UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Stochastic space–time models for the characterization of precipitation extreme values

A contribution to the study of the desertification phenomenon

Ana Cristina Marinho da Costa
(Master of Science)

Dissertation submitted for the degree of Doctor of Engineering Sciences

Advisor: Doutor Amílcar de Oliveira Soares

Juri

President: Reitor da Universidade Técnica de Lisboa

Members: Doutor Manuel José Vilares
Doutor Marco Octávio Trindade Painho
Doutor João Alexandre Medina Corte-Real
Doutor Henrique José de Figueiredo Garcia Pereira
Doutor António Jorge Gonçalves de Sousa
Doutor Amílcar de Oliveira Soares

April 2009

Registration n.º 954/2009

ABSTRACT

This study provides a qualitative classification of 107 daily rainfall series and evaluates spatiotemporal patterns in extreme precipitation by calculating climate indices at stations with records within the 1940–1999 period in the south of Portugal, where large areas have high susceptibility to desertification. Two procedures are proposed for the homogenization of climate time series. The direct sequential simulation and cosimulation algorithms, and the developed stochastic space-time models in particular, proved to be valuable procedures to deepen the knowledge on the space-time dynamics of precipitation extremes and to provide uncertainty assessments of the produced scenarios. The spatial patterns of precipitation extremes have become more homogenous during the last decades of the twentieth century, which is consistent with their decreasing relationship with elevation. On the other hand, climate variability is becoming greater in the time dimension. The most intense and more frequent rainfall events occur at areas of the Algarve region. The proposed simple *aridity intensity index* (*AII*) reflects increases in the magnitude of dryness. The southeast region is the most threatened by droughts and extreme dryness. Moreover, there is a tendency towards drier climatic conditions in coastal areas of the south and in the centre of the study region.

Key-words: desertification; homogenization; precipitation extremes; space-time dynamics; climatic trends; uncertainty.

RESUMO

Este estudo fornece uma classificação qualitativa de 107 séries diárias da precipitação e apresenta uma avaliação dos padrões espácio-temporais de valores extremos de precipitação baseada em índices climáticos – calculados para estações de monitorização com registos no período 1940–1999 no sul de Portugal, onde extensas áreas têm elevada susceptibilidade à desertificação. São propostos dois procedimentos para a homogeneização de séries temporais de elementos climáticos. Os algoritmos de simulação e co-simulação sequencial directa e, em particular, os modelos estocásticos espacio-temporais desenvolvidos, provaram ser procedimentos valiosos para aprofundar o conhecimento das dinâmicas espacio-temporais dos extremos de precipitação e para desenvolver indicadores de incerteza dos cenários produzidos. Os padrões espaciais das precipitações extremas tornaram-se mais homogéneos durante as últimas décadas do século vinte. Esta conclusão é consistente com a diminuição do seu relacionamento com a altitude. A variabilidade climática tem vindo a crescer na dimensão temporal. Os eventos de precipitações extremas mais intensos e frequentes ocorrem em áreas do Algarve. O índice *AII* proposto reflecte aumentos na magnitude da aridez. A região sudeste é a mais ameaçada por secas e aridez extrema. Observa-se também uma tendência para condições climáticas mais secas nas áreas costeiras do sul e no centro da região de estudo.

Palavras-chave: desertificação; homogeneização; extremos de precipitação; dinâmicas espacio-temporais; tendências climáticas; incerteza.

ACKNOWLEDGEMENTS

This work would have not been possible without the help and involvement of a large number of people to whom I wish to express my sincere gratitude:

- My supervisor, Professor Amílcar Soares, for his wise and astute suggestions. He not only improved the quality of the work, but also taught me a lot and brought my work into a new level of quality. I will always be grateful to him for all of this, and for giving me the opportunity to work on such an interesting and challenging research field.

- Rita Durão and Prof. Maria João Pereira for the variogram models' parameters of the R5D and R30 indices and for their valuable comments in several papers, which helped improving this work. I am also grateful for their support and enthusiasm.

- Prof. João Negreiros for his suggestions on the implementation of the geostatistical simulation approach that it is proposed for the homogenization of climate data.

- Prof. Pedro Cabral for helping me with the 'distance to the coastline' map.

- All the reviewers and referees that, in different conferences and journal papers, made valuable comments and helped improving the quality of this work.

- To Instituto Superior de Estatística e Gestão de Informação (ISEGI), Universidade Nova de Lisboa (UNL), for discharging me from lecturing classes during the 2007/08 academic year, thus providing me precious time to finish the thesis.

- Helena Guerra for her support, friendship and availability to help me with numerous assignments. She made my life easy at ISEGI–UNL in so many ways.

- Ana Marujo and all my colleagues for their encouraging support. A special word for my colleagues at Centro de Modelização de Reservatórios Petrolíferos (CMRP), Instituto Superior Técnico (IST), who are true wizards of geostatistics, for being always available to help.

- Finally, I would like to thank my dear parents for their love and support, especially my mother as she endured the most. Their confidence in me is not only a source of great comfort but most of all it is the drive to improve myself.

ACRONYMS AND ABREVIATIONS

ANN – Artificial neural network

CAI – Climatologically aided interpolation

CCD – Convention to Combat Desertification

cdf – Cumulative distribution function

CLIVAR – Climate Variability and Predictability project

CCoK – Collocated cokriging

coDSS – Direct sequential cosimulation

CoK – Ordinary cokriging

CV – Coefficient of variation

DEM – Digital elevation model

DSS – Direct sequential simulation

EGLS – Estimated generalized least squares

FAO – Food and Agriculture Organization of the United Nations

GEV – Generalized Extreme Values distribution

ECA – European Climate Assessment

ESAI – Environmentally Sensitive Areas Index (defined within the framework of the MEDALUS project to assess Environmentally Sensitive Areas to desertification in Europe)

GIS – Geographical information system

IDW – Inverse distance weighting

INAG – Portuguese Institute for Water (Instituto da Água)

KED – Kriging with an external drift

MAD – Median-absolute-deviation: is the median of the absolute values of the deviations of the sample values from the median

MAE – Mean absolute error

MASH – Multiple analysis of series for homogenisation: method developed in the Hungarian Meteorological Service

ME – Mean bias error

NAO – North Atlantic Oscillation

NAP – National Action Programme

NDVI – Normalized Difference Vegetation Index

OK – Ordinary kriging

OLS – Ordinary least squares

pdf – Probability density function

RMSE – Root mean square error

SGS – Sequential Gaussian simulation

SIS – Sequential indicator simulation

SK – Simple kriging

SKlm – Simple kriging with varying local means

SNHT – Standard normal homogeneity test

SNIRH – National System of Water Resources Information (Sistema Nacional de Informação de Recursos Hídricos), managed by INAG

STD – Standard-deviation

SUR – Seemingly unrelated regression

TAE – Temporal average of extremes series (moving average series of the extreme precipitation indices)

TVE – Temporal variability of extremes series (moving standard-deviation series of the extreme precipitation indices)

UNCCD – United Nations Convention to Combat Desertification

UNEP – United Nations Environmental Program

WMO – World Meteorological Organization

WMO-CCL – World Meteorological Organization Commission for Climatology

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: INTRODUCTION

1. Introduction

1.1 Problem statement and rationale

The most common definition of desertification is the adopted by the United Nations Convention to Combat Desertification (UNCCD 1994, p. 4) which defines desertification as "land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors, including climatic variations and human activities".

Climate has an important role on desertification processes through its impacts on dryland soils and vegetation, on dryland hydrology and, ultimately, on human land use (Benson *et al*., 1997; Geist and Lambin, 2004; Gringof and Mersha, 2006; Schreiber *et al*., 2008). The consequences of climate change, either caused by physical or anthropogenic factors, can be more significant at regional and local scales and have serious consequences (Correia, 2004). Therefore, research on climatic variability and trends is of major importance to understand desertification processes, especially at the regional and local scales.

The spatial, seasonal and inter-annual variability of rainfall follows a complex pattern in Mediterranean regions, where the environment is subject not only to droughts, but also to flooding and erosion phenomena caused by high intensity rainfalls. Those heavy downpours, which often occur after a very dry summer, and the high climatic fluctuations in short- and long-term have been pointed out as the main climatic characteristics affecting the vulnerability of the Mediterranean basin to erosion (Martínez-Casasnovas *et al*., 2002).

In Mediterranean climate regions, prolonged periods of unusually dry conditions reduce the availability of water resources and affect vegetation cover; while other areas can be affected by an increase in the number of heavy precipitation events, with an increase in the flood risk (e.g., Hidalgo *et al*., 2003a). Extreme precipitation situations such as drought and erosive rainfall events have been raising concern about the risks of land degradation and desertification in such regions (De Luís *et al*., 2001; Lázaro *et al*., 2001; Pereira *et al*., 2006). Drought periods amplify the soil erosion, damage the vegetation cover, reduce the water resources, increase the vulnerability to salinization, exhaust and degrade the agricultural lands, among other phenomena that cause land degradation (Correia, 2004). During wetter

periods, high intensity rainfall is the most important contributor to erosion in drylands (WMO, 2005).

Recent studies, based on climate models and past observed records, predict a drier climate in southern Europe as a result of increased evapotranspiration and a relatively slow decrease of rainfall amounts and precipitation frequency (Cubasch *et al*., 1996; Kostopoulou and Jones, 2005; IPCC, 2007; Vicente-Serrano and Cuadrat-Prats, 2007). The precipitation amounts are also projected to decrease in continental Portugal, especially in southern regions (Miranda *et al.*, 2006; Sillmann and Roeckner, 2008). These circumstances will potentially amplify the vulnerability of several Mediterranean regions to desertification by increasing their environmental problems, such as agro-forestry-grazing productivity, soil degradation, aquifers recharge, forest fires, biological diversity and composition (De Luís *et al*., 2001; Ceballos *et al*., 2004).

Portugal ratified the United Nations Convention to Combat Desertification on April 1, 1996. The National Action Programme (NAP) was approved by the Government on June 17, 1999 (PANCD, 1999). The susceptibility map of desertification of the NAP identifies vulnerable areas through an indicator combining four quality indices: climate, soil, vegetation, and management/land use (Rosário, 2004b). The susceptibility map of desertification shows that, under the mean climatic regime evaluated, the south of the country has extensive areas highly vulnerable to desertification.

The rainfall regime of the south of continental Portugal is Mediterranean with Atlantic influence, so it is highly variable in both the spatial and temporal dimensions. Accordingly, the climate is characterised by a dry and very hot season, and a very irregular distribution of precipitation over the wet season, as well as over the years, with very intense flood peaks and with frequent drought periods.

The climate quality component (aridity index map) of the Portuguese susceptibility map of desertification is not fully appropriate to map the areas susceptible to desertification because it is based on average computations (Pereira *et al*., 2006). Consequently, the development of indicators accounting for the dynamics of desertification processes and drought, particularly the trends and variability hidden behind "the means", should be further addressed (Rosário, 2004b; Pereira *et al*., 2006). Moreover, indicators of the uncertainty of the produced maps should also be developed (Rosário, 2004b).

In semi-arid and dry subhumid regions, such as the south of Portugal, research on the extent of dryness and space-time patterns of heavy rainfall events is an important contribution to evaluate desertification dynamics and to identify areas potentially at risk from land degradation. The impacts of climate change and variability on natural systems and human activities are usually experienced at the regional or local scale. However, studies focusing on the role of regional climate changing on erosive and aridity factors are lacking for this region, especially at the local scale (Rosário, 2004b; Pereira *et al*., 2006). Moreover, further work is necessary to obtain a complete view of the spatiotemporal variability of daily rainfall (Rodrigo and Trigo, 2007).

According to the previous discussion, the characterization of temporal trends and space-time patterns in extreme precipitation is expected to be a relevant contribution, not only for the assessment of local climate dynamics, but also to the study of the desertification phenomenon.

This thesis aims at increasing the knowledge on the time and space-time dynamics of extreme precipitation, in the south of Portugal, through the analysis of historical records of observations at meteorological stations.

1.2 Scientific background

Precipitation is one of the most important climate variables. Accurate quantification of its observed variability is required for a number of purposes. Information on the spatial variability of extreme precipitation is important for river basins management, flood hazards protection, studies related to climate change, erosion modelling and other applications for ecosystem and hydrological impact modelling.

Long-term series of *reliable* precipitation records are essential to develop these studies. The detection of temporal discontinuities in climate data is of major importance, because non-climatic factors may hide the true climatic signal and patterns, and thus potentially bias the conclusions of climate and hydrological studies. Therefore, the World Meteorological Organization (WMO) recommends that, besides routine quality control, the homogeneity testing of climate data should be evaluated before performing those studies (Aguilar *et al*., 2003).

Identification of changes in the occurrence of extreme weather events requires accurate and spatially consistent climatic time series with at least daily resolution (Easterling *et al.*, 1999; Peterson, 2005), since only daily and higher-resolution series account for the submonthly time scale nature of extreme weather events (Klein Tank *et al*., 2002). Precipitation measurements are particularly susceptible to irregularities that may affect the analysis of extreme precipitation events (Easterling *et al.*, 1999). For instance, station relocations may cause an artificial change in observed extremes, particularly in areas of heavy rain.

Climate extremes are events rarely observed and statistically correspond to the tails of the distribution of the climate variable. Changes in extremes correspond to changes in the distribution (location, scale and/or shape) of the variable. The analysis of changes in extremes can be performed by fitting appropriate theoretical distribution functions, named GEV (Generalized Extreme Values) distributions, to the observed daily climate data and then investigate the changes in the parameters of the distribution functions over time or space (e.g. Weisse and Bois, 2001; Beguería and Vicente-Serrano, 2006).

Another approach is based on the analysis of changes in climate indices, which are estimated from the empirical distribution of the daily observations. To gain a uniform perspective on observed changes in climate extremes, a core set of standardized indices was defined by the joint working group on climate change detection of the World Meteorological Organization – Commission for Climatology (WMO–CCL) and the Research Programme on Climate Variability and Predictability (CLIVAR, Peterson *et al*., 2001; Frich *et al*., 2002; Peterson, 2005). In general, these indicators represent events that occur several times per season or year giving them more robust statistical properties than measures of extremes that are far enough into the tails of the distribution so as not to be observed during some years (Frich *et al*., 2002; Alexander *et al*., 2006).

Each index describes particular aspects of climate extremes. Some indices are based on statistical quantities such as percentiles, which are more appropriate for regions that contain a broad range of climates (Haylock and Nicholls, 2000; Klein Tank and Können, 2003). Other indices involve fixed thresholds, such as the number of days per year with daily precipitation exceeding 10 mm or 20 mm (e.g. Klein Tank and Können, 2003; Kostopoulou and Jones, 2005). Indices based on fixed thresholds are beneficial for impact studies as they can be related with extreme events that affect human society or result in strong adverse effects on the

natural environment (Klein Tank and Können, 2003). Extreme impact events may be of short duration, but could also extend over several days, several months or perhaps even years (e.g. droughts).

An approach based on the later set of indices, and indices describing events with short return periods (moderate climate extremes), is suitable for the purposes of this research, because it allows assessing climate dynamics related to impact studies such as those concerning the desertification phenomenon, among other applications for ecosystem and hydrological impact modelling.

1.3 Objectives

The main objective of this research work is to characterize the time and space-time dynamics of precipitation indices, describing both wet and dry extreme values, at local scales in the south of Portugal.

A number of detailed objectives emerge from this general objective:

- To compile a daily rainfall database for the south of Portugal.
- To assess the homogeneity of the precipitation time series of the compiled dataset.
- To use the results of the homogenization analysis to provide a qualitative classification of the reliability of the daily rainfall series for studies on climatic variations and change as well as impact studies.
- To compute annually defined indices of precipitation extremes, able to provide information on the climatic 'wetness' and 'dryness', using the daily rainfall database for the south of Portugal.
- To investigate the trends and temporal dynamics of those indices of precipitation extremes.
- To produce space-time scenarios for the characterization of the frequency and magnitude of extreme precipitation in the last decades of the 20th century, aiming to
  - determine whether extreme precipitation events have distinct spatial patterns regarding the annual temporal evolution and variability in the southern region of

Portugal where the geographic and climate characteristics have a reasonable degree of diversity.

- To produce space-time scenarios for the characterization of dryness in the last decades of the 20th century, aiming to
  - determine whether dry conditions have distinct spatial patterns regarding the annual temporal evolution and variability in the southern region of Portugal where extensive areas are characterized by scarce precipitation and frequent drought periods.
- To provide an uncertainty evaluation of the produced space-time scenarios.
- To use those scenarios to produce an additional set of maps of indicators summarizing their underlying space-time dynamics.

1.4 Research questions

This thesis adds to our knowledge on the time and space-time dynamics of extreme precipitation by addressing the key question:

❖ How did the extremes of daily precipitation change, locally, in the south of Portugal through the last decades 20th century, and what can we learn from this?

A number of research questions emerge from this general question:

- Are the available observational datasets adequate to analyse precipitation extremes?
- Which annual trends are observed for the daily extremes of precipitation?
- How did the annual spatial patterns of extreme precipitation change through the last decades of the 20th century?
- Which are the patterns of spatial uncertainty associated with those space-time patterns?
- Do the produced space-time scenarios allow determining spatial patterns of annual trends in extreme precipitation?
- Do the produced space-time scenarios allow determining areas highly prone to frequent extreme precipitation events?

- Do the produced space-time scenarios allow determining areas highly prone to heavy precipitation events?
- Do the produced space-time scenarios allow determining which areas are more prone to extreme dryness?

1.5 Methodology and research outline

The research work was organized in three major stages. The first one comprised the homogenization analysis of the daily rainfall series. The second stage was dedicated to the analysis of the time component of extreme precipitation indices. Finally, the third stage corresponded to the development of stochastic space-time models for the characterization of indices of precipitation extremes. The following sections summarize the methodological framework of each research stage.

1.5.1 Homogenization of precipitation series

One of the hypotheses of this research was that the daily rainfall series, which were collected for the characterization of extreme precipitation values in the south of continental Portugal, could contain potential inhomogeneities. This assumption is recommended by Auer *et al.* (2005) even for studies that use previously quality controlled data.

Several techniques have been developed for non-climatic inhomogeneities detection and adjustment, i.e. homogenization. A review of different methods is presented by Peterson *et al*. (1998), and comparisons between procedures are provided by Ducré-Robitaille *et al*. (2003) and Reeves *et al*. (2007).

The homogeneity assessment of 107 daily precipitation series was developed through four major stages. The first one aimed at the identification of errors and suspicious daily precipitation records, which were flagged using several criteria. The emphasis of this stage was on the quality control of precipitation extremes. The following stages were dedicated to homogeneity testing and used as the testing variable the annual wet day count with 1 mm as threshold, which is expected to be representative of important characteristics of variation at the daily scale (Wijngaard *et al*., 2003). The second stage was an absolute approach

comprising the application of six statistical tests to the testing variable, at all locations. In the following stages, 62% of the long-term series were also checked through relative approaches (testing procedures that use records from neighbouring reference stations), which included the application of five homogeneity techniques that are capable of locating the year where a break is likely.

In the third stage, besides the application of three well-established statistical methods, a new procedure was proposed for the detection of non-climatic irregularities. The proposed technique is an extension of the Ellipse test (described by Allen *et al.*, 1998) that takes into account the contemporaneous relationship between several candidate series from the same climatic area (Costa and Soares, 2006). This procedure uses the residuals from a *seemingly unrelated regression equations* (SUR) model, thus it was named *SUR+Ellipse test*.

Finally, in the fourth stage, a geostatistical simulation approach based on the direct sequential simulation (DSS) algorithm (Soares, 2001) was proposed for the homogenization of climate time series. The proposed technique accounts for the joint spatial and temporal dependence between observations, and enhances the pre-eminence of the closer stations, in both spatial and correlation terms (Costa *et al.*, 2008a).

Following the hybrid approach proposed by Wijngaard *et al*. (2003), we did not attempt to remove non-climatic inhomogeneities from the daily precipitation series, but rather provide a qualitative classification of each station's records. Therefore, the results of the homogenization analysis were used to develop an overall classification of the daily series.

1.5.2 Trends in indices of precipitation extremes

The second research stage aimed at investigating the existence of trends and other temporal patterns in extreme precipitation indices, within the period 1955–1999, at 15 monitoring stations located in southern Portugal. This 45-year period was chosen to optimize data availability across the region, taking into consideration the homogenization analysis performed. Among the numerous indices of extreme precipitation described in recent literature (e.g., Frich *et al*., 2002; Kiktev *et al*., 2003; Klein Tank and Können, 2003; Kostopoulou and Jones, 2005; Moberg and Jones, 2005), we selected four of them (SDII, R5D, R30 and CDD) and developed two other indices (AII and FDD). Three of the indices (SDII, R5D and R30) provide information on the ‘wetness’, whereas the other three (CDD,

AII and FDD) characterize the ‘dryness’. All indicators are based on fixed thresholds and most of them describe moderate climate extremes.

The SDII monitors precipitation intensity on wet days, the R5D index is a measure of short-term precipitation intensity, and the R30 index characterizes the frequency of extremely heavy precipitation days. The CDD index characterizes the maximum length of a dry spell, the FDD index measures the frequency of dry spells, and the AII can be interpreted as a simple aridity index because it is a numerical indicator of the degree of dryness of the climate at a given location.

The six daily precipitation indices were subject to a number of diagnosis tests in order to verify the existence of autocorrelation and heteroscedasticity of the regression errors. Depending on the tests' results, trend estimation was performed using three different regression models, namely the *simple linear regression model*, the *autoregressive error model* and a *heteroscedastic linear model*. Moving window statistics (mean and standard deviation) of the precipitation indices were also computed to reduce random fluctuations and provide a clearer view of their underlying behaviour, such as non-linear trends or periods with distinct climatic variability.

1.5.3 Space-time dynamics of precipitation extremes

It is long recognized that topography and other geographical factors are responsible for considerable spatial heterogeneity of the precipitation distribution at the sub-regional scale (e.g., Martínez-Cob, 1996). Therefore, many authors attempted to incorporate elevation and other physiographic features into the spatial interpolation of rainfall fields. Some examples are multivariate geostatistics such as kriging with external drift or cokriging (Goovaerts, 2000; Nicolau, 2002; Lloyd, 2005), techniques combining distance weighting methods and regression (Faulkner and Prudhomme, 1998), splines (Hutchinson, 1995), and local regressions (Daly *et al*., 1994).

However, interpolation usually leads to a smoothing of the distribution inferred by the observations and thus to a loss of variance. To overcome this limitation, geostatistical stochastic simulation has become a widely accepted procedure to reproduce the spatial variability and uncertainty of highly variable phenomena (e.g., Bourennane *et al*., 2007). Among the sequential algorithms of geostatistical stochastic simulation, one advantage of

direct sequential simulation and cosimulation (Soares, 2001) is precisely the use of original variables instead of transformed ones: Gaussian (sequential Gaussian simulation) or indicator (sequential indicator simulation). Direct sequential simulation and cosimulation have been applied in several soil and air quality characterization studies (e.g., Franco *et al*., 2006; Russo *et al*., 2008).

For the last research stage, a subset of three indicators was selected for the characterization of the space-time dynamics of extreme precipitation in southern Portugal in the 1940–1999 period: two indices describing wet conditions (R5D and R30) and the proposed AII index, which characterizes dry conditions. For exploratory purposes, the space-time patterns of the R20 index were also analyzed for the 1970–1999 period, and uncertainty was assessed (Costa and Soares, 2007). This indicator characterizes the frequency of heavy precipitation days. Additionally, the spatial patterns of the CDD index were also investigated, but this index was discarded due to the spatial inconsistencies found.

For the spatial interpolation and uncertainty assessment of extreme precipitation, we explore the application of direct sequential cosimulation, which allows incorporating covariates such as elevation. The choice of cosimulation follows the premises that elevation and precipitation may interact differently not only in space, but also during drier and wetter periods (Goovaerts, 2000; Costa and Soares, 2007). Accordingly, the methodology not only accounts for local data variability by using stochastic simulation procedures, but also incorporates space-time models that allow capturing long-term trends of extreme precipitation, and local correlations between elevation and precipitation through time. Elevation was used as secondary information, but other physiographic features were also investigated.

The direct sequential cosimulation was performed for generating one map per year for the two wetness indices (R5D and R30), using 800 m × 800 m grid cells and elevation as exhaustive secondary information. For the dryness index (AII), direct sequential simulation was used instead, because no relevant correlations were found with physiographic features.

1.6 Thesis organization

The thesis manuscript is organized in six chapters and appendixes. The first chapter, which ends with a brief overview of each of the main chapters, discusses the motivation and

relevance of the research work, states the objectives and the research questions that the thesis attempts to address, and provides an indication of how the work progressed and which methods were used.

Chapter 2 is dedicated to the literature review. Taking into consideration the main motivation of the work, this chapter starts with an overview of a number of issues related to the desertification phenomenon (Section 2.1), including a discussion on the interactions of desertification and climate and the assessment of desertification in Portugal. Section 2.2 reviews the characteristics of the procedures that are most commonly used for the homogenization of climate time series, and discusses the problems that arise when the homogenization has to be performed for daily precipitation time series and precipitation extremes. Section 2.3 discusses a number of issues concerning the characterization of precipitation extremes, including an overview of the techniques commonly used in the spatial interpolation of rainfall fields, and a literature review on the analysis of extreme precipitation indices.

Chapter 3 describes the work developed during the first research stage, namely the homogenization analysis of the daily rainfall series. The study region and data are described in Section 3.1. The methodology is detailed in Section 3.2, which includes a brief overview of the basic quality control procedures applied to the daily data, and a thorough description of the statistical tests and techniques used to detect temporal discontinuities in the precipitation time series. Section 3.3 describes and discusses the results of this research stage.

Chapter 4 corresponds to the second research stage, which was dedicated to the evaluation of temporal trends in extreme precipitation by analysing a number of climate indices that are described in Section 4.1. The methodology is detailed in Section 4.2. Section 4.3 describes and discusses the results of this research stage, including the dynamic temporal evolution of precipitation extremes.

The work developed in the last research stage is described in Chapter 5 and corresponds to the characterization of the space-time dynamics of precipitation extremes. The study region and precipitation data are described in Section 5.1, which includes a discussion on the spatial inconsistencies found in the CDD index. The methodology is detailed in Section 5.2, which includes a description of the stochastic space-time models used to produce the scenarios for the indices of precipitation extremes, and the procedures used to produce maps of indicators

summarizing the space-time dynamics of precipitation extremes. Section 5.3 describes and discusses the results, including the space-time relationships between elevation and the indices, and the uncertainty assessment of the space-time scenarios produced.

Chapter 6 states the conclusions of this study by summarizing the main findings and contributions of the thesis, and by pointing out a number of suggestions for further research.

Chapter 2: LITERATURE REVIEW

2. Literature review

Desertification is one of the most complex environmental and socio-economical threatening events. Desertification is caused by multiple interacting factors of climatic, ecological, and socio-economic origin. This multilevel dynamic process can only be monitored, assessed and mapped by studying the dynamics of its individual components and their complex effects (Gringof and Mersha, 2006). Section 2.1 presents an overview of the desertification phenomenon. Despite human factors (e.g., demographics, poverty, market and trade systems, political stability) play a role on desertification processes, the review will be more centred on the physical processes and on the interactions of climate with desertification (Section 2.1.1). The assessment of desertification in Portugal is reviewed in Section 2.1.3.

The aim of climate data homogenization is to adjust observations, if necessary, so that the temporal variations in the adjusted data are caused only by climate processes. A great deal of effort has been made on the last two decades to develop procedures to identify and remove non-climatic inhomogeneities. Section 2.2 reviews the most commonly used homogenization procedures, and discusses the problems that arise when the homogenization has to be performed for daily precipitation time series and precipitation extremes.

The literature review on the characterization of extreme precipitation events is presented in Section 2.3, and mainly focuses works analysing the time and space-time patterns of indicators of extreme precipitation. The relationship between precipitation and physiographic features is discussed in Section 2.3.1, and an overview of the interpolation techniques used to map climate data is presented and discussed in Section 2.3.2. Finally, a review on extreme precipitation indices is presented in Section 2.3.3.

2.1 The desertification phenomenon

The concept of desertification dates to colonial western Africa in the 1920s and 1930s, and has been given more attention in the early 1970s in an attempt to understand a long series of drought years that brought a major economic, social and environmental problem to the African Sahel (Lonergan, 2005). In 1974, the United Nations General Assembly decided to

initiate concerted international action to combat desertification and assist the economic development of affected countries. For the international community, concerns about desertification reached an important threshold at the United Nations Conference on Environment and Development in Rio de Janeiro in 1992, when the Convention to Combat Desertification (CCD) was formulated and more than 60 countries have ratified the convention. On 17 March 2008, Serbia became the 193rd Party to the United Nations CCD (http://www.unccd.int, retrieved 27 March 2008).

The most common definition of desertification is the adopted by the United Nations Convention to Combat Desertification (UNCCD 1994, p. 4) and it is defined as "land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors, including climatic variations and human activities". *Arid, semi-arid and dry sub-humid areas* "means areas, other than polar and sub-polar regions, in which the ratio of annual precipitation to potential evapotranspiration falls within the range from 0.05 to 0.65". This means that desertification is not the natural expansion of existing deserts, but rather the degradation of land in arid, semi-arid, and dry sub-humid areas. Today, the terms desertification and land degradation are used almost interchangeably.

Enne and Zucca (2000) review the fundamental institutional stages and actions undertaken by the international community to understand and combat desertification. Although the issue of desertification has created a number of global environmental institutions and despite the efforts to collect more and better data and information, there are still divergent views regarding the causes and physical extent of desertification (Enne and Zucca, 2000; Lonergan, 2005; Gathara, 2006; Schreiber *et al*., 2008). The combination and interaction of many factors determine the status of desertification, but a number of driving forces are common to all definitions: desertification is a gradual process of soil productivity loss and degradation of the vegetation cover caused by human activities and climatic variations such as prolonged droughts and erosive rainfall.

According to Geist and Lambin (2004), a detailed understanding of the complex set of proximate causes and underlying driving forces affecting dryland-cover change in a given location is required before any assessment and policy intervention. Benson *et al*. (1997), Gathara (2006), Schreiber *et al*. (2008) analyse the causes and impacts of desertification and give a number of suggestions and recommendations to mitigate its effects. Boardman (2006)

reviews the current approaches for the assessment of soil erosion and discusses their limitations.

Pereira *et al*. (2006) have a holistic approach to the desertification phenomenon, based on the Portuguese environmental and socio-economic context: desertification is "considered as a man-induced long-term imbalance in the use of the land, the availability of water, and the opportunities for development, which affects the territory and the populations in arid, semi-arid and sub-humid climates, and combines damaged soil and degraded vegetation, inappropriate land use, mining and degradation of groundwater, increased flash flooding, and a deterioration of the carrying capacity of the ecosystems".

Even though the desertification phenomenon has numerous and complex causes, Schreiber *et al*. (2008) concluded that desertification is driven by a limited group of core variables, most prominently climatic, technological, political (both policy and institutional) and economic factors. However, according to Schreiber *et al*. (2008), the relative importance of physical processes and anthropogenic factors in affecting the global distribution of desertification are numerous and particular to pressures faced by people and agricultural systems in a specific location.

Desertification processes include accelerated soil erosion by wind and water, increasing salinization of soils and near-surface groundwater supplies, a reduction in soil moisture retention, an increase in surface runoff and stream flow variability, the reduction of organic elements in the soil, the plant cover degradation, and a reduction in the overall productivity in dryland ecosystems (Benson *et al*., 1997; Gringof and Mersha, 2006; Schreiber *et al*., 2008).

Gathara (2006) defines desertification as natural and anthropogenic processes involving all forms of degradation of land vulnerable to severe edaphic or climate aridity, thereby leading to the reduction or destruction of biological potential of the land, deterioration of the living standard and intensification of desert like conditions. Gathara (2006) discusses the four main human factors that may be directly involved on desertification processes: overgrazing, loss of vegetation cover, over-cultivation and poor irrigation practices. Regarding the natural causes of desertification, those authors point out the following factors: prolonged droughts (leading to vegetation destruction and loss in biological and economic productivity of the dryland); wind erosion; water erosion (extreme drought periods in many drylands are often followed by

extremely heavy rainfall occurrences that cause extension soil erosion); and climate change (permanent shift in the traditional patterns of climatic parameters).

