0.6 and ISF cycle CY23r4, while version 2.3 includes HIRLAM version 6.3.7 and cycle CY33r1. Due to the rapid increase in computer capacity over the years, these versions of RACMO have not only been applied to the Greenland and Antarctic Ice Sheets, but also at higher resolution to smaller areas such as Dronning Maud Land and Patagonia.

For the RACMO model in general the grids are defined over the equator and then rotated to the area of interest. Grid distance is defined in fraction of degrees, which results in near equidistant grid points as long as the domain is small enough. Note that the domain is thus not on a (polar) stereographic projection plane. In the vertical the model adopts a system of hybrid sigma levels, which evolve from terrain following sigma levels close to the surface to pure pressure levels at higher elevation. The actual number of horizontal grid points varies per model run; in most simulations 40 vertical layers were used.

Since RACMO is a regional model, it needs external information at the lateral boundaries and sea surface. At the lateral boundary zone of the model, the temperature, specific humidity, zonal and meridional wind components, and the surface pressure are relaxed towards the fields of a global model every 6 model hours, as are the sea surface temperature and sea ice concentration. RACMO is not forced at the model top. The interior of the model is not nudged towards observations and allowed to evolve freely.

Rossby Centre regional Atmospheric Climate model (RCA). RCA is a regional climate model for the atmosphere and its exchange with land surface. RCA has its origin from the numeric weather forecast model HIRLAM. A large part of the development has been in close cooperation with the HIRLAM-work including daily evaluation at the weather forecast service at SMHI and other meteorological institutes in Europe. Differences between RCA and HIRLAM regard especially the energy and water balances at the soil surface. The demand for good representation is far higher in long climate simulations preventing long-term drift in the climate.

Both HIRLAM and RCA are hydrostatic models performing calculations at a discrete grid net over an area.

The model runs founding the basis for the climate index calculations are made with a horisontal resolution of about 50 km and the area covers the larger parts of Europe and parts of the North Atlantic. To be able to make calculations with the regional model RCA, data on the atmospheric conditions outside the area and sea surface temperatures (in RCA2 also deep soil temperatures) has to be fed into the model. This information is delivered from a global climate model.

Different versions of RCA are used at the Rossby Centre. The latest version is RCA3 from 2004 (Kjellström et al., 2005). Earlier versions are RCA0 from 1998, RCA1 from 1999-2000 (Rummukainen et al., 2001) and RCA2 from 2002 (Samuelsson et al., 2011). Compared to RCA2, RCA3 has a completely new soil surface scheme as well as changes in the parametrisation of radiation, clouds, turbulence and precipitation processes.

Other models can be linked to RCA. Such an example is the *Rossby Centre ocean model* (RCO), describing sea and sea ice (Meier et al., 2003). When both models are coupled we talk about the RCAO model (Doscher et al., 2002).

REMO. The dynamical structure of REMO (REgional MOdel) is similar to the EM/DM system (Majewski, 1991). It is based on the primitive equations in a terrain-following hybrid coordinate system. The finite-difference equations are written in advective form on an Arakawa C-grid. Second order central differences in space are used and the vertical finite difference formulation follows Simmons and Burridge (1981) to conserve energy and angular momentum. To avoid numerical instabilities the vertical advection as well as the vertical turbulent fluxes are treated implicitly. The time-stepping is leap-frog with semi-implicit correction and Asselin-filter; the time step is 5 min.

The prognostic variables are surface pressure, horizontal wind components, temperature, specific humidity and cloud water. The vertical structure is the same as in the *Engineering Model* (EM) model with 20 model levels. The boundary conditions are specified at the top and the bottom of the model atmosphere, where the vertical velocity vanishes.

At the lateral boundaries, time-dependent values for all prognostic variables are specified from the global climate model ECHAM3. It is also possible to use model output from different models with different resolutions and to bring in data from analyses. A relaxation scheme according to Davies (1976) is used to adjust

the prognostic variables in a zone of 8 grid rows towards the prescribed boundary fields. The horizontal diffusion of momentum, temperature and moisture is fourth order with a space-dependent diffusion coefficient proportional to the total deformation of the horizontal wind field. The diffusion is performed on hybrid levels but correction terms are added in the temperature and moisture equation to account for the slope of the model layers.

Weather Research and Forecasting (WRF). The *Weather Research and Forecasting* (WRF) model is a *numerical weather prediction* (NWP) and atmospheric simulation system designed for both research and operational applications. WRF is supported as a common tool for the university/research and operational communities to promote closer ties between them and to address the needs of both. The development of WRF has been a multi-agency effort to build a next-generation mesoscale forecast model and data assimilation system to advance the understanding and prediction of mesoscale weather and accelerate the transfer of research advances into operations.

The *Advanced Research WRF* (ARW) is the ARW dynamics solver together with other components of the WRF system compatible with that solver and used in producing a simulation. Thus, it is a subset of the WRF modelling system that, in addition to the ARW solver, encompasses physics schemes, numerics/dynamics options, initialization routines, and a data assimilation package (WRF-Var).

The ARW solver shares the *WRF Software Framework* (WSF) with the *Non-hydrostatic Mesoscale Model* (NMM) solver and all other WRF components within the framework. Physics packages are largely shared by both the ARW and NMM solvers, although specific compatibility varies with the schemes considered. The association of a component of the WRF system with the ARW subset does not preclude it from being a component of WRF configurations involving the NMM solver. The major features of the ARW, Version 3, and reflects elements of WRF Version 3, which was first released in April 2008. This technical note focuses on the scientific and algorithmic approaches in the ARW, including the solver, physics options, initialization capabilities, boundary conditions, and grid-nesting techniques.

The climate data used in this study are shown in Table 6.2. In the first two models the driving model is the underlying global model. The models number 1 and 9 have shown anomalous values; in particular, these models are resulted unable to understand the seasonality of the precipitation, because of the abnormal season

cycle that is centred over Greece and Italy for these reason are not used in the next analysis.

6.6 Climate Extremes Indices

In order to achieve a preliminary evaluation of the performance of the RCMs selected in this study, a group of indices (listed in the following Table 6.3) have been used to evaluate the ability to model the extreme precipitation.

The *Expert Team on Climate Change Detection and Indices (ETCCDI)* has attempted to facilitate the analysis of such extremes over the last decade by defining a set of climate indices that provide a comprehensive overview of temperature and precipitation statistics focusing particularly on extreme aspects (Karl and Easterling, 1999, Klein Tank and Lenderink, 2009).

Most of the indices defined by the ETCCDI describe *moderate climate extremes* with recurrence times of a year or shorter, as compared to more extreme climate statistics such as 20 year return values of annual temperature and precipitation extremes as considered, for instance, in Kharin et al. (2007).

The indices are defined and described in detail in Klein Tank and Lenderink (2009) and Zhang et al. (2011). The indices fall roughly into four categories:

1. absolute indices, which describe, for instance, the hottest or coldest day of a year, or the annual maximum 1 day or 5 day precipitation rates;

2. threshold indices, which count the number of days when a fixed temperature or precipitation threshold is exceeded, for instance, frost days or tropical nights;

3. duration indices, which describe the length of wet and dry spells, or warm and cold spells;

4. percentile-based threshold indices, which describe the exceedance rates above or below a threshold which is defined as the 5^{th} or 95^{th} percentile derived from the 1972–2003 base period;

In particular, a subset of the standard IPCC indicators of extreme events were selected (Sillmann and Roeckner, 2008) related to precipitation, listed in Table 6.3.

In Figure 6.4, the values of the percentage over the total of precipitation due to 95th percentile events for each RCM (rcp95), derived from the 1972–2003 base

Table 6.2: The regional climate model data used in this study

ID	Domain	Institute	RCM name	Resolution	Driving model*	Driving experiment*	Driving ensemble	Period
1**	MED-CORDEX	CNRM	ALADIN v5.2	0.11 deg		Hist/RCP85	r8i1p1	1950-2005
2	MED-CORDEX	ICTP	RegCM 4.3	0.11 deg		His/RCP85	r1i1p1	1970-2005
3	EURO-CORDEX	KNMI	RACMO22E	0.11 deg	EC-EARTH	Hist/RCP85	r1i1p1	1950-2005
4	EURO-CORDEX	CLMcom	CCLM4-8-17	0.11 deg	CNRM-CM5-LR	Hist/RCP85	r1i1p1	1950-2005
5	EURO-CORDEX	CLMcom	CCLM4-8-17	0.11 deg	EC-EARTH	Hist/RCP85	r12i1p1	1950-2005
6	EURO-CORDEX	CLMcom	CCLM4-8-17	0.11 deg	MPI-ESM-LR	Hist/RCP85	r1i1p1	1950-2005
7	EURO-CORDEX	KNMI	RACMO22E	0.11 deg	HadGEM2-ES	Hist/RCP85	r1i1p1	1950-2005
8	EURO-CORDEX	CLMcom	CCLM4-8-17	0.11 deg	HadGEM2-ES	Hist/RCP85	r1i1p1	1950-2005
9**	EURO-CORDEX	IPSL-INERIS	WRF331F	0.11 deg	IPSL-CM5A-MR	Hist/RCP85	r1i1p1	1971-2005
10	EURO-CORDEX	SMHI	RCA4	0.11 deg	CNRM-CM5	Hist/RCP85	r1i1p1	1969-2005
11	EURO-CORDEX	SMHI	RCA4	0.11 deg	EC-EARTH	Hist/RCP85	r12i1p1	1969-2005
12	EURO-CORDEX	SMHI	RCA4	0.11 deg	IPSL-CM5A-MR	Hist/RCP85	r1i1p1	1969-2006
13	EURO-CORDEX	SMHI	RCA4	0.11 deg	HadGEM2-ES	Hist/RCP85	r1i1p1	1969-2005
14	EURO-CORDEX	SMHI	RCA4	0.11 deg	MPI-ESM-LR	Hist/RCP85	r1i1p1	1969-2005
15	EURO-CORDEX	DMI	HIRHAM5	0.11 deg	EC-EARTH	Hist/RCP85	r3i1p1	1950-2005

** Driving Model is the underlying global model.*

*** "The abnormal seasonal cycle is centred over Greece and Italy" - Samuel Somot.*

TABLE 6.3: IPPC Extreme Precipitation Indicator

Label	Description	Units
Cdd	consecutive dry days (<1 mm)	day
Cwd	consecutive wet days (>1 mm)	day
rx1day	maximum precipitation in 1 day	millimetre
rx5day	maximum precipitation in 5 days	millimetre
r10	number of days with precipitation over 10 mm/d	day
r20	number of days with precipitation over 20 mm/d	day
r95p	percentage over the total of precipitation due to 95th percentile events	percent
Map	Mean annual precipitation	millimetre
Adp	Average daily precipitation	millimetre

period, are shown. The RCMs reproduce well the spatial variability of the extreme events, indeed the highest values are shown in the east part of the Sicily.

The spatial resolution higher than GCM allows to represent well the variability of the precipitation in the region due to different factors, e.g. the complex morphology.

As a summary of all indices considered in Table 6.3, the Figure 6.5 shows the spatial correlation of each index (columns) for each RCMs before the procedure of correction. In such a plot, the good performance of a model in all indicators would show up as a tendency to see clear horizontal lines. On the contrary, the tendency of an indicator to be well represented by all models would show clear vertical lines.

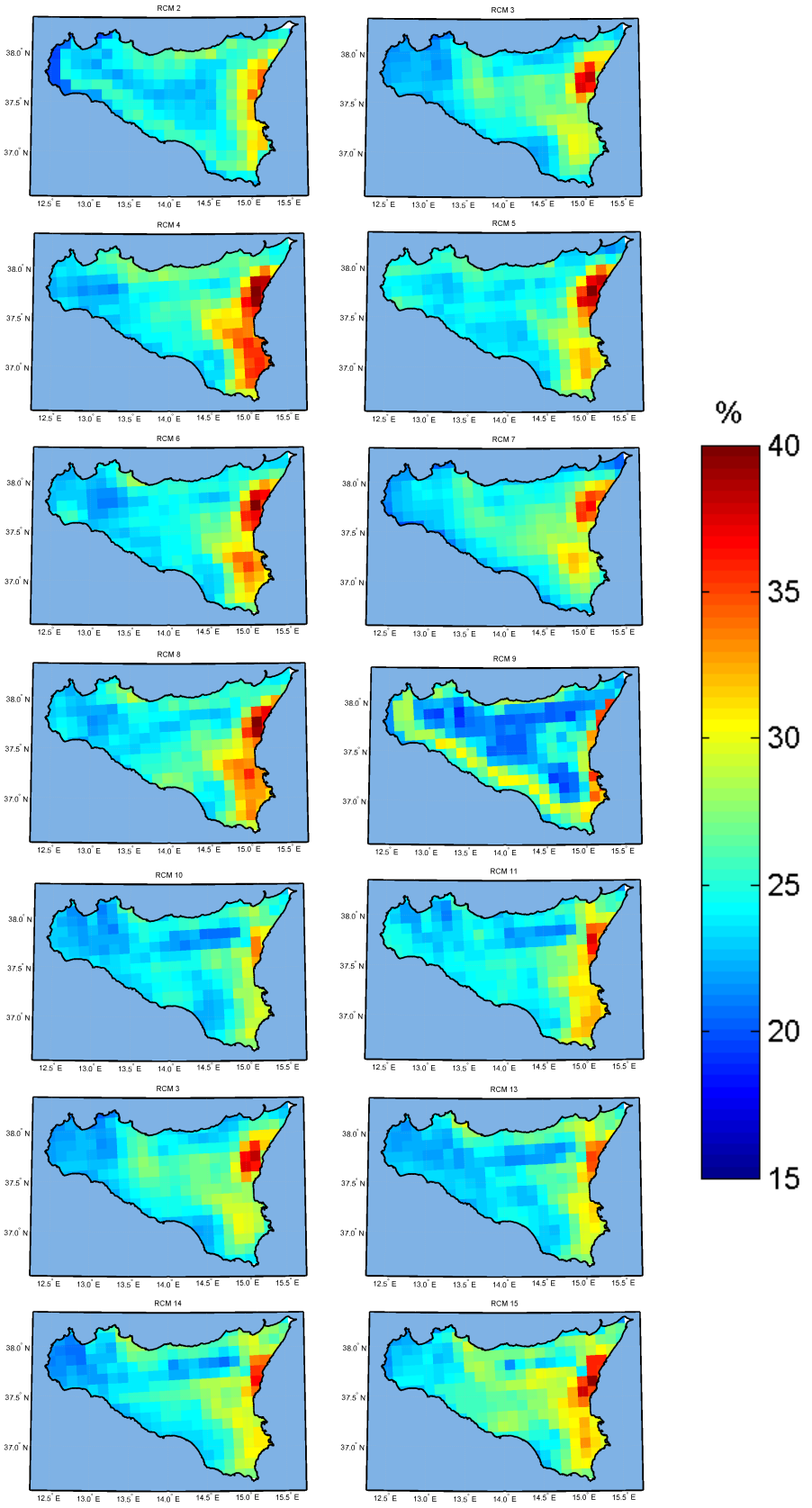
Analysing the Figure 6.5, it is possible to do a preliminary evaluation of the RCMs. In general, all RCMs have problems to model the *Cdd* and *Cwd*. The *r95(%)* is the best index represented by almost all RCMs which have shown high correlation value with the observed values. The RCMO2 and COSMO-CLM regional climate model have provided the highest performance, and among these, in terms of correlation the RMC3 has provided the best outcome.