Geist and Lambin (2004) analysed 132 case studies on the causes of desertification and concluded that it is driven by a limited suite of recurrent core variables, of which the most important at the underlying level are climatic factors, economic factors, institutions, national policies, population growth, and remote influences. These factors drive cropland expansion, overgrazing and infrastructure extension. For each location, a combination of causal factors, in combination with feedback mechanisms and regional land use, make up specific pathways of land change that could trigger desertification.

Gathara (2006) defines twelve key scientific issues in the mitigation of the desertification phenomenon:

- Assessment and mapping; regular assessment and mapping
- Monitoring: current status
- Prediction; future expectation projections
- Database: crucial to any mitigation plan [Web site addresses from a number of databases are provided by Schreiber *et al*. (2008)]
- Research: understanding of the processes
- Information exchange: learning from successes and failures of others
- Training and education
- Transfer of technology
- Funding for desertification and climate research seems to be decreasing. The associated problems are increasing e.g. sound environmental management and sustainable development
- Disaster preparedness: direct/indirect triggers of many desertification drivers
- Planning and management
- Regional and international co-operation Agenda 21 and the three conventions

2.1.1 Interactions of desertification and climate

Climate has an important role on desertification processes through its impacts on dryland soils and vegetation, on dryland hydrology and, ultimately, on human land use (e.g., Geist and Lambin, 2004; Gringof and Mersha, 2006; Schreiber *et al*., 2008). Desertification versus

climate processes are complex and highly interactive, involving various space-time dimensions (Gathara 2006). The analysis of climatic variability and trends is of major importance to understand the desertification processes, especially at the regional and local scales, because the consequences of climate change, either caused by physical or anthropogenic factors, can be more significant at these scales and have consequences that are more serious (Correia, 2004).

Gringof and Mersha (2006) discuss the impacts of drought and other extreme weather events (e.g., torrential rain, wind and water erosion, strong winds and dust storms) on desertification processes. When looking at the climatic factors that cause desertification, Gringof and Mersha (2006) make a distinction between arid climates and drought. In their understanding, "drought is a natural event of nature, caused by atmospheric circulation processes, entailing a prolonged absence of precipitation (or its significant reduction relative to the average norms over several years) in combination with increased air, soil and wind temperatures, leading to a sudden reduction in the corresponding humidity, the exhaustion of soil moisture reserves and disruption of the water balances of plants and animals". Gringof and Mersha (2006) argue that "an arid climate is the climate of an arid area in which, for climatic reasons, the probability of drought is more than 50%, as opposed to the situation in dry areas, where the humidity of atmospheric precipitation is never sufficient for dry-land farming".

Disregarding the possibility of climate change, aridity is a permanent climatic feature of regions of low precipitation, and usually of high temperature, while drought is a temporary feature affecting almost any precipitation or temperature regime around the world. Aridity is usually defined in terms of low-average precipitation, available water, or humidity. Drought, considered in the context of variability, is experienced only when precipitation falls appreciably below normal (Heim, 2002). A review on drought and aridity indicators is presented in the following section.

Figure 2.1 shows a schematic representation of how drought and aridity can favour the desertification susceptibility. Arid regions are particularly vulnerable to the natural phenomenon of drought. However, drought situations may cause the overexploitation of natural resources, through the incorrect use of land and water, resulting in desertification. On the other hand, desertification may in turn amplify the aridity situation, which initiates a cycle where the interaction of natural and human factors increases the land degradation and may

result in a permanent degradation situation (Correia, 2004). These interactions between aridity, drought and desertification are also shown in Figure 2.2, which points out the dynamic features of these desertification processes.

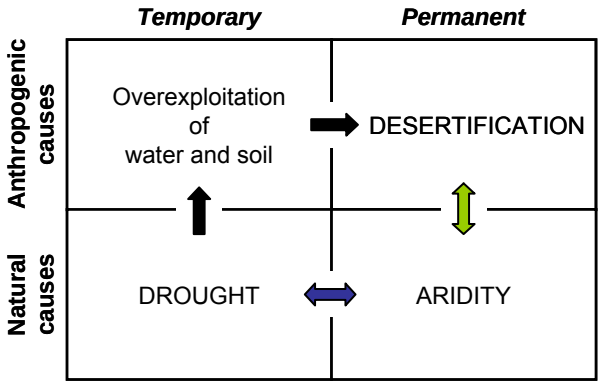

Figure 2.1 – Dynamic cycle showing drought and aridity as factors that favour the desertification susceptibility (adapted from Correia (2004) and Pereira *et al.* (2006))

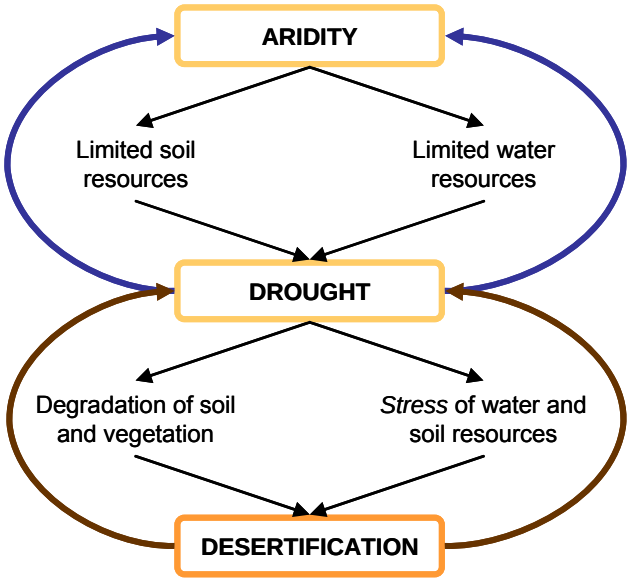

Figure 2.2 – Interactions between aridity, drought and desertification (adapted from Correia, 2004)

According to Geist and Lambin (2004), increased aridity is a robust proximate cause of desertification, both indirectly through greater rainfall variability and directly through prolonged droughts. The effects of increased aridity on land cover include increased fire frequencies and greater soil erosion triggered by more frequent oscillations between warmer, drier conditions and cooler, more humid conditions. However, Geist and Lambin (2004) verified that these effects are reported as causes of desertification only one-third as often as

prolonged drought periods. Climatic factors, mainly associated with a decrease in rainfall are prominent underlying driving forces of desertification. They operate both indirectly, through changes in land use resulting from variation in rainfall, and directly, affecting land cover in the form of prolonged droughts. Drought periods amplify the soil erosion, damage the vegetation cover, reduce the water resources, increase the vulnerability to salinization, exhaust and degrade the agricultural lands, among other phenomena that cause land degradation (Correia, 2004).

Dryland ecosystems are highly responsive to climatic variability. Surface evaporation from dryland soils may result in soil salinization. Removal or loss of vegetation cover results in an increased risk of soil erosion and land degradation, especially during drought periods, because the topsoil becomes more vulnerable to raindrop impact, surface runoff and wind. High rates of erosion reduce the soil profile depth, in turn reducing the total amount of soil moisture that can be stored. This means that the land becomes more prone to drought and floods (Schreiber *et al*., 2008).

During wetter periods, high intensity rainfall is the most important contributor to erosion in drylands (WMO, 2005). According to Kosmas *et al*. (1999), water erosion is controlled by a number of factors interacting with one another (Figure 2.3): climate, vegetation, soil properties and topography. Heavy rainfall events have a greater impact on the start of the growing season, when less protective vegetation exists (Schreiber *et al*., 2008). The rivers from dryland regions have extremely variable flow regimes, and both river discharge and sediment yield are highly sensitive to fluctuations in precipitation (Benson *et al*., 1997).

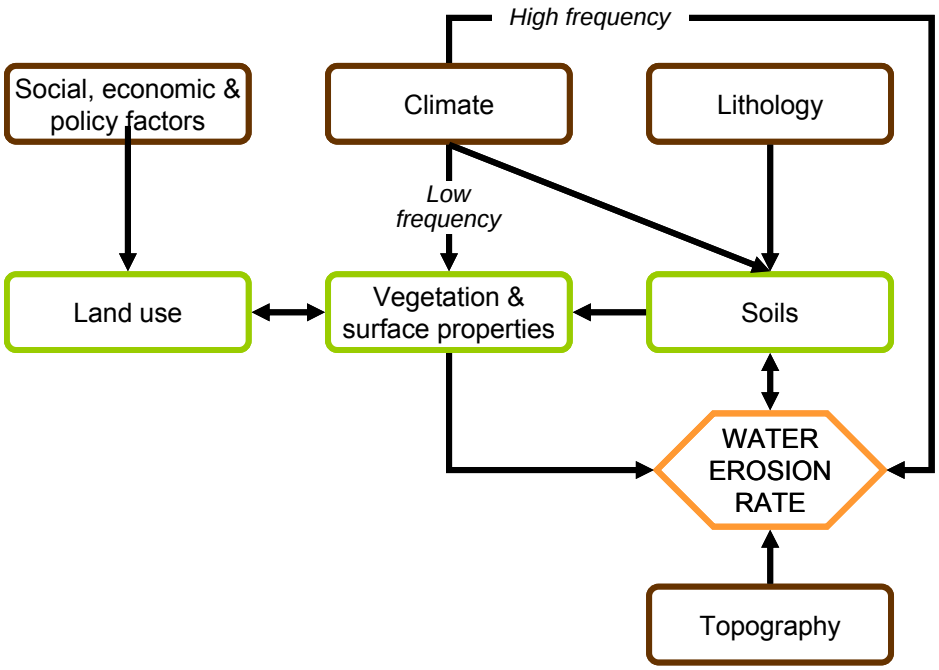

Figure 2.3 – Flow diagram showing the direct and indirect controls on water erosion rates (Kosmas *et al.*, 1999)

Climate change resulting from global warming may exacerbate the problem of desertification since a decrease in the total amount of rainfall in arid and semi-arid areas could increase the total area of dry lands worldwide, and thus the total amount of land potentially at risk from desertification. Drought and climate change strongly aggravate the process of desertification by increasing the pressure on limited natural resources and on vulnerable ecosystems (Pereira *et al*., 2006). The potential broad impacts of climate change in soil and water degradation are generally summarized in Figure 2.4.

Even if the concentrations of all greenhouse gases and aerosols had been kept constant at year 2000 levels, a further warming of about 0.1°C per decade would be expected (IPCC, 2007). An increase in temperature will most probably have the effect of increasing potential evapotranspiration rates in drylands, and in the absence of any large increases in precipitation, many drylands are accordingly predicted to become more arid (Benson *et al*., 1997).

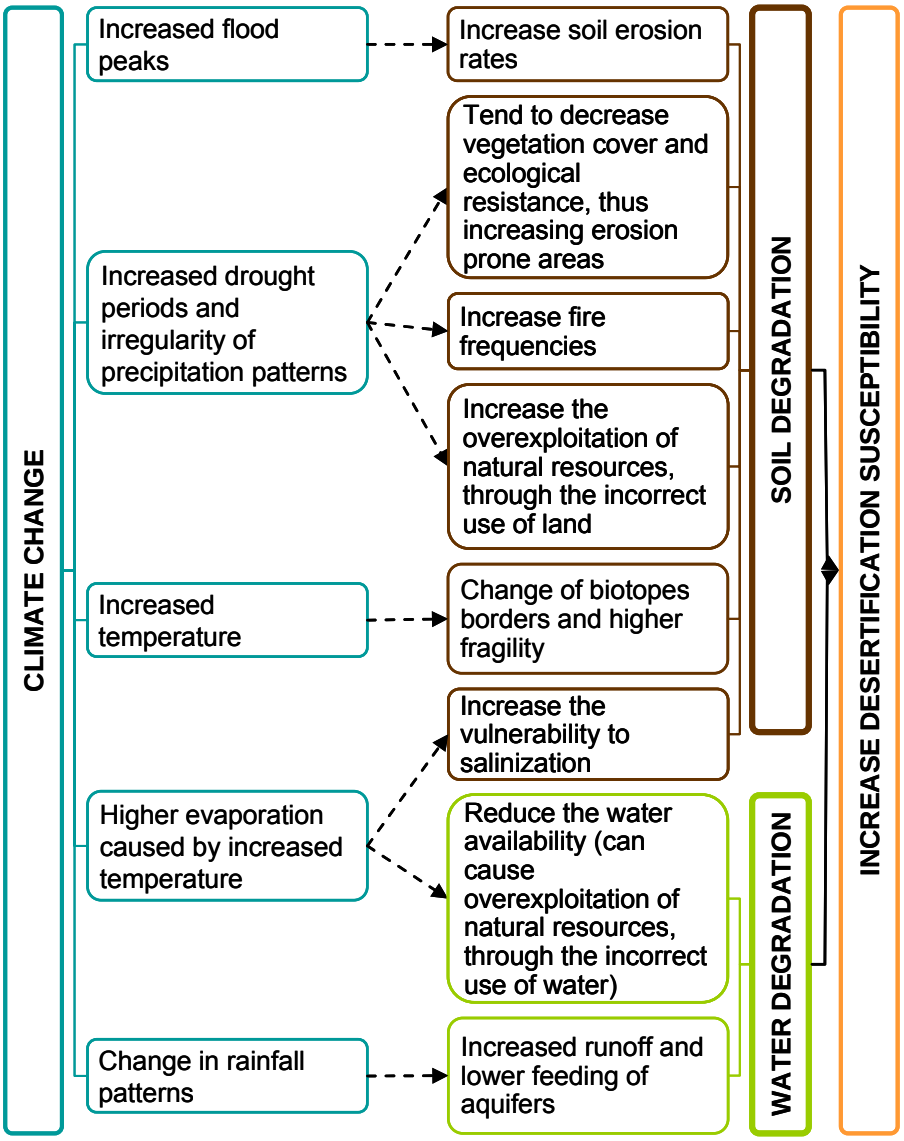

Figure 2.4 – Potential broad impacts of climate change in soil and water that may intensify the desertification susceptibility

Increased drying linked with higher temperatures and decreased precipitation has contributed to changes in drought (IPCC, 2007). Accordingly, drying has been observed in the Sahel, the Mediterranean, southern Africa and parts of southern Asia, which will most likely expand the desertification problem in these regions.

On the other hand, desertification processes (e.g., land use or vegetation losses) may in turn have an influence on climate (Benson *et al*., 1997; Millán *et al*., 2005). Sivakumar (2007) makes a thorough review on the interactions between climate and desertification, and discusses how human activities modify the surface characteristics and atmospheric composition of drylands and how these may influence local and regional dryland climates.

However, the role of land use and land cover changes on climatic variability and change is poorly understood, and will require more research (Herrmann and Hutchinson, 2005; Schreiber *et al*., 2008). Linkages to rainfall changes are even more complex, and the impact of human activities on rainfall remains a topic of considerable debate.

2.1.1.1 Drought and aridity characterization

Rainfall availability is fundamental in drylands. Recurrent low and scarce rainfall lead to drought, which is one of the main climatic driving factors of desertification. Droughts involve significant multiple impacts in terms of crop losses, urban water supply shortages, degradation and desertification, and forest fires. The literature on drought and aridity phenomena is extensive. Reviewing studies on drought classifications and indicators of drought and aridity are provided by Benson *et al*. (1997), Hisdal and Tallaksen (2000), Heim (2002), Gathara (2006). According to Heim (2002), no single index has been able to adequately capture the intensity and severity of drought because of its complexity.

Drought definitions are usually classified into four categories: meteorological, hydrological, agricultural and socio-economic. Drought indicators are variables describing the magnitude, duration, severity, and spatial extent of drought (Steinemann *et al*., 2005). Depending on the classification, drought indicators involve several variables, used either alone or in combination, such as precipitation, temperature, humidity, evaporation from free water, transpiration from plants, potential evapotranspiration, soil moisture, wind, river and stream flow, and plant condition.

There are numerous definitions of meteorological drought and, in general, depend on the climatic features of each region. Some definitions compare the degree of drought with an average amount evaluated over a long period (e.g. climate normal), while others are evaluated in relation to a relevant statistical parameter (e.g. percentiles). Meteorological drought is often expressed based on the degree of dryness and the duration of the dry period. Meteorological indices for drought analysis include elements based on meteorological and hydrological variables, such as precipitation intensity, potential evapotranspiration, potential water balance (rain-potential evapotranspiration) and combinations of the previous elements, among others.

Hydrological drought is defined in terms of the effects of dry spells on surface or ground water hydrology. Definitions of the frequency and severity of hydrological drought are often

based on its effect on catchments. Complex indices include elements based on meteorological and hydrological variables, such as precipitation, river and stream flow, reservoir storage, and combinations of the previous elements, among others. Hydrological droughts are more related with the effects of periods of precipitation shortfall on surface or subsurface water supply, which can be evaluated by a hydrological water balance model.

Agricultural drought is related to physiological drought, which is determined from conditions of natural vegetation, crops, livestock, pastures and other agricultural systems (Gathara, 2006). It is usually measured by the effects of water deficit in terms of economic losses to agriculturists. It is defined by measures of the availability of soil water to plants or animals, such as precipitation shortages, differences between actual and potential evapotranspiration, soil water deficits, radiation (heat), drying wind, etc. A soil-water balance model is often used to assess the agricultural drought because, in addition to meteorological conditions, it takes into account pedological and crop factors (Benson *et al*., 1997).

The definition of socio-economic drought associates the supply and demand of some economic good or service with elements of meteorological, hydrological, and agricultural drought (Hisdal and Tallaksen, 2000; Gathara, 2006).

Smakhtin and Hughes (2004) review a number of drought definitions and indicators, and analyse their applicability for drought prediction and management in the context of south Asia. A quite comprehensive list of complex and simple aridity and drought indices as well as statistical measures applied as drought indices can be found in Hounam *et al*. (1975), Benson *et al*. (1997), Heim (2000; 2002), Gringof and Mersha (2006).

Gringof and Mersha (2006) organize the various empirical expressions of drying coefficients into the following groups:

- Formulas that reflect the relationship between total atmospheric precipitation and total temperature for the year, or for various periods of the warm season;
- Formulas that reflect the relationship between total atmospheric precipitation and evaporation for a predetermined period of time;
- The relationships between total precipitation, temperature and relative humidity for a predetermined period of time;
- Formulas using the water deficit and evaporation deficit values;

- Aridity indices representing the relationship between moisture reserves in the soil layer within reach of the roots of plants and sum of air temperatures for a predetermined period.

While drought indicators are variables to detect and characterize drought conditions, drought triggers are indicator thresholds to define and activate levels of drought responses (Steinemann *et al*., 2005; Steinemann and Cavalcanti, 2006). Steinemann *et al*. (2005) review common indicators and triggers, their functions, and their strengths and limitations. Steinemann and Cavalcanti (2006) describe a process and analytic methods for the development, analysis, and evaluation of drought indicators and triggers.

McKee *et al*. (1993) developed the Standardized Precipitation Index (SPI) as an alternative to the well-known Palmer's indices (Palmer, 1965) for Colorado, although it is now widely used in different climate regimes. The SPI is the number of standard deviations that observed cumulative precipitation deviates from the climatological average, and can be calculated for any time scale (e.g., 3, 6 or 12 months). Soil moisture conditions respond to precipitation anomalies on a relatively short time scale, while ground water, stream flow, and reservoir storage reflect the longer-term precipitation anomalies. Hence, the different time scales for which the index is computed address the various types of drought: the shorter seasons for agricultural and meteorological drought, the longer seasons for hydrological drought (Heim, 2002). While Palmer's indices are water balance indices that consider water supply (precipitation), demand (evapotranspiration) and loss (runoff), the SPI is a probability index that considers only precipitation. Drought intensity, magnitude, and duration can be determined, as well as the historical data-based probability of emerging from a specific drought.

For exemple, Paulo *et al*. (2003) used the SPI (3 and 12-months time scales) and the theory of runs to characterize droughts in the Alentejo region (south of Portugal) for the period 1931/99, and show that droughts randomly affect the region. Additionally, their results suggest that drought occurrence and severity are higher during the second half of the study period, especially in spring periods, which highly influence the productivity of rain fed agriculture.

2.1.2 Desertification indicators

The Convention to Combat Desertification recommends the development and adoption of indicators systems by the National Action Programme (NAP) of each country. At a national level, those indicators should incorporate the physical and human factors that interact in desertification processes at regional levels.

The World Atlas of Desertification (UNEP, 1992) determines arid areas using an indicator representing the relationship between the annual average precipitation (P mm) and potential evapotranspiration (PET mm), which ranges from 0.05 to 0.65. The following criteria were used to determine the arid zones: hyper-arid ($P/PET<0.05$); arid ($0.05 \leq P/PET<0.20$); semi-arid ($0.20 \leq P/PET<0.50$); and dry sub-humid ($P/PET \geq 0.50$). The second edition of this atlas (UNEP, 1997) was improved by shifting the attention from only soil degradation to include also vegetation degradation.

Enne and Zucca (2000) argue that a Desertification Monitoring System should be created on the basis of impact indicators having a dual function of (i) providing a diagnosis, integrated in space and time, of the state of natural resources and of populations of the affected regions; (ii) supporting the decision-making process, providing information on environmental issues, both of a bio-physical and socio-economic nature.

Many indicator systems for desertification assessment were established following the United Nations Conference on Desertification, held in Nairobi in 1977. Even within European Mediterranean countries, the subject has been investigated in countless research projects (Enne and Zucca, 2000). However, a community or patch-based fine-scale indicator system in combination with high-resolution data and ecological models has not yet been established (Yang *et al*., 2005).

The first scientific effort for the identification of sensitive areas in the northern Mediterranean region is due to the European Commission MEDALUS (Mediterranean Desertification and Land Use) project (Kosmas *et al*., 1999; http://www.medalus.demon.co.uk, retrieved 26 March 2008). The methodology uses a multi-factor approach based on both a general and a local knowledge of the environmental processes that operate in European Mediterranean environments. The proposed indicators are appropriate at the European/National (regional desertification indicators) and regional (environmentally sensitive areas) scales. The

definition of environmentally sensitive areas to desertification requires the following information: soil data, vegetation data, climate data, and land management characteristics (Figure 2.5).

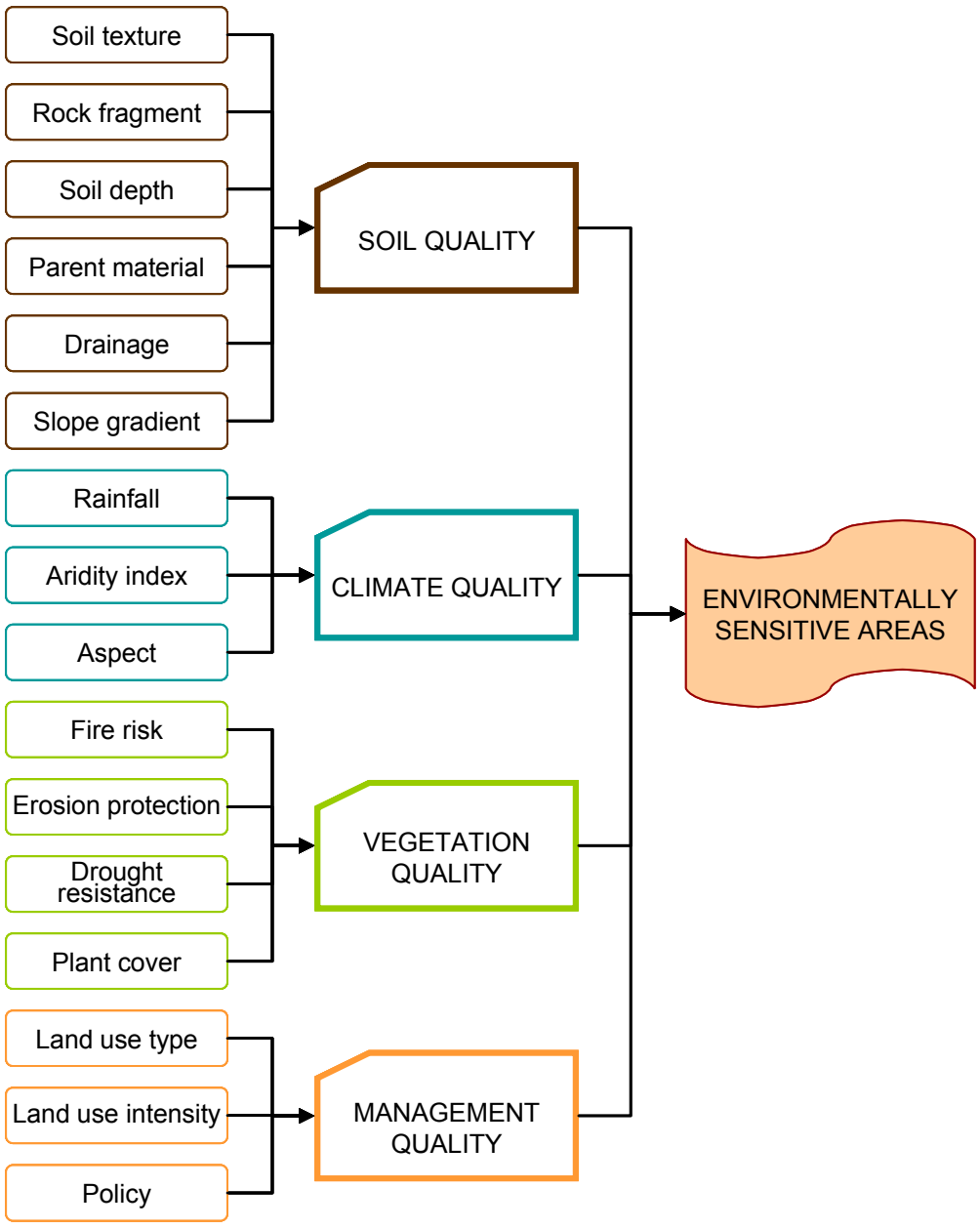

Figure 2.5 – Parameters used by the MEDALUS project for the definition and mapping of the Environmentally Sensitive Areas to desertification (Kosmas *et al.*, 1999)

For example, the following information is required to assess *climate quality*: mean monthly air temperature (°C), mean monthly precipitation amount (mm), mean monthly number of days with minimum temperature (°C), and mean monthly potential evapotranspiration (mm). Additionally, Kosmas *et al.* (1999) suggest the use of the Bagnouls-Gaussen bioclimatic

aridity index and the use of the slope aspect considered according to two classes (class one with NW and NE aspects, and class two with SE and SW aspects).

The indices for soil quality (SQI), climate quality (CQI), vegetation quality (VQI), and management quality (MQI) are first determined. Afterwards, the Environmentally Sensitive Areas Index (ESAI) is obtained as a geometrical average of the four quality indices: $ESAI=(SQI\times CQI\times VQI\times MQI)^{1/4}$. The range of ESAI values determines the types of environmentally sensitive areas to desertification (critical, fragile, potential and non-affected). Targeted areas in Greece (Lesvos Island), Italy (Agri basin) and Portugal (Alentejo region) were used to test and improve the methodology.

Enne and Zucca (2000) describe the methodology and indicators proposed for the assessment of desertification by the Italian Environment Protection Agency (ANPA). The indicators were selected from a number of sources: research programmes on the subject of desertification indicators in the Mediterranean environment or on a global, regional, or national level (e.g., the MEDALUS project, or the Portuguese NAP); interdisciplinary research programmes on desertification on a regional scale; and other programmes and research on related topics described in the literature. The proposed framework for the classification of the 19 indicators is summarized in Table 2.1.

Table 2.1 – Framework for indicators classification, proposed by Enne and Zucca (2000)

Criteria	**Classes**				
Operational objective	Prevention	Monitoring	Mitigation		
Position in the logical framework	Driving force	Pressure	State	Impact	Response
Space-scale	Station	Local	Sub-region	Region	European Mediterranean region
Time-scale	Daily or more	Monthly or seasonal	Annual	Less than annual	Single measure
Component of the system under consideration	Soil	Water resources	Vegetation	Climate	Socio-economic aspects
Nature of data	In data banks	Direct gathering	Remote sensing		

Other programmes and projects aiming to develop indicators, or related research, for the assessment of desertification in European and Mediterranean regions include:

- ACCELERATES – Assessing Climate Change Effects on Land use and Ecosystems; from Regional Analysis to the European Scale (http://www.geo.ucl.ac.be/accelerates, retrieved 20 March 2008);
- ATEAM – Advanced Terrestrial Ecosystem Analysis and Modelling (http://www.pik-potsdam.de/ateam/index.html, retrieved 20 March 2008);
- BioAridRisk – Space-Time Evaluation of the Risks of Climate Changes Based on an Aridity Index[1];
- CIDmeg – Construction of a Desertification Susceptibility Index for the Left Margin of Guadiana[2];
- CLEMDES – Clearing House Mechanism on Desertification for the Northern Mediterranean Region (http://www.inea.it/clemdes1, retrieved 20 March 2008);
- DESERTLINKS – Combating Desertification in Mediterranean Europe: Linking Science with Stakeholders (http://www.kcl.ac.uk/projects/desertlinks, retrieved 20 March 2008);
- DesertWatch – European Space Agency Desertification Monitoring Service http://dup.esrin.esa.it/desertwatch, retrieved 20 March 2008);
- DeSurvey – A Surveillance System for Assessing and Monitoring of Desertification (http://www.desurvey.net, retrieved 20 March 2008);
- DIS/MED – Desertification Information System to support National Action Programmes in the Mediterranean (http://dismed.eionet.europa.eu, retrieved 20 March 2008; DISMED Central Data Repository, http://cdr.dismed.eionet.europa.eu, retrieved 20 March 2008);
- GEORANGE – Geomatics in the Assessment and Sustainable Management of Mediterranean Rangelands (http://www.georange.org, retrieved 20 March 2008);
- INDEX – Indicators and Thresholds for Desertification, Soil Quality, and Remediation (www.soil-index.com, retrieved 20 March 2008);
- LADAMER – Land Degradation Assessment in Mediterranean Europe (http://www.ladamer.org, retrieved 20 March 2008);
- MEDACTION – Policies for Land Use to Combat Desertification (http://www.icis.unimaas.nl/medaction, retrieved 20 March 2008);

[1] Participants: Instituto Superior Técnico (Portugal); Universidade de Évora (Portugal); Universidade do Algarve (Portugal).

[2] Participant: Instituto Superior Técnico (Portugal).

- MEDRAP – Concerted Action to Support the Northern Mediterranean Regional Action Programme to Combat Desertification (http://nrd.uniss.it/medrap/index.htm, retrieved 20 March 2008);
- REACTION – Restoration Actions to Combat Desertification in the Northern Mediterranean (http://www.ceam.es/reaction/, retrieved 20 March 2008);
- RECONDES – Conditions for Restoration and Mitigation of Desertified Areas using Vegetation (http://www.port.ac.uk/research/recondes, retrieved 20 March 2008);
- REDMED – Restoration of Degraded Ecosystems in Mediterranean Regions (http://www.gva.es/ceam/redmed, retrieved 20 March 2008);
- SADMO – Système d'Évaluation et Contrôle de la Désertification dans la Méditerranée Occidentale[3].

In 2001, the UNEP launched the Land Degradation Assessment in Drylands (LADA) project (designed and executed by the FAO) aiming to develop and test an effective assessment methodology for land degradation in drylands at local, national, sub-regional and global levels (http://lada.virtualcentre.org, retrieved 25 March 2008). Currently, it is developing pilot projects in Argentina, China, Cuba, Senegal, South Africa and Tunisia.

Gringof and Mersha (2006) recommended the adoption of a *Total Threat of Desertification* formula, which sums up five aspects of desertification: present state, rate or speed of desertification; internal threat of desertification (including landscape characteristics that describe stability during degradation); the effect of livestock on the surrounding environment; and the population density. The degree of influence of each aspect, empirically determined and expressed in points, allows assessing the category of desertification (slight, average, severe and very severe).

Several processes may act as the contributing factors of desertification depending on the spatial scale in which the desertification phenomenon is analysed (Schreiber *et al*., 2008). The

[3] Participants: Instituto de Estudos Superiores de Recursos Naturais (Portugal, Lead partner); Instituto Superior Técnico (Portugal); Universidade de Jaén (Spain); Consejo Superior de Investigaciones Científicas / Estación Experimental de Zonas Aridas (Spain); Institut National de la Recherche Agronomique / Unité de Biométrie d'Avignon (France); Regione Toscana / Dipartimento dello Sviluppo Economico - Area delle Politiche Regionali dell'Innovazione e della Ricerca (Italy); Democritus University of Thrace / Laboratory I of Hydraulics (Greece); Comissão de Coordenação para o Desenvolvimento Regional do Alentejo (Portugal); Observatório do Sahara e do Sabel (Independent International Organization, Observer partner); Direction Général des Forest - Ministère de l'Agriculture (Algeria, Observer partner).

desertification criteria are of a regional character, and can include various characteristics. The need to identify indicators that are applicable to different spatial scales and to different environmental contexts is one of the fundamental problems of research on indicators (Enne and Zucca, 2000). Table 2.2 summarizes major causes and indicators of desertification at three different spatial scales.

Table 2.2 – Desertification causes and indicators at different spatial scale (Schreiber *et al*., 2008)

Spatial scale	Natural causes	Human causes	Indicators
Macro scale (2000–10000 km)	▪ Global climate change ▪ Increasing drought ▪ Shift of vegetation zones	▪ Large-scale migration ▪ Population increase	▪ Land use changes ▪ Reduced vegetation cover
Meso scale (2–2000 km)	▪ Local climate change ▪ Disturbed rainfall patterns ▪ Increasing temperatures	▪ Population increase ▪ Forced migration ▪ Settlement of herders ▪ Deforestation ▪ Urbanization	▪ Reduction in forest cover ▪ Decrease in grasslands ▪ Increase in cropland ▪ Declining yield statistics ▪ Sediment load in rivers ▪ Dust storm frequency
Micro scale (<2 km)	▪ Erratic rainfall pattern ▪ Increased temperatures ▪ More extreme events ▪ Disturbed water balances ▪ Increased erosion	▪ Poor land management ▪ Bad irrigation practices ▪ Soil nutrient depletion ▪ Tree removal ▪ Overgrazing	▪ Poor vegetation cover ▪ Low crop yields ▪ Water erosion features ▪ Wind erosion features ▪ Crusted soils ▪ Bare soils

Remote sensing methods (e.g. aerial and satellite images) are now used for the assessment and mapping of desertification processes (Lantieri, 2003). The signal recorded by the remote sensor is influenced by the types of predominant vegetation, the plant cover density and stage of development, as well as soil moisture, soil texture and relief, and anthropogenic processes (Lantieri, 2003; Schreiber *et al*., 2008). According to Lantieri (2003), it is possible to map on high-resolution imagery several land degradation features such as:

- wind erosion patterns, especially over large areas;
- salinization patterns in field crops of large irrigated schemes;
- overgrazing features;
- sedimentation of lakes or rivers which are consequent to soil erosion upstream;

- soil water erosion pattern only when of great size and over large areas (gullies);
- burnt areas or areas subject to bush fires.

The spectral radiance of soil and plant cover recorded by satellite information allows characterizing various desertification processes. Lantieri (2003) reviews different types of vegetation indices derived from remote sensing data, which have high potential for assessing the vegetation activity, because of both rainfall availability and land conditions. Vegetation coverage is influenced by climate change including short-term fluctuation (drought) and long-term variation (aridity). Yang *et al*. (2005) recommend the use of multi-temporal data to determine short-term vegetation fluctuation from long-term permanent change, because it may take 30–40 years before a permanent change in vegetation coverage becomes evident.

Lantieri (2003) reviews a number of research programmes making use of remote sensing information, from global environmental institutions, which are of interest to assess the causes and effects of the desertification phenomenon.

The Normalized Difference Vegetation Index (NDVI) is probably the most widely used indicator of the vegetation's density and productivity (e.g. Jabbar and Chen, 2006; Gouveia *et al*., 2008). Gringof and Mersha (2006) recommend that a NDVI with various resolutions should be used: globally, with a resolution of 8 kilometres; nationally, with a resolution of 1 kilometre; and regionally with a resolution of 200 meters or less.

Kogan (2000) describes a new numerical method for the early detection and monitoring of droughts and of their impact assessment from NOAA (National Oceanic and Atmospheric Administration, U.S.A.) operational environmental satellites. Unlike conventional methods that use the NDVI for vegetation monitoring, the new method is based on estimation of vegetation stress from indices derived from satellite images designed to monitor vegetation health, moisture, and thermal conditions. According to Kogan (2000), "this is the first globally universal technique to deal with such a complex phenomenon as drought".

A geographical information system (GIS) enables analysis of combinations of different indicators (data layers), which may result in a better understanding of land degradation problems, causes and consequences (Diodato and Ceccarelli, 2004; Schreiber *et al*., 2008). Therefore, many studies of land degradation combine satellite remote sensing information with other spatial data, such as climate characteristics, soils and land use, into a GIS (e.g.,

Kosmas *et al*., 1999; Rosário, 2004b; Mendicino and Versace, 2007). For example, Lantieri (2003) suggests that, in addition to comparing the present NDVI with the one of a reference year and/or with rainfall data, NDVI images could be superimposed with land cover maps and/or biomass maps to assess and monitor land degradation. These combinations of spatial information are enabled by GIS.

2.1.3 Assessment of desertification in Portugal and European Mediterranean regions

As an outcome of the combination and interaction of climatic and social driving forces, desertification is not new in the Mediterranean. Puigdefábregas and Mendizabal (1998) reviewed the major climatic and land use fluctuations during the past 500 years in the Iberian Peninsula, and identified three critical stages. The first stage, in the 16th–17th centuries, was associated with the joint effects of the Little Ice Age, political changes and requirements of American colonization. The second phase, at the beginning of the 20th century, was mainly associated with the demographic saturation of rural areas. The third stage started in the 1960s and it was driven by social and technological changes in rural life.

Enne and Zucca (2000) review a number of initiatives and indicators that aimed to assess desertification in European Mediterranean countries. According to those authors, desertification in the European Mediterranean regions is linked to the following general characteristics of the region:

- Particular climatic and geomorphologic characteristics, which combined with often poorly adapted use of land, have resulted in a highly vulnerable environment.
- The strong human pressure was aggravated after the 1950s, because of the intensification and mechanisation of agro-pastoral practices, the strong increase in water demand also linked to urban and tourist development, and the appearance of soil and water pollution.

The spatial, seasonal and inter-annual variability of rainfall follows a complex pattern in Mediterranean regions, where the environment is subject not only to droughts, but also to flooding and erosion phenomena caused by high intensity rainfalls. Those heavy downpours, which often occur after a very dry summer, and the high climatic fluctuations in short- and long-term have been pointed out as the main climatic characteristics affecting the vulnerability of the Mediterranean basin to erosion (Martínez-Casasnovas *et al*., 2002).

In western Mediterranean, the major driving force of desertification is the degradation of irrigated lands by overexploitation of water resources, which caused the exhaustion and deterioration of aquifers, soil salinization and damage to downstream fluvial and wetland systems (Puigdefábregas and Mendizabal, 1998). On the other hand, according to De Luís *et al*. (2001), fire and water erosion are frequently considered major causes of soil degradation and desertification in Mediterranean ecosystems.

There is evidence of a trend towards a drier climate in southern Europe as a result of increased evapotranspiration and a relatively slow decrease of rainfall amounts and precipitation frequency (Cubasch *et al*., 1996; Kostopoulou and Jones, 2005; IPCC, 2007; Vicente-Serrano and Cuadrat-Prats, 2007). These circumstances will potentially amplify the vulnerability of several Mediterranean regions to desertification by increasing their environmental problems, such as agro-forestry-grazing productivity, soil degradation, aquifers recharge, forest fires, biological diversity and composition (De Luís *et al*., 2001; Ceballos *et al*., 2004).

In the context of the MEDALUS project, Kosmas *et al*. (1999) proposed a methodological approach to identify and map environmentally sensitive areas to desertification, based on a choice of appropriate indicators at relevant scales (see Section 2.1.2 *Desertification indicators*). In this context, the physical loss of soil by water erosion, and the associated loss of soil nutrient status, was identified as the dominant problem, whereas wind erosion and salinization problems were considered less significant.

2.1.3.1 Implementation of the Convention to Combat Desertification in Portugal

Portugal ratified the Convention to Combat Desertification (CCD) on April 1, 1996. The National Action Programme (NAP) elaboration process occurred between December 1997 and June 1998, and was approved by the Government on June 17, 1999, Resolution of Cabinet number 69/99 (PANCD, 1999).

The strategic objectives of the NAP can be summarized as follows: conservation of soils and water; settling of active population in agricultural areas; recovery of affected areas; population awareness of the desertification problem; consider the fight against desertification in the general and sector politics. The operation guidelines established in the NAP were developed and executed based on interdisciplinary scientific and technical information given by different organisations and with the support of the civil society.

Accordingly, the NAP was based on a number of projects and initiatives developed with government bodies, universities and scientific research centres, among other activities taken by public and private organisms. For example, a number of partnerships have been established in the framework of several projects, such as:

- CLEMDES – Clearing House Mechanism on Desertification for the Northern Mediterranean Region (http://www.inea.it/clemdes1/, retrieved 27 March 2008);
- DESERTLINKS – Combating Desertification in Mediterranean Europe: Linking Science with Stakeholders (http://www.kcl.ac.uk/projects/desertlinks, retrieved 27 March 2008);
- DesertWatch – European Space Agency Desertification Monitoring Service http://dup.esrin.esa.it/desertwatch, retrieved 27 March 2008);
- DIS/MED – Desertification Information System to support National Action Programmes in the Mediterranean (http://dismed.eionet.europa.eu, retrieved 27 March 2008; DISMED Central Data Repository, http://cdr.dismed.eionet.europa.eu, retrieved 27 March 2008);
- MEDACTION – Policies for Land Use to Combat Desertification (http://www.icis.unimaas.nl/medaction, retrieved 27 March 2008).

By the time of the NAP design, a first attempt was made to identify and map the susceptible areas to desertification using three categories of indices: climate, soil loss and drought. At this stage, the climate index was defined as the ratio of the annual average precipitation to the average annual potential evapotranspiration calculated by the Penman method. The soil loss index combined four determinant factors of the erosive process: erosion by precipitation, soil type, vegetation cover and slope of hillsides. The drought index was based on the proportion of years with annual precipitation below the 0.01 quantile of the lognormal distribution. These three indices were then combined, using a geographical information system (GIS), in order to obtain the desertification susceptibility index.

According to Pimenta *et al*. (1997), one of the weaknesses of this approach is the climate index, although proposed by the UNEP (1992) as a first aridity indicator, because it does not reflect the water stresses attached to the soil moisture. Pimenta *et al*. (1997) developed an extension of that methodology, assumed by the DISMED Portuguese Group, which used a redefinition of the climate index based on the monthly water balance. The combination of the three improved indices (climate, soil loss and drought) allowed to conclude that the most

vulnerable areas to desertification were located in the southeast (Alentejo region), and in a few areas in the north. It was also estimated that 60% of the Portuguese continental territory had a moderate risk of desertification.

The group of indicators proposed by Enne and Zucca (2000), described in Section 2.1.2 *Desertification indicators*, includes the indices used by the Portuguese system that were developed by Portuguese DISMED, and described in the work of Pimenta *et al*. (1997).

A second stage of the Portuguese DISMED, under the coordination of the Focal Point[4], aimed to update the susceptibility map of desertification of the NAP (Pimenta *et al*., 1997), and answer the demands at the level of the Mediterranean implementation (Rosário, 2004a). In this phase, the approach was inspired by the methodology developed in the framework of the MEDALUS project (see Section 2.1.2 *Desertification indicators*) and on the set of indices proposed by Enne and Zucca (2000). As proposed by Kosmas *et al*. (1999), the final susceptibility map of desertification of the NAP identifies the vulnerable areas through an indicator combining four quality indices (Figure 2.5 of Section 2.1.2): soil, climate, vegetation, and management/land use. It is important to point out that Kosmas *et al*., (1999) proposed to merge the four layers by computing the geometric average of the four indices, whereas the Portuguese DISMED overlapped the four composite indices without any additional computations.

The maps of the four composite indices and the final susceptibility map of desertification of the NAP (Rosário, 2004b), are available at http://panda.igeo.pt/pancd (retrieved 30 March 2008). Moreover, even though socio-economic factors were not directly accounted for in mapping the susceptibility to desertification, additional sets of social and economic indicators, reflecting cause or effect relationships with the biophysical features of desertification, were also adopted by the Portuguese DISMED and included in the NAP (Rosário, 2004b).

Rosário (2004b) describes the definitions and characteristics of the four composite indices, which are summarized in Table 2.3. For example, the index of climate quality corresponds to the aridity index defined as the ratio of the annual average precipitation to the annual average potential evapotranspiration. The annual average precipitation for the period 1959/60 –

[4] President of the Coordinating Committee of Portuguese NAP.

1990/91 (hydrological years) was mapped through kriging using elevation as external drift, with a spatial resolution of 1 km ×1 km, by Nicolau (2002). This work will be further described in Section 2.3.2 *Overview of interpolation techniques* (Table 2.6). The annual average potential evapotranspiration for the period 1961–1990 was mapped through ordinary kriging. The aridity index map corresponds to the ratio of these two maps.

The susceptibility map of desertification allows concluding that, under the mean climatic regime evaluated, 36% of the Portuguese continental territory is susceptible to desertification. Nevertheless, there are a significant number of regions, on the remaining 64% of the territory, with soils ranging from susceptible to highly susceptible to drought and desertification, regardless of the climatic conditions (Rosário, 2004b). The most vulnerable regions correspond to extensive areas of the Algarve and Alentejo regions, in the south of the country, and a few more areas in the centre-east and northeast (inland regions). These results agree with the preliminary findings of Correia (2004).

Table 2.3 – Parameters, indicators and indices of desertification susceptibility in continental Portugal (adapted from Rosário, 2004b)

DESERTIFICATION SUSCEPTIBILITY			
Aridity index	**Soils susceptibility index**	**Vegetation quality index**	**Land use quality index**
▪ Annual average precipitation (1959/60 – 1990/91) ▪ Annual average potential evapotranspiration (1961–1990)	▪ Thickness ▪ Permeability ▪ Structural stability ▪ Rockiness ▪ Drainage ▪ Slope gradient	▪ Fire risk ▪ Drought resistance ▪ Erosion protection ▪ Vegetation cover (% horizontal cover) ▪ Structural cover (presence of vertical strata) ▪ Climax proximity (degree of natural potential vegetation)	▪ Urban, industrial and tourist areas (actual and projected) ▪ Wet areas (e.g., lakes) ▪ Irrigated land

According to Rosário (2004b), several issues should be further addressed:

- The development of indicators and maps with higher spatial resolution in order to support the regional/national planning;

- The development of indicators of the dynamics of desertification processes and drought, at national and regional levels, capturing the trends and variability hidden behind "the means";
- The development of indicators of the uncertainty of the produced maps, which should be based on the quality of the different information levels, on the lack of knowledge of the different physical phenomena, and on the typical variability of those phenomena connected to desertification;
- The organization and development of work processes at Iberian level and with Spanish institutions aiming to establish coherent and consistent results along the country border, even if with lower spatial resolution;
- The institutional (internal and external) extension of the DISMED process by reinforcing the network of associations, information and work that was established through DISMED.

Pereira *et al*., (2006) describe the identification of the areas susceptible to desertification in Portugal and discuss the insufficiency of the used indicators, mainly in relation to the climatic driving forces and the need to include socio-economic indicators. These authors argue that the aridity index map is not fully appropriate to map the areas susceptible to desertification because it is based on average computations and, consequently, it does not account for droughts, which are part of the climate driving forces influencing desertification. Moreover, in their opinion, such a climate index should also be an indicator of water resources availability, which is impacted by droughts.

2.1.3.2 Causes and impacts of desertification in Portugal

According to Pereira *et al*. (2006), the desertification processes that affect the south and the inland regions of Portugal correspond to large areas where water resources are scarce, soil resources are often poor and non-agricultural vegetation is far from climax. Moreover, in these regions, population density is very low and the respective growth rate is highly negative, aging is increasing with the dependency on aged people, illiteracy is above average, and the purchasing power of the population is much lower than average.

In southern Portugal, a combination of population growth and a cereal self-sufficiency policy increased the desertification phenomenon since the beginning of the twentieth century (Puigdefábregas and Mendizabal, 1998; Roxo *et al*., 1999). In recent years, as in other

southern regions of Europe, a disinvestment in labour caused agriculture and land abandonment, and the soils left behind were exhausted and degraded (Puigdefábregas and Mendizabal, 1998; Geist and Lambin, 2004).

The Algarve and Alentejo regions, in the south of the country, have large areas with high susceptibility to desertification (Correia, 2004; Rosário, 2004b). The southern part of the territory, especially the Alentejo region, is a drought prone area that is characterized by scarce precipitation, little runoff and water availability. Several studies characterized the local and regional droughts of these regions (e.g., Paulo *et al*., 2003; Paulo *et al*., 2005; Moreira *et al*., 2006) and, more generally, the drought phenomenon over Portugal (e.g., Santos, 1998; Santo *et al*., 2005).

Soil erosion is one of the major environmental problems to be faced in these regions (Loureiro and Coutinho 1995, 2001). Water deficits are of great ecological and agronomic importance, especially during the dry season, and an irregular precipitation regime highly influences the productivity of rain fed agriculture.

As other southern European regions, the rainfall regime in southern Portugal is Mediterranean, and so highly variable in both the spatial and temporal dimensions. Accordingly, the climate is characterised by a dry and very hot season, and a very irregular distribution of precipitation over the wet season, as well as over the years, with very intense flood peaks and with frequent drought periods. As discussed previously, whenever the precipitation variability is associated with extreme phenomena, such as intensive rainfall events or drought situations, it may cause soil degradation and vegetation loss that contribute to the desertification of the most vulnerable regions (Santo *et al*., 2004).

One particularly relevant feature of the rainfall regime in southern Portugal is the occurrence of short but very intensive rainfall events that may lead to significant damages, by causing flash floods that affect small drainage basins (Ramos and Reis, 2002), and high rates of soil erosion (Pimenta, 1998; Ó and Roxo, 2001). Fragoso and Gomes (2008) concluded that the most southern region (Algarve) is the one where episodes of heavy rainfall are most frequent and exhibits the strongest torrential character. The Alentejo area, north of Algarve, is mainly an agro-silvo-pastoral region and the most affected by desertification and drought (e.g., Roxo *et al*., 1999; Santo *et al*., 2004).

The application of the methodology developed within the MEDALUS project (see Section 2.1.2 *Desertification indicators*) to the Mértola municipality (Alentejo) resulted on a map of environmentally sensitive areas to desertification, providing an identification of affected and fragile areas that need carefully defined management strategies in terms of land use and human intervention (Roxo *et al*., 1999). The core desertification problem identified in this region was the physical and chemical soil degradation caused by the destruction of natural vegetation cover and the continuous and intensive use of soil resource. The critical areas exist somehow all over the municipality and correspond to areas of greater agricultural activity or with severely eroded soils. According to Roxo *et al*. (1999), the climatic characteristics of this region, associated with topographic and land use factors, have favoured desertification for a long time in the past, with very negative consequences for natural resources such as soil, water and vegetation and ecosystems present.

2.2 Homogenization of climate time series

A homogeneous climate time series is defined as one where variations are caused only by variations in climate (Aguilar *et al.*, 2003). Non-climatic factors (monitoring stations relocations, changes in instrumentation, changes of the surroundings, instrumental inaccuracies, changes of observational and calculation procedures, etc.) may hide the true climatic signal and patterns, and thus potentially bias the conclusions of climate and hydrological studies.

Unfortunately, few long-term climate time series are free of irregularities (e.g. Reek *et al.*, 1992; Peterson *et al.*, 1998; Allen and DeGaetano, 2000; Tuomenvirta, 2001; Auer *et al.*, 2005). Consequently, it is an important task to assess the homogeneity of long climate records before they can be reliably used, and as a recent World Meteorological Organization (WMO) publication recommends, 'it is important, therefore, to remove the inhomogeneities or at least to determine the possible error they may cause' (Aguilar *et al.*, 2003).

Several techniques have been developed for non-climatic inhomogeneities detection and adjustment, i.e. homogenization. If the identified irregularities are due to non-climatic factors then adjustments are performed to compensate for the biases produced by the inhomogeneities. The approaches underlying the homogenization techniques are quite

different and typically depend on the type of element (temperature, precipitation, pressure, evaporation, etc.), the temporal resolution of the observations (annual, seasonal, monthly or sub-monthly), the availability of metadata (station history information) and the monitoring station network density (spatial resolution). The inherently high (temporal and spatial) variability of precipitation makes homogenization of precipitation records more difficult to accomplish than other elements (e.g. temperature).

Some tests rely on metadata (direct methods) while others use a variety of statistical techniques and may be used for homogeneity testing series for which station history is poorly documented or metadata are totally missing (indirect methods). When the station network is dense enough, a common approach is to use difference (temperature, pressure) or ratio (precipitation) series between a *candidate* (or *target*) station and its neighbouring stations (*reference* stations). Afterwards, statistical tests are applied to these composite series for inhomogeneities detection (e.g., Menne and Williams, 2005).

Even though it is becoming very common to develop homogenized climate databases (Tayanç *et al.*,1998; Easterling *et al.*, 1999; Eischeid *et al.*, 2000; Tuomenvirta, 2001; Wijngaard *et al.*, 2003; Feng *et al.*, 2004) or to perform homogenization procedures prior to analyzing data (e.g.; Klein Tank *et al.*, 2002; Llasat and Quintas, 2004; Begert *et al.*, 2005), there are some arguments against homogenizing (Peterson *et al.*, 1998; Auer *et al.*, 2005) and two of them are particularly relevant for the analysis of extreme precipitation. A first argument states that when the monitoring station network is large, the inhomogeneities are expected to become random and hence they can be neglected. Another argument is the danger of smoothing the existing spatial variability by transmitting distant climatic signals from one (or few) reference series to many homogenized series.

As stated before, most of long-term climate time series have inhomogeneities, which can potentially bias the conclusions of climate studies. Furthermore, precipitation measurements are particularly susceptible to irregularities (e.g. station relocations) that may affect the analysis of extreme precipitation events. Therefore, as expressed in the WMO recommendations (Aguilar *et al.*, 2003) and by several authors (e.g. Tayanç *et al.*,1998; Tuomenvirta, 2001; Auer *et al.*, 2005), at least some basic homogeneity testing should be performed before using climate databases, but any kind of adjustment should be carefully considered. It is advisable to seek confirmation of detected inhomogeneities using station

history information before proceeding with the adjustments (Begert *et al.*, 2005). If the irregularities are due to non-climatic factors then the climate series should be rejected from the analysis or properly adjusted.

The second argument against homogenizing is also related with adjustments and is particularly relevant for studies on climate extremes. Some adjustment methods, such as regression-based techniques that are known to smooth the existing spatial and temporal variability of data, should be used cautiously, especially when extreme precipitation totals are to be studied.

Further difficulties arise when attempting to homogenize data with high temporal resolution. Most of the statistical procedures, including nonparametric tests, require serially independent data. Their application may be suitable for annual data, as the independency assumption can be relaxed. Nevertheless, a common approach when homogenizing monthly data is to form individual time series for each month separately, and thus cope with serial dependence and seasonality. However, the existence of serial correlation in daily time series will affect the ability of the tests to assess correctly the significance of trends or shifts in the mean (or median).

More than one undocumented inhomogeneity may be present in a climate time series. According to Reeves *et al*. (2007), in the ideal case, all possible breakpoints should be identified jointly before their mean shift magnitudes are estimated. However, the number of multiple breakpoints detection procedures is limited, thus this is an active current area of statistical research (Reeves *et al*., 2007).

Moreover, for precipitation series with high temporal resolution (sub-monthly data) finding the proper adjustment technique is not straight forward and new approaches will have to be developed (Easterling *et al.*, 1999; Aguilar *et al.*, 2003; Wijngaard *et al.*, 2003; Auer *et al.*, 2005).

2.2.1 General issues

Non-climatic inhomogeneities in climate time series may be introduced by an abrupt change (i.e. break, ‘jump’, step or shift in the mean), by a gradual trend or by a jump superimposed on a trend. A break could result from a recalibration of an instrument or station relocation; a

linear trend could result from a gradual but constant degradation of a sensor; and a non-linear trend could result from vegetative growth around the instruments. The homogenization procedures aim to determine the occurrence and magnitude of such irregularities in order to avoid erroneous and inconsistent inferences from the records.

Two groups of homogeneity testing techniques can be distinguished and are usually referred to as *absolute methods* and *relative methods*. In the first set of procedures, the statistical tests are applied to each station data separately. In the second one, the testing procedures use records from neighbouring stations (named reference stations) which presumably are homogeneous.

While both approaches are worthwhile and valid, they each have drawbacks. Using only data from an individual station is problematical because it is difficult to determine if changes or lack of changes result from non-climatic or climatic influences (Peterson *et al.*, 1998). However, this problem can be reduced, or even solved, by making use of station history information.

Relative methods intend to isolate the non-climatic influences. They assume that within a geographical region, climatic patterns will be identical and that observations from all sites within the region will reflect this identical pattern. Data collected at all sites within the same climatic region should be highly correlated, have similar variability, and differ only by scaling factors and random sampling variability.

Regarding the relative approach, problems arise when the inhomogeneities in the climate data series are caused by simultaneous changes in the observational network, such as simultaneous changes in the measuring technique, as relative tests become insensitive since all series are affected at the same time (Lanzante, 1996; Tuomenvirta, 2001; Wijngaard *et al.*, 2003). Furthermore, ambiguous conclusions are possible when several neighbouring stations do have inhomogeneities themselves (Tayanç *et al.*, 1998; Boissonnade *et al.*, 2002).

Another disadvantage of relative methods is that the use of composite reference series (see Section 2.2.1.1 for details) depends on the proper selection and weighting of individual series, which requires a rather dense station network. In this sense, the absolute tests are more widely applicable, even though they are more dependent on the availability of metadata.

On the other hand, Auer *et al.* (2005) argue that the choice of a specific test method is of minor importance and discuss seven principles that are of major relevance for the homogenization process of monthly data. We quote five of them that are appropriate regardless of the data temporal resolution:

1. Ignore any previous homogeneity work undertaken for any of the series (i.e. assume that all series contain potential breaks).
2. Test in small, well-correlated subregions (a maximum of 10 series tested against each other results in a 10 x 10 matrix, which enables most breaks detected to be assigned to a most likely candidate series).
3. Choose the most appropriate reference series with a non-affected subinterval for the adjustment of each break detected (i.e. different reference series can be used for each break detected in a candidate series).
4. Attempt to determine support for homogeneity adjustments when few metadata are available (i.e. contact data providers for more information in difficult cases).
5. Give preference to good metadata rather than mathematical methods in all cases, especially where adjustment factors can be calculated directly from sufficiently long series of parallel measurements.

Menne and Williams (2005) evaluated three hypothesis test statistics[5] to ascertain whether multiple tests can be combined to improve overall confidence in undocumented inhomogeneities detection. These authors also evaluated different composite reference series formulations. Using Monte Carlo case studies, Menne and Williams (2005) concluded that for reasonably well correlated time series and if the reference series are homogeneous, the choice of reference series formulation has relatively little impact on candidate series inhomogeneities detection skill and, consequently, the choice of the test statistic has a greater impact. However, Menne and Williams (2005) argue that those circumstances are probably rare in practice, thus the choice of reference series formulation has implications that are more important in breakpoints detection than the choice of the test statistic.

[5] Methods evaluated by Menne and Williams (2005): (1) SNHT for a single break, (2) two-phase regression, and (3) two-phase regression without slope.

2.2.1.1 Relative approaches issues

As pointed out before, if an absolute test detects a break in a station's time series it may indicate an inhomogeneity or it may simply indicate an abrupt change in the regional climate. In order to account for regional climate changes and to isolate the effects of station irregularities, many techniques use data from surrounding stations inside the same climatic region (reference stations).

The most common approach for selecting reference stations is to form Pearson correlation matrices between the candidate site and neighbouring stations' data, which presumably are homogeneous, and to take as reference the highest correlated ones (e.g. Tayanç *et al.*, 1998; Boissonnade *et al.*, 2002). Other approaches extract principal components from the whole data network, or use an independent data source thought to be homogeneous (Aguilar *et al.*, 2003).

Some procedures search for breakpoints or artificial trends in a composite reference series (or alternatively in the data when a suitable composite series cannot be built), while some statistical tests compare the candidate data series with data from reference stations (e.g., tests for the difference between medians). Using composite reference series – ratio series for precipitation and difference series for temperature – is a standard procedure in the detection of non-climatic homogeneities. The assumption from this approach is that the composite reference series includes the regional climate trends and fluctuations present in the data of the candidate, but does not contain discontinuities itself during the period of analysis of the discontinuity in the candidate station. Composite reference series attempt to reduce, if not to eliminate, most of the climatic signals, and thus diminish some of the problems that were previously discussed concerning relative methods.

Composite reference series, or simply reference series, are computed as a weighted average of data from neighbouring stations by using some measure of statistical similarity (usually the correlation coefficient or an inverse function of the distance) between them. Romero *et al.* (1998) proposed a combined use of those measures in order to increase the contribution of the records from closer stations, both in spatial and correlation terms. Alexandersson and Moberg (1997) proposed the construction of ratio (difference) reference series, which are generally used in precipitation (temperature) studies. If the candidate has no inhomogeneities then the

resulting ratio (difference) series will oscillate around one (zero); otherwise, one or more irregularities due to non-climatic factors will be uncovered.

The most usual approach to obtain adjustment factors is to calculate separate averages on the difference or ratio series for the two sections defined by a breakpoint (Aguilar *et al.*, 2003). When abrupt changes are identified in the time series, the obtained means are compared by calculating their ratio or difference and the obtained factor is applied to the inhomogeneous part. When dealing with gradual trends or breakpoints superimposed on trends, the inhomogeneous section is de-trended using the slope calculated on the difference or ratio time series.

Creating and using composite reference time series may encounter two major problems:

1. lack of data to build them, because the monitoring stations network was, or is, too sparse to find enough neighbouring stations to construct a reliable reference;
2. a number of series from the neighbouring stations have inhomogeneities.

The common period of observations between the candidate series and time series from neighbouring stations might be too short to properly select and weight the individual series, and thus construct a reliable reference. Moreover, if too many distant (or less correlated) neighbouring stations are used, the resulting reference may not reflect properly the true climatic signal of the candidate station (Boissonnade *et al.*, 2002). Regarding the second problem, as mentioned before, ambiguous conclusions are possible when several neighbouring stations do have inhomogeneities themselves.

These difficulties may increase dramatically with the increase in spatial variability of the data caused by the inherent variability of the element (e.g. precipitation), the time series resolution (e.g. sub-monthly data) or the network location. Aguilar *et al.* (2003) provide a clarifying example: ‘it is intuitively very understandable that it is easier to create a good reference time series for annual averaged temperatures for a station at the equator than building a reference for August precipitation for a station in the Mediterranean’. Further problems arise when attempting to build a precipitation ratio series in the Mediterranean, where zero-precipitation summer months typically occur.

2.2.2 Procedures for monthly, seasonal and annual data

Several widely used techniques for inhomogeneity detection and homogenization are summarized. A few descriptions and references were obtained from Peterson *et al*. (1998) and Aguilar *et al.* (2003) who give a comprehensive review of methods. The listing of procedures is presented in alphabetical order, without distinguishing between absolute and relative methods, since absolute tests can also be used in relative approaches by applying the test to composite reference series.

Subjective methods (e.g. Boissonnade *et al.*, 2002), such as the well known *double-mass analysis* (Kohler, 1949), will not be presented despite the fact that they may be particularly helpful in exploratory data analysis. Other common homogenization procedures mentioned in the literature, but not described here, are: the *Craddock test* (Craddock, 1979; Auer *et al.*, 2005; Begert *et al.*, 2005); the *Potter's method* (Potter, 1981; Begert *et al.*, 2005); the *Rank-order change point test* (or *L-method*) proposed by Lanzante (1996); and the Mann-Whitney U test, also called the Mann-Whitney-Wilcoxon, Wilcoxon rank-sum test, or Wilcoxon-Mann-Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947; Lanzante, 1996; Lee and Maeng, 2003). The procedures described subsequently are intended to illustrate the variety of approaches that are commonly used. The literature is replete with techniques, but most of them are similar or variations of the methodologies described here.