6.7 Bias correction methods

RCM simulations are typically affected by systematic and random model errors. Misestimated climate variables in general, incorrect seasonal variations of precipitation (Christensen et al., 2008, Terink et al., 2009, Teutschbein and Seibert, 2010) and the simulation of too-many drizzle (i.e., low intensity rain) days

TABLE 6.4: Percentage over the total of precipitation due to 95th percentile events for each RCM (rcp95)

Regional climate models - rcp95



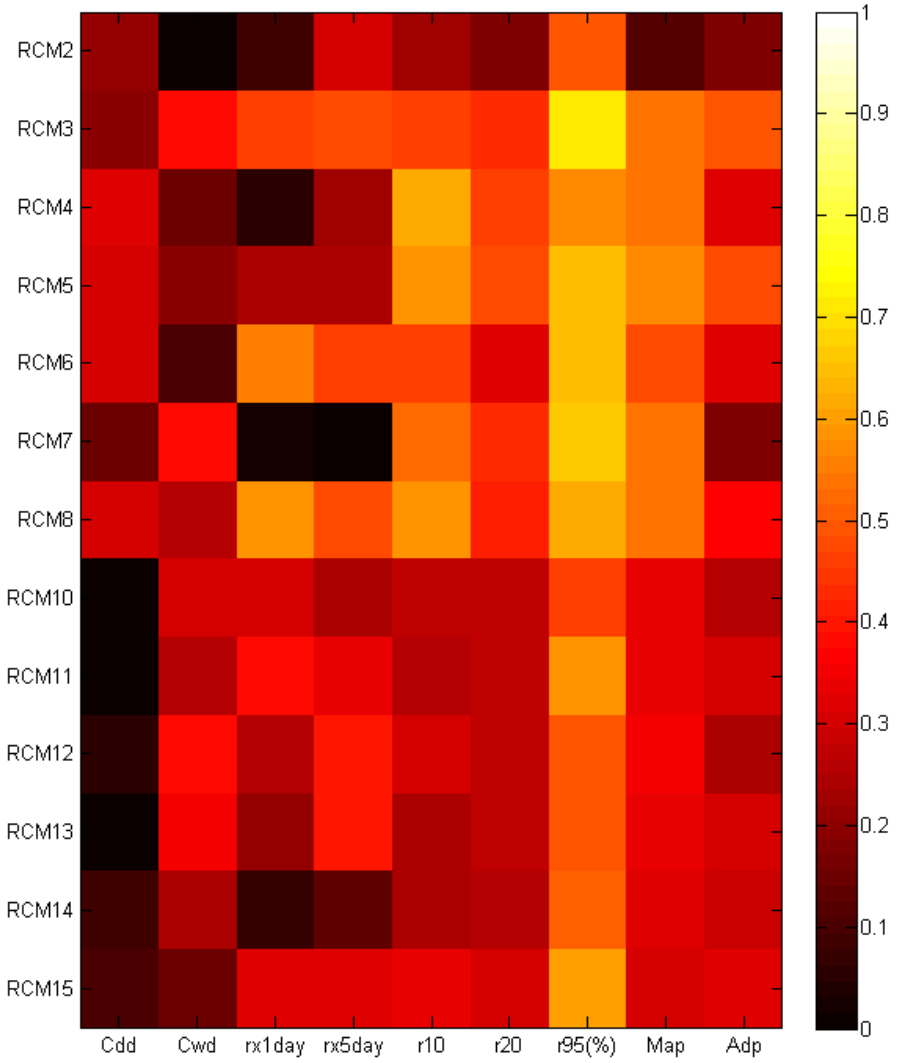


FIGURE 6.5: Spatial correlations of each indicator for each RCM with respect to observation. Shaded values range from 0 to 1.

(Ines and Hansen, 2006) are just a few examples of common systematic errors (biases). In other words, climate variables simulated by individual RCMs often do not agree with observed time series.

This poses a problem for using these simulations as input data for hydrological impact studies. One possible solution is to use an ensemble of RCM simulations (Déqué et al., 2007, Giorgi, 2006) that provides principally two advantages: (1) the spread of individual ensemble members covers a more realistic range of uncertainty and (2) the ensemble median may fit observations better (Jacob et al., 2007), which is especially true for temperature simulations. However, for precipitation simulations even the ensemble median often deviates considerably from observations and is not able to capture the variability in the observations. This

shows that it is not enough to only employ an RCM ensemble, but that additional correction procedures are needed.

Several bias correction methods have already been applied in weather forecasting under the name *model output statistics* (MOS) about four decades ago (Glahn and Lowry, 1972, Klein and Glahn, 1974). In the context of correcting RCM output, however, it is today a controversial subject (Ehret et al., 2012, Muerth et al., 2013): despite their advantageous ability to reduce errors in climate model output, most correction methods are criticized to diminish the advantages of climate models (Ehret et al., 2012) and to not have much added value in a complex modelling chain when considering other sources of uncertainty (Muerth et al., 2013).

Typical correction approaches aim at correcting the systematic error (bias) in RCM-simulated climate variables by employing a transformation algorithm and are therefore called *bias correction methods*. The concept is based on the identification of possible biases between observed and simulated climate variables, which is the starting point for correcting both control and scenario RCM runs. It should be noted that there is a risk of not only correcting systematic errors (biases) but also unintentionally modifying simulations due to unsystematic (random) model errors (Maraun et al., 2008).

A common assumption of most bias correction methods is *stationarity*, or *time invariance*, of the model errors. This implies that the empirical relationships in the correction algorithm and its parametrization for current climate conditions do not change over time and are also valid for future conditions. This assumption is, however, likely not met under changing climate conditions (Ehret et al., 2012, Maraun, 2012, Maraun et al., 2010, Vannitsem and Nicolis, 2008).

6.7.1 Statistical transformations

Statistical transformations attempt to find a function h that maps a modelled variable RCM such that its new distribution equals the distribution of the observed variable obs.

In order to correct the daily precipitation fields, the quantile matching are applied to the RCM output. The quantile matching adjusts all moments of the *probability distribution function* (PDF) of any variable of the model (Hagemann et al., 2011, Piani et al., 2010) by using the PDF of observations, integrating both PDFs to *cumulative distribution functions* (CDFs) and construct a transfer function.

Following Piani et al. (2010) this transformation can in general be formulated as:

$$obs = h(RCM) \tag{6.1}$$

Statistical transformations are an application of the probability integral transform (Angus et al., 1994) and if the distribution of the variable of interest is known, the transformation is defined as

$$obs = F_o^{-1}(F_m(RCM)) \tag{6.2}$$

where F_m is the CDF of RCM and F^{-1} o is the inverse CDF (or quantile function) corresponding to obs.

Parametric transformations. The quantile–quantile relation can be modelled directly using parametric transformations. Here, the following parametric transformations were used:

$$PAR - LIN \rightarrow \hat{obs} = a + b(RCM) \tag{6.3}$$

$$PAR - POW \rightarrow \hat{obs} = b(RCM - x_0)^c \tag{6.4}$$

where, $\hat{obs}$ indicates the best estimate of obs and a, b, c and x_0 are free parameters that are subject to calibration.

In the linear case (Eq. 6.3) b is simply a multiplicative correction factor and a is an additive correction factor. The a and b parameters are also related to the dry day correction factor $x_0 = -a/b$, which is not used in Eq. (6.3). x_0 is the value of precipitation below which modelled precipitation is set to zero. This is done to equate the number of modelled and observed dry days. In conventional bias correction techniques (e.g., Chen et al. (2000), Roeckner et al. (1999)), x_0 is derived directly as the difference in the number of dry days between observations and model output. Eq. (6.4) is a power law with an explicit dry day correction x_0.

The transformation Equations (6.3), (6.4) were all used by Piani et al. (2010) and some of them have been further explored in follow up studies (Dosio and Paruolo, 2011, Rojas et al., 2011). Following Piani et al. (2010), all parametric transformations were fitted to the fraction of the CDF corresponding to observed

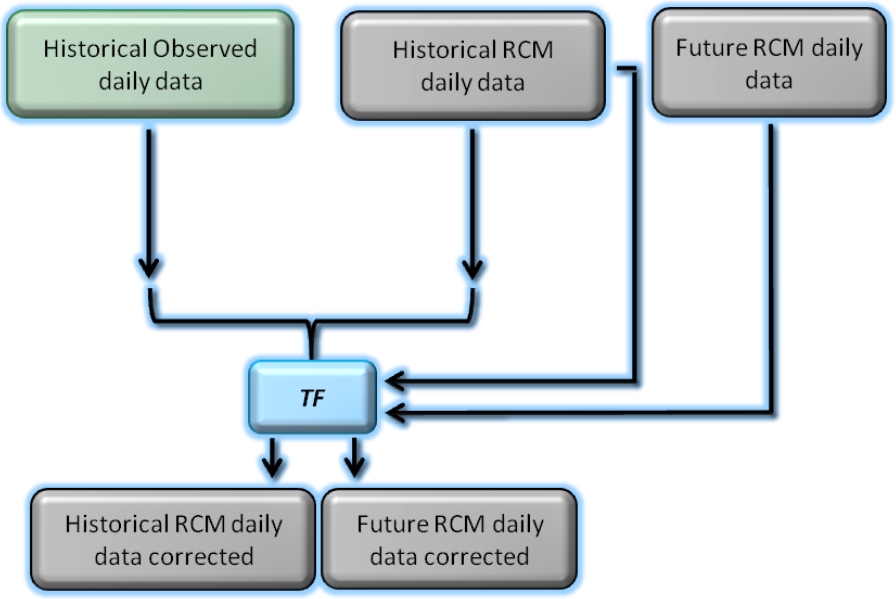


FIGURE 6.6: Daily correction procedure

wet days ($obs > 0$) by minimising the residual sum of squares. Modelled values corresponding to the dry part of the observed empirical CDF were set to zero.

Nonparametric transformations *(NON-PAR).* A common approach is to solve Equation (6.2) using the empirical CDF of observed and modelled values instead of assuming parametric distributions. Following the procedure of Boé et al. (2007) the empirical CDFs are approximated using tables of empirical percentiles. Values in between the percentiles are approximated using linear interpolation. If new model values (e.g. from climate projections) are larger than the training values used to estimate the empirical CDF, the correction found for the highest quantile of the training period is used Boé et al. (2007), Jakob Themeßl et al. (2011).

6.7.2 Daily correction

Both the methods of bias correction beforehand described were applied for the daily correction between the daily data observed and estimated by RCMs. The transfer functions (TFs) were applied inside the range from 1972 to 2003, because it is the range with greater number of serviceable stations. It was hypothesized absence of variability in the future, then at the future RCMs was applied the same correction. In the Figure 6.6 is schematically shown the daily correction procedure.

In order to derive the bias correction parameters, equal time lengths of observed and simulated daily data are required. In what follows, it will assume to deal with a time interval of daily precipitation data. There is no day-to-day correspondence between simulated and observed data, the time coordinates in the model data are purely internal. The initial datasets for each simulated RCM values and for observed data can be characterized as having the following form:

$$X_{obs} = x_{obs}(\varphi, \theta, d, m, y) \tag{6.5}$$

$$x = x(\varphi, \theta, d, m, y) \tag{6.6}$$

where x_{obs} and x are observed and modelled precipitation respectively, φ and θ are the longitude and latitude of the grid point, respectively, and d, m and y are indexes of day, month and year respectively.