Ducré-Robitaille *et al.* (2003) compared eight homogenization methods[6] using simulated series reproducing a vast range of possible situations (homogeneous series and series having one or more breakpoints). Their results show that the most reliable techniques for the identification of homogeneous series are the Standard normal homogeneity test (SNHT) for a single break, the multiple linear regression and the Bayesian approach with reference series. Moreover, Ducré-Robitaille *et al.* (2003) concluded that the SNHT for a single break is the best approach to identify a single small step, to detect a random number of irregularities, and to identify the correct number of breakpoints. The methods examined by Reeves *et al*. (2007) include the SNHT, the Wilcoxon's nonparametric test, two-phase regression procedures,

[6] Methods evaluated by Ducré-Robitaille *et al.* (2003): (1) standard normal homogeneity test (SNHT) for a single break; (2) SNHT with trend, (3) multiple linear regression, (4) two-phase regression, (5) Wilcocon rank-sum test, (6) sequential testing for equality of means, (7) Bayesian approach without reference series, and (8) Bayesian approach with reference series.

among other inhomogeneity tests and various variants thereof. Their results indicate that the common trend two-phase regression and Sawa's Bayes criteria procedures seem optimal for most climate time series, whereas the SNHT procedure and its nonparametric variant are probably best when trend and periodic effects can be diminished by using homogeneous reference series.

It is also important to mention that several homogenization techniques described subsequently, particularly tests for single break detection, are some times used iteratively by systematically dividing the tested series into smaller segments when a break is detected, and then performing the test on those segments. Techniques that use series from surrounding stations, some times run the test once, relying the reference to be homogeneous, or engage in an iterative procedure in which all stations in the dataset are seen consecutively as candidates and references. Procedures based on test iteration such as those are powerful but computationally intensive, and thus can be time consuming and exacting work.

2.2.2.1 Buishand range test

The Buishand range test (Buishand, 1982) is a parametric test and supposes, under the null hypothesis, that the values of the testing variable are independent and identically normally distributed. Under the alternative hypothesis, it assumes that a step-wise shift in the mean (a break) is present. Wijngaard *et al.* (2003) extend the table of critical values for the test given by Buishand (1982) and also provide a mathematical description. This test is capable of locating the period (year/month) where a break is likely, but it is more sensitive to breaks in the middle of a time series (Wijngaard *et al.*, 2003).

For example, Wijngaard *et al.* (2003) applied this test (among other homogenization techniques) to three variables, with annual resolution, derived from the daily series of the European Climate Assessment dataset for the period 1901-1999: the annual mean of the diurnal temperature range; the annual mean of the absolute day-to-day differences of the diurnal temperature range; and, the wet day count (threshold 1 mm).

Another example of the application of the Buishand range test can be found in Feng *et al.* (2004). These authors describe the development of a climate dataset that contains 10 daily variables with data from 726 stations in China from 1951 to 2000: maximum and minimum surface air temperatures, mean surface air temperature, skin surface temperature, surface air

relative humidity, wind speed, wind gust, sunshine duration hours, precipitation, and pan evaporation. The homogenization procedures include the application of the Buishand range test to annual reference series.

2.2.2.2 Caussinus-Mestre technique

The method proposed by Caussinus and Mestre (1996) is based on the premise that between two breaks, a time series is homogeneous and these homogeneous sections can be used as reference series. This approach allows accounting for the detection of multiple breaks.

Each single series (candidate) is tested for discontinuities by means of the differences (temperature, pressure) or ratios (precipitation) series that are constructed using neighbouring stations from the same climatic area. When a detected break remains constant throughout the set of comparisons of a candidate station with its neighbours, the break is attributed to the candidate time series.

Instead of comparing a series to an artificial reference, Caussinus and Lyazrhi (1997) developed a new technique based on the comparison of several perturbed series. They formulated it as a problem of testing multiple hypotheses, and provided a Bayes invariant optimal multi-decision rule for detecting a set of an unknown number of change-points and outliers based on a penalized log-likelihood statistic. The penalty term curbs the increase in the likelihood and picks the solution with the right number of breaks most of the time (Peterson *et al*., 1998).

2.2.2.3 Kruskal-Wallis test

The Kruskal-Wallis test (Kruskal, 1952; Kruskal and Wallis, 1952) is a well known nonparametric (or distribution free) test used to compare two[7] or more independent groups of sampled data. One of the assumptions of the Kruskal-Wallis test is that the observations are drawn randomly and independently from their respective populations. This test is an alternative to the independent group ANOVA F test (which compares the means of several groups), when the assumption of normality is not met.

[7] When comparing only two populations, the nonparametric Mann-Whitney U test is sometimes preferred.

This, like many nonparametric tests, uses the ranks of the data rather than their raw values to calculate the test statistic. By using ranks, the impact of any outliers or skewness in the data is greatly diminished, allowing for a test that is not dependent on the data coming from a normal distribution.

The null hypothesis for the Kruskal-Wallis test is that all the samples come from identical populations. The alternative hypothesis is that not all of the samples come from identical populations. If all of the population distributions have the same shape (normal or not), these hypotheses are also sometimes written as the testing of the equality of the central tendency of the populations (i.e. testing whether all the independent samples have been drawn from populations possessing equal medians). The exact distribution of the Kruskal-Wallis statistic under the null hypothesis depends on all the sample sizes, so tables are awkward. However, the test statistic is distributed approximately as a chi-square distribution when the null hypothesis holds and when the sample sizes are not too small (greater than 5), making it easy to obtain p-values.

The Kruskal-Wallis test gives little information about the probable date for a shift in the median and no information about the magnitude of the break.

Tayanç *et al.* (1998) tested the efficiency of the Kruskal-Wallis test using artificially generated time series. These authors also used it in the homogenization process of 82 annual temperature series of Turkish stations for the period 1951-1990. Another application of this test in the homogenization context was also performed by Türkes (1999).

2.2.2.4 Mann-Kendall test

The Mann-Kendall test is a nonparametric test for the detection of trend in a time series. Since the first proposals of the test by Mann (1945) and Kendall (1975), the test was extended in order to include seasonality (Hirsch *et al.*, 1982; Hirsch and Slack, 1984), multiple time series (Lettenmaier, 1988) and covariates (Libiseller and Grimvall, 2002). The Mann-Kendall test requires data to be serially independent. Yue and Wang (2004) discuss several approaches that use the effective sample size to modify the test statistic in order to eliminate the effect of serial correlation.

An assumption of trend tests is that trends are consistently increasing or decreasing, otherwise known as monotonic changes. The Mann-Kendall test is traditionally used to test randomness

against (monotonic) trend. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that a stochastic trend exists. The rank-based test statistic is distributed approximately as a normal distribution when the null hypothesis holds and when the sample size is not too small (greater than 8).

For example, Tarhule and Woo (1998) used the Mann-Kendall ranked τ statistic to analyse the occurrence of trends in several rainfall characteristics using data collected at 25 locations in northern Nigeria for the period 1931-1996. Another application of the Mann-Kendall test in the homogenization of precipitation series can be found in Santos and Henriques (1999).

2.2.2.5 Multiple analysis of series for homogenisation (MASH)

The MASH method was developed in the Hungarian Meteorological Service (Szentimrey, 1994, 1995a, 1995b, 1996, 1999), and is a relative homogeneity test procedure that does not assume that the reference series are homogeneous. MASH is a multiple break points detection technique that takes into account the significance and the efficiency of the test. Moreover, it provides not only estimated break points and shift values, but the corresponding confidence intervals as well, and hence the series can be adjusted by using the point and interval estimates. Another feature of MASH is that an additive or a cumulative model can be used depending on the climate elements (e.g. temperature, precipitation). The theoretical basis of the method can be found in Szentimrey (1999).

Possible break points and shifts can be detected and adjusted through mutual comparisons of series within the same climatic area. The role of series (candidate or reference series) changes step by step in the course of the procedure.

An interesting feature of the software developed for this method (MASH system) is that the probable dates of break points provided by metadata information can be used automatically. In case of having monthly series for all the 12 months, the MASH system also allows the monthly, seasonal and annual series to be homogenized together. More recently, Szentimrey (2003) introduced in the MASH system a new procedure to evaluate the homogenization results. The verification procedure evaluates the quality of the homogenized series by the joint comparative mathematical examination of the original and the homogenized series systems.

For other applications with the MASH system, see for example Griffiths *et al*. (2003), Květoň and Žák (2003), Piccarreta *et al.* (2004) and Auer *et al.* (2005).

2.2.2.6 Pettit test

Pettit (1979) developed a nonparametric test that is capable of locating the period (month or year) where a break is likely. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that a step-wise shift in the mean (a break) is present. The test statistic is related to the Mann-Whitney statistic. A mathematical description and the significance level can also be found in Tarhule and Woo (1998).

This test is based on the ranks of the elements of a series rather than on the values themselves, thus it is less sensitive to outliers than other methods. Like other tests, the Pettit test is more sensitive to breaks in the middle of a time series (Wijngaard *et al.*, 2003).

Wijngaard *et al.* (2003) use the Pettit test, among others, to assess the homogeneity of the European Climate Assessment dataset for the period 1901-1999, and provide a table of critical values based on simulations, plus a mathematical description.

2.2.2.7 Regression-based methods

This section reviews several regression-based techniques that have been proposed for the homogenization of climate time series. Most of them can only be appropriately used with annual data, as serial correlation negatively affects the precision of parameter estimates[8]. Nevertheless, we call attention to the regression-based approach used by Feng *et al.* (2004) to detect spatial outliers in daily series (including temperature and precipitation).

A two-phase regression technique for detecting a change point in the trend of a time series is described by Solow (1987). In this method, the regression lines before and after the year that is being tested are constrained to meet at that point. Easterling and Peterson (1995a, 1995b) developed a variation on this technique in which the regression lines are not constrained to meet, and where a linear regression is fitted to the part of the reference series before the year being tested and another one after the year being tested. This test is repeated for all years of the time series (with a minimum of 5 years in each section), and the year with the lowest residual sum of the squares is considered the year of a potential discontinuity. The time series is then divided into two at that year and both sub-series are similarly tested. This subdividing

[8] If serial correlation is present in data, the least squares estimator will still be unbiased, but no longer the best linear unbiased estimator (B.L.U.E). Moreover, in the case of positive serial correlation, estimates of standard errors will be biased downward.

process continues until no significant breaks are found or the time series are too short to test. Reeves *et al*. (2007) discuss a number of variants of the two-phase regression technique. See for example Mekis and Vincent (2003) for applications in the homogenization of precipitation and temperature series in Canada.

Lanzante (1996) presents two resistant regression approaches, namely the three-group resistant regression and the pairwise slopes method. In the first one, the sample is subdivided into three groups based on the abscissas. The median coordinates of the left and right groups are used to define a line which serves as a starting point for an iterative process; a special procedure insures convergence to a final solution. The second technique involves the computation of the slopes defined by all possible pairs of points; the final slope estimate is the median of these values. Significance tests are not provided for these regression techniques, thus the significance of the Spearman correlation coefficient is used instead. Lanzante (1996) argues that the pairwise slopes method seems preferable because of its greater efficiency and resistance, even though being more computationally expensive.

Vincent (1998) proposed a multiple linear regression approach based on the application of four regression models to determine whether the tested series is homogeneous, has a trend, a single step, or trends before and/or after a step. The dependent variable is the series of the candidate station and the independent variables are the series of a number of surrounding stations. The first model determines whether the candidate series is homogeneous (in this case, the remaining models are not used). The series is considered homogeneous if the residuals from the regression are independent normal variables with zero mean and constant variance. If there is significant autocorrelation in the residuals (assessed by generalized Durbin-Watson tests and by the correlogram), then a second regression is calculated in which a linear trend is included. If autocorrelation in the residuals of this second model exist, then the model is discarded and a third model is examined. The third regression is calculated for sequential increases in the time at which a step can occur. The minimum residual sum of squares from these regressions identifies the time of a break. If autocorrelations exist in the residuals from the regression with the step, a fourth model is considered. The last regression accounts for trends before and after the identified step. The existence of trend provides an indication of multiple inhomogeneities in the candidate series. In this case, the series is subdivided at the position in time of the identified step and each segment is tested separately starting with the first model. The estimated parameters corresponding to steps and trends

provide the magnitude of each inhomogeneity. Adjustments are then applied to bring each segment into agreement with the most recent homogeneous part of the series. The method proposed by Vincent (1998) has been recently improved by Reeves *et al*. (2007).

Allen *et al.* (1998) describe a method of cumulative residuals (*Ellipse test* or A*ccumulated residual method*) that tests if a weather data set is homogeneous using the cumulative residuals from the linear regression between the candidate series (dependent variable) and data from a neighbouring station (independent variable), or the average observations of several surrounding stations inside the same climatic region. The residuals from the regression should be considered homoscedastic and independent random variables with mean zero. The candidate series can be considered homogeneous if the cumulative residuals are not biased. The bias hypothesis can be tested using an ellipse defining the confidence limits. Plotting the cumulative residuals against time, using the time scale (interval) of the variable under analysis, the accumulated residual curve is obtained. If all the cumulative residuals lie inside the ellipse then the hypothesis of homogeneity is not rejected for the significance level considered. This test is capable of locating the period (year) where a break is likely to occur. An application of the accumulated residual method in the homogenization of annual precipitation series can be found in Santos and Henriques (1999).

Feng *et al.* (2004) used a linear regression approach in order to detect spatial outliers by comparing data from neighbouring stations. Correlation coefficients are computed for each month between daily data from the candidate station and the 10 nearest stations. Series with large positive correlation coefficients are used to create simple linear regressions between the candidate station and each neighbouring station. If more than five neighbouring stations have significant correlation with the candidate station at a specific month, than the five neighbouring stations with the lowest root-mean-square error of the regressions are chosen. A daily value is flagged as suspicious if it falls outside the specified regression-based confidence intervals for all pairs of stations chosen. Estimation of missing and suspicious values is also based in this regression approach.

2.2.2.8 Standard normal homogeneity test (SNHT)

The Standard normal homogeneity test (SNHT) is a parametric test developed by Alexandersson (1986) that is capable of locating the period (month or year) where a break is likely. The null hypothesis is that the data are independent, identically normally distributed

random quantities and the alternative is that a step-wise shift in the mean (a break) is present. The SNHT is a likelihood ratio test and it is usually performed on a ratio or difference series between the candidate station and reference series. The SNHT detects breaks near the beginning and ending of series relatively easily (Ducré-Robitaille *et al.*, 2003; Wijngaard *et al.*, 2003).

There are now variations of this test to account for more than one discontinuity, testing for inhomogeneous trends rather than just breaks, and inclusion of change invariance (Alexandersson and Moberg, 1997).

The standard normal homogeneity test is one of the most widely used homogeneity tests (e.g. Tuomenvirta, 2001; Hidalgo *et al.*, 2003b; Mekis and Vincent, 2003; Müller-Westermeier, 2003; Wijngaard *et al.*, 2003; Feng *et al.*, 2004; Auer *et al.*, 2005; Begert *et al.*, 2005).

2.2.2.9 Von Neumann ratio test

Von Neumann (1941) proposed a statistic defined as the ratio of the mean square successive (year-to-year) difference to the variance. The Von Neumann ratio test is distribution free and is not location specific, which means that it gives no information about the date of the break. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that the time series is not randomly distributed. When the sample is homogeneous, the expected value of the test statistic is equal to two.

The Von Neumann ratio test was one of the statistical tests used by Wijngaard *et al.* (2003) to verify the homogeneity of the European Climate Assessment (ECA) dataset for the period 1901-1999. Those authors also provide a table of critical values and a mathematical description. For other applications, see for example Rodriguez *et al.* (1999), Llasat and Quintas (2004).

2.2.2.10 Wald-Wolfowitz runs test

The Wald-Wolfowitz runs test is a well-known nonparametric test for randomness (Wald and Wolfowitz, 1943), i.e. tests the assumption that the data collected constitute a random sample so that each observation or measurement is drawn randomly and independently from its population. This test is sensitive to shifts and trend, but gives little information about the probable dates for breaks (Tayanç *et al.*, 1998). The null hypothesis is that the process that

generates the set of numerical data is random (with respect to the median) over time. For the two-tailed test, the alternative hypothesis is that the data set is not randomly distributed.

A sequence is formed by assigning one of two symbols to each observation, depending on whether its measurement falls above or below a certain value (usually the median). A *run* is defined as a consecutive series of similar symbols that are bounded by symbols of a different type or by beginning or ending of the sequence. The test statistic is the number of runs present in the data, and it is distributed approximately as a normal distribution when the null hypothesis holds and when the sample size is not too small (greater than 40).

The underlying idea of this test is that, if the sequence is randomly generated, the symbol of an observation will be independent both of its position in the sequence and of the values of the observations that precede it and follow it. For a one-tailed test, the alternative hypothesis is that a trend effect is present in the data. In this case, the null hypothesis is rejected if too few runs occur. On the other hand, for a one-tailed test in which the alternative hypothesis is that a systematic or periodic effect is present in the data, the null hypothesis is rejected if too many runs occur.

Tayanç *et al.* (1998) tested the efficiency of the Wald-Wolfowitz runs test using artificially generated time series. These authors also used it in the homogenization process of annual temperature series of Turkish stations for the period 1951-1990. For other applications see for example Santos and Henriques (1999), Lee and Maeng (2003).

2.2.2.11 Data adjustments

Aguilar *et al.* (2003) recommend the adoption of a reverse chronological approach to adjust annual (monthly) series experiencing more than one discontinuity in which the most recent homogeneous period is used as a standard and earlier periods are adjusted to reflect these current conditions. By doing so, incoming data in the future will still be homogeneous unless further changes occur in the monitoring station. Moreover, even if additional changes take place, another advantage of this strategy is that it allows for easier updating (Auer *et al.*, 2005).

Allen and DeGaetano (2000) argue that it is also reasonable to base adjustments on the longest stationary homogeneous period within a station's record and then proceed chronologically but with the decision to adjust earlier or more recent periods again based on

the series length. One advantage of this approach is that the quantity of data that is subject to adjustment is minimized.

Regardless of which approach is taken to define the sequence of adjustments, the adjustment factors are usually obtained using the following methods (Aguilar *et al.*, 2003). If a series must be adjusted for a sudden shift, a common approach is to calculate separate averages on ratio series (for precipitation) for the two sections defined by the breakpoint. Then, the obtained means are compared by calculating their ratio and the resulting factor is then applied to the inhomogeneous part. When gradual inhomogeneities are detected, the usual approach is to de-trend the inhomogeneous section using the slope calculated on the ratio time series.

Several authors (e.g. Peterson *et al.*, 1998; Tuomenvirta, 2001) use Student's t-tests to determine whether the sample of estimated/adjusted temperature values is significantly different from the original data. However, precipitation distributions generally exhibit positive skewness, particularly in arid and semi-arid areas, and thus the usual parametric t-tests are not appropriate. For this reason, Eischeid *et al.* (2000) employed the simple ratio test to compare the observed versus estimated precipitation values.

2.2.3 Homogenization of sub-monthly data

As described in the previous section, there are a number of tests available for the homogenization of climate series with low temporal resolution. However, well-established statistical methods for the homogeneity testing of sub-monthly precipitation data are lacking (Easterling *et al.*, 1999; Aguilar *et al.*, 2003; Wijngaard *et al.*, 2003; Auer *et al.*, 2005).

Furthermore, adjusting daily and hourly data is not straightforward, thus the WMO makes no recommendations regarding adjusting sub-monthly data. As an alternative, the WMO advises that data should be carefully evaluated for the impacts of inhomogeneities, and that portions of time series with homogeneity problems be excluded from the analysis before using sub-monthly data in long-term climate change analysis (Aguilar *et al.*, 2003).

Both parametric and nonparametric tests that were described above look at one or a few of the characteristics of a frequency distribution. These characteristics do not include nonlinear effects nor do they consider non-climatic influences that affect data in a nonuniform manner, such as only during certain weather events, seasons, etc. The daily data may reflect more of a

mixture of populations, and are more likely to be affected by nonlinear and nonuniform weather events than data that are averaged over a longer time interval. Another drawback is that most procedures, including nonparametric approaches, are not appropriate because of the assumption that data should be serially independent.

Common homogenization procedures also fail to detect non-climatic irregularities in sub-monthly precipitation series because of the higher variability of daily or sub-daily records. Assuming that the magnitudes of the levels of change that can be detected in monthly or annual data also apply to daily data, the inherent high variability of the daily data leads to rather large ranges of undetectable inhomogeneities.

In order to overcome these difficulties, the homogenization of high temporal resolution climate databases is usually performed by using traditional procedures with monthly or annual totals, or other variables, derived from the daily series (e.g., Wijngaard *et al.*, 2003; Feng *et al.*, 2004), or solely by the use of metadata (Guttman, 1998).

Wijngaard *et al.* (2003) developed a hybrid method by compiling from the daily set an annual resolution set of variables representing important characteristics of variation at the daily scale, and then applying to these testing variables four established statistical tests: the Standard normal homogeneity test (SNHT) for a single break (Alexandersson, 1986), the Buishand range test (Buishand, 1982), the Pettit test (Pettit, 1979), and the Von Neumann ratio test (Von Neumann, 1941). For precipitation, the testing variable used was the wet day count using 1 mm as threshold. Due to the sparse spatial distribution of the station network, relative methods could not be used. Wijngaard *et al.* (2003) did not try to adjust the daily series for the inhomogeneities detected. Instead, the results of the different tests were condensed into three classes (‘useful’, ‘doubtful’, and ‘suspect’) and a qualitative interpretation of the classification was given, as well as recommendations for the use of the labelled series in trend analysis and variability analysis of weather extremes.

Feng *et al.* (2004) estimated missing data and suspicious records that were previously screened by several basic quality control procedures using a regression-based approach, before proceeding with the homogenization of daily meteorological data. The homogeneity check was performed using annual reference time series that were built using the annual average value of each variable (except for precipitation for which the method was adapted and the annual total precipitation was used). After creating reference series, three statistical tests

were used: the moving t-test (Peterson *et al.*, 1998), the Standard normal homogeneity test (SNHT) for a single break (Alexandersson, 1986), and the Buishand range test (Buishand, 1982). Whenever the test results from these methods had discrepancies and station's metadata was not available, the series were visually evaluated and it was subjectively decided whether the changes detected had occurred.

Feng *et al.* (2004) adjusted the daily series of all variables except precipitation and wing gust. Monthly corrections were obtained for the twelve individual months, and the daily adjustments were derived from the monthly corrections using a linear interpolation between midmonth 'target' values that were objectively chosen so that the average of the daily adjustments over a given month was equal to the monthly correction.

2.2.4 Homogenization of precipitation extremes

Precipitation is one of the most important climate variables. Accurate quantification of its observed variability is required for a number of purposes. These include: the climate changes monitoring; the validation of numerical weather prediction models, general circulation models and regional climate models; the modelling of erosion, runoff and pollutant transport; the design and management of irrigation systems, farm management systems and water supplies, among other applications for ecosystem and hydrological impact modelling.

Accurate long-term precipitation series, with at least daily resolution, are required for several of those situations, especially those that involve the analysis of extreme precipitation events. However, as stated before, most of long-term climate time series have inhomogeneities that can potentially bias the conclusions of climate and hydrological studies.

Furthermore, precipitation measurements are particularly susceptible to irregularities that may affect the analysis of extreme precipitation events (Easterling *et al.*, 1999). For instance, station relocations may cause an artificial change in observed extremes, particularly in areas of heavy rain.

Most of the rain gauge errors result in less rain being measured by the gauge than what actually has fallen. These errors include wind-induced undercatch, heavy rain splash out, and gauge wetting. At some point in time, changes in instrument mounting and sheltering might have been introduced in order to increase the amount of precipitation caught, and thus

decrease the measurement error. However, this may have a large effect on the incidence of extreme precipitation events and introduces a break in the data time series.

Undercatch is an issue concerning just about every rain gauge, including those ones incorporating windshields and splashguards. Tipping-bucket gauges are well known to undercatch precipitation during heavy rainfall events when the precipitation rate exceeds the capability of the tipping mechanism to keep up with the water flowing through the gauge. Thus, the precipitation gauge type might have changed, causing a huge impact on data homogeneity.

Naturally, the quality control of extreme precipitation values is particularly relevant for this study but, unfortunately, it is especially difficult to accomplish. Extremes are rare events that frequently arise from a unique set of weather conditions. Determining the proper homogeneity adjustment for these unique conditions can be difficult since few extreme records are available for the assessment (Aguilar *et al.*, 2003).

Further difficulties arise when the homogenization of precipitation extremes has to be performed in arid and semi-arid regions, where localized convective storms can often give large amounts of precipitation at one observing station and a nearby station receives nothing (Easterling *et al.*, 1999). Common homogenization techniques based on comparisons with neighbouring stations may therefore exclude genuine extremes from the data series.

Furthermore, most of the homogenization procedures described in the literature (see Section 2.2.1.1) are only suitable for low temporal resolution data sets. Consequently, the homogenization of extremes usually relies on metadata and quality control checks for outliers identification (Lanzante, 1996; Easterling *et al.*, 1999; Klein Tank *et al.*, 2002; Wijngaard *et al.*, 2003; Feng *et al.*, 2004).

Although not especially relevant for this work, it is worthwhile mentioning the method developed by Allen and DeGaetano (2000) to homogenize non-climatic discontinuities in temperature extreme exceedence series, as their approach seems promising for the homogenization of daily temperature datasets.

2.3 Characterization of extreme precipitation events

Climate extremes are events rarely observed and statistically correspond to the tails of the distribution of the climate variable. Changes in extremes correspond to changes in the distribution (location, scale and/or shape) of the variable. The analysis of changes in extremes can be performed by fitting appropriate theoretical distribution functions, named GEV (Generalized Extreme Values) distributions, to the observed daily climate data and then investigate the changes in the parameters of the distribution functions over time or space (e.g. Weisse and Bois, 2001; Durrans and Kirby, 2004; Beguería and Vicente-Serrano, 2006). Another approach allowing to identify changes in climate extremes is based on the analysis of climate indices. Indicators characterizing extreme events can be estimated from the empirical distribution of the daily observations, and their changes analysed directly.

Hundecha and Bárdossy (2005) point out that, using the GEV approach, the assumption made about the distribution function may not be met by many stations, which is a serious practical limitation when analysing data from many stations, thus the results of the analysis may lead to a wrong interpretation of the changes in the extremes. Considering this and other arguments, Hundecha and Bárdossy (2005) argue that nonparametric approaches based on extreme climate indicators can be more appropriate to analyse changes in climate extremes than the GEV approach. In general, these indicators represent events that occur several times per season or year giving them more robust statistical properties than measures of extremes that are far enough into the tails of the distribution so as not to be observed during some years (Alexander *et al*., 2006).

In fact, as revealed by Frich *et al*. (2002), the WMO Commission for Climatology (WMO-CCL) and the Climate Variability and Predictability (CLIVAR) Joint Working Group on Climate Change Detection held a meeting in Geneva in November 1999 and recommended the development of indices focusing on indicators of changing extremes. Later, Frich *et al*. (2002) realised that time series based on just a few extreme events per year or season, or very rare events with large return periods, would rarely provide the robustness needed on traditional seasonal and regional scales. Hence, those authors proposed a "selection of less extreme, and therefore less noisy, but hopefully more robust indicators".

Klein Tank and Können (2003) showed that assessing trends in very rare weather events is difficult because the detection probability decreases the rarer the event. Consequently, many studies focus on indices for rather ‘"moderate" extremes as suggested by Frich *et al*. (2002).

Numerous extreme precipitation indices are described and analyzed in recent literature (e.g., Peterson *et al*., 2001; Frich *et al*., 2002; Kiktev *et al*., 2003; Klein Tank and Können, 2003; Haylock and Goodess, 2004; Kostopoulou and Jones, 2005; Moberg and Jones, 2005). There are three main categories of extreme climate indicators: percentile-, threshold- or duration-based indices. The first category of indices is based upon statistical quantities such as percentiles, so the tails of the statistical distribution are examined and days exceeding (not exceeding) a given high (low) percentile are counted. Klein Tank and Können (2003) argue that the effect on the indices of using either empirical methods for percentile calculations or parametric methods relying on distributions is small. Indices based on percentile thresholds have a clear advantage for climate-change detection studies as they compare the changes in the same parts of the precipitation distributions and thus can be used in studies of wide regions containing a broad range of climates (Haylock and Nicholls, 2000; Brunetti *et al.*, 2001; Griffiths *et al.*, 2003; Klein Tank and Können, 2003).

The indices of the second category are based on counts of days crossing a specified fixed value (e.g. the number of days per year with daily precipitation exceeding 20 mm). Indices based on absolute thresholds are beneficial for impact studies as they can be related with extreme events that affect human society and the natural environment (Klein Tank and Können, 2003). Duration-based indices (e.g. the highest consecutive 5-day precipitation total) allow the characterization of the magnitude of wet/dry spells or heat/cold waves.

The literature review on precipitation extremes will mainly focus works analysing the time and space-time patterns of extreme precipitation indices. First, the relationship between precipitation and physiographic features is discussed (Section 2.3.1), and a review on the interpolation techniques used to map climate data is presented in Section 2.3.2. The evaluation of estimation accuracy is discussed in Section 2.3.2.1. Section 2.3.2.2, not only summarizes the strengths and weaknesses of major spatial interpolation techniques, but also presents a literature survey on mapping of precipitation fields. Finally, a review on extreme precipitation indices is presented in Section 2.3.3.

2.3.1 Relationship between precipitation and physiographic features

It is long recognized that topography and other geographical factors are responsible for considerable spatial heterogeneity of the precipitation distribution at the sub-regional scale (e.g., Martínez-Cob, 1996; Faulkner and Prudhomme, 1998; Prudhomme and Reed, 1998; Brunsdon *et al*., 2001; Daly, 2006). A comprehensive review on the complex relationship between precipitation, airflow and physiographic features of mountainous regions is presented by Johansson and Chen (2003), and Smith and Barstad (2004). Several authors (e.g., Prudhomme and Reed, 1998; Drogue *et al*., 2002) verified that, in general, no more than four morpho-topographic parameters are necessary to reach a good explanation of the spatial variability of rainfall fields in complex mountainous terrain. In fact, according to Daly (2006), the main physiographic features affecting spatial patterns of climate are terrain (i.e., orography) and water bodies. This author also points out slope and aspect, riparian zones, and land use/land cover as spatial climate-forcing factors that are also important at scales of less than 1 km, but that are traditionally not accounted for in climate spatial interpolation.

The relationship between elevation and precipitation is complex and highly variable in space, but in general, precipitation increases with elevation, mainly because of the orographic effect of mountainous terrain (e.g., Prudhomme and Reed, 1998; Goovaerts, 2000; Johansson and Chen, 2003). On the windward side, forced lifting of approaching air masses causes the release of rainfall and an increase in precipitation with elevation. Depending on the mountain size and the efficiency of the release processes, precipitation will decrease on the leeward side, hence the leeward slopes are drier and warmer (Föhn effect) than windward slopes. Moreover, it has also been noticed in several studies that the correlation between elevation and precipitation is stronger for averaged elevation over a larger area (usually a window with square shape) surrounding the observation point, than the effective elevation (e.g., Prudhomme and Reed, 1998; Diodato, 2005; Kyriakidis *et al*., 2001). On the other hand, the correlation between elevation and precipitation decreases with increasing time resolution, thus it is less useful for estimation purposes (Faulkner and Prudhomme, 1998; Lloyd, 2005; Haberlandt, 2007).

Interpolation of climate data making use of physiographic information has been a subject of much research in hydrologic and climatic studies. Areas of great topographic complexity and regions with contrasting atmospheric or oceanic influences present more problems than flatter areas or regions with constant atmospheric patterns (Vicente-Serrano *et al*., 2003).

Interpolation methods performance depends strongly on the region, the variable under study, the data's spatial configuration and density, etc. Consequently, the superiority of a particular interpolation method is difficult to establish, since an interpolation method may be the 'best' for some specific situation and not for others (Isaaks and Srivastava, 1989; Martínez-Cob, 1996).

For example, Lloyd (2005) mapped monthly precipitation for 1999, in Great Britain, using five interpolation schemes and concluded that kriging using elevation as external drift provided the most accurate estimates from March to December, whereas for January and February ordinary kriging performed better. Note that when elevation was used as a secondary variable the accuracy of estimating precipitation was increased for most months, but the increase of complexity introduced in the estimation method did not payoff in all situations.

Accordingly, it is commonly accepted that interpolation techniques that make use of the relationship between existing station data and explanatory physiographic variables (e.g., elevation or distance to the coastline) have the potential to better represent the actual climatic patterns, especially in mountainous areas and in regions with complex atmospheric influences (Prudhomme and Reed, 1998; Daly, 2006). In areas of complex terrain, univariate techniques do not display the spatial richness of climate at local scales and do have higher prediction errors than methods using external physiographic variables (e.g., Prudhomme and Reed, 1999; Goovaerts, 2000; Kyriakidis *et al*., 2001; Vicente-Serrano *et al*., 2003). On the other hand, the work of Haberlandt (2007) shows that the information content of elevation for interpolation of sub-daily precipitation plays only a minor role, although interpolation methods using secondary information, such as radar observations or daily precipitation of a denser network, produced the best results.

2.3.2 Overview of interpolation techniques

A number of methods have been proposed for spatial interpolation of precipitation data. The simplest approach assigns to the unsampled location the nearest sampled value (Thiessen, 1911). This method amounts at drawing around each sampled location a polygon of influence with the boundaries at a distance halfway between sampled pairs.

Inverse distance weighting (IDW) is another simple method that assigns weights in an averaging function based on the inverse of the distance (raised to some power) to every data points located within a given search radius centred on the point of estimate. The underlying principle is that climatic values are more alike between the nearest points than between distant points. In the most applied form, the averaging function is based on the inverse of the distance raised to a power of two, thus this technique is named *inverse square distance*.

Geostatistical estimators, known as kriging – named after its first practitioner (Krige, 1951), are a family of generalized least-squares regression algorithms that provide statistically unbiased estimates of surface values from a set of observations at recorded locations, using the estimated spatial and temporal covariance model of the observed data. When developing the kriging equations the model of spatial covariances, or variogram (inverse function of the spatial covariances), is assumed known. This is a key function of geostatistics and characterizes the variability of the spatial (and temporal) patterns of physical phenomena. Typically, a mathematical variogram model is selected from a small set of authorised ones (e.g. exponential or spherical) and is fitted to experimental semivariogram values calculated from data for given angular and distance classes. The way in which the variogram models are chosen and their parameters are estimated – using automatic or manual fitting procedures that make use of expert knowledge – is controversial (Goovaerts, 1997, pp. 97-107).

Examples of univariate geostatistical techniques commonly used are *simple kriging* (SK) and *ordinary kriging* (OK). Simple kriging assumes a known stationary mean, i.e. the mean of the property of interest is assumed constant across the study region. In ordinary kriging, the unknown mean is estimated as part of the kriging procedure using samples in the local neighbourhood of the point being estimated. OK is usually preferred to SK because it requires neither knowledge nor stationarity of the mean over the entire study region.

As discussed in the previous section, information on distance alone can be insufficient to produce accurate spatial estimates in topographically complex regions. A major advantage of geostatistical prediction is that sparsely sampled observations of the primary variable can be complemented by secondary attributes that are more densely sampled (e.g., Goovaerts 1999, 2000).

Simple kriging with varying local means (SKlm) is a multivariate geostatistical technique that replaces the known stationary mean in the simple kriging estimate by known varying means

derived from secondary information. Depending on the secondary information available (e.g. categorical or continuous data), different estimates of the primary local mean can be used. If the secondary information relates to a continuous attribute, the primary local mean is usually estimated by linear regression. Then, the regression residuals are interpolated by simple kriging and are added to the trend estimate (i.e., the regression-derived estimate), thus this approach is sometimes called *detrended kriging* (Kyriakidis *et al*., 2001).

Unlike the simple kriging procedures, in *kriging with an external drift* (KED) the primary local mean is not estimated through a calibration or regression process prior to the kriging of the primary attribute, but rather derived as part of the kriging procedure using exhaustive secondary information. In SKlm, the trend coefficients are derived once and independently of the kriging system, whereas in the KED approach the regression coefficients of the trend component (called drift) are implicitly estimated through the kriging system within each neighbourhood (Goovaerts, 1997, p. 198).

A popular multivariate geostatistical technique is *ordinary cokriging* (CoK), or simply *cokriging*, which explicitly accounts for the spatial cross-correlation between primary and secondary attributes. The cokriging estimate is a linear combination of neighbouring primary and secondary data. Unlike CoK, in *collocated (ordinary) cokriging* (CCoK), the cokriging system only retains the secondary datum collocated with the location being estimated, but the secondary attribute must be known at all locations being estimated. Although cokriging is very effective when the secondary data are highly correlated with the prediction variable, Goovaerts (1997, pp. 235-240) discusses the advantages of collocated cokriging over cokriging when the secondary data are much more densely sampled than the primary attribute.

Note that in the SKlm and KED approaches the secondary datum provides information only about the primary trend at the location being estimated, whereas cokriging approaches incorporate the secondary datum directly into the computation of the estimate and account for spatial cross-correlation between primary and secondary variables. Interested readers should refer to geostatistical textbooks (e.g., Isaaks and Srivastava, 1989; Goovaerts, 1997) for detailed descriptions of univariate and multivariate geostatistical interpolation methods. There are numerous successful applications of kriging interpolation described in the literature. Goovaerts (1999, 2000), Nicolau (2002) and Nicolau *et al*. (2002) compared the application

of some of these techniques to precipitation fields in Portugal (see Table 2.6 of the 2.3.2.2 *Summary* section for details and other examples of applications).

Another set of methods, allowing incorporating physiographic factors, combines distance weighting interpolation techniques (usually IDW or OK) with regression. A linear regression is fitted between the primary variable and the secondary variable(s), and the residuals are calculated at stations' locations. Under the assumption that predicted and residual values of the regression are uncorrelated, an interpolation method is then applied to the residuals. If the secondary variable(s) is known everywhere in the studied region, the regression prediction can be directly calculated everywhere. The final map of the primary variable is then obtained by combining the (regression) map of predicted values with the (IDW or OK) map of the residuals. This procedure incorporates the local variations through the spatial interpolation of regression residuals. Ninyerola *et al*. (2007) compared different forms of this approach using annually and monthly averaged rainfall for the whole Iberian Peninsula (see Table 2.6 of the 2.3.2.2 *Summary* section for details and other examples of applications).

Other methods fit mathematical spline functions to stations' data points. Smoothing or tension parameters can be introduced into the models, giving more or less smoothed maps. These techniques can also take into account physiographic factors. Spline algorithms are mathematically quite complex but are standard in current GIS software (Vicente-Serrano *et al*., 2003). The software package ANUSPLIN (Hutchinson, 1995) fits thin-plate splines (usually second- or third-order polynomials) through station records in three dimensions: latitude, longitude, and elevation. Recent applications using ANUSPLIN include Price *et al*. (2000), Boer *et al*. (2001) and Hijmans *et al*. (2005). See Table 2.6 of the 2.3.2.2 *Summary* section for details and other examples of the splines approach.

Local regression models are based on multiple regressions, or polynomials, conducted within a moving window or search radii centred on the point of estimate (e.g. Brunsdon *et al*., 2001). While multiple regression models include physiographic variables, polynomials usually just include the nearest neighbours' data of the point being estimated. Thornton *et al*. (1997) proposed a local regression model, named DAYMET, which develops local linear regressions between climate and elevation for each grid cell on a digital elevation model (DEM), using data from surrounding stations. The method is based on the spatial convolution of a truncated Gaussian-weighting filter with the set of station locations. More recently, Hasenauer *et al*.

(2003) developed a DAYMET point version that interpolates daily weather for any location, as it is needed to link existing field observations with missing weather data. This adaptation of DAYMET allows producing weather data (daily minimum and maximum temperature, precipitation, solar radiation, and vapour pressure) for any day and location within Austria starting in 1960.

The Precipitation-elevation Regression on Independent Slopes Model (PRISM) is based on local climate-elevation regression functions (Daly *et al*., 1994). In PRISM, each DEM cell is assigned to a topographic facet by assessing slope orientation. Precipitation is estimated by regression of precipitation and elevation for nearby stations within a DEM grid cell's topographic facet. More recently, PRISM uses weighting functions to incorporate gauge data of neighbouring topographic facets for regressions, which involves a sophisticated parameterization (Daly *et al*., 2002). Station weights are calculated on the basis of an extensive spatial climate knowledge base accounting for spatial variations in climate caused by elevation, terrain orientation, effectiveness of terrain as a barrier to flow, coastal proximity, moisture availability, a two-layer atmosphere, and topographic position (Daly, 2006).

Guan *et al*. (2005) describe a specific model called Atmospheric Effects Detrended Kriging (ASOADeK) which combines local regressions with ordinary kriging of the correspondent residuals to map long-term monthly averaged precipitation in a mountainous region of New Mexico (USA). ASOADeK uses a multivariate linear regression approach conditioned on gauge data to autosearch regional and local climatic settings (i.e., infer the spatial gradient in atmospheric moisture distribution and the effective moisture flux direction) and local orographic effects (the effective terrain elevation and aspect). The observed gauge precipitation data are then spatially detrended by the autosearched regression surface. The spatially detrended gauge data are further interpolated by ordinary kriging to generate a residual precipitation surface. The precipitation map is then constructed by adding the regression surface to the kriged residual surface. Guan *et al*. (2005) show that this methodology gives better estimates than precipitation kriging and precipitation-elevation cokriging, and also that ASOADeK produces maps comparable to the PRISM products for the case study considered. Hijmans *et al*. (2005) compare global climate databases developed using ANUSPLIN, PRISM and DAYMET, while Daly (2006) discusses the methodology used to develop those, and other, products.

Regional regression techniques develop a single, domain-wide, multivariate regression function between climate and physiographic variables. Local variations are dismissed as random, and the climatic map is created based on the general structure of the climate variable at all available points. This approach often explains a large proportion of the climate variability within small domains, or across topographically simple regions. See Table 2.6 of the 2.3.2.2 *Summary* section for details and other examples of regression approaches.

The climatologically aided interpolation (CAI) is a hybrid approach also used to produce spatial climate data sets (Willmott and Robeson, 1995). This method uses an existing spatial climate data set to improve the interpolation of another data set. Daly (2006) discusses the most common forms of this approach. Examples of data sets developing using CAI include New *et al*. (2000) and Daly *et al*. (2004).

Interpolation usually leads to a smoothing of the distribution inferred by the observations and thus to a loss of variance. For example, it is well known that kriging is locally accurate in the minimum error variance sense, but does not provide representations of spatial variability given the "smoothing" effect of kriging (Yamamoto, 2005). Interpolation typically leads to an overestimation of small values and underestimation of large ones. Moreover, the smoothing depends on the local data configuration, since it is minimal close to the data locations and increases as the location being estimated gets farther away from data locations. Such conditional bias is undesirable when trying to detect patterns of extreme attribute values (Goovaerts, 1997, p. 370). The smoothing effect in precipitation data is a serious shortcoming considering the modelling of floods or other extreme hydrological processes (Haberlandt, 2007). To overcome this limitation, geostatistical stochastic simulation has become a widely accepted procedure to reproduce the spatial variability and uncertainty of highly variable phenomena in geosciences (e.g., Franco *et al*., 2006; Bourennane *et al*., 2007). Geostatistical simulation methods describe local data variability based on many, equally probable, realizations of the phenomenon, consistent with the data and its statistical characteristics.

Sequential simulation is a widely used geostatistical tool for obtaining a set of equiprobable simulated realizations of variables from natural phenomena, conditional to observed data, honouring their spatial distribution and uncertainty. *Sequential Gaussian simulation* involves the generation of many independent realizations of a Gaussian (or in case of *sequential indicator simulation*, binary) random field (Deutsch and Journel, 1997; Goovaerts, 1997;

Emery, 2004). While these procedures require the transformation of original variables, direct sequential simulation (DSS) has been proposed (Journel, 1994) for simulating directly in the original data space and does not rely on multi-Gaussian assumptions.

Journel (1994) showed that for the sequential simulation algorithm to reproduce a specific covariance model it suffices that simulated values are drawn from local distributions centred at the simple kriging estimates with a variance corresponding to the simple kriging estimation variance. This result guarantees that the spatial covariance, and the global sample mean and variance, of the original variable are reproduced but not the histogram. To overcome this limitation, Soares (2001) proposed a *direct sequential simulation* (DSS) algorithm that uses the local simple kriging estimates of the mean and variance, not to define the local cumulative distribution function (cdf) but to sample from the global cdf. Oz *et al*., (2003) compare this algorithm to a similar one, and briefly discuss other sequential simulation algorithms that have been proposed to guarantee histogram reproduction. More recently, Robertson *et al*. (2006) proposed two nonparametric approaches and compared them to the algorithm introduced by Oz *et al*., (2003). Robertson *et al*. (2006) concluded that, overall, there is very little difference between the three approaches, with the algorithms resulting in similar histogram and semivariogram reproduction.

There are also a number of geostatistical simulation approaches for the joint simulation of interdependent attributes, such as *sequential Gaussian cosimulation* (Verly, 1993), *stepwise conditional transformation for Gaussian simulation of multiple* variables (Leuangthong and Deutsch, 2003), *unconditional Gaussian cosimulation* (Oliver, 2003), *successive cokriging of indicators* (Vargas-Guzmán and Dimitrakopoulos, 2003), among other forms of *multivariate indicator simulation* (e.g., Emery, 2004). Soares (2001) extended the DSS algorithm for the joint simulation of different variables, thus named *direct sequential cosimulation* (coDSS) algorithm. Instead of simulating all variables simultaneously, this approach simulates each variable in turn conditioned to the previous simulated variable.

2.3.2.1 Accuracy evaluation and assessment of uncertainty

The accuracy and uncertainty of gridded data sets is difficult to assess because the field that is being estimated is unknown between data points. Spatial interpolation errors are interdependent functions of the station-network distribution, the efficacy of the interpolation procedure, and the real (but unknown) spatial distribution of the underlying climatic field

(Willmott and Matsuura, 2006). Error estimates based on model assumptions are useful in a relative sense only, and cannot be compared to those of other models (Daly, 2006).

Cross-validation is a widely used model evaluation method that allows comparing different interpolation techniques (e.g., Hevesi *et al*., 1992; Goovaerts, 2000; Kyriakidis *et al*., 2001), even though the comparison is valid only when all of the parameters of the interpolation – the domain, input data, grid resolution, etc. – are identical (Daly, 2006). In *jackknife cross-validation*, also known as "leave-one-out" cross-validation, sample values are deleted from the dataset, one at the time, and then the interpolation method is applied to estimate the missing value using the remaining observed values[9]. Once the process is complete, the estimation errors are calculated as the differences between estimated and observed values. Ideally, the distribution of these errors is centred in zero.

At each station's location, the bias of estimation can be calculated as the ratio of the estimation error to the observation multiplied by 100 (Prudhomme and Reed, 1999). Overall error statistics commonly computed include the mean bias error (ME) or the mean absolute error (MAE) that check if the estimation is biased, and the root mean square error (RMSE) which provides a measure of accuracy of the method. Haberlandt (2007) used the ratio of the variance of estimated values to the variance of observed values to assess the ability of the interpolation method to preserve the variance, because interpolation usually leads to a smoothing of the observations and thus to a loss of variance.

Willmott and Matsuura (2006) describe an approach to spatial cross-validation, and examine three average-error statistics with respect to their abilities to evaluate spatial interpolators: the root-mean-square error (RMSE), the mean absolute error (MAE), and the mean bias error (ME). Their analysis indicates that the RMSE is an inappropriate measure of average error because it is a function of three characteristics of a set of errors, rather than of one (the average error). Moreover, Willmott and Matsuura (2006) concluded that MAE and ME are the most natural measures of spatial-average interpolation error, and that (unlike RMSE) they are unambiguous measures of spatial-average error.

---

[9] The common kriging cross-validation procedure recalculates the kriging weights for each cross-validated point using the same variogram model, which is previously defined using all of the data points.

A disadvantage of cross-validation is that it is susceptible to clustering. If observations appear in clusters, removing a single observation has little effect on prediction performance at the location of the removed observation since nearby observations provide most of the information for prediction. The clustering disadvantage of cross-validation can be overcome, whenever the monitoring stations network is dense enough, by *removing samples* procedures (Martínez-Cob, 1996). In these approaches, a set of validation observations are (randomly or subjectively) select to be excluded from the interpolation process, which is then used to estimate the missing values based on the remaining observations (e.g., Boer *et al*., 2001; Drogue *et al*., 2002; Vicente-Serrano *et al*., 2003). Like in cross-validation, the estimation errors are calculated as the differences between estimated and observed values. The obvious disadvantage of both approaches is that estimation error statistics are limited to locations for which stations exist.

As stated before, geostatistical simulation methods generate a set of alternative realizations of the spatial distribution of an attribute. The series of simulated maps can be post-processed and the spatial uncertainty summarized using probability maps, quantile maps, and maps of spread (Goovaerts, 1997, pp. 431-436). Hence, the uncertainty at an unsampled location can be evaluated through spread measures, such as the variance or the interquartile range, derived from the corresponding local histogram.

2.3.2.2 Summary

This section summarizes the previous ones by presenting an overview of strengths and weaknesses of major interpolation techniques used to map climate data (Table 2.4), and by summarizing (Table 2.5) and describing (Table 2.6) a number of applications on mapping of precipitation fields.

A literature review on the spatial interpolation of precipitation for Portugal is provided by Nicolau (2002). In her Ph.D. thesis, Nicolau (2002) evaluated the performance of several univariate and multivariate methods for mapping the spatial variability of *averaged* (annual and monthly) precipitation in Portugal. The maximum annual precipitation for 100 and 2 years return periods were also interpolated. Nicolau (2002) verified that kriging using elevation as external drift, performed over a moving window, provided the closer predictions for all the precipitation fields. Further information on this work is provided in Table 2.6.

Table 2.4 – Summary of strengths and weaknesses of major interpolation techniques used to map climate data. If an entry is a specific model, the general interpolation approach it employs is given in parenthesis after the name (extended and adapted from Daly, 2006)

Group of interpolation techniques	Description	Strengths	Weaknesses
Simple methods	Methods that assign to the unsampled location the nearest sampled value.	▪ Readily available ▪ Very easy to apply	▪ Very simple, providing poor estimates ▪ Abrupt spatial discontinuities arise in the values when passing from one polygon to another ▪ Tessellation pattern depends on distribution of data ▪ Spatial dependence between observations is not accounted for ▪ Does not account for secondary information ▪ No errors assessment, only one data point per polygon
Inverse distance weighting (IDW)	Methods that assign weights in averaging function based on the inverse of the distance (raised to some power) to every data points located within a given search radius centred on the point of estimate.	▪ Readily available ▪ Very easy to apply ▪ Quick interpolation from sparse data	▪ Very simple, thus only adequate for a small region with a simple topography and a very dense monitoring stations network ▪ Accounts for distance relationships only
Univariate kriging	Geostatistical methods that provide statistically unbiased estimates of surface values using an estimated spatial and temporal covariance model of the observed data. Weights in the kriging equations are determined such as to minimize the estimation variance.	▪ Readily available ▪ Relatively easy to apply ▪ Accounts for spatial and temporal continuity changes as a function of the distance and direction between any pair of points in space and time ▪ Binary and nominal data can be easily interpolated (e.g. with indicator kriging)	▪ Requires domain-wide semivariogram, which limits size and heterogeneity of domain ▪ Does not account for secondary information

Group of interpolation techniques	Description	Strengths	Weaknesses
Multivariate kriging	Geostatistical methods making use of secondary attributes that are more densely sampled than the primary variable.	▪ Accounts for spatial and temporal continuity changes as a function of the distance and direction between any pair of points in space and time ▪ May account for effects of multiple variables (usually just elevation) on climate patterns - SKlm provides an easy way to incorporate several secondary variables - SKlm and KED introduce the residual spatial correlation to the mapping of the primary variable - CoK accounts for changes in cross-correlation across the study area - CoK is most effective when the covariate is highly correlated with the prediction variable	▪ Not readily available ▪ Requires domain-wide semivariogram, which limits size and heterogeneity of domain ▪ Generally, variogram models have to fit the "linear model of co-regionalization" - In KED, the relationship between primary trend and secondary variable must be linear. If not, an appropriate transformation of the secondary variable is needed (Goovaerts, 1999)
Combination of distance weighting methods and regression	This is a two-stage process where a regression model is fitted to the primary data using the secondary information, followed by the interpolation of the corresponding residuals (e.g., using OK or IDW). Then, the regression map and the residual surface are added together to get the final gridded map.	▪ Readily available ▪ Relatively easy to apply ▪ Accounts for effects of multiple variables (usually latitude, longitude, and elevation) without requiring a co-regionalization model ▪ Regression map accounts for global pattern and residuals surface accounts for local variations (Perry and Hollis, 2005)	▪ Spatial dependence between observations is not accounted for in the regression map ▪ Violation of theoretical assumptions (e.g., building a variogram on residuals from an OLS fitting procedure, i.e. assuming that the residual values are spatially uncorrelated)
Splines	Methods that fit a mathematical spline function to the data points, and can be used for exact interpolation or for "smoothing".	▪ Readily available ▪ Relatively easy to apply ▪ Quick interpolation ▪ Performs better when dense, regularly-spaced data are available	▪ Difficulty handling sharp spatial gradients in relationship because of smoothing

Group of interpolation techniques	Description	Strengths	Weaknesses
Local regression models	▪ Local interpolators break the full sample area into smaller pieces that are each evaluated individually by a particular function. The method is usually based on local regressions of climate versus physiographic variables and/or nearest neighbours.	▪ May account for effects of multiple variables (usually just elevation) on climate patterns ▪ Local regression accounts for spatially varying elevation relationships	▪ Not readily available ▪ Results depend on the fit of the regression model and the quality and detail of the input data surfaces
Regional regression	Regional models use a single interpolation function that is mapped across the entire area of concern. Usually, a domain-wide multivariate regression between climatic and physiographic variables is developed.	▪ Readily available ▪ Relatively easy to apply ▪ Accounts for effects of multiple variables (usually latitude, longitude, and elevation) on climate patterns ▪ Stable statistical relationship	▪ A single, domain-wide relationship limits size and heterogeneity of modelling domain ▪ Results depend strongly on the fit of the regression model and the quality and detail of the input data surfaces ▪ Spatial dependence between observations is not accounted for ▪ May not reproduce station values
ANUSPLIN (thin plate splines: Hutchinson, 1995)	Specific model that fits smoothing splines to the station data in three dimensions: latitude, longitude, and elevation.	▪ Readily available ▪ Relatively easy to apply ▪ Accounts for spatially varying elevation relationships	▪ Simulates elevation relationship only ▪ Difficulty handling sharp spatial gradients in relationship
DAYMET (local regression: Thornton *et al*., 1997)	Specific model that fits local linear regressions of climate versus elevation.	▪ Local regression accounts for spatially varying elevation relationships	▪ Not readily available ▪ Simulates elevation relationship only ▪ Cannot handle nonlinear and nonmonotonic elevation relationships

Group of interpolation techniques	Description	Strengths	Weaknesses
PLUVIA (combination of distance weighting and regression: Drogue *et al*., 2002)	Specific model that determines the best regional linear regression of annual and monthly rainfall amounts versus several morpho-topographic predictors. The interpolated residual fields are produced by inverse square distance.	▪ Accounts for spatially varying elevation relationships as a function of the distance and direction ▪ Also accounts for large- and local-scale topographic effects on the spatial distribution of rainfall	▪ Not readily available ▪ Seems to be less accurate for months with high spatial variability ▪ Spatial analysis can be a little redundant because of the nested windows (but the interdependence and collinearity of parameters observed for small windows decrease when the inter-windows gap increases)
PRISM (local regression: Daly *et al*., 1994, 2002)	Specific model that fits local linear regressions of climate versus elevation, with slopes that vary with elevation. In the current version, station weights are calculated based on an extensive spatial climate knowledge base.	▪ Local regression accounts for spatially varying elevation relationships ▪ Also accounts for effectiveness of terrain as barriers, terrain-induced climate transitions, cold air drainage and inversions, and coastal effects	▪ Not readily available ▪ Requires significant effort to take advantage of full capability
Geostatistical conditional simulation	▪ Describe local data variability based on many, equally probable, realizations of the phenomenon, consistent with the data and its statistical characteristics.	▪ Accounts for spatial and temporal continuity changes as a function of the distance and direction between any pair of points in space and time ▪ May account for effects of multiple variables (usually just elevation) on climate patterns ▪ Reproduces the variance of the sampled data ▪ Provides an estimate of the range of possible values of an attribute at unsampled locations ▪ Spatial uncertainty can be summarized using probability maps, quantile maps, and maps of spread ▪ Allows producing measures of uncertainty, at all grid points, that do not depend on data distribution but rather on local values - DSS simulates directly in the original data space and does not rely on multi-Gaussian assumptions	▪ Not readily available ▪ Computationally intensive and complex

Table 2.5 – Summary of applications on general interpolation techniques used to map precipitation fields (if an entry is a specific model, it is referred under the general interpolation approach it employs)

Reference / example data set URL	Thiessen Polygons	Inverse distance weighting (IDW)	Ordinary kriging (OK)	Simple kriging with varying local means (SKlm)	Kriging with external drift (KED)	Cokriging (CoK)	Other forms of kriging	Combination of distance weighting methods and regression	Splines	Local regression	Regional regression	Geostatistical conditional simulation
Boer *et al*. (2001)			X			X		X	ANUSPLIN			
Carrera-Hernández and Gaskin (2007)			X		X		X					
Daly *et al*. (1994)			X			X		X		PRISM		
Daly *et al*. (2002)										PRISM		
Diodato (2005)			X			X						
Drogue *et al*. (2002)					X	X		PLUVIA				
Faulkner and Prudhomme (1998)								X				
Goovaerts (1999)				X	X	X					X	
Goovaerts (2000)	X	X	X	X	X	X					X	
Guan *et al*. (2005)			X			X		ASOADeK				
Haberlandt (2007)	X	X	X		X		X					
Hasenauer *et al*. (2003)										DAYMET		
Hevesi *et al*. (1992)		X	X			X					X	
Hijmans *et al*. (2005)									ANUSPLIN			
Hundecha and Bárdossy (2005)					X							
Kyriakidis *et al*. (2001)			X	X	X							
Lloyd (2005)		X	X	X	X					X		
Marquínez *et al*. (2003)											X	
Martínez-Cob (1996)			X			X		X				
Nicolau (2002)	X	X	X		X	X			X	X		
Nicolau *et al*. (2002)			X		X					X		

Reference / example data set URL	Thiessen Polygons	Inverse distance weighting (IDW)	Ordinary kriging (OK)	Simple kriging with varying local means (SKlm)	Kriging with external drift (KED)	Cokriging (CoK)	Other forms of kriging	Combination of distance weighting methods and regression	Splines	Local regression	Regional regression	Geostatistical conditional simulation
Ninyerola *et al*. (2007)		X						X	X			
Perry and Hollis (2005)								X				
Price *et al*. (2000)		X							ANUSPLIN			
Prudhomme and Reed (1998)											X	
Prudhomme and Reed (1999)			X					X				
Thornton *et al*. (1997)										DAYMET		
Vicente-Serrano *et al*. (2003)	X	X	X			X	X	X	X		X	
Weisse and Bois (2001)			X					X				
Xia *et al*. (1999)		X									X	
Web site developed to help distribute global climate data sets, documentation and related publications produced by Willmott, Matsuura and collaborators at the Center for Climatic Research, University of Delaware. http://climate.geog.udel.edu/~climate/ (retrieved 11 March 2008)		CAI	CAI									
The DAYMET U.S. Data Center provides a source for daily surface weather data and climatological summaries for the USA. http://www.daymet.org (retrieved 11 March 2008)										DAYMET		
The Digital Climatic Atlas of the Iberian Peninsula provides a set of digital climatic maps of mean air temperature (minimum, mean and maximum), precipitation and solar radiation with spatial resolution of 200 m and monthly and annual temporal resolution. http://www.opengis.uab.es/wms/iberia/en_index.htm (retrieved 11 March 2008)								X				
Met Office gridded climate data sets for the UK (28 weather parameters). http://www.metoffice.gov.uk/research/hadleycentre/obsdata/ukcip/index.html (retrieved 12 February 2008)								X				
The PRISM Group produces and distributes gridded climate data sets of the USA. http://www.prism.oregonstate.edu/ (retrieved 11 March 2008)										PRISM		
WorldClim is a set of global climate grids with a spatial resolution of a square kilometre. http://www.worldclim.org (retrieved 6 March 2008)									ANUSPLIN			

Table 2.6 – Literature survey on mapping of precipitation fields, and best model found whenever comparisons are available

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Boer *et al*. (2001)	▪ **Study domain** - Jalisco State, Mexico ▪ **Grid cells resolution** - Not specified ▪ **Primary data** - Monthly mean precipitation (at least 19 years of records within 1940/90) ▪ **Secondary information** - Elevation	▪ ANUSPLIN - bivariate thin plate spline - partial thin plate spline - trivariate thin plate spline ▪ OK ▪ CoK ▪ Regression-kriging ▪ Trivariate regression-kriging	▪ Overall error statistics computed after *removing samples* procedure (5 validation sets of stations selected by the authors): MSE (mean square error) and MPE (maximal prediction error). ▪ Best models: trivariate thin plate splines and trivariate regression-kriging
Carrera-Hernández and Gaskin (2007)	▪ **Study domain** - Basin of Mexico ▪ **Grid cells resolution** - 200 m ▪ **Primary data** - Daily precipitation (June 1978 and June 1985) ▪ **Secondary information** - Elevation	▪ Ordinary kriging in a global neighbourhood (OK) ▪ Ordinary kriging in a local neighbourhood (OKl) ▪ Block kriging with external drift (BKED) ▪ Kriging with external drift in a global neighbourhood (KED) ▪ Kriging with external drift in a local neighbourhood (KEDl).	▪ RMSE (root mean square error) computed after jackknife cross-validation ▪ The monthly-accumulated maps derived from daily interpolations were sampled at the location of each climatological station and the absolute difference between them and the accumulated point values was computed to obtain the estimation errors. Overall error statistics computed: correlation, mean, standard deviation, minimum and maximum. ▪ Best model: KEDl
Daly *et al*. (1994)	▪ **Study domain** - Western United States ▪ **Grid cells resolution** - Approx. 6 km x 9 km ▪ **Primary data** - Monthly and annual precipitation ▪ **Secondary information** - Elevation	▪ PRISM (local regressions) ▪ OK ▪ CoK ▪ Combination of distance weighting methods (OK?) and regional regression	▪ Overall error statistics computed after jackknife cross-validation of (log of) averaged annual precipitation for the period 1982/88 in Willamette River basin: ME (mean error) and MAE (mean absolute error) ▪ Best model: PRISM

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Diodato (2005)	▪ **Study domain** - Benavento province, Italy ▪ **Grid cells resolution** - 500 m ▪ **Primary data** - Logarithm transformation of averaged annual and seasonal precipitation (computed from monthly records within 1955/99) ▪ **Secondary information** - Elevation derived from 3x3 km DEM - Topographic index (vegetation cover factor multiplied by the square root of elevation)	▪ OK ▪ CoK	▪ Overall error statistics computed after jackknife cross-validation: ME (mean error), RMSE (root mean square error), ASE (average kriging standard error), mean standard error, RMSSE (root-mean-square standardized error). In addition, hypotheses test to compare changes in RMSE and ASE. ▪ Best model: CoK with topographic index
Drogue *et al*. (2002)	▪ **Study domain** - Eastern part of the Rhine-Meuse basin, France and Switzerland ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Annual and monthly rainfall (the data were decimal log-transformed; different periods of analysis within 1971/90) ▪ **Secondary information** - Log-transformation of several morpho-topographic directional predictors calculated on a 1x1 km DEM - Stations' elevation - Stations' geographical coordinates	▪ PLUVIA (specific model combining IDW of power 2 and a different regional regression model for each analysis period) ▪ Kriging using elevation as external drift (KED) ▪ Ordinary Cokriging with elevation (CoK)	▪ Overall error statistics computed after *removing samples* procedure (10 validation sets of stations, selected by the authors): MAE (mean absolute error expressed as a percentage), ME (mean error) and standard deviation of estimation errors. ▪ Best model: CoK