Bias corrections, i.e. TFs, can be derived for every month of the year. In order to calculate monthly transfer function, the year label y in Eqs. (6.5) and (6.6) was removed grouping the data according to calendar month. The data are sorted at each grid point by depth precipitation:

$$X_{obs} = x_{obs}(\varphi, \theta, m, i) \tag{6.7}$$

$$x = x(\varphi, \theta, m, i) \tag{6.8}$$

such that $x(\varphi, \theta, m, i+1) \geq x(\varphi, \theta, m, i)$ and i spans the entire population of daily precipitation values for that month and all years of the decade. For fixed (φ, θ, m) the emerging TF is defined discretely for all i indexes as:

$$TF_{\varphi,\theta,m}(x(\varphi, \theta, m, i)) = x_{obs}(\varphi, \theta, m, i) \tag{6.9}$$

Only the portion of the emerging TF which corresponds to observed wet days is fitted. A wet day is evaluated during the evaluation of transfer function for each grid. The wetter days correspond to the portion of the TF on the right side of the intersection with the x-axis (Figure 6.7). This method was applied both historical climate data and future climate data for each RCM. The transfer functions can only be constructed at grid points where almost one observation station is situated.

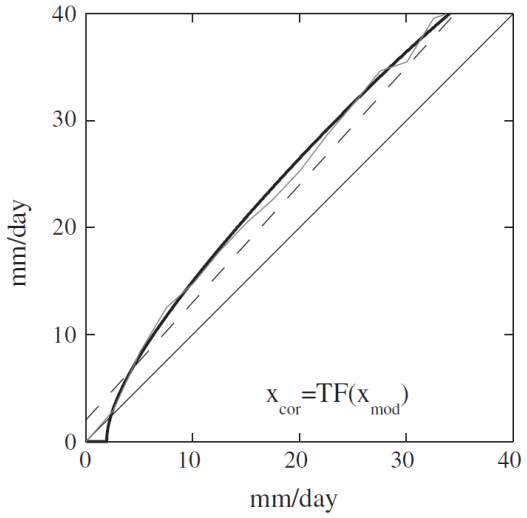


FIGURE 6.7: Transfer function derived from CDFs (continuos thick line) superimposed on "perfect" transfer function derived by re-sorting and plotting precipitation values directly (continous line). Also shown is the linear fit to "perfect" trasfer function (dashed line).

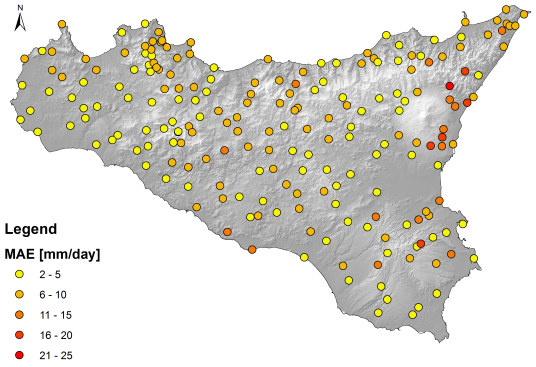
6.7.3 Comparison between different methods

In order to assess the performance of the different methods, a set of scores is needed that quantifies the similarity of the observed and the (corrected) modelled empirical CDF. Overall performance is measured using the *mean absolute error* (MAE) that is a quantity used to measure how close the corrected empirical CDF are to the eventual the observed empirical CDF.

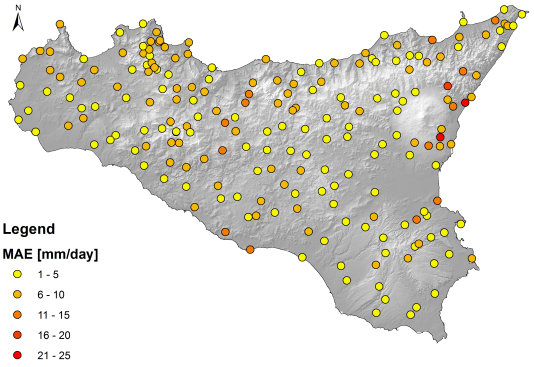
Statistical transformations, as any statistical technique, quietly assume that the modelled relation holds if confronted with new data. In the context of climate impact assessment this assumption is critical as it has to be expected that climate variables exceed the observed range in future periods.

Further, highly adaptable methods, such as the non-parametric techniques used in this study, are prone to overfitting the data. Both issues require that model error is quantified using data that have not been used for calibration.

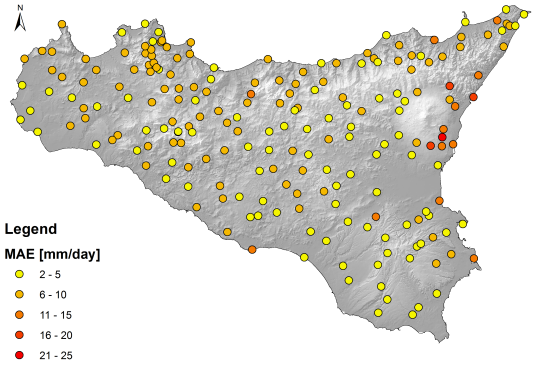
The suitability of the different statistical transformations to correct model precipitation from the CORDEX project was tested using observed daily precipitation rates of 208 stations in Sicily, all covering the 1972–2003 time interval. The dataset of the observed and modelled data was divided in two subsets. The first sub-set have been used in the calibration procedure to fit the parameters of different statistical transformation with 1972-1987 time interval while the second sub-set have been used for the validation of the statistical methods with 1988-2003 time interval.



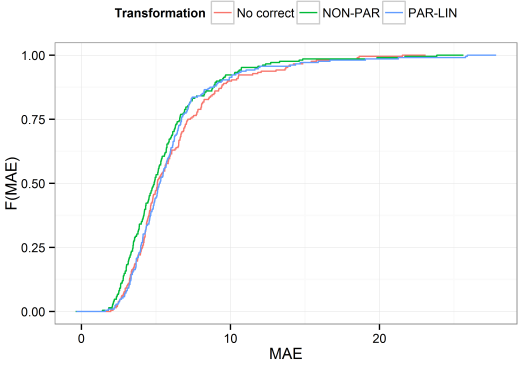
(a) Uncorrected modelled



(b) NON-PAR correction: statistical trasformation based on empirical quantile



(c) PAR-LIN correction: statistical trasformation based on parametric function



(d) Cdf of the MAE

FIGURE 6.8: Mean absolute error (MAE) between the observed and modelled empirical CDF of daily precipitation for different statistical transformation obtained from the mean value for all RCMs, in the validation period.

The MAE for all stations and two methods under consideration and the respectively CDF of the MAE are shown in Figure 6.8. The statistical transformation with power function is not shown because it has achieved values of MAE very hight and for this reason, this transformation was excluded. For the uncorrected model output, MAE has pronounced geographic variations. The largest errors are found along the east coast, where the RCMs cannot model the orographic effect on precipitation with sufficient detail. This analysis are confirmed to the value of the CDF of the MAE where the NON-PAR result be the better transformation function.

In Table 6.5 are reported the total mean absolute error for each RCM averaged over all stations. The NON-PAR transformation have achieved results better than PAR-LIN transfer function for almost all models, for this reason the value correct in this way are used in subsequently analysis.

TABLE 6.5: Total MAE, averaged over all stations.

Id RCM	Uncorrected values	NON-PAR	PAR-LIN
RCM2	5.21	4.47	4.44
RCM3	5.20	3.86	4.04
RCM4	5.20	3.86	4.04
RCM5	5.33	6.45	10.17
RCM6	4.02	4.14	3.99
RCM7	7.89	5.51	6.37
RCM8	9.16	8.34	7.59
RCM10	5.30	4.95	5.00
RCM11	4.75	4.52	4.72
RCM12	5.63	5.41	7.57
RCM13	7.55	6.90	6.88
RCM14	7.23	7.75	7.67
RCM15	5.83	4.92	4.60

Prior to application of statistical transformations and related post-processing methods, it is important to recall that these correction techniques are designed with a limited scope: to adjust the simulated climate variable such that its distribution (or some aspects of it) matches the distribution of observed values.

The validity of this assumption cannot be fully assessed, as the variable of interest may exceed the observed range in a changing climate. Further, numerical experiments on the global scale have shown that uncertainty related to the choice of calibration period is small compared to uncertainties related to choice of climate model and emission scenario Chen et al. (2011).

The Figure 6.9 shows the spatial correlation of each index (columns) for each RCMs after the procedure of correction. In general, the procedure of correction enhances the performance of all the models; the RCM6 have provided the greater performance while RCM13 have provided high value of correlation for some indicators; with regard to the extreme indices, only the consecutive dry day (Cdd) and the consecutive wet day (Cwd) do not show significant improvement.

6.7.4 Model capability to rappresent seasonality

In this section the skill of the different RCMs were analysed in reproducing the mean precipitation regime and the seasonal cycle. The historical and future (two different scenarios) mean annual precipitation for the ensemble median of the RCMs are shown in Figure 6.10.

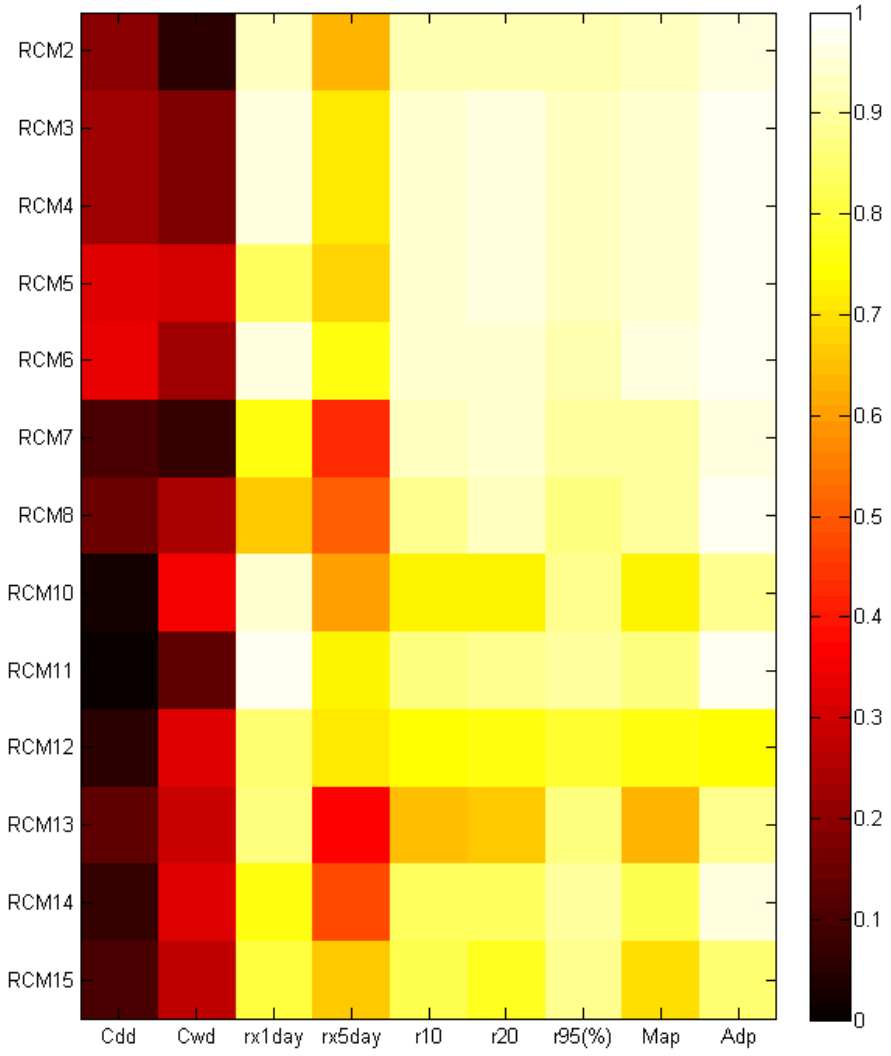


FIGURE 6.9: Spatial correlations after the correction of each indicator for each RCM with respect to observation. Shaded values range from 0 to 1.

In order to evaluate the capability to depict the monthly variability, the monthly averaged value for each RCM was derived and analysed over the regions defined in the Chapter 5. The different monthly precipitation for each region, before and after the bias correction, are illustrated in Figure 6.11 and Figure 6.12 respectively. For a better comparison, the y-axis scale use to rappresent the for precipitation is the same in all the plots (ranging from 0 to 260 mm).

All the regions present a seasonal cycle with their minimum values in July and August. The black line shows the observed monthly (spatially averaged) precipitation for each region while the other lines show the monthly averaged provided by each RCM.

In general, the models are able to understand the monthly variability, before the correction (Figure 6.11); only for the region 3 the RCMs provide values

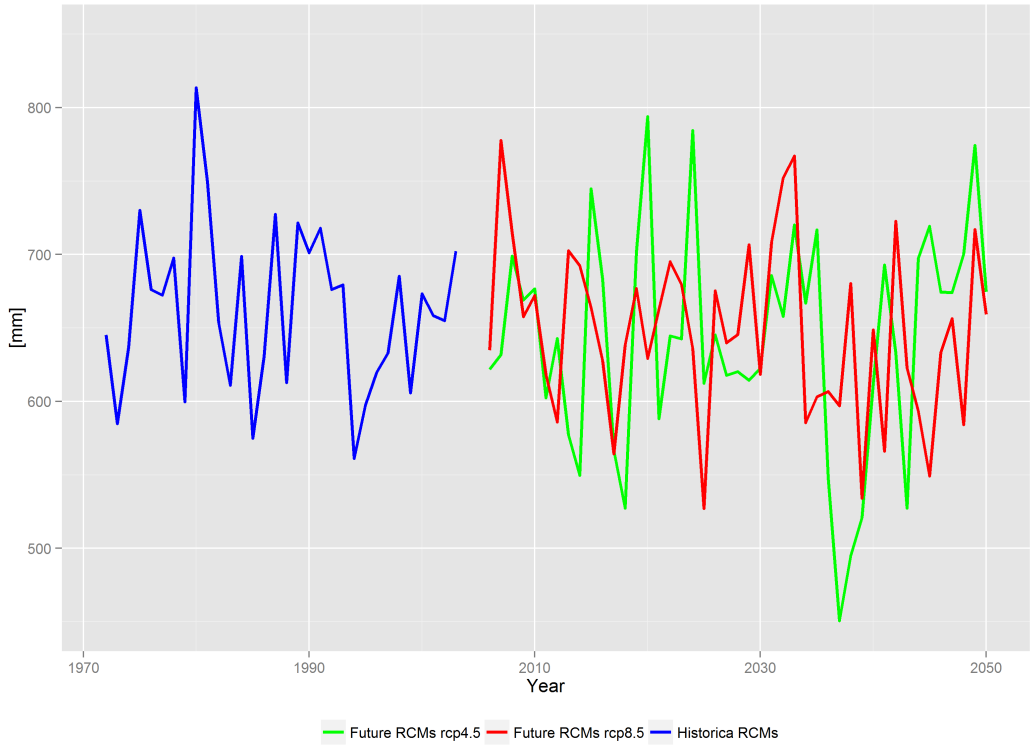


FIGURE 6.10: Mean annual precipitation for historical RCMs and for two different scenarios.

rather different respect to the observed value, while in the other region the RCMs have shown better performances. After the correction (Figure 6.12) the difference among observation and RCMs are strongly reduced, even in the region 3.