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Faulkner and Prudhomme (1998)	▪ **Study domain** - United Kingdom ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - RMED: median of the annual maximum rainfall, computed using different rainfall durations (1 hour to 8 days) ▪ **Secondary information** - Topographic variables considered by Prudhomme and Reed (1998), plus new averaged variables	▪ Combination of distance weighting methods (OK) and regional regression	▪ Best model for long-duration RMED (1 to 8 days): 3-parameter regression model ▪ Best model for short-duration RMED (1 to 12 hours): 6-parameter regression model
Goovaerts (1999)	▪ **Study domain** - Algarve region, Portugal ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Monthly and annual erosive storm empirical index (computed using monthly rainfall, monthly rainfall for days where precipitation exceeds 10 mm, and monthly number of days where precipitation exceeds 10 mm; January 1970-March 1995) ▪ **Secondary information** - Elevation	▪ SKlm ▪ KED ▪ CCoK ▪ Regional regression	▪ Overall error statistics computed after jackknife cross-validation: MAE (mean absolute error) ▪ Best model: CCoK

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Goovaerts (2000)	▪ **Study domain** - Algarve region, Portugal ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Annual and monthly rainfall averaged over the period January 1970-March 1995 ▪ **Secondary information** - Elevation	▪ Thiessen polygons ▪ Inverse square distance ▪ OK ▪ SKlm ▪ KED ▪ CCoK ▪ Regional regression	▪ Overall error statistics computed after jackknife cross-validation: MSE (mean square error) ▪ Best model: SKlm yield slightly better results than KED and CCoK
Haberlandt (2007)	▪ **Study domain** - 25000 km^2 in southeast Germany ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Hourly rainfall (storm period of the 10th to the 13th of August 2002) ▪ **Secondary information** - Radar observations - Daily precipitation of a denser network - Elevation	▪ Thiessen polygons ▪ Inverse square distance ▪ OK ▪ Ordinary indicator kriging (IK) ▪ KED ▪ Indicator kriging with external drift (IKED)	▪ Overall error statistics computed after jackknife cross-validation: ME (mean error), RMSE (root mean square error normalized with the observed average), CORR (coefficient of correlation), ratio of the variance of estimated values to the variance of observed values ▪ Best model: all additional information used simultaneously with KED
Hevesi *et al*. (1992)	▪ **Study domain** - Yucca Mountain, Nevada, USA ▪ **Grid cells resolution** - Not specified ▪ **Primary data** - Natural logarithm transformation of averaged annual precipitation, multiplied by 1000 (8-53 years records length) ▪ **Secondary information** - Elevation	▪ OK ▪ CoK ▪ Neighbourhood averaging ▪ 3 different inverse distance methods ▪ 2 regression equations	▪ Overall error statistics computed after jackknife cross-validation: PAEE (percent average estimation error), RMSE (relative mean square error), SMSE (standardized mean square error). Calculated estimation variances were assumed consistent with observed RMSE if SMSE was within the interval $[1 \pm 2(2/n)]^{½}$. ▪ Best model: CoK

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Kyriakidis *et al*. (2001)	▪ **Study domain** - Northern California coastal region, USA ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Precipitation measurements representing the seasonal [November–December–January (NDJ)] average of daily rainfall for 1 November 1981–31 January 1982 ▪ **Secondary information** - 13×13 km^2 window averaged elevation - Lower-atmosphere state variable related to specific humidity - Lower-atmosphere state variable related to vertical wind - Variables characterizing the interactions between those three features	▪ OK ▪ SKlm ▪ KED	▪ Overall error statistics computed after jackknife cross-validation: RMSE (root mean square error), correlation between cross-validation estimates and true (sample) values, correlation between cross-validation errors and true (sample) values. In addition, traditional regression diagnosis analysis between predictors and cross-validation errors. ▪ RMSE computed after *removing samples* procedure (15 stations with high precipitation values) ▪ Best model: SKlm using as predictors: vertical wind, interaction between humidity and elevation, and interaction of humidity with elevation and vertical wind.
Lloyd (2005)	▪ **Study domain** - Great Britain ▪ **Grid cells resolution** - 661.1 m ▪ **Primary data** - Monthly rainfall for 1999 ▪ **Secondary information** - Elevation	▪ Inverse square distance ▪ OK ▪ SKlm ▪ KED ▪ Local regression	▪ Overall error statistics computed after jackknife cross-validation: ME (mean error), RMSE (root mean square error) ▪ Best model: KED for all months from March to December, and OK for January and February

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Marquínez *et al*. (2003)	▪ **Study domain** - Autonomous Region of Asturias in northern Spain ▪ **Grid cells resolution** - 200 m ▪ **Primary data** - Annual precipitation averaged over 1966/90 - Mean monthly precipitation for the dry season (June/September) averaged over 1966/90 - Mean monthly precipitation for the wet season (rest of the year) averaged over 1966/90 ▪ **Secondary information** - Elevation - Euclidean distance from the coastline - Shortest distance to an arbitrary line further west than any point in the area - Elevation derived over sub-basins - Slope derived over sub-basins	▪ Regional regression	▪ Traditional regression diagnosis analysis and overall error statistics computed after *removing samples* procedure (33 stations, selected by the authors): ME (mean error), MAE (mean absolute error expressed as a percentage). ▪ Best model: 5-parameter regression model
Martínez-Cob (1996)	▪ **Study domain** - Aragón, northeast Spain ▪ **Grid cells resolution** - 5 km ▪ **Primary data** - Base-ten logarithm transformation of long-term mean values of total annual precipitation (10-20 years records length for most weather stations, and up to 50 years for some of them) ▪ **Secondary information** - Base-ten logarithm transformation of elevation	▪ OK ▪ CoK ▪ Combination of distance weighting methods and regression	▪ Overall error statistics computed after *removing samples* procedure (randomly selected stations): MAE (mean absolute error), MSE (mean squared error), EEV (estimation error variance). Calculated EEV were assumed consistent with true errors if SMSE (standardized mean square error) was within the interval $[1 \pm 2(2/n)]^{½}$) ▪ Best model for the precipitation variable: CoK

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Nicolau (2002)	▪ **Study domain** - Continental Portugal ▪ **Grid cells resolution** - 100 m, 250 m, 500 m, 1000 m - 25 m, 50 m, 100 m, 250 m, 500 m, 1000 m for methods using elevation as only secondary data ▪ **Primary data** - Annual and monthly precipitation averaged over 1959/60 – 1990/91 (hydrological years) - Precipitation in dry and wet years averaged over 1959/60 – 1990/91 (hydrological years) - Maximum annual precipitation for 100 and 2 years return periods (determined for stations with at least 30 years of observations and measured until the hydrological year of 1994/95) ▪ **Secondary information** 125 variables related to: - Easting and northing - Elevation - Slope - Distance to the coastline - Dominant orientation of the hillsides - Counting of blockages to the advance of the air masses - Altimetry platforms reached since the coastline - Altimetry barriers in the neighbourhood of each cell (Different radii and directions were also considered)	▪ Thiessen polygons ▪ Delaunay triangulation ▪ Polynomial interpolation ▪ Thin plate splines ▪ IDW (powers of 1, 2 and 3) ▪ OK ▪ Local regression ▪ KED (using 1, 2 or 3 aux. variables) ▪ CoK with elevation	▪ Conclusions related to secondary information: - The maximum annual precipitation variables were not strongly correlated with the secondary variables. - The correlation between precipitation and elevation was significantly higher when it was evaluated in a local neighbourhood. - The overall conclusion, after a local correlation analysis, was that elevation was the most important variable to explain the phenomena when analyzed on restricted neighbourhoods. ▪ Conclusions related to grid cells resolution: - The 1 km grid square resolution proved to be the best one for mapping the spatial variability of 14 of the 17 precipitation fields. ▪ Overall error statistics computed after jackknife cross-validation: Pearson's correlation coefficient between estimated and real values, ME (mean error), MAE (mean absolute error), MAE expressed as a percentage, MSE (mean square error), and an overall error indicator that was proposed to summarize all of those previous ones. ▪ Best model: KED with elevation using1 km grid square resolution.

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Nicolau *et al*. (2002)	▪ **Study domain** - Continental Portugal ▪ **Grid cells resolution** - 500 m ▪ **Primary data** - Annual and monthly precipitation averaged over 1959/60 – 1990/91 (hydrological years) - Precipitation in dry and wet years averaged over 1959/60 – 1990/91 (hydrological years) ▪ **Secondary information** - Elevation	▪ OK ▪ KED ▪ Local regression	▪ Overall error statistics computed after jackknife cross-validation: ME (mean error), MAE (mean absolute error expressed as a percentage), MSE (mean square error). ▪ Best model: KED
Ninyerola *et al*. (2007)	▪ **Study domain** - Iberian Peninsula ▪ **Grid cells resolution** - 200 m ▪ **Primary data** - Annual and monthly rainfall averaged over 1950/99 ▪ **Secondary information** - Elevation - Latitude - Different types of distance from the sea (linear, logarithmic and quadratic) - Terrain curvature - Solar radiation	▪ Combination of distance weighting methods (splines and inverse square distance) and regional regression (whole Peninsula) ▪ Combination of distance weighting methods (splines and inverse square distance) and "local" regressions (main drainage basins) ▪ Splines (as univariate interpolator) ▪ Inverse square distance	▪ 40% of the meteorological stations were randomly selected and model comparisons were based on the determination coefficients from simple regressions between independent observed data and predicted values. ▪ Best model (based on the mean of the determination coefficients): regional regression with splines interpolation of the residuals.

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Perry and Hollis (2005)	▪ **Study domain** - United Kingdom ▪ **Grid cells resolution** - 5 km - One grid for each year and one for each month of the 1961/2000 period, derived independently ▪ **Primary data** Monthly elements (records within 1961/2000): - Precipitation total - Nº days with rainfall ≥0.2mm - Nº days with rainfall ≥1mm - Nº days with rainfall ≥10mm Annual elements (records within 1961/2000): - Greatest five-day precipitation total - Rainfall intensity: average precipitation in wet days (rainfall ≥1mm) - Maximum number of consecutive dry days (rainfall <0.2mm) ▪ **Secondary information** - Easting and northing - Elevation - Terrain shape (mean altitude over a 5 km radius centred 10 km to the north, east, south and west of the station, or alternatively the mean altitude within a 5 km radius of the station) - Coastal effect (percentage of open water, including lakes and sea, within a 5 km radius of the station) - Urban effect (percentage of urban land use within a 5 km radius of the station) (Different radii and directions were also considered)	▪ Combination of distance weighting methods (IDW) and regional regression	▪ Overall error statistics computed after *removing samples* procedure (randomly selected stations): ME (mean error) and RMSE (root mean square error), among others. ▪ Best regression predictors: - Precipitation total → easting and northing - Nº days with rainfall ≥0.2mm → easting and northing, elevation, terrain shape - Nº days with rainfall ≥1mm → easting and northing - Nº days with rainfall ≥10mm → easting and northing, elevation, terrain shape - Greatest five-day precipitation total → easting and northing, elevation, terrain shape, coastal effect - Rainfall intensity: average precipitation in wet days (rainfall ≥1mm) → easting and northing, elevation, terrain shape - Maximum number of consecutive dry days (rainfall <0.2mm) → easting and northing, elevation, coastal effect

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Prudhomme and Reed (1998)	▪ **Study domain** - North west Highlands, Scotland ▪ **Grid cells resolution** - Stations' locations only ▪ **Primary data** - RMED: median of the annual maximum of daily rainfall (stations with at least 10 years of daily records) ▪ **Secondary information** - 14 topographic variables derived from a DTM in a 1x1 km grid (including elevation, geographical position represented by easting and northing in km of gauge grid point, distance from the sea considering different cardinal directions, and slope)	Regional regression	▪ The model fitted to the north west Highlands was tested in three regions: (i) the Highlands, as a whole, and its two sub-regions: the north west Highlands (fitting region) and the Grampian Mountains; (ii) southern Scotland; (iii) the whole of Scotland. Evaluation used the traditional measures of regression diagnosis. ▪ A 4-parameter regression model has been chosen and fitted to the Highlands area. The final model estimates the inverse of RMED using a mixture of geographical parameters (average distance from the sea in opposing directions) and of topographical parameters (obstruction against the prevailing wind, and roughness between the main moisture source and the gauge).
Prudhomme and Reed (1999)	▪ **Study domain** - North west Highlands, Scotland ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - RMED: median of the annual maximum of daily rainfall (stations with at least 10 years of daily records) ▪ **Secondary information** - Average distance from the sea in a 90° sector centred on SW direction - Average distance from the sea in a 90° sector centred on NE direction - Variable representing obstruction - Roughness index	▪ OK ▪ Combination of distance weighting methods (OK) and regional regression	▪ Overall error statistics computed after jackknife cross-validation and *removing samples* procedure (validation stations selected from difficult areas): ME (mean error), RMSE (root mean square error) and bias. ▪ Best model: combination of OK and regional regression

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Vicente-Serrano *et al*. (2003)	▪ **Study domain** - Middle Ebro Valley, Spain ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Natural logarithm transformation of averaged annual precipitation (stations with at least 20 years of annual records within 1950/2000) ▪ **Secondary information** - 21 topographic variables (longitude, latitude, distance to Mediterranean Sea, distance to Cantabrian Sea, incoming solar radiation, elevation, among other variables related to these ones)	▪ Thiessen polygons ▪ Inverse distance weighting (power=1, 2 and 3) ▪ Splines ▪ SK ▪ OK ▪ Block kriging ▪ Directional kriging ▪ Universal kriging ▪ CoK with elevation ▪ Combination of distance weighting methods (inverse square distance and splines) and regional regression ▪ Regional regressions (linear and non-linear)	▪ Overall error statistics computed after *removing samples* procedure (randomly selected stations): ME (mean error), MAE (mean absolute error), RMSE (root mean square error), among others. ▪ Best models for precipitation based on MAE and RMSE: block kriging, followed closely by CoK

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Weisse and Bois (2001)	▪ **Study domain** - French Alps ▪ **Grid cells resolution** - 525, 1050, 1575, 2100, and 2625 m ▪ **Primary data** - Statistical parameters of 1, 2, 3, 6, 12, and 24 h precipitation events with recurrence frequencies of 10 (10-yr rainfall) and 100 yr (100-yr rainfall) Local values of these statistical parameters were estimated for time steps ranging from 1 h to 1 day by fitting a Gumbel distribution using the moments method for samples of seasonal monthly maxima ▪ **Secondary information** - Stations' altitude - Topographic variables corresponding to 10 principal components (obtained by PCA of 25 variables based on the altitudes surrounding a point) - Local variables describing height, exposure, tangents of a site, slope, and azimuths and radii of principal curvature - Regional variables: X and Y coordinates; distance to the Mediterranean; distance to the sea and to the Rhône; shape of the Alps; barrier effect	▪ OK ▪ Combination of distance weighting methods (OK) and regional regression	▪ Error statistics computed after *removing samples* procedure (randomly selected stations): mean and standard deviation of the estimated and real values; and determination coefficient between the estimated and real values. ▪ Best model: combination of OK and regional regression for time steps of 3 h or less; OK for the remaining time steps

Reference	Study domain and data	Interpolation technique(s)	Method used to compare the techniques and best model found
Xia *et al*. (1999)	▪ **Study domain** - Bavarian forest, Germany ▪ **Grid cells resolution** - 1 km ▪ **Primary data** - Monthly precipitation averaged over 1966/95 ▪ **Secondary information** - Elevation - x-direction and y-direction distance (Gauss-Krueger coordinates)	▪ Barnes (Barnes, 1973) ▪ Cressman (Cressman, 1959) ▪ Optimum interpolation ▪ Simple arithmetic average ▪ IDW ▪ Regional regression	▪ MAE – mean absolute error calculated at three validation stations located in typical meteorological zones in the study area (representing a range of terrain and elevation conditions). ▪ Best model for precipitation: Barnes interpolation combined with empirical transfer functions

2.3.3 Indices of precipitation extremes

The joint working group on climate change detection of the World Meteorological Organization – Commission for Climatology (WMO–CCL) and the Research Programme on Climate Variability and Predictability (CLIVAR, Peterson *et al*., 2001; Frich *et al*., 2002; Peterson, 2005) defined a set of key indicators of changing extremes representing a wide variety of climate aspects. The proposed indices should be statistically robust with fairly short return periods. The definitions of some of those indices, as well as their calculations, were reconsidered by some authors mainly because of the estimation of percentiles (Moberg *et al*., 2006; Sillmann and Roeckner, 2008). Table 2.7 presents the major indicators of daily precipitation extremes described and analyzed in recent literature and a few indices providing information about mean conditions, which are sometimes used for inter-comparisons.

Numerous studies of changes in extreme weather events focus on linear trends in the indices, aiming to determine whether there has been a statistically significant shift in such indices of extremes. Table 2.8 summarizes the literature review on extreme precipitation indices.

Table 2.7 – Indicators of daily precipitation extremes (no distinction made between annually, seasonally or monthly specified indices). Indices recommended by Peterson *et al.* (2001) and Frich *et al.* (2002) are denoted in bold

Index	Description and rationale	Units
SDII – *Simple daily intensity index*	It is defined as the average precipitation per wet day. A wet day is usually defined as a day with at least 1 mm of precipitation (R≥1mm). The SDII is particularly defined to examine scenarios of future climate change related with heavy precipitation events, as it monitors precipitation intensity on wet days (Kostopoulou and Jones, 2005).	mm
R30 or R30mm	It is defined as the number of days with at least 30 mm of precipitation. The R30 index characterizes the frequency of extremely heavy precipitation days.	days
R20 or R20mm	It is defined as the number of days with at least 20 mm of precipitation. The R20 index characterizes the frequency of very heavy precipitation days.	days
R10 or R10mm	It is defined as the number of days with at least 10 mm of precipitation. The R10 index characterizes the frequency of heavy precipitation days. This indicator is highly correlated with total annual and seasonal precipitation in most climates (Peterson *et al*., 2001).	days
R1day or RX1day	It is defined as the highest 1-day precipitation total, thus it is an absolute extreme.	mm

Index	Description and rationale	Units
5D-total	It is defined as the highest (not consecutive) 5-day precipitation total. The 5D-total index characterizes the magnitude of extremely heavy precipitation days.	mm
R5D or RX5day	It is defined as the highest consecutive 5-day precipitation total. The R5D index is a measure of short-term precipitation intensity, thus it is a potential flood indicator.	mm
R10D	It is defined as the highest consecutive 10-day precipitation total. The R10D characterizes the magnitude of strong precipitation events.	mm
CWD	It is defined as the maximum number of consecutive wet days. A wet day is usually defined as a day with at least 1 mm of precipitation (R≥1mm). The CWD index characterizes the maximum length of a wet spell.	days
CDD	It is defined as the maximum number of consecutive dry days. A dry day is usually defined as a day with less than 1 mm of precipitation (R<1mm). The CDD index characterizes the maximum length of a dry spell, thus it is a potential drought indicator. Real drought conditions, however, are caused by more complex conditions than captured by CDD, e.g. interactions of precipitation deficits as well as soil and land use characteristics (Tebaldi *et al*., 2006; Sillmann and Roeckner, 2008). It may be relevant for assessing climate changes effects on vegetation and ecosystems. A decrease in this indicator would reflect a wetter climate if change were due to more frequent wet days (Peterson *et al*., 2001; Frich *et al*., 2002).	days
R99p	It is calculated by first determining the 99th percentile threshold (site-specific value) of all events greater than 1 mm over the baseline period, and then counting the number of events above this threshold. The R99p index characterizes the frequency of extremely heavy precipitation days.	days
R99p*	It is calculated by first determining the 99th percentile threshold (mean value) of all events greater than 1 mm over the baseline period, and then counting the number of events above this threshold. The R99p index characterizes the frequency of extremely heavy precipitation days.	days
R95p or R95%	It is calculated by first determining the 95th percentile threshold (site-specific value) of all events greater than 1 mm over the baseline period, and then counting the number of events above this threshold. The R95p index characterizes the frequency of very wet days.	days
R90p or R90N	It is calculated by first determining the 90th percentile threshold (site-specific value) of all events greater than 1 mm over the baseline period, and then counting the number of events above this threshold. The R90p index characterizes the frequency of very heavy precipitation days. Increasing (decreasing) trends of R90p are indicators of a change in the mean of the precipitation distribution towards a wetter (drier) climate (Kostopoulou and Jones, 2005).	days

Index	Description and rationale	Units
R75p or R75%	It is calculated by first determining the 75th percentile threshold (site-specific value) of all events greater than 1 mm over the baseline period, and then counting the number of events above this threshold. The R75p index characterizes the frequency of moderate wet days.	days
R95T or R95pTOT or R95%tot	It is determined as the proportion (percentage) of total precipitation from events greater than long-term 95th percentile of wet days. It is computed as R_w/R_{tot}, where R_w is the sum of daily precipitation amounts above the 95th percentile of the distribution of daily precipitation totals at days with 1 mm or more precipitation in the baseline period; and R_{tot} is the annual precipitation total. The R95T index characterizes the proportion of total precipitation falling during extreme precipitation events. This index may be highly correlated with the number of extreme events, because a year with more events above the threshold will usually show a larger proportion of the total rainfall from these events simply because there are more events (Frich *et al*., 2002). The index R95%tot allows exploring the supposed augmented response of the extreme precipitation events relative to the change in total amount, and it is not sensitive to changes in the number of wet days (Klein Tank and Können, 2003). At stations where the annual amount increases, positive R95%tot trends are indicative of a disproportionate large contribution of the extremes to this wetting. On the other hand, at stations where the annual amount decreases, positive R95%tot trends indicate that the very wet days are less affected than the other wet days. Negative R95%tot trends indicate a smaller than proportional contribution of very wet days to wetting or drying (Klein Tank and Können, 2003).	%
R90T	It is determined as the proportion (percentage) of total precipitation from events greater than long-term 90th percentile of wet days. It is computed as R_w/R_{tot}, where R_w is the sum of daily precipitation amounts above the 90th percentile of the distribution of daily precipitation amounts at days with 1 mm or more precipitation in the baseline period; and R_{tot} is the annual precipitation total. The R90T index characterizes the proportion of total precipitation falling during very intense events.	%
PREC95p	This index is determined as the total precipitation from events greater than the 95th percentile of wet day amounts for a specified baseline period. The PREC95p index characterizes the upper tail of the precipitation distribution providing an indicator of the magnitude of very strong precipitation events.	mm
PREC90p or RQ90	This index is determined as the total precipitation from events greater than the 90th percentile of wet day amounts for a specified baseline period. The PREC90p index characterizes the magnitude of strong precipitation events.	mm

Index	Description and rationale	Units
PREC90p*, PREC95p*, PREC98p*	These indices are determined as the total precipitation from events greater than the 90th, 95th and 98th percentiles, respectively, of all daily precipitation amounts (both wet and dry days) for a specified baseline period. These indices characterize the magnitude of strong to extremely strong precipitation events (Moberg *et al*., 2006).	mm
R4a	It is determined as the average precipitation amount of the highest 4 events. The R4a index characterizes the magnitude of extremely heavy precipitation events by describing changes in the upper percentiles. If R95aw has a stronger positive trend than R4a, it can be explained by considering a rainfall distribution whose shape remains constant over time but whose population (rain days) decreases. As the number of rain days decreases, the 4th highest events would be expected to decrease (Haylock and Nicholls, 2000).	mm
R95aw	It is calculated by first determining the 95th percentile threshold of all events greater than 1 mm over the baseline period, and then averaging the daily rainfall amounts above this threshold. The R95aw characterizes the magnitude of heavy precipitation events. A reduction (increase) in the number of days with low rainfall will lead to an increase (decrease) in this index. This is because the number of rain days will change, which will affect the threshold used in the calculation of the index. With a reduction in the population but no change in the frequency of higher events, the upper percentiles would be expected to rise. This is not the case if the extreme intensity is calculated using all days [R95a index] (Haylock and Nicholls, 2000).	mm
R95a	It is calculated by first determining the 95th percentile threshold of all events over the baseline period, and then averaging the daily rainfall amounts above this threshold. The R95a characterizes the magnitude of heavy precipitation events.	mm
R4a%	It is determined as the ratio between the R4a index and the total rainfall. The R4a% index measures how much of the total rain comes from extremely heavy precipitation events. It is an indicator of changes in the shape of the rainfall distribution.	%
R95aw%	It is determined as the ratio between the R95aw index and the total rainfall. The R95aw% index measures how much of the total rain comes from extreme events. It is an indicator of changes in the shape of the rainfall distribution.	%
R95a%	It is determined as the ratio between the R95a index and the total rainfall. The R95a% index measures how much of the total rain comes from heavy precipitation events. It is an indicator of changes in the shape of the rainfall distribution.	%
C1,…, C10 quantile classes (1st and 10th classes are extreme)	Proportion of daily precipitation falling into 10 precipitation class-intervals in a period compared with the corresponding total precipitation.	%

Index	Description and rationale	Units
Rmean – *Mean climatological precipitation*	If R_{ij} corresponds to the daily precipitation total for day *i* of period *j*, then the mean climatological precipitation in period *j* is $$R_{mean_j} = \sum_{i=1}^{I} R_{ij} / I$$ This is not an extreme index, but rather an indicator of change of the normal moisture availability within a region (Kostopoulou and Jones, 2005).	mm
PCI – *Precipitation concentration index*	If R_{ij} corresponds to the daily precipitation total for day *i* of period *j*, then the precipitation concentration index is $$PCI_j = 100 \frac{\sum_i \left(R_{ij}^2\right)}{\left(\sum_i R_{ij}\right)^2}$$ The PCI is used to analyse and compare the concentration of rainfall due to its emphasis on the relative distribution of rainfall irrespective of the total rainfall received. This is not an extreme indicator. PCI monthly values lower than 10 indicate a uniform distribution of the monthly rainfall along the year; values between 11 and 20 point to a certain seasonal trend, and values above 20 indicate an appreciable variability in the monthly distribution of annual rainfall (Ceballos *et al*., 2004).	%
WD	It is determined as the number of wet days (usually, R>=1mm) This is not an extreme indicator.	days
TP	It is determined as the annual/seasonal/monthly precipitation total. This is not an extreme indicator.	mm
TPw	It is determined as the total precipitation amount on wet days (usually, R>=1mm) This is not an extreme indicator.	mm

About 50 indices for monitoring and analysing changes in climate extremes in Europe were developed under the framework of the project *European Climate Assessment & Dataset* (ECA project, e.g. Klein Tank *et al*., 2002; Klein Tank and Können, 2003). The indices were calculated following the definitions recommended by the WMO–CCL/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (http://eca.knmi.nl/, retrieved 17 March 2008). A dictionary of 40 indices is also provided on the ECA project Web site, giving an overview of which indices have been used by different research groups.

Further development on the definition of extremes indices has been undertaken in a project named *Statistical and Regional Dynamical Downscaling of Extremes for European Regions* (STARDEX project, e.g. Haylock and Goodess, 2004). This project is aimed to "provide a rigorous and systematic inter-comparison and evaluation of statistical, dynamical and

statistical-dynamical downscaling methods for the construction of scenarios of extremes" (STARDEX, http://www.cru.uea.ac.uk/cru/projects/stardex/, retrieved 17 March 2008).

Under the framework of another project, named *European and North Atlantic Daily to Multidecadal Climate Variability* (EMULATE project, e.g. Moberg *et al*., 2006), the list of indices calculated for STARDEX has been extended and the EMULATE catalogue contains 64 climate indices. This project aims to extend the availability of daily historic records of air pressure over the extratropical Atlantic and Europe, for the period 1850 to the present. "Once the record has been extended, atmospheric circulation patterns will be derived and relationships between the circulation, sea-surface temperatures and surface temperature and precipitation patterns across Europe evaluated, and compared with model simulations." (http://www.cru.uea.ac.uk/cru/projects/emulate/, retrieved 17 March 2008).

Table 2.8 – Literature review on extreme precipitation indices (for definitions see Table 2.7)

Reference	Precipitation indices	Methodology / Approach
Alexander *et al*. (2006)	▪ **Study period & region** - Global, 1901/2003 ▪ **Index** - SDII - R1day - R5D - R10 - R20 - CDD - CWD - PREC90p - PREC95p - R95T - TP	▪ The indices were computed using the RClimDex and FClimDex software packages (http://cccma.seos.uvic.ca/ETCCDMI/software.shtml, retrieved 12 March 2008). In the estimation of percentiles, a bootstrapping method was applied to avoid inhomogeneities at the boundaries of baseline periods for all percentile-based indices. ▪ Seasonal and annual indices for the period 1951/2003 were gridded (cells with 2.5 degrees of latitude by 3.75 degrees of longitude) using a modified version of Shepard's angular-distance weighting algorithm. ▪ Trends in station and grid point data were computed and tested for statistical significance. Trends were estimated using the nonparametric Kendall's tau based slope estimator. An iterative procedure was adopted to compute the magnitudes of trends and to test their statistical significance (5% level) taking into consideration the serial correlation in the residuals. The grid boxes used to calculate trends had data ending no earlier than 1999 and had at least 80% of the records for the period under consideration. ▪ A bootstrap technique, similar to the one used by Kiktev *et al*. (2003), was used to test if the pattern of trends estimated from the actual station data was due to climate noise. Since this method is computationally expensive, field significance was only calculated for the annual indices. ▪ Empirical probability distributions of the indices were derived from approximately 600 precipitation stations, with near-complete data for 1901/2003 and covering a very large region of the northern hemisphere mid latitudes and parts of Australia, and then compared for the periods 1901/50, 1951/78 and 1979/2003. The comparisons were done using a 2-tailed Kolmogorov-Smirnov test with a null hypothesis that two cumulative distribution functions computed for two periods are identical.
Bonaccorso *et al*. (2005)	▪ **Study period & region** - Sicily, 1920/2000 ▪ **Index** - R1day	▪ The t-test was used for linear trend detection, and the Mann-Kendall test was used for non-linear trend detection. The trend significance was determined by the traditional asymptotic distributions of the statistics, as well as by a bootstrap approach. ▪ The return period of one storm (21-22 Nov, 2003) was determined using the Gumbel distribution. The effect of trend in the index on this assessment was investigated.

Reference	Precipitation indices	Methodology / Approach
Brunetti *et al*. (2001)	▪ **Study period & region** - Italy, 1951/96 ▪ **Indices** - SDII - C1,…, C10 classes - WD - TP	▪ The regional and sub-regional average series were calculated by simply averaging the series over all the stations in the area. ▪ The slope of the trend was calculated by ordinary least squares (OLS) and the trend significance was determined using the Mann–Kendall nonparametric test.
Ceballos *et al*. (2004)	▪ **Study period & region** - Duero Basin, Spain, 1967/2000 ▪ **Index** - PCI	▪ The time series of the monthly index and the average monthly distribution of annual rainfall were smoothed and, afterwards, the Spearman rank correlation test was applied to assess the trend significance. ▪ PCI monthly values lower than 10 indicate a uniform distribution of the monthly rainfall along the year; values between 11 and 20 point to a certain seasonal trend, and values above 20 indicate an appreciable variability in the monthly distribution of annual rainfall.
Frich *et al*. (2002)	▪ **Study period & region** - Global (except Africa and south America), 1946/99 ▪ **Indices** - SDII - R10 - R5D - CDD - R95T	▪ Each individual station record was divided in half, and the average of one post-1946 multi-decadal period was compared to another by means of a t-test (mathematical description in paper appendix). ▪ Aiming to spatially average the results of each indicator time series, the anomalies in each year of the indicator time series were calculated from a base period of 1961/90 (mathematical description in paper appendix). The trend of each has been calculated by weighting the anomalies according to the number of stations available from the network each year and has been tested for significance using weighted linear regression analysis. ▪ Probability density functions were calculated by giving each indicator a number of 'bins' across its range and calculating the corresponding frequency, similar to producing histograms. The result was then normalised to sum to one to give estimated probabilities rather than a frequency distribution.
Goodess and Jones (2002)	▪ **Study period & region** - Iberian Peninsula, 1958/97 ▪ **Indices** - Rmean - WD - TPw - C1,…, C10 classes	▪ The slope of the trend was calculated by OLS and the trend significance (10% level) was determined using an F-test. ▪ Principal component analysis was performed on the trends of the C1,…, C10 quantile classes. ▪ Links between an automated circulation classification scheme and the North Atlantic Oscillation (NAO) were analysed, as were the direct links between rainfall and the NAO. ▪ Trends in rainfall and circulation-type frequency are compared by linear regression analyses using circulation-type frequencies as predictor variables for rainfall totals for winter months.

Reference	Precipitation indices	Methodology / Approach
Griffiths *et al*. (2003)	▪ **Study period & region** - South Pacific, 1961/2000 ▪ **Indices** - WD - TP - 5D-total - CDD - R99p* - R4a - R4a%	▪ The slope of the trend was calculated by OLS and the trend significance (5% level) was determined using the Mann-Kendall ranked t test. ▪ The relationship between the indices was investigated through several procedures. Correlations between the indices were computed and their significance was assessed using p-values adjusted by the Bonferroni correction. A principal component analysis and a hierarchical cluster analysis were also performed. ▪ A t-test using the Bonferroni correction was used to verify differences between orographic extremes (stations located in mountainous regions) and convective extremes (stations located in flat regions). ▪ The Pettitt test was used to identify change points (climatic jumps) in the indices at each individual station. Whenever a break point was found, the trend was also tested before and after the break.
Haylock and Goodess (2004)	▪ **Study period & region** - Europe, 1958/2000 ▪ **Indices** - R10D - CDD - R90p values normalized for the number of missing days - PREC90p	▪ The indices were computed using the STARDEX software (http://www.cru.uea.ac.uk/cru/projects/stardex/, retrieved 12 March 2008). ▪ The slope of the trend was calculated by a three-group resistant line method and the trend significance (5% level) was assessed by the Kendall-tau test. ▪ A principal component analysis, using Monte Carlo simulations to find the number of components to retain for rotation, was performed to determine the major modes of inter-annual variability. ▪ The relationship between each index and mean atmospheric circulation was investigated by - looking at the relationship between the principal components of the index and other surface and upper-air variables; - canonical correlation analysis to quantify the relationship between the index and sea-level pressure.

Reference	Precipitation indices	Methodology / Approach
Haylock and Nicholls (2000)	▪ **Study period & region** - Australia, 1910/98 ▪ **Indices** - R95p - R4a - R95aw (a cut-off of 15 days was used as the minimum number of rain days for which the 95th percentile could be calculated) - R95a - R4a% - R95aw% - R95a%	▪ The slope of the trend was calculated by OLS and the trend significance (5% level) was determined using the Kendall-tau test. ▪ The problems associated with determining statistical significance of trends of series that are averages of other series have been addressed using an alternative test of significance. Each component series was normalized by replacing each value in the series with its rank, then applying the Kendall-tau test to the average of the ranked series. All trends that were found to be significant using the Kendall-tau test on the average of the raw series were also significant using this 'rank' test. ▪ The relationship between each index and total rainfall was assessed by calculating correlations between them.
Hundecha and Bárdossy (2005)	▪ **Study period & region** - German part of the Rhin basin, 1958/2001 ▪ **Indices** - SDII - 5D-total - CDD - R90p - R90T - PREC90p	▪ The indices were computed using the STARDEX software (http://www.cru.uea.ac.uk/cru/projects/stardex/, retrieved 12 March 2008). ▪ The Kendall-tau test was used with permutation to assess the presence of trend and to calculate the corresponding significance level. The trend magnitude was determined by OLS. ▪ Daily precipitation was interpolated on a 5km×5km grid through external drift kriging using as secondary variable the square root of topographic elevation. Afterwards, the indices were calculated on grids of 5, 10, 25, 50 km to investigate the effect of upscaling on trends, and to make comparisons with trend analysis at point scale.
Kiktev *et al*. (2003)	▪ **Study period & region** - Global, 1950/95 ▪ **Indices** - SDII - R10 - R5D - CDD	▪ The slope of the trend was calculated by OLS for stations with more than 35 years of data. ▪ A spatial autocorrelation function was determined as the region-mean correlation estimated for each 100–km interval, and an exponential decay function was fitted by OLS (applied to indices and trend). ▪ The trends were gridded by a modified version of Shepard's angular-distance weighting algorithm. ▪ Bootstrap was used to estimate the significance (10% level for SDII, and 5% level for R10, R5D, CDD) of trend at each grid point. ▪ Bootstrap was used to test if the pattern of trends estimated from the actual station data was due to climatic noise. ▪ The gridded trend estimates were compared with those simulated by a suite of climate model runs using probability distribution functions by estimating similarity measures (centred pattern correlation, congruence, regression, amplitude).

Reference	Precipitation indices	Methodology / Approach
Klein Tank *et al*. (2002)	▪ **Study period & region** - Europe, 1946/99 ▪ **Indices** - SDII - WD - TP	▪ The slope of the trend was calculated by OLS and the trend significance was determined using a t-test (5% and 25% levels).
Klein Tank and Können (2003)	▪ **Study period & region** - Europe, 1946/99 ▪ **Indices** - R20 - R10 - R1day - R5D - R75p - R95p - R95T	▪ Propose an expression to determine indices' return periods, and show that the selected indices are expressions of events with return periods of 5–60 days. ▪ The slope of the trend was calculated by OLS and the trend significance was determined using a Student's t-test (5% and 25% levels). ▪ The probability of detecting trends in time series depends on the trend magnitude, the record length, and the statistical properties of the variable of interest, especially the variance. Therefore, the signal-to-noise ratio was proposed to detect the relative trend with a given probability of *q%*. ▪ The dependence of signal-to-noise ratio on the significance level for different detection probabilities was analysed. Moreover, the relative trend that is required for a given *q%* detection probability (significance level 5%) was also investigated.
Kostopoulou and Jones (2005)	▪ **Study period & region** - Eastern Mediterranean, 1958/2000 ▪ **Indices** - SDII - Rmean - R10 - R5D - CDD - R90p - R90T - PREC90p	▪ The indices were computed using the STARDEX software (http://www.cru.uea.ac.uk/cru/projects/stardex/, retrieved 12 March 2008). ▪ The slope of the trend was calculated by OLS and the trend significance was determined using the Kendall-tau test (5% level). ▪ The time series of regional standardised anomalies were determined using the methodology described by Frich *et al*. (2002).
Martínez-Casasnovas *et al*. (2002)	▪ **Study period & region** - Alt Penedès–Anoia region, Catalonia, Spain, 19??/2000 ▪ **Index** - PCI	▪ This study characterizes an extreme rainfall event (10 June 2000) that caused soil erosion in the region.

Reference	Precipitation indices	Methodology / Approach
Moberg and Jones (2005)	▪ **Study period & region** - Central and western Europe, 1901/99 ▪ **Indices** - SDII - R5D - CDD - R90p - R90T - PREC90p	▪ The indices were computed using the ClimDex software (http://cccma.seos.uvic.ca/ETCCDMI/software.shtml, retrieved 12 March 2008). ▪ The slope of the trend was calculated by OLS and the resistant method. ▪ The trend significance (5% level) was determined using the methodology described by Kiktev *et al*. (2003).

Reference	Precipitation indices	Methodology / Approach
Moberg *et al*. (2006)	▪ **Study period & region** - Europe, 1901/2000 ▪ **Indices** - SDII - PREC90p* - PREC95p* - PREC98p* - TP	▪ The indices were computed for the winter (Dec.-Feb.) and summer (Jun.-Aug.) seasons using the software developed for the EMULATE project (http://www.cru.uea.ac.uk/cru/projects/emulate/, retrieved 17 March 2008). ▪ In the estimation of percentiles, a bootstrapping method was applied to avoid inhomogeneities at the boundaries of baseline periods for all indices that count the number of occurrences above (or below) a percentile-based threshold. ▪ The slope of the trend was calculated by OLS and the trend significance (5% level) was determined using a two-tailed t-test where the degrees of freedom were reduced to account for serial correlation (lag-1 autocorrelation). ▪ Trend analysis was performed in three different ways for each index: (i) comparing Europe-average trends in the means with trends in the percentiles; (ii) analysing the spatial distribution of statistically significant trends for different indices; (iii) comparing the pattern of trends in the various indices in six selected sub-regions. i) For stations satisfying completeness criteria, the average trends were calculated for each index to obtain Europe-wide averages. Estimates of the uncertainties in these averages were provided by classical 95% confidence intervals using the t-distribution. ii) A number of maps showing trends in indices were produced using more stations because the completeness criteria were relaxed. iii) Precipitation trends were expressed as percentages of the 1961/90 climatological average for each respective index. To account for the problem of very rare precipitation events at some stations, the 1961/90 seasonal mean values of the precipitation indices must exceed 1 mm for inclusion in the trend analyses. ▪ The linear (Pearson) correlation coefficient was calculated to see how similar, or how different, the index time series for mean conditions are compared with those for the various percentiles.

Reference	Precipitation indices	Methodology / Approach
Sillmann and Roeckner (2008)	▪ **Study period & region** - Global (most of the world) and regional analysis for three European regions (southern, central and northern Europe) - Indices based on observational data correspond to the 1951/2003 period (HadEX indices) ▪ **Indices** - SDII - R1day - R5D - R10 - R20 - CDD - CWD - PREC95p (denoted by *R95p* in the paper) - TPw (denoted by *PRCPTOT* in the paper)	▪ The "model-based indices" were calculated on the basis of the global climate model ECHAM5/MPI-OM simulations of the twentieth century and SRES A1B and B1 emission scenarios for the twenty-first century, using the FClimDex software (http://cccma.seos.uvic.ca/ETCCDMI/software.shtml, retrieved 17 March 2008). ▪ All observation-based indices were first calculated for all weather stations and then interpolated onto the latitude-longitude grid (3.75°×2.5°, 96×73 grid boxes), whereas the model-based indices were calculated from the variables representative for the whole gridbox area. The authors did not specify which interpolation technique was used. ▪ For model evaluation, the model-based indices representing the present climate were compared with the observation-based indices using global maps and time series corresponding to spatial averages of the indices for the three European regions. The global maps allow a comparison of the large-scale patterns of the individual indices. ▪ The present-day climate state (1971–2000) was derived from the 20C ensemble, the future one (2071–2100) from the three ensemble members of the scenario runs A1B and B1, respectively. To detect changes in the indices, the time means for these periods were compared. The statistical significance of the differences between these climate states was assessed through a non-parametric test.

2.3.3.1 Spatial interpolation of extreme precipitation

The number of studies analyzing space-time patterns of extreme precipitation indices at the regional and local scales is very limited. The literature review of the previous sections shows that the large majority of studies only focus on the temporal linear trends of the indices (e.g., Table 2.6 and Table 2.8). Although desirable, a spatial analysis is sometimes not feasible due to the sparse number of monitoring stations over large study regions (Klein Tank and Können, 2003; Moberg *et al*., 2006).

Faulkner and Prudhomme (1998) and Prudhomme and Reed (1998, 1999) analysed an index of extreme rainfall, named RMED – median of the annual maximum rainfall, which can be computed considering different rainfall durations (e.g., 1 hour or 1 day). Prudhomme and

Reed (1998) found that the spatial distribution of RMED (1 day) did not depend on elevation in a simple way in mountainous regions of Scotland, but rather reflected more complex relationships with relief and position relative to moisture sources. For this reason, Prudhomme and Reed (1999) thought very unlikely that the cokriging method, using just one external variable for interpolation of RMED (1 day), would improve the mapping over Scotland. Therefore, Prudhomme and Reed (1999) compared the application of ordinary kriging and modified residual kriging (uses ordinary kriging to interpolate the residuals of a regional regression using four explanatory topographical variables). Not surprisingly, the later was concluded as the most suitable for mapping the median of annual maximum daily rainfall in Scotland. Faulkner and Prudhomme (1998) used the modified residual kriging procedure to interpolate the RMED, computed using different rainfall durations (from 1 hour to 8 days), over the United Kingdom. These authors concluded that the relationship between topographic features and precipitation decreases with increasing time resolution, making it more difficult to model.

Hundecha and Bárdossy (2005) interpolated the station values of daily precipitation, over a 5km x 5km grid for the German part of the Rhin basin, through external drift kriging using as secondary variable the square root of topographic elevation. Afterwards, several extreme precipitation indices (Table 2.8) were calculated on grids of 5, 10, 25 and 50 km^2 to investigate the effect of upscaling on trends, and to make comparisons with trend analysis at point scale.

Perry and Hollis (2005) used a combination of distance weighting methods (IDW) and regional regression to produce gridded data sets (5 km grid cells) of several extreme precipitation indices (Table 2.6), one for each year and one for each month of the 1961/2000 period, for the United Kingdom.

In order to compare observed indices with global climate model simulations, Kiktev *et al*. (2003) gridded some of the Frich *et al*. (2002) indices data (Table 2.8) onto a regular latitude-longitude grid, using a modified version of Shepard's angular distance weighting (ADW) algorithm. This method was also used by Alexander *et al*. (2006) to produce global grids for a number of extreme precipitation indices (Table 2.8), but the trends computed on the grids showed little significance. Kiktev *et al*. (2003) and Alexander *et al*. (2006) used coarse resolution grids of the world.

Sillmann and Roeckner (2008) computed several indices for precipitation extremes (Table 2.8) based on the global climate model ECHAM5/MPI-OM simulations of the twentieth century and SRES A1B and B1 emission scenarios for the twenty-first century. These model-based indices were calculated from the variables representative for the whole gridbox area, whereas observation-based indices were first calculated for all weather stations and then interpolated onto the latitude-longitude grid. Sillmann and Roeckner (2008) did not specify which interpolation technique was used. The model-based indices representing the present climate were compared with the observation-based indices, and projected changes were analysed (see Table 2.8 for further methodological details). This comparison showed that, depending on the index and region under consideration, the model was able to realistically capture the observed climatological large-scale patterns of temperature and precipitation indices. However, the model biases were substantial for the extreme precipitation indices in southern Europe. Sillmann and Roeckner (2008) argue that, to some extent, this might be due to methodological differences in the computation of observation-based and model-based indices, because the later were computed on coarse resolution grids that lead to a stronger smoothing of extremes. Moreover, according to Santos *et al*. (2007), the direct comparison between the simulated precipitation in a specific gridbox and the precipitation observed at a single station within that gridbox can be misleading, because the numerical models simulate area-averaged values rather than local values.

2.3.3.2 Portugal and Mediterranean regions

Brunetti *et al*. (2001) analysed the trends in the daily intensity of precipitation in Italy from 1951 to 1996 (Table 2.8), and concluded that the trend for the number of wet days per year is significantly negative throughout Italy, stronger in the north than in the south. Their results also show that there is a tendency toward an increase in precipitation intensity. However, they verified that this increase is globally less strong and significant than the decrease in the number of wet days. Moreover, Brunetti *et al*. (2001) concluded that, in northern Italy, the increase in precipitation intensity is mainly owing to a strong increase in the heaviest events, while in central–southern Italy, it depends on a larger part of the distribution of wet day's amounts. The analysis of the evolution of the class-interval contributions (see Table 2.7 for a description) shows that the positive trend of the heaviest events starts in the 1970s, as does the negative trend of lightest events.

Similarly, the results from Kostopoulou and Jones (2005), for the period 1958/2000, show significant positive trends towards intense rainfall events and greater amounts of precipitation on the western part of the study region, which was represented by Italian stations. In contrast, the eastern half shows negative trends in all precipitation indices indicating drier conditions in recent times. Significant positive trends were revealed for the index of maximum number of consecutive dry days (CDD), especially for stations in southern regions.

Haylock and Goodess (2004) analysed the trends in several extreme precipitation indices, for the winter months (Dec.-Feb.) of the period 1958/2000, over Europe (Table 2.8). In the northwest of the Iberian Peninsula, the results showed a small decrease in CDD, while the rest of the peninsula has seen large increases in this index. In contrast, the frequency of very heavy precipitation days (R90p index) showed a decrease over most of the peninsula (including the northwest), but a slight increase in the southeast. Haylock and Goodess (2004) have also analysed the influence of the NAO[10] on the indices. A canonical correlation analysis of each of the two indices (CDD and R90p) with mean sea-level pressure has revealed that the NAO is an important influence on extreme rainfall over Europe. Their results suggest that the observed trends in CDD and R90p are mainly due to changes in the NAO. Furthermore, they found similar results for two other indices: R10D and PREC90p (see Table 2.7).

Moberg *et al*. (2006) analysed the trends in several extreme precipitation indices, for the winter (Dec.-Feb.) and summer (Jun.-Aug.) seasons of the period 1901/2000, over Europe (Table 2.8). The recent observed decrease in Mediterranean winter precipitation in other studies (e.g., Klein Tank and Können, 2003; Kostopoulou and Jones, 2005) did not emerge in a comparable manner on the centennial timescale considered by Moberg *et al*. (2006). Moreover, a dominance of insignificant winter precipitation trends was found on the Iberian Peninsula. Trends in precipitation percentiles in summer were difficult to assess for the Mediterranean climate, because the number of wet days was often so small that trends became difficult to interpret.

Under the framework of the ECA project (http://eca.knmi.nl, retrieved 17 March 2008), about 50 indices of extreme events have been calculated for the 1946/2006 period over Europe, and

[10] Haylock and Goodess (2004) calculated the NAO using the Gibraltar–Iceland pressure difference, and used a Dec.-Feb. average of this index in their analysis.

analysed further by several authors (e.g., Klein Tank *et al*., 2002; Klein Tank and Können, 2003; Moberg and Jones, 2005; Alexander *et al*., 2006). Note that the series used for those indices calculation are "blended series", i.e. most missing values were replaced by using nearby stations' data and updated using synoptical messages. Klein Tank and Können (2003) searched for trends in several annual indices of precipitation extremes for the period 1946/99, over Europe (Table 2.8), and found no significant trends (5% level) in the indices calculated for the southern region of Portugal. Considering the annual precipitation indices available at the ECA project Web site (http://eca.knmi.nl, retrieved 17 March 2008) for the whole 1946/2006 period, the conclusions for the southern region of Portugal are identical, i.e. there are no significant trends in the annual precipitation indices. However, the seasonal analysis reveals a few significant trends (5% level) and regional contrasts in the indices:

- The maximum number of wet days (CWD) has decreasing trends in Lisboa and Tavira stations in the winter-half season (Oct.-Mar.), as well as in Beja station in spring (Mar.–May). In winter (Dec.-Feb.), the decreasing trend in Lisboa and Tavira stations is stronger.
- The frequency of heavy precipitation days (R10) has a negative trend in Tavira in spring (Mar.–May), and Lisboa has a positive trend in autumn (Sep.-Nov.).
- The frequency of very heavy precipitation days (R20) has a negative trend in Beja in spring.
- The highest 1-day precipitation total (R1day) has a positive trend in Beja in the summer-half season (Apr.-Sep.), whereas it has a negative trend in Lisboa in winter and a negative trend in Tavira in spring.
- The intensity of short-term precipitation (R5D) has negative trends in Lisboa and Tavira in spring.
- The frequency of moderate wet days (R75p) has a negative trend in Tavira in spring.
- The frequency of very wet days (R95p) has negative trends in Beja and Tavira in spring.
- The precipitation fraction due to very wet days (R95T) has a negative trend in Lisboa in winter and a negative trend in Tavira in spring.

- The frequency of extremely wet days (R99p) has a negative trend in Tavira in spring, whereas it has a positive trend in summer (June-Aug.) in Beja.

Moreover, a negative trend in spring precipitation, especially in March, has been detected in the south of Portugal (Corte-Real *et al*., 1998; Santo *et al*., 2004; Trigo and DaCamara, 2000).

Rodrigo and Trigo (2007) investigated annual and seasonal trends in five precipitation variables, aiming to analyse the behaviour of daily rainfall in the period 1951–2002, using data from 22 stations scattered across the Iberian Peninsula. Among these stations, five of them are located in southern Portugal: Lisboa, Grândola, Serpa, Relíquias and Monforte. The few significant trends (determined through the Mann–Kendall test at the 5% level) found for these five stations can be summarized as follows:

- The total amount of rainfall has annual negative trends in Relíquias, as well as in winter (Dec.–Feb.) and spring (Mar.–May). Decreasing trends are also reported for Monforte in spring and for yearly values.
- The number of wet days (0.1 mm threshold) has a positive trend in spring for Relíquias, while it has a negative signal in Monforte in spring, summer (Jun.–Aug.) and for yearly values.
- The average precipitation per wet day (*daily intensity index*) has negative trends in Relíquias in all seasons and for yearly values. A decreasing is trend is also reported in Grândola in the summer season.
- The 95th percentile[11] and the percentage of rain falling on days with rainfall above the 95th percentile both have negative trends in Relíquias in all seasons and for yearly values. These variables have decreasing trends in Grândola in spring and summer, whereas Serpa has increasings trends for yearly values.

Based on an *F*-test, Goodess and Jones (2002) identified a significant (at the 10% level) decrease of the number of rain days during the winter season in Mértola for the period 1958/94. Moreover, this decrease was not compensated by an increase of the wet day amount (no trend present). The mean precipitation total also showed a decreasing tendency in winter,

[11] The 95th percentile was determined by fitting an appropriate gamma distribution to the data.

although not significant. For spring and autumn, no significant trends were found for the number of rain days, wet day amount and mean seasonal rainfall in Mértola. Regarding the trend analysis of extreme rainfall amounts, those authors proceeded as follows. For each station and month, all wet-day amounts for the period 1958/97 were sorted into ascending order and then grouped into ten classes. In order to identify any spatially coherent patterns in the data, a principal component analysis of the quantile trends was undertaken. However, the leading winter principal component, for example, explains only 26% of the variance over the peninsula, hence the identified trends are very weak and the results have little consistency.

Tebaldi *et al*. (2006) focused primarily on changes of extremes in future climate projections, whereas Sillmann and Roeckner (2008) not only analysed projected changes but also compared model-based indices with observation-based indices for temperature and precipitation extremes (Table 2.8). The simulations results found by Sillmann and Roeckner (2008) are, in general, consistent with the multi-model study of Tebaldi *et al*. (2006).

Sillmann and Roeckner (2008) verified that the ensemble-mean of the model-based indices underestimates the extreme precipitation evaluated by most of the observation-based indices in southern Europe. The exceptions are the consecutive wet day (CWD) index, which is well captured by the model, and the consecutive dry day (CDD) index that it is overestimated. In southern Europe, the model overestimates CDD by more than 20 days, except in the 1950s when the simulated CDD is closer to the observed one. However, Sillmann and Roeckner (2008) argue that the abrupt change in the simulated CDD is not a real effect but can be attributed to a changed masking of missing values in the data used in the observation-based index. In what concerns the future climate simulations, the CDD index is projected to increase significantly in regions around the Mediterranean Sea, including Portugal. The short-term precipitation intensity (R5D index) decreases slightly in the twenty-first century, whereas the magnitude of very strong precipitation events (PREC95p index) shows hardly any trend in Mediterranean regions. In contrast to the wet extreme precipitation events, CDD is projected to increase substantially in southern Europe: the longest dry period within a year is projected to be prolonged by 1 (1.5) months at the end of this century in B1 scenario (A1B scenario). Nevertheless, in Mediterranean regions, the model biases can be substantial and climate change shows high heterogeneity at the local scale.

In summary, Sillmann and Roeckner (2008) concluded that, in general, the differences between humid and arid climate zones of the world tend to increase under global warming. In southern Europe, extensive irrigation will be required because of higher temperatures, less precipitation, and prolonged dry spells in the future climate. These changes in climate extremes will have a severe impact on living conditions, water supply, and agriculture in the regions around the Mediterranean Sea. Accordingly, regions endangered by desertification are projected to grow. In particular, the Mediterranean regions will face longer drought conditions lasting for more than 3 months, especially in summer.

Chapter 3: HOMOGENIZATION OF PRECIPITATION TIME SERIES

3. Homogenization of precipitation time series

Precipitation is one of the most important climate variables. Accurate quantification of its observed variability is required for a number of purposes. Long-term series of reliable precipitation records are essential for climate change monitoring, general circulation models and regional climate models, modelling of erosion, runoff and pollutant transport, among other applications for ecosystem and hydrological impact modelling.

However, high quality data seldom exist because in reality many types of non-climatic factors (e.g. stations relocations, changes of the surroundings, different observational and calculation procedures, etc.) can cause time series discontinuities which may hide the true climatic signal and patterns, and thus potentially bias the conclusions of climate and hydrological studies (e.g. Peterson *et al.*, 1998; Tuomenvirta, 2001; Auer *et al.*, 2005). Therefore, it is recommended that, besides routine quality control, the homogeneity testing of data to be evaluated before performing those studies (Aguilar *et al.*, 2003).

Several techniques have been developed for detecting inhomogeneities in time series of weather elements. The approaches underlying the homogenization techniques are quite different and typically depend on the type of element (temperature, precipitation, pressure, evaporation, etc.), the temporal resolution of the observations (annual, seasonal, monthly or sub-monthly), the availability of metadata (station history information) and the monitoring station network density (spatial resolution). A review of different statistical methods is presented by Peterson *et al*. (1998), and comparisons between procedures are provided by Ducré-Robitaille *et al*. (2003) and Reeves *et al*. (2007).

The research developed along this thesis focus on the south of continental Portugal, which is one of the most vulnerable regions to desertification in Portugal (e.g., Correia, 2004; Rosário, 2004b). For the homogenization analysis, a set of 45 long-term and 62 short-term series of daily precipitation were compiled.

Before being collected, the precipitation series had already been subject to several basic quality-control procedures and several statistical homogeneity tests, as discussed in Section 3.1.1. Nevertheless, we assumed that the 107 daily precipitation series could contain potential

inhomogeneities, as recommended by Auer *et al.* (2005), and thus several homogenization procedures were applied to all of them. Furthermore, two new homogenization approaches are proposed (Costa and Soares, 2006; Costa *et al.*, 2008a).

The homogeneity assessment of the precipitation time series was accomplished through four major stages. The first one aimed at the identification of errors and suspicious daily precipitation records, and the following stages were dedicated to homogeneity testing. The second stage is an absolute approach that was implemented in order to select a subset of series with quality data, including a set of reference series that are presumed homogeneous by the relative procedures.

The third stage is a relative approach in which, besides the application of several well-established statistical tests, a new procedure is proposed for the detection of non-climatic irregularities. The proposed technique is an extension of the Ellipse test (described by Allen *et al.*, 1998) that takes into account the contemporaneous relationship between several candidate series from the same climatic area (Costa and Soares, 2006).

Finally, in the fourth stage, a geostatistical simulation approach, using the direct sequential simulation algorithm (Soares, 2001), is proposed for inhomogeneities detection in precipitation time series. This relative technique accounts for the joint spatial and temporal dependence between observations, and enhances the pre-eminence of the closer stations, in both spatial and correlation terms (Costa *et al.*, 2008a).

Section 3.1 describes the study domain and the available precipitation data, plus its previous quality control analysis and homogenization assessment. The inhomogeneities detection methodology is detailed in Section 3.2, including the description of the basic quality control procedures developed (Section 3.2.1), the six statistical tests applied (Section 3.2.2), and the proposed extension of the Ellipse test (Section 3.2.3). Several issues related to the relative procedures used are discussed in Section 3.2.4. The proposed geostatistical simulation approach is detailed in Section 3.2.5. Finally, the main results from the homogenization analysis are presented and discussed in Section 3.3, and some conclusions are drawn.

3.1 Study domain and precipitation data

There are two main reasons for the selection of the south of continental Portugal as the study region of this work. First, accurate long-term precipitation series with at least daily resolution are required to characterize precipitation extremes (Easterling *et al.*, 1999; Klein Tank and Können, 2003). For a sub-monthly time scale, such high quality data was not available, at least to us, for the complete continental territory of Portugal. The homogenization and adjustment of all the obtainable climatic series is out of the scope of this work and was not completed due to time constrains. Second, recent studies show that this is one of the most vulnerable regions to desertification in continental Portugal (e.g. Correia, 2004; Rosário, 2004b).

The daily precipitation series analysed were compiled from the European Climate Assessment (ECA) dataset and the National System of Water Resources Information (SNIRH – Sistema Nacional de Informação de Recursos Hídricos) database[12], and are available through free downloads from the ECA website (http://eca.knmi.nl) and the SNIRH website (http://snirh.inag.pt), respectively. The compiled precipitation series were downloaded during the first semester of 2004.

For the homogenization analysis, the study domain is defined by the Arade, Guadiana, Mira, Ribeiras do Algarve and Sado basins. Despite being outside the study domain, data from Lisbon and Badajoz (Spain) stations were also considered. All stations with at least 30 years with less than 5% of observations missing were selected. Shorter series with at least 10 years lacking a maximum of 5% of data were also chosen, and hence the series with too many gaps were discarded. Using those criteria, 45 long-term and 62 shorter series of daily precipitation were accepted for the homogenization analysis. Data from those 107 weather stations were checked in order to verify the length of the series and the occurrence of gaps (Figure 3.1).

Even though the beginning and ending of the series from the SNIRH database is highly variable, 44 long-term series have a common period of observation of 20 years, located in the 1964-1983 interval. Most of the long-term series (more than 90%) cover the standard normal period 1961/1990, and 33% of them extend back to 1931 (Figure 3.2).

[12] Managed by Instituto da Água (INAG).

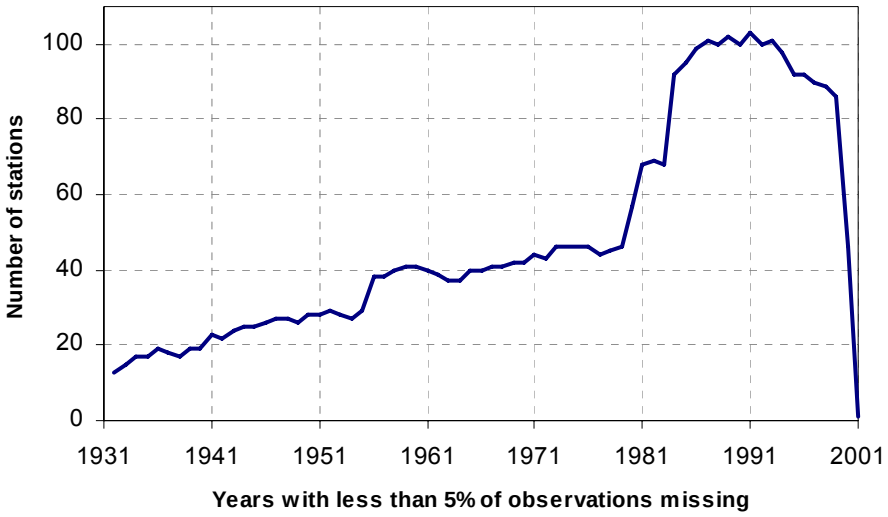

Figure 3.1 – Precipitation series by years lacking a maximum of 5% of data

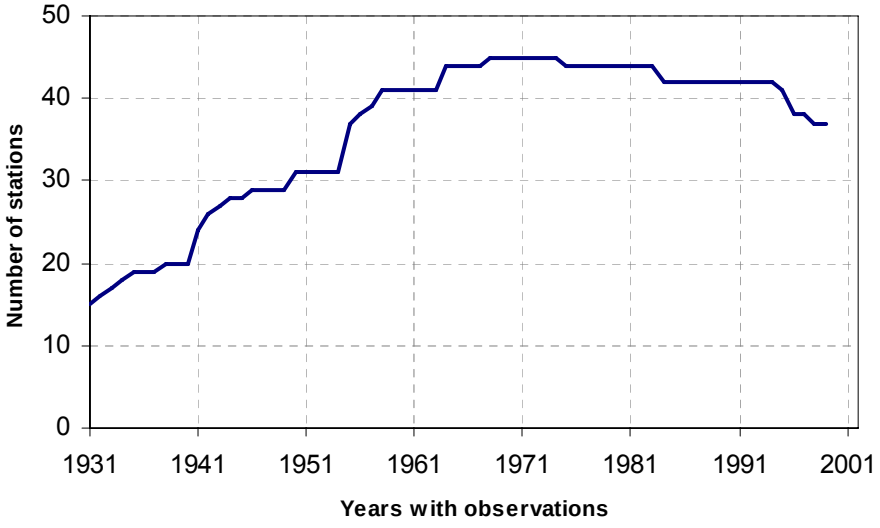

Figure 3.2 – Long-term precipitation series by years

Long-term series have, in average, approximately 51 years of length with less than 5% of observations missing (Figure 3.3) and most of them cover the standard normal period 1961/1990 without too many gaps (Figure 3.4).