6.8 Temporal downscaling

In order to obtain the quantiles of precipitation for sub-daily duration, a temporal downscaling is needed. Starting from the daily value provided from the RCMs, the sub daily have been carried out.

Several approaches for statistical temporal downscaling of precipitation time series have been suggested in the literature, including methods based on a point-process model (Glasbey et al., 1995, Koutsoyiannis and Onof, 2001, Marani and Zanetti, 2007, Rodriguez-Iturbe et al., 1988). Koutsoyiannis and Onof (2001) have developed a disaggregation methodology for the generation of hourly data that aggregate up to given daily totals using the Bartlett-Lewis model. Marani and Zanetti (2007) presented a temporal downscaling method based on a point-process model that employs theoretically based estimates of rainfall variability on

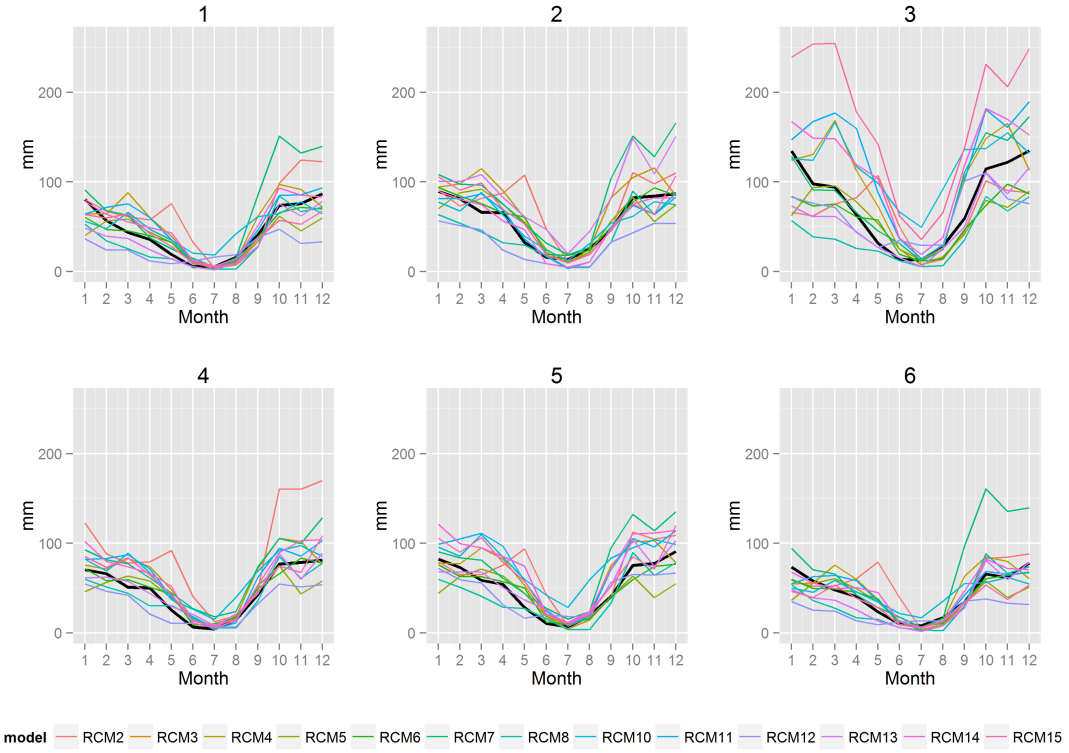


FIGURE 6.11: Monthly distribution of the spatially averaged precipitation for each region before the procedure of bias correction.

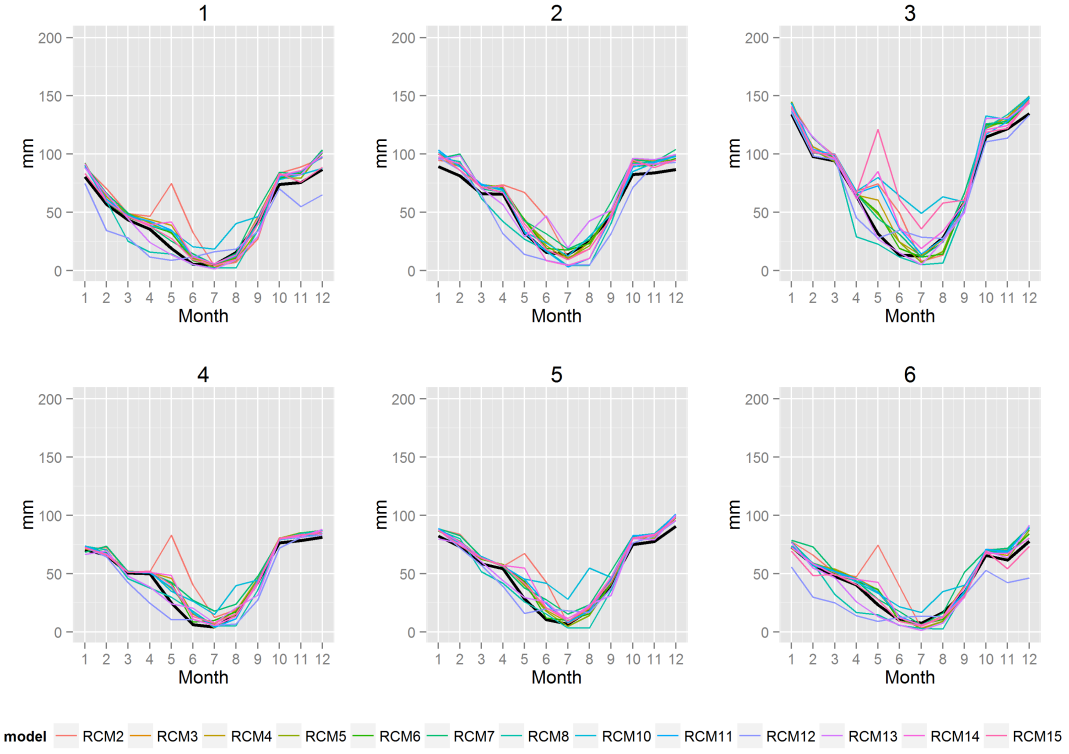


FIGURE 6.12: Monthly distribution of the spatially averaged precipitation for each region after the procedure of bias correction.

an hourly scale derived from daily statistics to produce approximately unbiased estimates of rainfall variance on an hourly time scale.

In this study, in order to obtain the maximum sub-daily precipitation from the daily time series for each RCMs, a simple method of temporal downscaling was proposed starting the idea developed by Srivastav et al. (2014) that have presented a methodology based on *equidistance quantile matching* (EQM) for updating the DDF curves under climate change.

Analysing the correlation among the observed precipitation for different duration (mean maximum daily and sub-daily) has carried out high values. The flow chart of proposed methodology is shown in Figure 6.13.

From the each RCMs the annual maximum daily value of precipitation was extracted. At this point, a non-parametric transfer function (NON-PAR), with the same approach illustrated in the section 6.7.2, was obtained among the max-daily and max sub-daily observed values using the data for the period 1972-2003. The parameters of transfer function, obtained by a comparison among the empirical CDF, was evaluated from the duration, e.g. daily-24h, 24h-12h, etc. The values of correlation coefficient, that illustrates the quantitative measure of some type of correlation and dependence, were carried out among the max-daily and max sub-daily observed values, and has returned high value near the unit (Table 6.6), confirming the hypothesis of correlation between the duration.

Five transfer functions derived from the observed value, for the couple of duration shown in the first columns in Table 6.6.

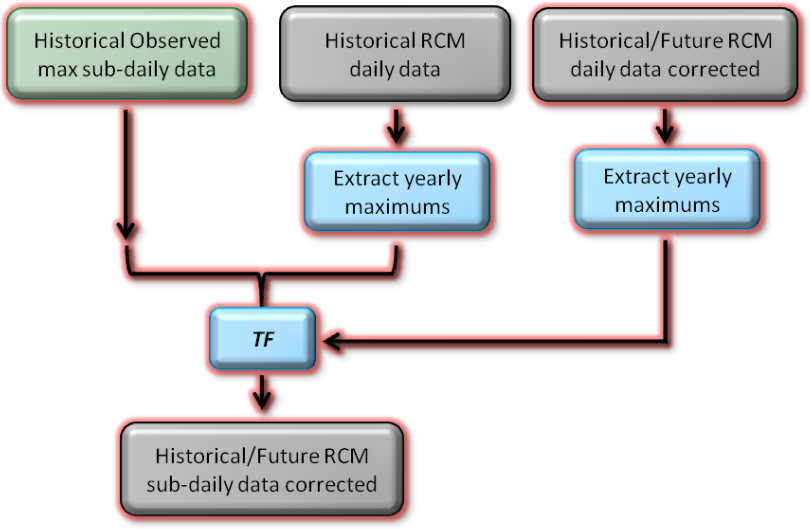


FIGURE 6.13: Flow chart of temporal downscaling

TABLE 6.6: Mean correlation coefficient between the maximum value of precipitation and different duration

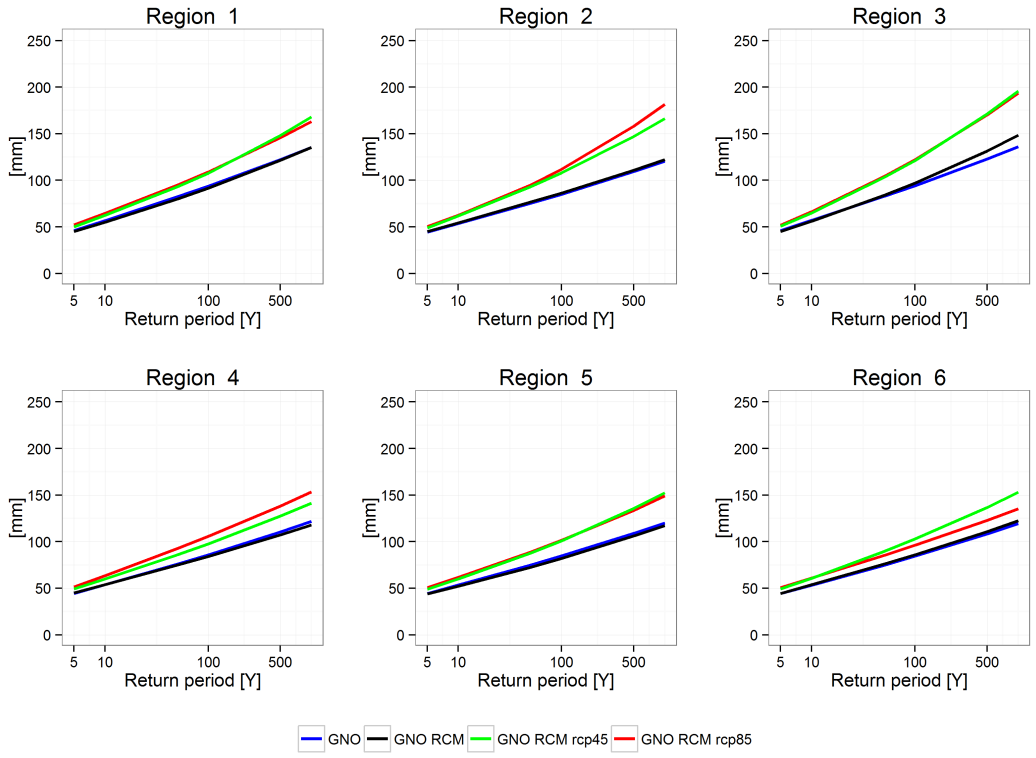
Duration	Correlation coefficient
Day-24h	0.759
24h-12h	0.925
12h-6h	0.905
6h-3h	0.900
3h-1h	0.837

Follow a "cascade method", wherein first the pair of highest duration values (i.e. Day-24h) were corrected and subsequently that of low duration (i.e. 3h-1h). At this point, for each RCMs, the maximum sub daily values are available.

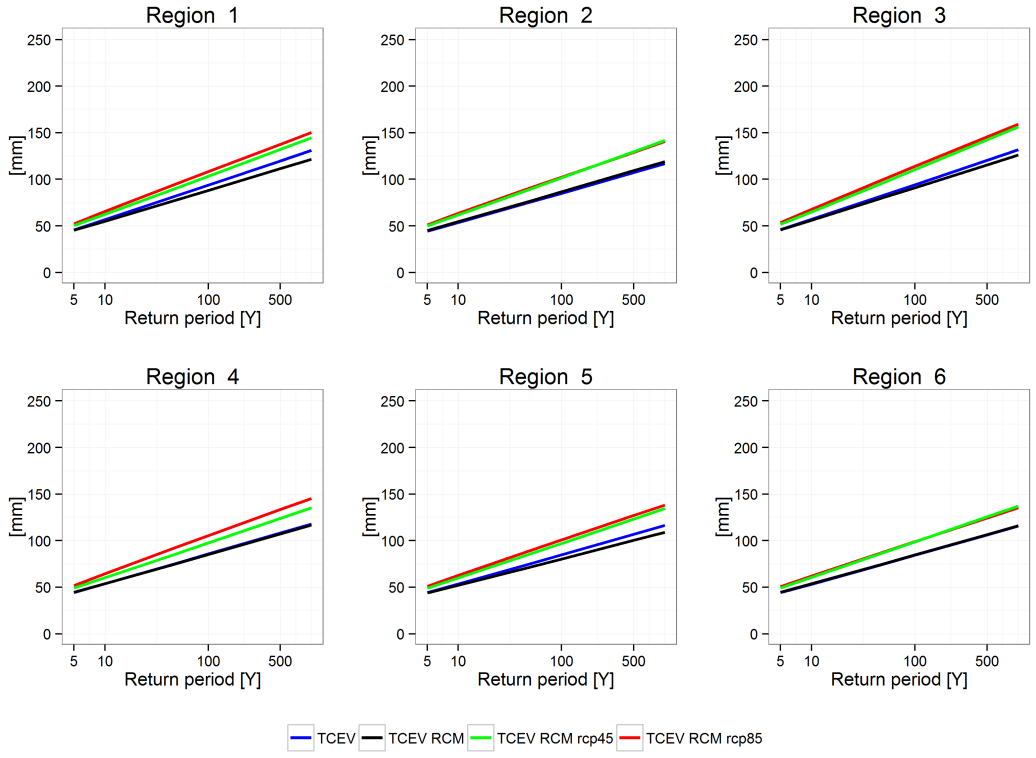
From the maximum sub-daily values obtained from the temporal downscaling, the growth curves for two different scenario were achieved. In order to permit a comparison have been used only the RCMs with values for both scenario (RCM 10, 11, 12, 13, 14, and 15). The calculation is carried out with the same procedure illustrated in the chapter 5 realising a comparison with historical growth curves obtained from the analysis in the previous chapter. The growth curves for two duration (3h - 24h) and distributions, that have provided the better performances in the previous analysis (GNO, TCEV), are shown in Figure 6.14 - 6.15. The figures show historical curves obtained from the observation (blue), the historical corrected curve obtained from the RCMs after the temporal downscaling (black), and the future curve for the different scenario, *RCP4.5* (green) and *RCP8.5* (red). From the comparison of the curves, it is possible to see as the GNO provides the lower quantiles than TCEV. The black lines result almost perfectly overlapping observed regional growth curves, and this behaviour confirm the goodness of the temporal downscaling method. Only the region 3 (north-east of Sicily) shows a small gap among the growth curves confirming the difficult of the RCMs to model the precipitation in this region.

In order to provide a quantitative assessment of the future change of the quantiles, the relative variation of the $\Delta(\%)$ has been obtained as the ratio between $Q_{RCP45/85} - Q_{hist}$ and Q_{hist}. The values obtained (by RCM 10, 11, 12, 13, 14, and 15) are shown in the Figure 6.16 and 6.17.

In the Figure 6.16 and 6.17, the $\Delta(\%)$ are shown for different scenario, RCP4.5 (continuous line) and RCP8.5 (dashed line), and different distributions. Generally, the precipitation for the lower duration has shown an increase higher than

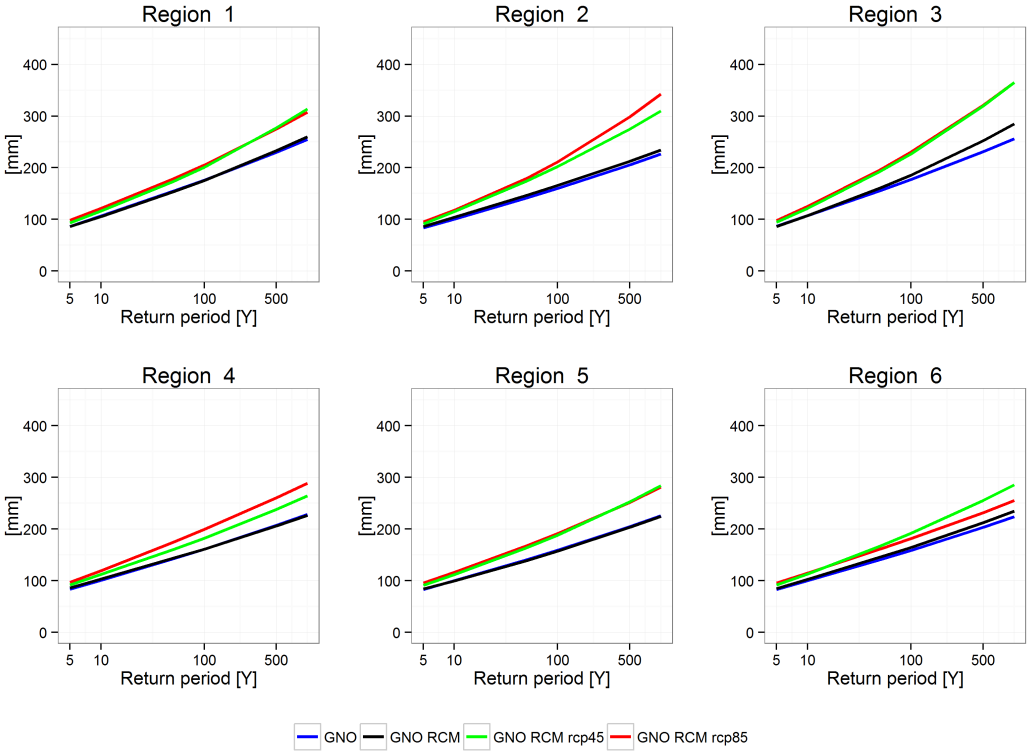


(a) GNO

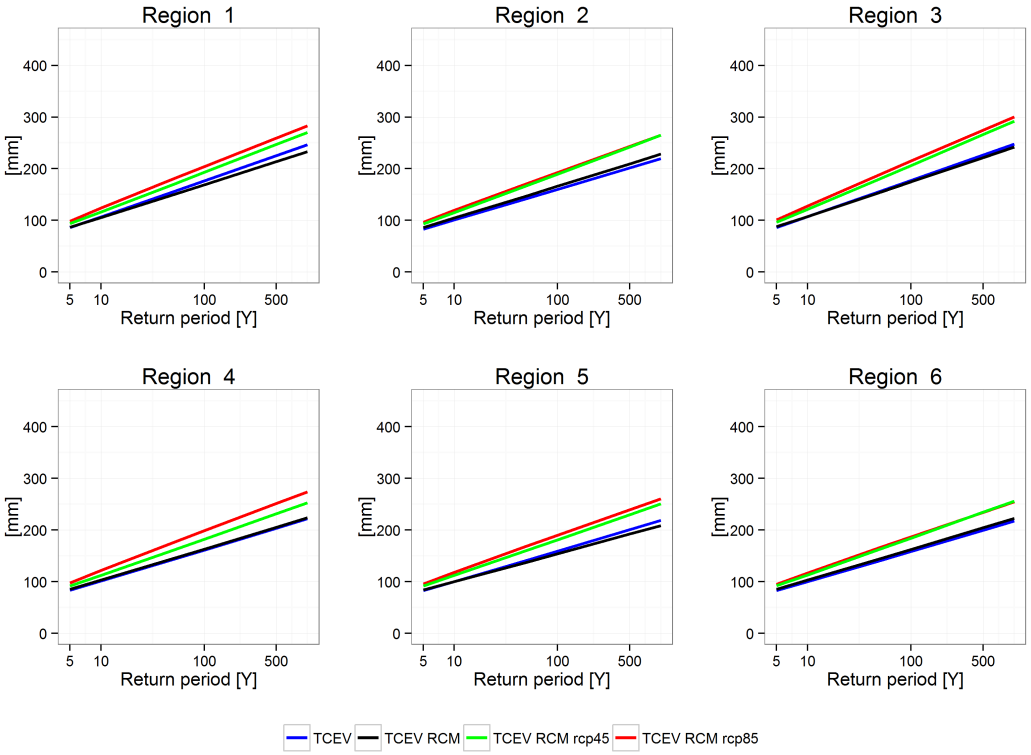


(b) TCEV

FIGURE 6.14: Growth curves for the each distribution for the duration of 3h (2006-2050).



(a) GNO



(b) TCEV

FIGURE 6.15: Growth curves for the each distribution for the duration of 24h (2006-2050).

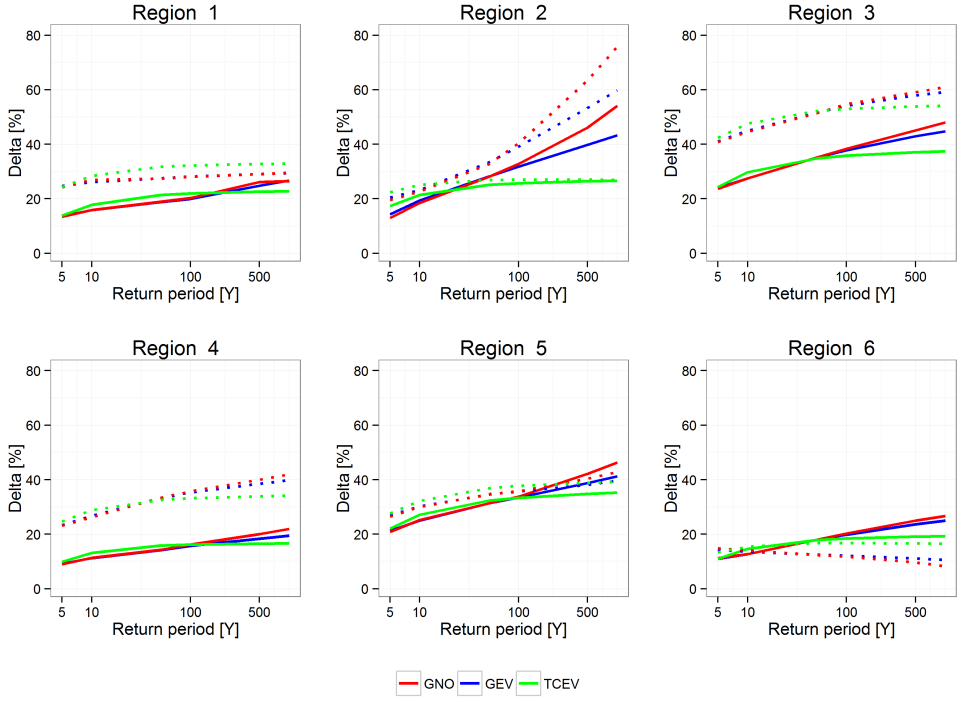


FIGURE 6.16: $\Delta(\%)$ change 3h (2006-2050).

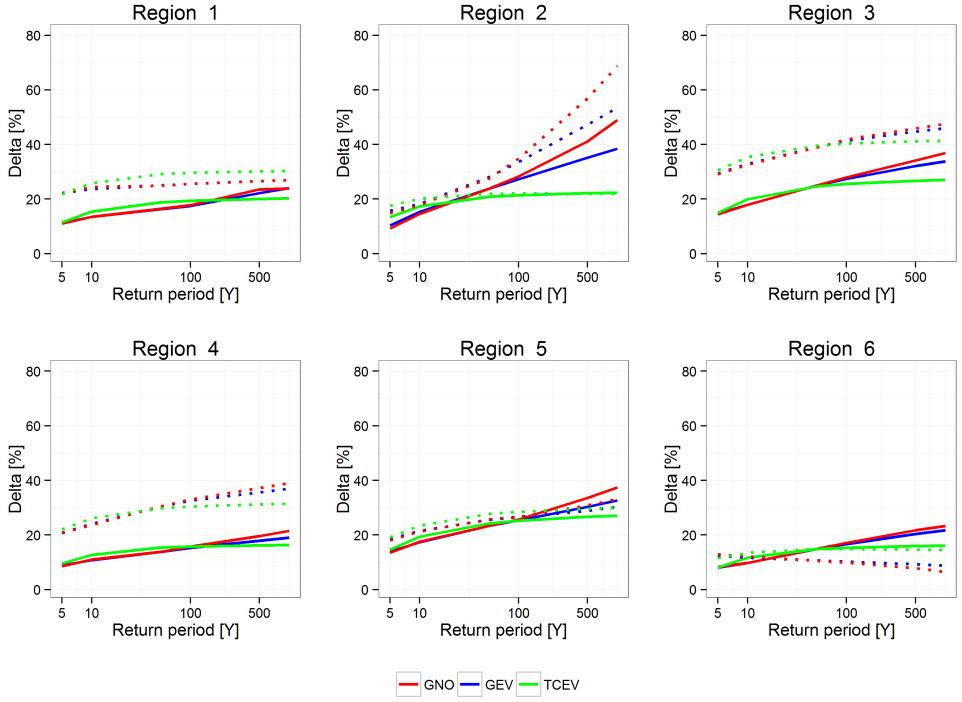


FIGURE 6.17: $\Delta(\%)$ change 24h (2006-2050).

precipitation for greater duration.

The behaviour of the $\Delta(\%)$ is different among the regions. The regions 1, 3 and 4 show a similar trend with the scenario RCP4.5 provides values higher than scenario RCP8.5; the value increased with return period, and in the region 3, the values overtake the 50% for the high return period while in the other two regions this value is not exceeded.

A behaviour totally different is shown in the region 2 and 6. In the first, both scenarios provide an exponential increment with an exclusion of TCEV distribution; opposite in the region 6, where the $\Delta(\%)$ have shown a reduction with the return period, and in this case the RCP4.5 have provided values higher than RCP8.5. The TCEV has provided systematic increment while the LN3 and GEV show a strong increase for high return period.

This behaviour is confirmed by many works of the scientific community (Nikulin et al., 2011, Sillmann et al., 2013), that underline this trend. The indications of climate change show that the future years will be driest of the past with an increase of the extreme rainfall events.

In conclusion, these results show again the complexity of the extreme rainfall in Sicily as previously said. The reason is linked to the complexity of morphology and the geographical position in the centre of the Mediterranea area with a mixture of climatology characteristics.

Conclusions

Many practical problems require knowledge of the behaviour of extreme values. For example, high precipitation amounts and resulting streamflows may affect sewerage systems, dams, reservoirs and bridges with different impacts. The motivation for analysing extremes is often to find an optimum balance between adopting high safety standards that are very costly on the one hand, and preventing major damage to equipment and structures from extreme events that are likely to occur during the useful life of such infrastructure on the other hand.

The evaluation of the extreme precipitation is a fundamental branch of the hydrologic application. Extreme precipitation events pose an increasing threat to society and infrastructure, especially under global warming. The research of more accurate tool to build the Depth-Duration- Frequency (DDF) curves, is justified by the engineering applications, that need to achieve more reliable and correct estimates.

In the case of extreme events, our major interest is not in what has occurred, but the likelihood that further extreme and damaging events will occur at some point in the future. The occurrence of many extreme events in hydrology cannot be forecasted on the basis of deterministic information with sufficient skill and lead time, and, in such cases, a probabilistic approach is required. If the occurrences of extreme events can be assumed to be independent in time, (i.e., the timing and magnitude of an event bear no relation to preceding events) then frequency analysis can be used to describe the likelihood of any event or a combination of events over the time horizon of a decision.

Changes in extreme weather and climate events have significant impacts and are among the most serious challenges to society in coping with a changing climate. Indeed, "confidence has increased that some extremes will become more frequent, more widespread and/or more intense during the 21st century" (IPCC, 2007). As a result, the demand for information services on weather and climate extremes

is growing. The sustainability of economic development and living conditions depends on our ability to manage the risks associated with extreme events.

The implementation of the regional frequency analysis, that allows to use more complex statistical distribution and the evaluation of the parameters, may be realized with a more wide datasets as an alternative to the classical approach.

In this thesis, a new regional frequency analysis based on *L-moments approach* has been implemented for the Sicily, Italy. Sicily shows a variety of morphological and climatic characteristics that result appropriate for the application of RFA. The proposed RFA has been carried out integrating the meteo-climatic information with directional statistics, number of dry day and ratio between summer rainfall and winter ranfall, since this kind of variables are able to provide information on the timing and seasonality of the extreme rainfall events.

The several steps of frequency analysis have been carried out through other statistical tools. In the first steps, the variables selected from the rainfall dataset were analysed with the *principal components analysis* which is able to analyse the data representing the observations described by several dependent variables. Its goal is to extract the significant information from the data table removing the ground noise, through a set of new orthogonal variables called *principal components* (PCs). The PCs obtained were used as input in a clustering method. In this study, the partition method called k-means algorithm has been implemented. The k-means needed the k value, that represents the number of the clusters to be obtained. A range k of values was selected considering the previous works developed in Sicily. The cluster analysis was performed for this range of values and the best value was selected using the *silhouette method*.

The six homogeneous regions obtained by k-mean algorithm were analysed by a test of homogeneity. The test of homogeneity is necessary to evaluate if the regions obtained do not present station with abnormal behaviour respect to the other stations over the region. Two tests of homogeneity have been performed with the target to identify the stations with abnormal behaviour. The stations individuated by discordance measure, in the first attempt were moved in another region, and if the resulted yet discordant, these were deleted.

In the next steps, different probability distributions were fit on the homogeneous region previously obtained. In order to select the frequency distribution, a goodness-of-fit procedure was used and applied to five candidate distributions. Two frequency distributions, the *three-parameter lognormal* (LN3) and *generalised*

extreme value (GEV), were selected. In order to evaluate the result with the previous works, the *two component extreme value* (TCEV) distribution was added. A typical three levels hierarchical regional approach has been adopted for the parameter estimation procedure based on the *L-moments method*.

In order to compare the quantiles, the "index value method" was developed using the regional growth curves carried out for each homogeneous region. The validity of the proposed approach was verified by the analysis of the values of regional L-moments and confirmed by the test of homogeneity for the highest statistics moments.

At the first levels the higher statistics moments, i.e., the L-skewness and L-kurtosis, were considered constant over the Sicily as a whole. At the second level, for each obtained homogeneous region, the value of L-CV was considered constant. At the final level, the "index value", that rappresent the mean value of precipitation, was evaluated for all Sicily by a spatial interpolation developed through *universal kriging* algorithm using the values of precipitation at gauged stations.

The growth curves for the sub-daily durations (1, 3, 6, 12, and 24 h) were estimated through the assessment of the quantiles for different duration and return period. In order to evaluate the performance of the regional frequency analysis, an assessment of accuracy was achieved. For each distribution, the BIAS and RMSE have been evaluated for different return periods. The analysis has shown that the TCEV and the LN3 provide the best performances.

In the second part of this work, the response of precipitation extremes under climate change was studied using ensembles of climate models provided by the *Coordinated Regional Climate Downscaling Experiment* (CORDEX). The climate data for the period 2006-2050 were obtained from *regional climate models* (RCMs). The climate models have provided daily values of precipitation. A bias correction method was applied on daily precipitation provided from the RCMs by comparing three different kinds of transfer functions. The best solution, evaluated in terms of *mean absolute error* (MAE), was achieved by the non-parametric transformation.

In order to obtain sub-daily values of precipitation, a straightforward method of downscaling was implemented. The temporal downscaling uses a cascade method, in the first step the precipitation for the duration of 24 h were corrected, and after the precipitation for the lower durations. After the temporal downscaling was accomplished, the growth curves for each distribution have been estimated and an

evaluation on the future sub-daily precipitation was obtained for different return periods.

The precipitation for the highest duration has shown an increase lower than that relative to the precipitation for the shortest duration. This trend, combined with the reduction of the mean annual precipitation for the period 2006-2050 follows the results of many works in the scientific community (Frei et al., 2006, Meehl et al., 2000).

In conclusion, this thesis has provided an updated tool for the modelling of extreme precipitation for the area of Sicily, with different features respect to previous works both in terms of the definition of homogeneous zones and in terms of parameters of the frequency distribution. Meteo-climatic information and the seasonality of extreme events retrieved from the dataset have been exploited in the analysis and have provided a better characterization of the extreme rainfall events.

The analysis of climate change of extreme precipitation has shown an increase of the quantiles; for this reason an increase of the attention on the management and planning of the future hydraulic infrastructures is needed. In the practical applications, a careful evaluation of the DDF taking into account affects of climate change to use in the design of the infrastructures could be essential, to reduce the damage and loss of human lives in the future. Indeed, the most of the works on the climate change show an increment of the extreme precipitation (Nikulin et al., 2011, Sillmann et al., 2013). For this reason, during the planning of the infrastructures, the use of the DDF that take account the possible increase of precipitation must be indispensable, in particular for major infrastructures designed for high return periods.

Further efforts could be made in the future, in order to improve the evaluation of climate models.

Bibliography

Alexander, L., Zhang, X., Peterson, T., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., et al. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D5).

Alila, Y. (1999). A hierarchical approach for the regionalization of precipitation annual maxima in canada. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 104(D24):31645–31655.

Angus, J., Gardner, P., Kirkegaard, J., and Desmarchelier, J. (1994). Biofumigation: isothiocyanates released frombrassica roots inhibit growth of the take-all fungus. *Plant and soil*, 162(1):107–112.

Arnell, N. and Beran, M. (1987). Testing the suitability of the two-component extreme value distribution for regional flood estimation.

Ashkar, F. and Ouarda, T. B. (1996). On some methods of fitting the generalized pareto distribution. *Journal of hydrology*, 177(1):117–141.

Bacchi, B. (1996). Metodi di ristrutturazione e progetto delle reti pluviometriche. *Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche, Presidenza del Consiglio dei Ministri, Dipartimento della Protezione Civile*.

Barnett, V. and Lewis, T. (1994). *Outliers in statistical data*, volume 3. Wiley New York.

Bayliss, A. C. and Jones, R. C. (1993). *Peaks-over-threshold flood database*. Institute of Hydrology.

Beguería, S. and Vicente-Serrano, S. M. (2006). Mapping the hazard of extreme rainfall by peaks over threshold extreme value analysis and spatial regression techniques. *Journal of applied meteorology and climatology*, 45(1):108–124.

Benjamin, J. R. and Cornell, C. A. (1970). Probability, statistics and decision theory for civil engineers.

Blenkinsop, S., Fowler, H., Dubus, I., Nolan, B., and Hollis, J. (2008). Developing climatic scenarios for pesticide fate modelling in europe. *Environmental Pollution*, 154(2):219 – 231.

Boberg, F., Berg, P., Thejll, P., Gutowski, W. J., and Christensen, J. H. (2010). Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ensembles models. *Climate dynamics*, 35(7-8):1509–1520.

Boé, J., Terray, L., Habets, F., and Martin, E. (2007). Statistical and dynamical downscaling of the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, 27(12):1643–1656.

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D. (2006). Precipitation-frequency atlas of the united states. *NOAA atlas*, 14(2).

Bureau, U. W. (1962). Rainfall-frequency atlas of the hawaiian islands. *Technical Paper*, 43.

Burn, D. H. (1997). Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of Hydrology*, 202(1):212–230.

Cannarozzo, M., D'asaro, F., and Ferro, V. (1995). Regional rainfall and flood frequency analysis for sicily using the two component extreme value distribution. *Hydrological sciences journal*, 40(1):19–42.

Cao, C., Piga, E., Salis, M., and Sechi, G. (1991). Valutazione delle piene in sardegna. *Rapporto Regione Sardegna CNR-GNDCI, LINEA1, Istituto di Idraulica Università degli Studi di Cagliari*.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, 1(2):245–276.

Cavicchia, L., von Storch, H., and Gualdi, S. (2014). Mediterranean tropical-like cyclones in present and future climate. *Journal of Climate*, 27(19):7493–7501.

Chen, C., Haerter, J. O., Hagemann, S., and Piani, C. (2011). On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle. *Geophysical Research Letters*, 38(20).

Chen, D., Cane, M. A., Zebiak, S. E., Canizares, R., and Kaplan, A. (2000). Bias correction of an ocean-atmosphere coupled model. *Geophysical Research Letters*, 27(16):2585–2588.

Chow, V. (1964a). Statistical and probability analysis of hydrologic data, part i, frequency analysis. *Handbook of Applied Hydrology, McGraw-Hill Book Co., New York, NY*, pages 8–1.

Chow, V. T. (1964b). Runoff. *Handbook of applied hydrology*, pages 14–1.

Chowdhury, J. U., Stedinger, J. R., and Lu, L.-H. (1991). Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research WRERAQ*, 27(7):1765–1776.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20).

Christensen, J. H., Christensen, O. B., Lopez, P., van Meijgaard, E., and Botzet, M. (1996). The hirham4 regional atmospheric climate model.

Christensen, O. B., Christensen, J. H., Machenhauer, B., and Botzet, M. (1998). Very high-resolution regional climate simulations over scandinavia-present climate. *Journal of Climate*, 11(12):3204–3229.

Cocke, S. and LaRow, T. (2000). Seasonal predictions using a regional spectral model embedded within a coupled ocean-atmosphere model. *Monthly Weather Review*, 128(3):689–708.

Coelho, C., Ferro, C., Stephenson, D., and Steinskog, D. (2008). Methods for exploring spatial and temporal variability of extreme events in climate data. *Journal of Climate*, 21(10):2072–2092.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

Colin, J., Déqué, M., Radu, R., and Somot, S. (2010). Sensitivity study of heavy precipitation in limited area model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. *Tellus A*, 62(5):591–604.

Cong, S., Li, Y., Vogel, J. L., and Schaake, J. C. (1993). Identification of the underlying distribution form of precipitation by using regional data. *Water Resources Research*, 29(4):1103–1111.

Copertino, V., Fiorentino, M., Sole, A., Valanzano, A., Claps, P., and Galasso, D. (1992). Carte tematiche della raccolta.

Cunderlik, J. M. and Ouarda, T. B. (2009). Trends in the timing and magnitude of floods in canada. *Journal of Hydrology*, 375(3):471–480.

Cuomo, A. and Guida, D. (2010). Definizione gis based delle barriere orografiche dellšappennino campano-lucano (italia meridionale). *XXXII Convegno Nazionale di Idraulica e Costruzioni Idrauliche*.

D'Agostino, R. B. (1986). *Goodness-of-fit-techniques*, volume 68. CRC press.

Dales, M. and Reed, D. (1989). *Regional flood and storm hazard assessment*. Institute of Hydrology.

Dalrymple, T. (1960). Flood-frequency analyses, manual of hydrology: Part 3. Technical report, USGPO,.

Davies, H. (1976). A lateral boundary formulation for multi-level prediction models.[numerical weather forecasting.

Déqué, M., Dreveton, C., Braun, A., and Cariolle, D. (1994). The arpege/ifs atmosphere model: a contribution to the french community climate modelling. *Climate Dynamics*, 10(4-5):249–266.

Déqué, M., Rowell, D., Lüthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., and van den Hurk, B. (2007). An intercomparison of regional climate simulations for europe: assessing uncertainties in model projections. *Climatic Change*, 81(1):53–70.

Di Piazza, A., Conti, F. L., Noto, L., Viola, F., and La Loggia, G. (2011). Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for sicily, italy. *International Journal of Applied Earth Observation and Geoinformation*, 13(3):396–408.

Dickinson, R. E., Errico, R. M., Giorgi, F., and Bates, G. T. (1989). A regional climate model for the western united states. *Climatic Change*, 15(3):383–422.

Doscher, R., Willén, U., Jones, C., Rutgersson, A., Meier, H. M., Hansson, U., and Graham, L. P. (2002). The development of the regional coupled ocean-atmosphere model rcao. *Boreal Environment Research*, 7(3):183–192.

Dosio, A. and Paruolo, P. (2011). Bias correction of the ensembles high-resolution climate change projections for use by impact models: Evaluation on the present climate. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 116(D16).

Drago, A., Cartabellotta, D., Lo Bianco, B., and Monterosso, I. (2002). Atlante climatologico della sicilia-seconda edizione. *Regione Siciliana, Assessorato Agricoltura e Foreste, Palermo*.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J. (2012). Hess opinions "should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, 16(9):3391–3404.

Elderton, W. (1953). Frequency curves and correlation cambridge university. *New York*.

Emanuel, K. (2005). Genesis and maintenance of" mediterranean hurricanes". *Advances in Geosciences*, 2(2):217–220.

Estivill-Castro, V. and Yang, J. (2000). Fast and robust general purpose clustering algorithms. In *PRICAI 2000 Topics in Artificial Intelligence*, pages 208–218. Springer.

Farda, A., Déué, M., Somot, S., Horányi, A., Spiridonov, V., and Tóth, H. (2010). Model aladin as regional climate model for central and eastern europe. *Studia Geophysica et geodaetica*, 54(2):313–332.

Faulkner, D. (1999). *Flood estimation handbook, volume 2: Rainfall frequency estimation*. Institute of Hydrology Wallingford.

Ferro, V. and Porto, P. (1988). Regional analysis of rainfall-depth-duration equation for south of italy.

Fiorentino, M., Gabriele, S., Rossi, F., and Versace, P. (1987). Hierarchical approach for regional flood frequency analysis. *Regional flood frequency analysis*, 3549.

Fisher, N. (1993). Statistical analysis of circular data (cambridge up, cambridge, uk).

Fowler, H., Ekström, M., Kilsby, C., and Jones, P. (2005). New estimates of future changes in extreme rainfall across the uk using regional climate model integrations. 1. assessment of control climate. *Journal of Hydrology*, 300(1):212–233.

Fowler, H. and Kilsby, C. (2003a). A regional frequency analysis of united kingdom extreme rainfall from 1961 to 2000. *International Journal of Climatology*, 23(11):1313–1334.

Fowler, H. J. and Kilsby, C. G. (2003b). Implications of changes in seasonal and annual extreme rainfall. *Geophysical Research Letters*, 30(13):n/a–n/a. 1720.

Fraley, C. and Raftery, A. E. (1998). How many clusters? which clustering method? answers via model-based cluster analysis. *The computer journal*, 41(8):578–588.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L. (2006). Future change of precipitation extremes in europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D6).

Gabriele, S. and Arnell, N. (1991). A hierarchical approach to regional flood frequency analysis. *Water Resources Research*, 27(6):1281–1289.

Gabriele, S. and Chiaravalloti, F. (2013). Using the meteorological information for the regional rainfall frequency analysis: An application to sicily. *Water resources management*, 27(6):1721–1735.

GELEYN, J.-F. (1988). Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus A*, 40(4):347–351.

Giorgi, F. (1990). Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate*, 3(9):941–963.

Giorgi, F. (2006). Climate change hot-spots. *Geophysical research letters*, 33(8).

Giorgi, F. and Mearns, L. O. (1991). Approaches to the simulation of regional climate change: a review. *Reviews of Geophysics*, 29(2):191–216.

Giorgi, F. and Mearns, L. O. (1999). Introduction to special section: Regional climate modeling revisited. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 104(D6):6335–6352.

Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (mos) in objective weather forecasting. *Journal of applied meteorology*, 11(8):1203–1211.

Glasbey, C., Cooper, G., and McGechan, M. (1995). Disaggregation of daily rainfall by conditional simulation from a point-process model. *Journal of Hydrology*, 165(1):1–9.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15(5):1049–1054.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C. (2011). Impact of a statistical bias correction on the projected hydrological changes obtained from three gcms and two hydrology models. *Journal of Hydrometeorology*, 12(4):556–578.

Haktanir, T. (1992). Comparison of various flood frequency distributions using annual flood peaks data of rivers in anatolia. *Journal of Hydrology*, 136(1):1–31.

Han, C. and Carlin, B. P. (2001). Markov chain monte carlo methods for computing bayes factors. *Journal of the American Statistical Association*, 96(455).

Han, J., Kamber, M., and Pei, J. (2011). *Data mining: concepts and techniques: concepts and techniques*. Elsevier.

Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Applied statistics*, pages 100–108.

Helsel, D. R. and Hirsch, R. M. (1992). *Statistical methods in water resources*, volume 49. Elsevier.

Herrmann, M., Somot, S., Calmanti, S., Dubois, C., and Sevault, F. (2011). Representation of spatial and temporal variability of daily wind speed and of intense wind events over the mediterranean sea using dynamical downscaling: impact of the regional climate model configuration. *Natural Hazards and Earth System Sciences*, 11:1983–2001.

Hobbs, P. V. and Locatelli, J. D. (1978). Rainbands, precipitation cores and generating cells in a cyclonic storm. *Journal of the atmospheric Sciences*, 35(2):230–241.

Holton, J. (1992). An introduction to dynamic meteorology (international geophysics series, san diego, new york.

Homar, V., Romero, R., Stensrud, D., Ramis, C., and Alonso, S. (2003). Numerical diagnosis of a small, quasi-tropical cyclone over the western mediterranean: Dynamical vs. boundary factors. *Quarterly Journal of the Royal Meteorological Society*, 129(590):1469–1490.

Hosking, J. and Wallis, J. (1988). The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research*, 24(4):588–600.

Hosking, J., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124.

Hosking, J. R. M. and Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, 29.

Hosking, J. R. M. and Wallis, J. R. (2005). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

Hostetler, S. (1994). Hydrologic and atmospheric models: the (continuing) problem of discordant scales. *Climatic Change*, 27(4):345–350.

Hostetler, S., Bates, G., and Giorgi, F. (1993). Interactive coupling of a lake thermal model with a regional climate model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 98(D3):5045–5057.

Houze Jr, R. A., Hobbs, P. V., Biswas, K. R., and Davis, W. M. (1976). Mesoscale rainbands in extratropical cyclones. *Monthly Weather Review*, 104(7):868–878.

Ines, A. V. and Hansen, J. W. (2006). Bias correction of daily gcm rainfall for crop simulation studies. *Agricultural and forest meteorology*, 138(1):44–53.

Jacob, D., Bärring, L., Christensen, O. B., Christensen, J. H., de Castro, M., Deque, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., et al. (2007). An inter-comparison of regional climate models for europe: model performance in present-day climate. *Climatic change*, 81(1):31–52.

Jacob, D., Elizalde, A., Haensler, A., Hagemann, S., Kumar, P., Podzun, R., Rechid, D., Remedio, A. R., Saeed, F., Sieck, K., et al. (2012). Assessing the transferability of the regional climate model remo to different coordinated regional climate downscaling experiment (cordex) regions. *Atmosphere*, 3(1):181–199.

Jakob Themeßl, M., Gobiet, A., and Leuprecht, A. (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, 31(10):1530–1544.

Janusz, N. (1982). Areal intensity-duration-frequency curves for short term rainfall events in lund. *Nordic Hydrology*, 13(4):193–204.

Ji, Y. and Vernekar, A. D. (1997). Simulation of the asian summer monsoons of 1987 and 1988 with a regional model nested in a global gcm. *Journal of climate*, 10(8):1965–1979.

Jolliffe, I. (2002). Principal components in regression analysis. *Principal component analysis*, pages 167–198.

Jolliffe, I. T. (1972). Discarding variables in a principal component analysis. i: Artificial data. *Applied statistics*, pages 160–173.

Jones, M. R., Blenkinsop, S., Fowler, H. J., and Kilsby, C. G. (2014). Objective classification of extreme rainfall regions for the uk and updated estimates of trends in regional extreme rainfall. *International Journal of Climatology*, 34(3):751–765.

Jones, P., Jonsson, T., and Wheeler, D. (1997). Extension to the north atlantic oscillation using early instrumental pressure observations from gibraltar and south-west iceland. *International Journal of Climatology*, 17(13):1433–1450.

Karl, T. R. and Easterling, D. R. (1999). Climate extremes: Selected review and future research directions. In *Weather and Climate Extremes*, pages 309–325. Springer.

Kendall, M. (1975). *Multivariate analysis*. Charles Griffin.

Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the ipcc ensemble of global coupled model simulations. *Journal of Climate*, 20(8):1419–1444.

KIDA, H., KOIDE, T., SASAKI, H., and CHIBA, M. (1991). A new approach for coupling a limited area model to a gcm for regional climate simulations. *Journal of the Meteorological Society of Japan*, 69(6):723–728.

Kjellström, E., Bärring, L., Gollvik, S., Hansson, U., Jones, C., Samuelsson, P., Rummukainen, M., Ullerstig, A., Willén, U., and Wyser, K. (2005). A 140-year simulation of european climate with the new version of the rossby centre regional atmospheric climate model (rca3).

Klein, W. H. and Glahn, H. R. (1974). Forecasting local weather by means of model output statistics. *Bulletin of the American Meteorological Society*, 55(10):1217–1227.

Klein Tank, A. and Lenderink, G. (2009). Climate change in the netherlands; supplements to the knmi'06 scenarios. *De Bilt, Koninklijk Nederlands Meteorologisch Instituut*.

Kottegoda, N. T. and Rosso, R. (1997). Probability, statistics, and reliability for civil and environmental engineers.

Koutsoyiannis, D. and Onof, C. (2001). Rainfall disaggregation using adjusting procedures on a poisson cluster model. *Journal of Hydrology*, 246(1):109–122.

Kupiainen, M., Samuelsson, P., Jones, C., Jansson, C., Willén, U., Hansson, U., Ullerstig, A., Wang, S., and Döscher, R. (2011). Rossby centre regional atmospheric model, rca4. *Rossby Centre Newsletter, June*.

Laio, F. (2004). Cramer–von mises and anderson-darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*, 40(9).

Lettenmaier, D. P. and Potter, K. W. (1985). Testing flood frequency estimation methods using a regional flood generation model. *Water Resources Research*, 21(12):1903–1914.

Lettenmaier, D. P., Wallis, J. R., and Wood, E. F. (1987). Effect of regional heterogeneity on flood frequency estimation. *Water Resources Research*, 23(2):313–323.

Ligda, M. G. (1951). Radar storm observation. *Compendium of meteorology*, pages 1265–1282.

Limpert, E., Stahel, W. A., and Abbt, M. (2001). Log-normal distributions across the sciences: Keys and clues on the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question. *BioScience*, 51(5):341–352.

Lo Conti, F., Noto, L., La Loggia, G., and Cannarozzo, M. (2007). Regional frequency analysis of extreme precipitation in sicily, italy. *Variability in space and time of extreme rainfalls, floods and droughts*.

Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction.

Louis, J.-F. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, 17(2):187–202.

Madsen, H., Mikkelsen, P. S., Rosbjerg, D., and Harremoës, P. (2002). Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics. *Water Resources Research*, 38(11):21–1.

Madsen, H., Pearson, C. P., and Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. regional modeling. *Water Resources Research*, 33(4):759–769.

Magilligan, F. J. and Graber, B. E. (1996). Hydroclimatological and geomorphic controls on the timing and spatial variability of floods in new england, usa. *Journal of Hydrology*, 178(1):159–180.

Majewski, D. (1991). The europa-modell of the deutscher wetterdienst. In *ECMWF Seminar on numerical methods in atmospheric models*, volume 2, pages 147–191.

Marani, M. and Zanetti, S. (2007). Downscaling rainfall temporal variability. *Water resources research*, 43(9).

Maraun, D. (2012). Nonstationarities of regional climate model biases in european seasonal mean temperature and precipitation sums. *Geophysical Research Letters*, 39(6).

Maraun, D., Osborn, T., and Gillett, N. (2008). United kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006. *International Journal of Climatology*, 28(6):833–842.

Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48(3).

Mardia, K. V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 349–393.

Mariotti, A., Struglia, M. V., Zeng, N., and Lau, K. (2002). The hydrological cycle in the mediterranean region and implications for the water budget of the mediterranean sea. *Journal of climate*, 15(13):1674–1690.

Meehl, G. A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., and Whetton, P. (2000). Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change*. *Bulletin of the American Meteorological Society*, 81(3):427–436.

Meier, H., Döscher, R., and Faxén, T. (2003). A multiprocessor coupled ice-ocean model for the baltic sea: Application to salt inflow. *Journal of Geophysical Research: Oceans (1978–2012)*, 108(C8).

Miller, J. (1972). Physiographically adjusted precipitation-frequency maps: distribution of precipitation in mountainous areas. *WMO Publ*, 326:264–277.

Mills, G. F. (1995). Principal component analysis of precipitation and rainfall regionalization in spain. *Theoretical and Applied Climatology*, 50(3-4):169–183.

Moisello, U. (1999). *Idrologia tecnica*. La goliardica pavese.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., et al. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756.

Muerth, M., Gauvin St-Denis, B., Ricard, S., Velázquez, J., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R., and Turcotte, R. (2013). On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences*, 17(3):1189–1204.

Nakicenovic, N. and Swart, R. (2000). Special report on emissions scenarios. *Special Report on Emissions Scenarios, Edited by Nebojsa Nakicenovic and Robert Swart, pp. 612. ISBN 0521804930. Cambridge, UK: Cambridge University Press, July 2000.*, 1.

Neal, R. A. and Phillips, I. D. (2009). Summer daily precipitation variability over the east anglian region of great britain. *International Journal of Climatology*, 29(11):1661–1679.

NERC, N. E. R. C. (1975). *Flood studies report*, volume I—Hydrologic studies. NERC, London.

Nguyen, V., Nguyen, T., and Ashkar, F. (2002). Regional frequency analysis of extreme rainfalls. *Water Science & Technology*, 45(2):75–81.

Nguyen, V.-T.-V., In-na, N., and Bobee, B. (1989). New plotting-position formula for pearson type-iii distribution. *Journal of Hydraulic Engineering*, 115(6):709–730.

Nikulin, G., Kjellström, E., Hansson, U., Strandberg, G., and Ullerstig, A. (2011). Evaluation and future projections of temperature, precipitation and wind extremes over europe in an ensemble of regional climate simulations. *Tellus A*, 63(1):41–55.

Norbiato, D., Borga, M., Sangati, M., and Zanon, F. (2007). Regional frequency analysis of extreme precipitation in the eastern italian alps and the august 29, 2003 flash flood. *Journal of hydrology*, 345(3):149–166.

NRCC (1998). *Decade-to-century-scale climate variability and change: a science strategy*. National Academies Press.

Peel, M. C., Wang, Q., Vogel, R. M., and McMAHON, T. A. (2001). The utility of l-moment ratio diagrams for selecting a regional probability distribution. *Hydrological sciences journal*, 46(1):147–155.

Piani, C., Haerter, J., and Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over europe. *Theoretical and Applied Climatology*, 99(1-2):187–192.

Piazza, A. D., Conti, F. L., Viola, F., Eccel, E., and Noto, L. V. (2015). Comparative analysis of spatial interpolation methods in the mediterranean area: application to temperature in sicily. *Water*, 7(5):1866–1888.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

Pilgrim, J. M., Fang, X., and Stefan, H. G. (1998). Stream temperature correlations with air temperatures in minnesota: implications for climate warming. *Journal of the American Water Resources Association*, 34(5):1109–1122.

Potter, K. W. and Lettenmaier, D. P. (1990). A comparison of regional flood frequency estimation methods using a resampling method. *Water Resources Research*, 26(3):415–424.

Pozo-Vázquez, D., Esteban-Parra, M., Rodrigo, F., and Castro-Diez, Y. (2001). The association between enso and winter atmospheric circulation and temperature in the north atlantic region. *Journal of Climate*, 14(16):3408–3420.

Price, C., Stone, L., Huppert, A., Rajagopalan, B., and Alpert, P. (1998). A possible link between el nino and precipitation in israel. *Geophysical Research Letters*, 25(21):3963–3966.

Prudhomme, C. and Reed, D. W. (1999). Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in scotland. *International Journal of Climatology*, 19(12):1337–1356.

Reale, O. and Atlas, R. (2001). Tropical cyclone-like vortices in the extratropics: observational evidence and synoptic analysis. *Weather and forecasting*, 16(1):7–34.

Reed, D. (1994). Plans for the flood estimation handbook. In *Proceedings of The MAFF Conference of Ri6er and Coastal Engineers*, pages 8–3.

Reed, D. et al. (1999). *Procedures for flood freequency estimation, Volume 3: Statistical procedures for flood freequency estimation*. Institute of Hydrology.

Rockel, B., Will, A., and Hense, A. (2008). The regional climate model cosmo-clm (cclm). *Meteorologische Zeitschrift*, 17(4):347–348.

Rodó, X., Baert, E., and Comin, F. (1997). Variations in seasonal rainfall in southern europe during the present century: relationships with the north atlantic oscillation and the el niño-southern oscillation. *Climate Dynamics*, 13(4):275–284.

Rodo, X., Pascual, M., Fuchs, G., and Faruque, A. (2002). Enso and cholera: a nonstationary link related to climate change? *Proceedings of the national Academy of Sciences*, 99(20):12901–12906.

Rodriguez-Iturbe, I., Cox, D., and Isham, V. (1988). A point process model for rainfall: further developments. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 417, pages 283–298. The Royal Society.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., et al. (2003). The atmospheric general circulation model echam 5. part i: Model description.

Roeckner, E., Bengtsson, L., Feichter, J., Lelieveld, J., and Rodhe, H. (1999). Transient climate change simulations with a coupled atmosphere-ocean gcm including the tropospheric sulfur cycle. *Journal of Climate*, 12(10):3004–3032.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D. (2011). Improving pan-european hydrological simulation of extreme eventsthrough statistical bias correction of rcm-driven climatesimulations. *Hydrology and Earth System Sciences*, 15(8):2599.

Rossi, F., Fiorentino, M., and Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, 20(7):847–856.

Rossi, F. and Villani, P. (1994). Valutazione delle onde di piena in campania. *Rapporto Regionale Campania, CNR-GNDCI.*

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.

Rummukainen, M., Räisänen, J., Bringfelt, B., Ullerstig, A., Omstedt, A., Willén, U., Hansson, U., and Jones, C. (2001). A regional climate model for northern europe: model description and results from the downscaling of two gcm control simulations. *Climate Dynamics*, 17(5-6):339–359.

Salas-Mélia, D., Chauvin, F., Déqué, M., Douville, H., Gueremy, J., Marquet, P., Planton, S., Royer, J., and Tyteca, S. (2005). Description and validation of the cnrm-cm3 global coupled model. *Climate Dynamics*, 103:1–36.

Samuelsson, P., Jones, C. G., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E., Nikulin, G., and Wyser, K. (2011). The rossby centre regional climate model rca3: model description and performance. *Tellus A*, 63(1):4–23.

Sanchez-Gomez, E., Somot, S., and Mariotti, A. (2009). Future changes in the mediterranean water budget projected by an ensemble of regional climate models. *Geophysical Research Letters*, 36(21).

Schaefer, M. (1990). Regional analyses of precipitation annual maxima in washington state. *Water Resources Research*, 26(1):119–131.

Scholz, F. W. and Stephens, M. A. (1987). K-sample anderson–darling tests. *Journal of the American Statistical Association*, 82(399):918–924.

Sevruk, B. (1985). Systematischer niederschlagsmessfehler in der schweiz. *Beiträge zur Geologie der Schweiz–Hydrologie*, 31:65–75.

Sillmann, J., Kharin, V., Zwiers, F., Zhang, X., and Bronaugh, D. (2013). Climate extremes indices in the cmip5 multimodel ensemble: Part 2. future climate projections. *Journal of Geophysical Research: Atmospheres*, 118(6):2473–2493.

Sillmann, J. and Roeckner, E. (2008). Indices for extreme events in projections of anthropogenic climate change. *Climatic Change*, 86(1-2):83–104.

Simmons, A. J. and Burridge, D. M. (1981). An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Monthly Weather Review*, 109(4):758–766.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., Powers, J. G., et al. (2008). A description of the advanced research wrf version 3. *NCAR technical note*, 475:113.

Srivastav, R. K., Schardong, A., and Simonovic, S. P. (2014). Equidistance quantile matching method for updating idfcurves under climate change. *Water resources management*, 28(9):2539–2562.

Stedinger, J. and Vogel, R. (1993). E., foufoula-georgiou, 1993, frequency analysis of extreme events. *Handbook of hydrology*, pages 18–1.

Stephenson, D. B., Diaz, H., and Murnane, R. (2008). *Definition, diagnosis, and origin of extreme weather and climate events*, volume 348. Cambridge University Press: New York.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2009). A summary of the cmip5 experiment design. *PCDMI Rep*, 33.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

Tebaldi, C., Hayhoe, K., Arblaster, J. M., and Meehl, G. A. (2006). Going to the extremes. *Climatic change*, 79(3-4):185–211.

Terink, W., Hurkmans, R., Torfs, P., and Uijlenhoet, R. (2009). Bias correction of temperature and precipitation data for regional climate model application to the rhine basin. *Hydrology and Earth System Sciences Discussions*, 6(4):5377–5413.

Terry, M. E. (1952). Some rank order tests which are most powerful against specific parametric alternatives. *The Annals of Mathematical Statistics*, pages 346–366.

Teutschbein, C. and Seibert, J. (2010). Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies. *Geography Compass*, 4(7):834–860.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, 84(9):1205–1217.

Tropeano, R. and Furcolo, P. (2005). The effect of orography on extreme rainfall: a simplified meteo-morphological model. In *Geophysical Research Abstracts*, volume 7, page 10298.

Unden, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., et al. (2002). Hirlam-5 scientific documentation.

Valle, S., Li, W., and Qin, S. (1999). Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to

other methods. *Industrial and Engineering Chemistry Research*, 38(11):4389–4401. cited By 182.

van Meijgaard, E., Van Ulft, L., Lenderink, G., De Roode, S., Wipfler, E. L., Boers, R., and van Timmermans, R. (2012). *Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe*. Number KVR 054/12. KVR.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., et al. (2011). The representative concentration pathways: an overview. *Climatic change*, 109:5–31.

Van Vuuren, D. P., Meinshausen, M., Plattner, G.-K., Joos, F., Strassmann, K. M., Smith, S. J., Wigley, T. M., Raper, S., Riahi, K., De La Chesnaye, F., et al. (2008). Temperature increase of 21st century mitigation scenarios. *Proceedings of the National Academy of Sciences*, 105(40):15258–15262.

Vannitsem, S. and Nicolis, C. (2008). Dynamical properties of model output statistics forecasts. *Monthly Weather Review*, 136(2):405–419.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernández, J., García-Díez, M., Goergen, K., Güttler, I., et al. (2013). The simulation of european heat waves from an ensemble of regional climate models within the euro-cordex project. *Climate dynamics*, 41(9-10):2555–2575.

Viglione, A., Laio, F., and Claps, P. (2007). A comparison of homogeneity tests for regional frequency analysis. *Water Resources Research*, 43(3).

Villi, V., Bacchi, B., and nazionale per la difesa dalle catastrofi idrogeologiche. Linea 1, G. (2001). *Valutazione delle piene nel Triveneto*. s. n.

Vogel, R. M. and Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water Resources Research*, 29(6):1745–1752.

von Storch, H., Langenberg, H., and Feser, F. (2000). A spectral nudging technique for dynamical downscaling purposes. *Monthly weather review*, 128(10):3664–3673.

Wald, A. and Wolfowitz, J. (1943). An exact test for randomness in the non-parametric case based on serial correlation. *The Annals of Mathematical Statistics*, 14(4):378–388.

Wallis, J. R., Schaefer, M. G., Barker, B. L., and Taylor, G. H. (2007). Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for washington state. *Hydrology and Earth System Sciences*, 11(1):415–442.

Wallis, J. R. and Wood, E. F. (1985). Relative accuracy of log pearson iii procedures. *Journal of Hydraulic Engineering*, 111(7):1043–1056.

Wang, Q. (1991). The pot model described by the generalized pareto distribution with poisson arrival rate. *Journal of Hydrology*, 129(1):263–280.

Washington, W. M. and Parkinson, C. L. (2005). *An introduction to three-dimensional climate modeling*. University science books.

Watt, W., Lathem, K., Neill, C., Richards, T., and Rousselle, J. (1989). Hydrology of floods in canada: A guide to planning and design national research council canada. *Ottawa, Ont.*

Waymire, E., Gupta, V. K., and Rodriguez-Iturbe, I. (1984). A spectral theory of rainfall intensity at the meso-scale. *Water Resour. Res*, 20(10):1453–1465.

Weart, S. (2011). Global warming: How skepticism became denial. *Bulletin of the atomic scientists*, 67(1):41–50.

Wigley, T. M., Briffa, K. R., and Jones, P. D. (1984). On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology*, 23(2):201–213.

Wilks, D. S. (1993). Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. *Water Resources Research*, 29(10):3543–3549.

Wilks, S. S. (1963). Multivariate statistical outliers. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 407–426.

Willems, P. (2001). A spatial rainfall generator for small spatial scales. *Journal of Hydrology*, 252(1):126–144.

Willems, P. (2009). A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. *Environmental Modelling & Software*, 24(3):311–321.

Willems, P., Guillou, A., and Beirlant, J. (2007). Bias correction in hydrologic gpd based extreme value analysis by means of a slowly varying function. *Journal of Hydrology*, 338(3):221–236.

WMO, C. f. h. (2008). *The Guide to Hydrological Practices (WMO No.168)*, volume II. World Meteorological Organization, sixth edition edition.

Xoplaki, E. (2002). Climate variability over the mediterranean. *PhD, University of Bern, Switzerland*.

Xoplaki, E., Gonzalez-Rouco, J., Luterbacher, J., and Wanner, H. (2003). Mediterranean summer air temperature variability and its connection to the large-scale atmospheric circulation and ssts. *Climate Dynamics*, 20(7-8):723–739.

Yakir, D. (1996). Fluxes of 602 and water between. *Nature*, 380:11.

Yue, S., Pilon, P., Phinney, B., and Cavadias, G. (2002). The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16(9):1807–1829.

Zafirakou-Koulouris, A., Vogel, R. M., Craig, S. M., and Habermeier, J. (1998). L moment diagrams for censored observations. *Water resources research*, 34(5):1241–1249.

Zalina, M., Desa, M., Nguyen, V., and Kassim, A. (2002). Selecting a probability distribution for extreme rainfall series in malaysia. *Water Science & Technology*, 45(2):63–68.

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.

Zwiers, F. W. and Kharin, V. V. (1998). Changes in the extremes of the climate simulated by ccc gcm2 under co2 doubling. *Journal of Climate*, 11(9):2200–2222.