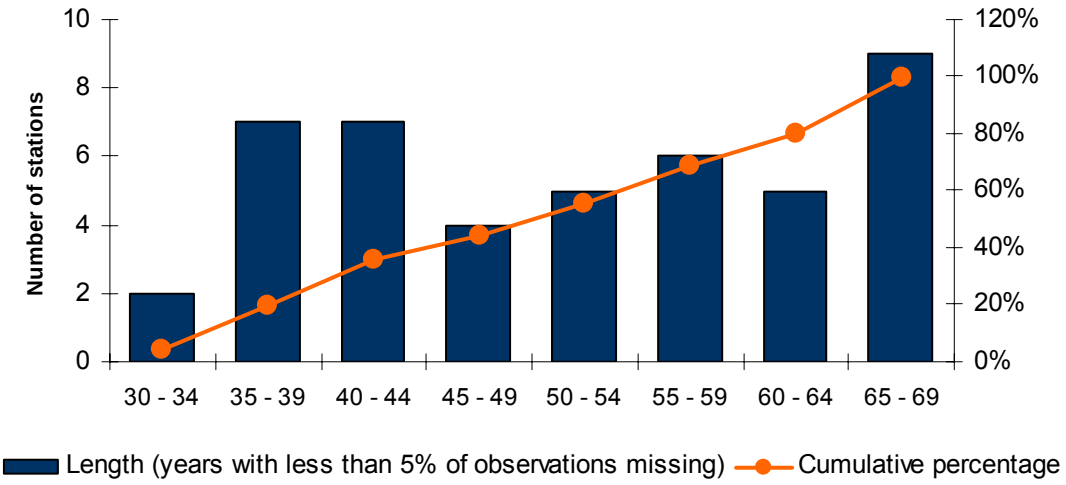

Figure 3.3 – Distribution of long-term series by length (years lacking a maximum of 5% of data)

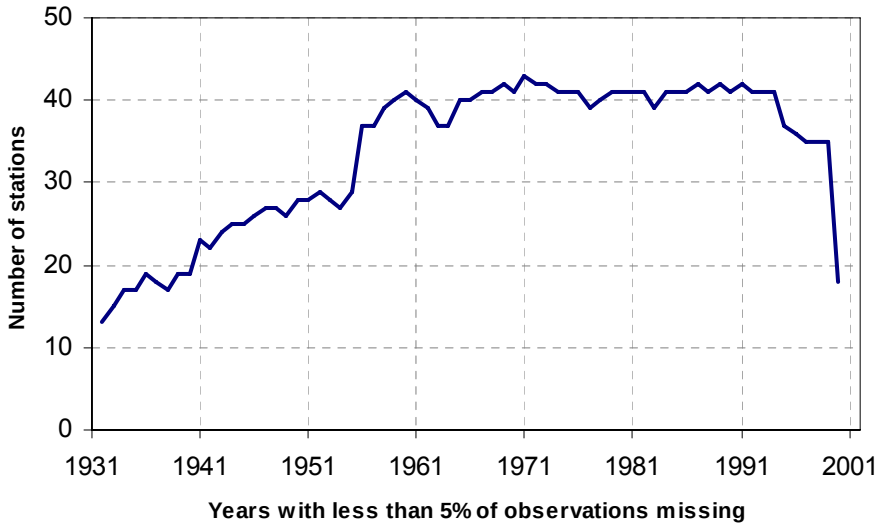

Figure 3.4 – Long-term precipitation series by years lacking a maximum of 5% of data

Figure 3.5 shows the study domain and the geographical distribution of stations for which daily time series have been selected. The data are spatially representative of the study domain that covers approximately 25200 km^2.

Details of all 107 series are presented in Appendix I, including the station name and ID, altitude, location coordinates, available data periods, and series length (years lacking a maximum of 5% of data).

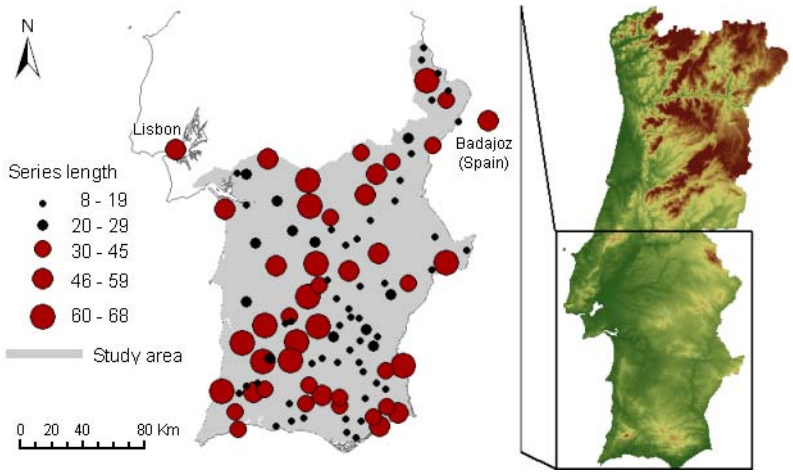

Figure 3.5 – Study area and stations with daily precipitation series. Station dots are scaled with the length (years lacking a maximum of 5% of data) of the time series. Red dots: long-term series. Black dots: short-term series.

The Portuguese border and the Arade, Guadiana, Mira, Ribeiras do Algarve and Sado basins themes used in the maps that are presented throughout the thesis were collected from the "Atlas do Ambiente Digital - Instituto do Ambiente".

3.1.1 Previous quality control analysis

Before being collected for this study, the daily series of the ECA dataset had already been subject to several basic quality-control procedures and four statistical homogeneity tests: the Standard normal homogeneity test (SNHT) for a single break (Alexandersson, 1986), the Buishand range test (Buishand, 1982), the Pettit test (Pettit, 1979), and the Von Neumann ratio test (Von Neumann, 1941). The homogeneity tests were applied to an annual resolution set of variables representing important characteristics of variation at the daily scale. For precipitation, the testing variable used was the wet day count using 1 mm as threshold. Because of the sparse density of the ECA station network, absolute tests were applied rather than relative tests, i.e. testing a candidate station series relative to neighbouring stations' series.

The ECA project used historic metadata information to find supporting evidence of changes in observational routines that may have triggered the irregularities detected. The ECA daily series were not adjusted for the inhomogeneities identified. Instead, the results of the different tests were grouped in an overall classification (‘useful’, ‘doubtful’, and ‘suspect’). The four long-term precipitation series (from Beja (666), Lisboa Geofísica (675), Tavira (681) and

Badajoz Talavera (709)) compiled from the ECA dataset for this study were all marked as 'useful', as the four homogeneity tests did not reject the homogeneity hypothesis, at the 1% level. For further details see Klein Tank *et al.* (2002), Wijngaard (2003), Wijngaard *et al.* (2003), and the ECA project website (http://eca.knmi.nl).

Like most long-term climate time series, the daily precipitation series from the SNIRH database are expected to suffer from several non-climatic factors, although some homogeneity testing of the annual precipitation totals has already been carried out by Nicolau (1999, 2002), for the period 1959/60 – 1990/91. This author performed a double-mass analysis (Kohler, 1949), and also considered an absolute approach using three homogeneity tests: the Wald-Wolfowitz runs test (Wald and Wolfowitz, 1943), the Wilcoxon-Mann-Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947), and a nonparametric ranking test that was not specified. All of the annual precipitation series of the monitoring stations considered here were considered homogeneous. Nevertheless, when the daily series were compiled for this study (first semester of 2004), in the SNIRH website (http://snirh.inag.pt, Dados de Base) it was mentioned that data quality verification had not been concluded.

In summary, the compiled series from the ECA dataset were neither adjusted or tested using relative procedures, and the series from the SNIRH database are expected to suffer from several irregularities since *objective* relative methods were not performed and the full length of each series was not analysed.

Taking into consideration the previous discussion, we assumed that the selected 107 daily precipitation series could contain potential breaks, as recommended by Auer *et al.* (2005), and thus several homogenization procedures were applied to all of them.

3.2 Methodology

As discussed previously in the *Literature review* chapter (Section 2.2.3), there are a number of tests available for the homogenization of climate series with low temporal resolution. However, well-established statistical methods for homogeneity testing sub-monthly precipitation data are lacking (Easterling *et al.*, 1999; Aguilar *et al.*, 2003; Wijngaard *et al.*, 2003; Auer *et al.*, 2005). Furthermore, adjusting daily and hourly data is not straightforward,

thus the WMO makes no recommendations regarding adjusting sub-monthly data (Aguilar *et al.*, 2003).

In order to overcome those limitations and taking into consideration the previous quality control analysis of the selected ECA series (see Section 3.1.1), the homogeneity assessment followed the hybrid approach proposed by Wijngaard *et al.* (2003) for the ECA dataset. Hence, the homogeneity procedures used as the testing variable the annual wet day count with 1 mm as threshold, which is expected to be representative of important characteristics of variation at the daily scale. The results of the different procedures implemented were then used to develop an overall classification of the daily series.

The homogeneity assessment of the precipitation time series was developed through four major stages (Figure 3.6). The first one comprises several basic quality control procedures that aim to identify errors and suspicious daily precipitation records. The emphasis of this stage was on the quality control of precipitation extremes. The second stage is dedicated to absolute homogeneity testing and aims to select a subset of series with quality data, including a set of reference series that are presumed homogeneous by relative procedures. The third stage is a relative approach in which, besides the application of several well-established statistical tests, a new procedure is proposed for the detection of non-climatic irregularities. Finally, in the fourth stage, a geostatistical simulation approach is proposed for inhomogeneities detection in precipitation time series.

The absolute approach comprises the application of six statistical tests to the testing variable, at all locations: the Mann-Kendall test (Mann, 1945; Kendall, 1975), the Wald-Wolfowitz runs test (Wald and Wolfowitz, 1943), the Von Neumann ratio test (Von Neumann, 1941), the Standard normal homogeneity test (SNHT) for a single break (Alexandersson, 1986; Alexandersson and Moberg, 1997), the Pettit test (Pettit, 1979), and the Buishand range test (Buishand, 1982).

The relative approach comprises the application of these last three homogeneity tests (which are capable of locating the year where a break is likely) to long-term composite reference series, and the application of a new procedure to the testing variable. The proposed technique is an extension of the Ellipse test (described by Allen *et al.*, 1998) that takes into account the contemporaneous relationship between several candidate series from the same climatic area

by using the residuals from a *Seemingly unrelated regression equations* (SUR) *model*. The methodology and results from this approach are described by Costa and Soares (2006).

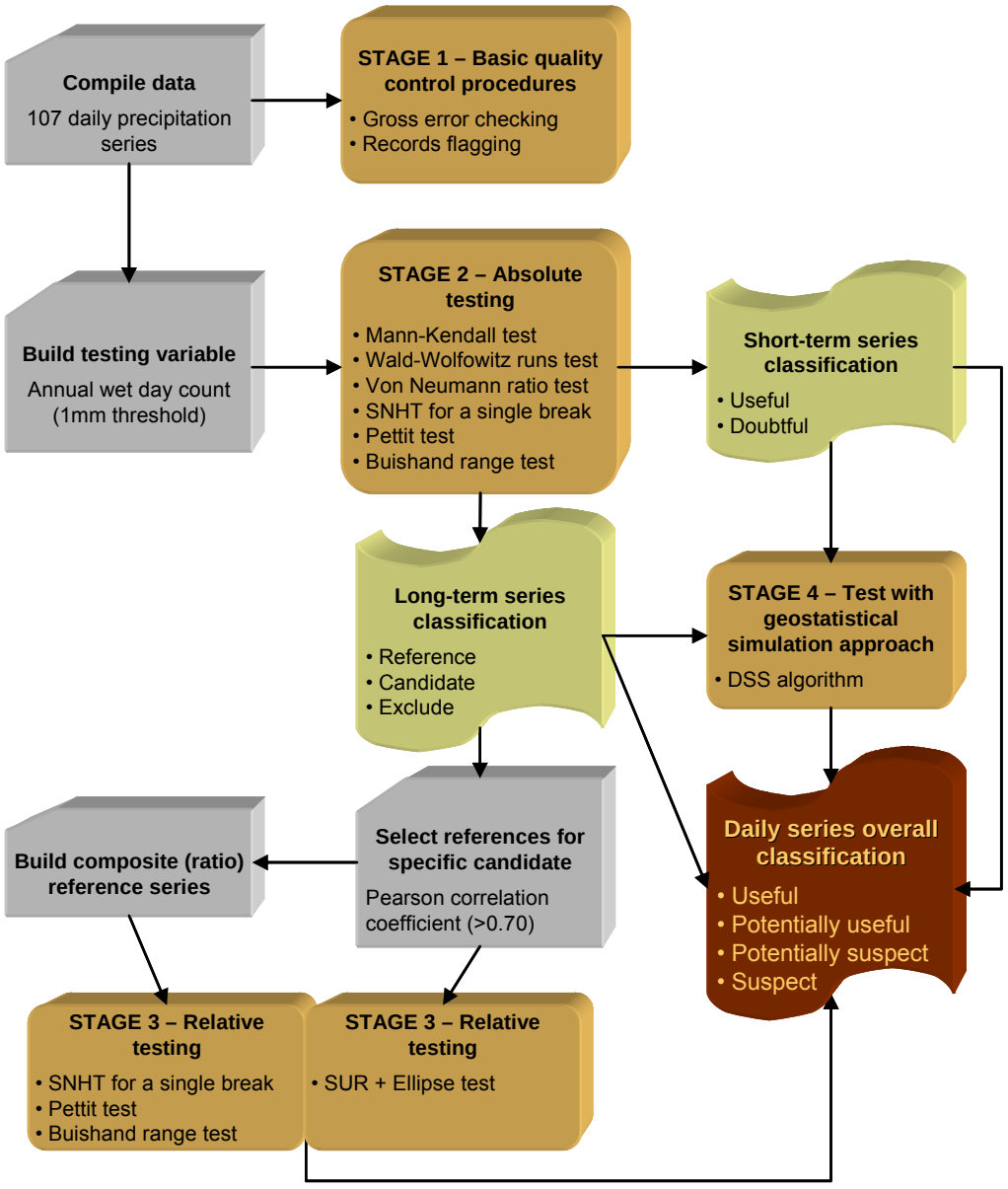

Figure 3.6 – Schematic representation of the methodology for the homogeneity assessment of the precipitation time series

It is important to point out that the homogenization techniques considered were used iteratively by systematically dividing the tested series into smaller segments when a break was detected, and then performing the test on those segments. Parameter-specific characteristics of the inhomogeneities detected through relative testing were evaluated, namely their magnitude, their frequency distribution, and their temporal occurrence.

The geostatistical approach proposed for the detection of non-climatic irregularities (applied to the testing variable) uses the direct sequential simulation (DSS) algorithm (Soares, 2001) to determine local probability density functions at a candidate station's location, by using spatial and temporal neighbourhood observations. Costa *et al.* (2008a) describe the methodology and results from this relative approach.

Section 3.2.1 overviews the basic quality control procedures, and details the quality control of precipitation extremes. The six statistical tests for inhomogeneities detection are described in Section 3.2.2, and the proposed extension of the Ellipse test is presented in Section 3.2.3. Section 3.2.4 discusses several issues related to the relative approach, including the selection of reference series and the construction of composite reference series. Finally, the geostatistical simulation approach proposed is described in Section 3.2.5.

3.2.1 Basic quality control procedures

Several basic quality control procedures were used to identify errors and suspicious daily precipitation records. A first set of procedures checks to eliminate gross errors and a second one flags questionable data.

3.2.1.1 Gross error checking

The aim of gross error check is to verify if the values are within acceptable range limits (e.g. check negative precipitation), and to perform basic time checks (e.g. check non-existent dates). Besides this routine quality control, several robust location and scale estimates were computed for outlier detection, by using all records from the time series, and by computing estimates for each year.

The asymmetric pseudo-standard deviation consists of an upper and a lower pseudo-standard deviation. The upper (lower) pseudo-standard deviation is defined as twice the distance between the upper (lower) quartile and the median divided by 1.349. These can be used to construct an asymmetric confidence interval for outlier flagging (Lanzante, 1996).

Other robust location and scale estimates can be computed using the biweight method, namely the biweight estimates of the mean and standard deviation. The biweight estimate is a weighted average such that weighting decreases nonlinearly (to zero) going away from the

centre of the distribution (Lanzante, 1996). The computation of weights uses the median-absolute-deviation (MAD), which is the median of the absolute values of the deviations of the sample values from the median.

3.2.1.2 Flagging of suspicious records

Several criteria (data outlying pre-fixed thresholds and graphical analysis) were used to flag the daily precipitation records with the following classification scheme: (1) 'useful', (2) 'doubtful', (3) 'suspect', and (4) 'erroneous'. It should be noted that flagging a record as (2) or (3) does not necessarily imply erroneous data as it can very well be because of extreme weather conditions.

The first criterion was the following: records greater than the 99th percentile were flagged as (2) 'doubtful'; records greater than 100 mm were flagged as (3) 'suspect'; and all others as (1) 'useful'. It is important to point out that the smallest 99th percentile value is equal to 20 mm when considering all 107 precipitation series. The second criterion used was a subjective evaluation of data previously flagged as (3) 'suspect'. If at least two monitoring stations had daily precipitation records greater than 100 mm on the same day, or within a one day range, their flag was set to (2) 'doubtful'.

The final flagging classification was achieved after plotting the data against time. When a peak in the graph seemed suspicious, even if that value was previously classified as (1) 'useful' or (2) 'suspect', a closer look was taken by plotting the data against time together with highly correlated stations (Pearson's correlation coefficient greater than 0.70 or highly significant Spearman rank-order correlation coefficient) for the three month period centred in the suspicious day. After a subjective analysis of all the graphs, several records were reclassified. Afterwards, the Portuguese Institute for Water (INAG – Instituto da Água), which is responsible for the SNIRH database, was contacted in order to clarify if data flagged as (4) 'erroneous' were outliers or a result of extreme weather phenomena.

The last quality control procedure used was a 'flat line' check, which identifies data of the same value for at least three (Feng *et al.* (2004) considered seven) consecutive days (not applied to zero precipitation data). For those detected records, the first occurrence was flagged as (0) 'useful', and the following records as (1) 'suspect'. All other records were flagged as (0) 'useful'.

3.2.2 Homogeneity tests

3.2.2.1 Mann-Kendall test

The nonparametric Mann-Kendall test (Mann, 1945; Kendall, 1975) is traditionally used to test randomness against (monotonic) trend and requires data to be serially independent. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that a stochastic trend exists.

The Mann-Kendall test statistic *S* is defined as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}\left(x_j - x_i\right) \tag{3.1}$$

where the x_j are sequential data values, n is the length of the data set, and

$$\mathrm{sgn}(\theta) = \begin{cases} 1 & \text{if} \quad \theta > 0 \\ 0 & \text{if} \quad \theta = 0 \\ -1 & \text{if} \quad \theta < 0 \end{cases} \tag{3.2}$$

Under the null hypothesis, the *S* statistic is approximately normally distributed when $n \geq 8$, and the mean and variance, corrected for ties, are given by

$$E[S] = 0 \tag{3.3}$$

$$V[S] = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{i=1}^{n} i(i-1)(2i+5) \right] \tag{3.4}$$

where *i* denotes the multiplicity of tied values.

Therefore, provided that $n \geq 8$, the critical regions and p-values of the Mann-Kendall test statistic (*S*) can be approximately determined using the standardized *S* statistic and the appropriate quantiles of the standard normal distribution, for a given significance level α.

For the one-tailed test, positive and negative *S* values indicate upward and downward trends, respectively.

3.2.2.2 Wald-Wolfowitz one-sample runs test

The Wald-Wolfowitz runs test is a nonparametric test for randomness (Wald and Wolfowitz, 1943), i.e. tests the assumption that the data collected constitute a random sample so that each observation or measurement is drawn randomly and independently from its population.

The null hypothesis is that the process that generates the set of numerical data is random (with respect to the median) through time. For the two-tailed test, the alternative hypothesis is that the data set is not randomly distributed.

A sequence is formed by assigning one of two symbols to each observation, depending on whether its measurement falls above or below the median:

$$\begin{cases} x_i = S & \text{if} \quad X_i \geq \text{median} \\ x_i = F & \text{if} \quad X_i < \text{median} \end{cases}, \qquad i = 1, \ldots, n$$

where the x_j are sequential data values; n is the length of the data set; S is a symbol denoting a success and F denotes a failure. If n_1 and n_2 are the number of successes and failures of series, respectively, then n_1+n_2 equals n.

A *run* is defined as a consecutive series of similar symbols that are bounded by symbols of a different type or by beginning or ending of the sequence. The test statistic U is the total number of runs present in the data.

For $n_1<20$ or $n_2<20$, critical values for the U statistic are given by the Wald-Wolfowitz U table (*e.g.* Berenson and Levine, 1996). Under the null hypothesis, the U statistic is approximately normally distributed when $n \geq 40$, and the mean and variance are given by

$$E[U] = \frac{2n_1n_2}{n} + 1 \tag{3.5}$$

$$V[U] = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)} \tag{3.6}$$

Therefore, provided that $n \geq 40$, the critical regions and p-values of the Wald-Wolfowitz test statistic (U) can be approximately determined using the standardized U statistic and the quantiles of the standard normal distribution, for a given significance level α.

3.2.2.3 Von Neumann ratio test

Von Neumann (1941) proposed a statistic defined by:

$$N = \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{3.7}$$

where n is the length of the data set; the x_j are sequential data values and $\bar{x}$ is the average value.

The Von Neumann ratio test is distribution free and the null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that the time series is not randomly distributed. Under the null hypothesis of a constant mean, the expected value of the test statistic is equal to 2 (Buishand, 1982). Critical values for the N statistic can be found in Wijngaard *et al.* (2003).

3.2.2.4 Standard normal homogeneity test for a single break

The Standard normal homogeneity test (SNHT) for a single break is a parametric test developed by Alexandersson (1986) that is capable of locating the period (year) where a break is likely. The null hypothesis is that the data are independent, identically normally distributed random quantities and the alternative is that a step-wise shift in the mean (a break) is present.

At any position k of the series, the $T(k)$ statistic compares the mean of the first k years of the record with that of the last $n–k$ years:

$$T(k) = k\bar{z}_1^{\,2} + (n-k)\bar{z}_2^{\,2}, \qquad k = 1,\ldots,n \tag{3.8}$$

where n is the length of the data set; $\bar{z}_1$ and $\bar{z}_2$ are defined as

$$\bar{z}_1 = \frac{1}{k}\sum_{i=1}^{k} (x_i - \bar{x})/s \tag{3.9}$$

$$\bar{z}_2 = \frac{1}{n-k}\sum_{i=k+1}^{n} (x_i - \bar{x})/s \tag{3.10}$$

where, the x_j are sequential data values; $\bar{x}$ is the mean and s the standard deviation of the series to be tested.

If a break is likely to be located at the year K, then $T(k)$ reaches a maximum near the year $k=K$, thus the test statistic is

$$T_0 = \max_{1\le k<n}\{T(k)\} \tag{3.11}$$

If T_0 is larger than a certain critical level the series should be classified as non-homogeneous. For 1% and 5% significance levels, critical values for the T_0 statistic can be found in Wijngaard *et al.* (2003).

3.2.2.5 Pettit test

Pettit (1979) developed a nonparametric test that is capable of locating the period (month or year) where a break is likely. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that a step-wise shift in the mean (a break) is present.

This test is based on the ranks of the elements of a series rather than on the values themselves. At any position k of the series, the X_k statistic is defined as

$$X_k = 2\sum_{i=1}^{k} R_i - k(n+1), \qquad k = 1,\ldots,n \tag{3.12}$$

where R_i is the rank of the ith observation when the values x_1, x_2, …, x_n in the series are arranged in ascending order.

If a break is likely to be located at the year E, then X_k is maximal or minimal near the year $k=E$, thus the test statistic is

$$X_E = \max_{1\le k\le n}\{|X_k|\} \tag{3.13}$$

If X_E is larger than a certain critical level the series should be classified as non-homogeneous. For 1% and 5% significance levels, critical values based on simulations for the X_E statistic can be found in Wijngaard *et al.* (2003). The statistical significance of the change point can also

be assessed by comparing the calculated value of X_E with its theoretical value, at probability level α, given as (Pettit, 1979):

$$X_{E,\alpha} = \left[-(\ln\alpha)\left(n^3 + n^2\right)/6\right]^{1/2} \quad (3.14)$$

where n is the length of the data set.

Taking into consideration the previous quality control analysis of the selected ECA series for this study (see Section 3.1.1), the table of critical values given by Wijngaard *et al.* (2003) was preferred.

3.2.2.6 Buishand range test

The Buishand range test (Buishand, 1982) is a parametric test and assumes, under the null hypothesis, that the values of the testing variable are independent and identically normally distributed. Under the alternative hypothesis, it assumes that a step-wise shift in the mean (a break) is present.

The test is based on the adjusted partial sums or cumulative deviations from the mean:

$$S_0^* = 0; \qquad S_k^* = \sum_{i=1}^{k}(x_i - \bar{x}), \qquad k = 1,\ldots,n \quad (3.15)$$

where, the x_j are sequential data values; $\bar{x}$ is the mean of the series to be tested, and n is the length of the data set.

For a homogeneous series, one may expect that the S_k^* values fluctuate around zero since there is no systematic pattern in the deviations of the records from their average value. If a break is likely to be located at the year K, then S_k^* reaches a maximum (negative shift) or a minimum (positive shift) near the year $k=K$. The test statistic is defined as $R/\sqrt{n}$, where

$$R = \left(\max_{0\le k\le n} S_k^* - \min_{0\le k\le n} S_k^*\right)\Big/ s \quad (3.16)$$

and s is the standard deviation.

For 1% and 5% significance levels, Wijngaard *et al.* (2003) extend the table of critical values for the test given by Buishand (1982).

3.2.3 Proposed extension of the Ellipse test

Allen *et al.* (1998) describe a method of cumulative residuals (*Ellipse test*) that tests if a weather data set from a monitoring station is homogeneous using the cumulative residuals from the linear regression between the candidate series (dependent variable) and data from a neighbouring station (independent variable), or the average observations of several surrounding stations inside the same climatic region. We propose an extension of this method that takes into account the contemporaneous relationship between several candidate series from the same climatic area. Instead of using the residuals from a linear regression model, the proposed technique uses the residuals from a *Seemingly unrelated regression equations* (SUR) *model*.

Zellner (1962) proposed the SUR approach for situations where at least two equations are being estimated and the error terms are contemporaneously but not serially correlated. In a general specification of *M* seemingly unrelated regression equations the *i*th equation is given by

$$y_i = X_i\beta_i + e_i \,, \qquad i = 1,...,M \tag{3.17}$$

where y_i is a vector of dimension (T x 1) containing all the observations on the *i*th dependent variable (*T* observations of the *i*th candidate series); X_i is a matrix of dimension (T x K_i) containing all the observations and all the explanatory variables (neighbouring stations' series) including the constant term; β_i is a vector, of dimension (K_i x 1), of unknown coefficients to be estimated; and e_i is a vector of dimension (T x 1) containing the error terms for all observations.

Note that each equation involves K_i regressors, meaning that each equation does not have to have the same number of explanatory variables. However, if all equations have identical explanatory variables, then generalized least squares is equivalent to equation by equation ordinary least squares (Greene 2003, p. 343). Thus, for those situations the proposed approach is equivalent to the method of cumulative residuals (*Ellipse test*) described by Allen *et al.* (1998).

Combining all equations into one model yields

$$(3.18)\qquad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_M \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_M \end{bmatrix}$$

Or, alternatively,

$$(3.19)\qquad y = X\beta + e$$

where the definitions of y, X, β and e are obvious from Equation (3.18) and their dimensions are, respectively, (MT x 1), (MT x K), (K x 1), and (MT x 1), with $K = \sum_{i=1}^{M} K_i$.

Given that e_{it} is the error for the ith equation in the tth time period, the assumption of contemporaneous disturbance correlation, but no correlation over time, implies that the covariance matrix for the complete error vector can be written as

$$(3.20)\qquad W = E[ee'] = \Sigma \otimes I_T$$

where

$$(3.21)\qquad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & \sigma_{MM} \end{bmatrix}$$

and $\otimes$ denotes the Kronecker product, indicating that each element of Σ is multiplied by an identity matrix. The matrix Σ is symmetric, so that $\sigma_{ij} = \sigma_{ji}$ and it is non-singular, and thus has an inverse.

The *generalized least squares estimator* of β, typically denoted by $\hat{\beta}$, is the best linear unbiased estimator (Griffiths *et al*. 1993, p. 570), but assumes that Σ is known. In practice, the variances and covariances are usually unknown and must be estimated, thus the *feasible generalized least squares estimator* $\hat{\hat{\beta}}$ is generally used (Greene 2003, p. 344).

To compute a feasible generalized least squares estimator, Σ must be replaced by a consistent estimator (Ruud 2000, p.704). To estimate the σ_{ij}, each equation is first estimated by ordinary

least squares in order to obtain the least squares residuals $\hat{e}_i$. Consistent estimates of the variances and covariances are then given by

$$\hat{\sigma}_{ij} = \frac{1}{T}\hat{e}_i'\hat{e}_j = \frac{1}{T}\sum_{t=1}^{T}\hat{e}_{it}\hat{e}_{jt} \tag{3.22}$$

If we define $\hat{\Sigma}$ as the matrix Σ with the unknown σ_{ij} replaced by $\hat{\sigma}_{ij}$, then the feasible generalized least squares estimator for β can be written as

$$\hat{\hat{\beta}} = \left[X'\left(\hat{\Sigma}^{-1}\otimes I_T\right)X\right]^{-1}X'\left(\hat{\Sigma}^{-1}\otimes I_T\right)y \tag{3.23}$$

The proposed procedure (SUR+Ellipse test) uses the cumulative residuals from such a SUR model to identify inhomogeneities in several candidate series from the same climatic area. A candidate series can be considered homogeneous if the cumulative residuals are not biased. The bias hypothesis can be tested using an ellipse defining the confidence limits (Allen *et al.*, 1998). For each equation *i* (*i*th candidate series), the axes of the ellipse are defined by

$$\begin{cases}\alpha_i = T/2\\ \beta_i = \dfrac{T}{\sqrt{T-1}}z_p S_{e,i}\end{cases} \tag{3.24}$$

where z_p is the standard normal variate for the desired probability p (confidence level), and $S_{e,i}$ is the standard deviation of the residuals of the *i*th equation. Thus, the parametric equation of the *i*th ellipse is

$$\begin{cases}X_i = \alpha_i\cos(\theta)\\ Y_i = \beta_i\sin(\theta)\end{cases} \tag{3.25}$$

with θ [rad] varying from 0 to 2π.

Plotting the cumulative residuals against time, using the time scale (interval) of the variable under analysis, the accumulated residual curve is obtained. If all the cumulative residuals lie inside the ellipse then the hypothesis of homogeneity is not rejected for the significance level considered. This test is then capable of locating the period (year) where a break is likely to occur.

3.2.4 Relative approach issues

As pointed out before in the *Literature review* chapter, if an absolute test detects a break in a station's time series it may indicate an inhomogeneity or it may simply indicate an abrupt change in the regional climate. Historic metadata support is then essential for evaluating the breaks detected. Unfortunately, these metadata was not available, at least to us, thus the results from the absolute approach were used to provide a qualitative classification of the series to be used in a relative approach.

In order to account for regional climate change and to isolate the effects of station irregularities, many techniques use data from nearby stations (reference series). Searching for breakpoints or artificial trends in a composite reference series (ratio for precipitation) is a standard methodology in the detection of non-climatic homogeneities. However, such a relative approach for the homogenization of all series could not be used, as it would require an iterative procedure in which *all* stations in the data set were seen consecutively as candidates and references. Consequently, with the objective of finding, at least, a subset of high quality long-term series, the results from the absolute approach were used to determine which series were more appropriate to be selected as reference (presumed homogeneous). Short-term series were not tested through relative procedures, but the absolute tests results were used to provide a global quality classification. The next section details the candidate and reference series selection, and the criteria used to classify the short-term series. Section 3.2.4.2 describes the construction of composite reference series.

3.2.4.1 Reference series selection

The results from the absolute approach were used to determine which long-term series were more appropriate to be selected as references, which presumably are homogeneous. Nevertheless, an iterative procedure in which several stations in the data set were seen consecutively as candidates and references was used.

The absolute approach comprised the application of six statistical tests (see Section 3.2.2) to the testing variable of each precipitation series. Their outcomes were then grouped together, and the criteria to determine which long-term series were more appropriate to be selected as references relied on the number of tests rejecting the homogeneity hypothesis at the 5% significance level (Table 3.1).

Table 3.1 – Criteria for long-term and short-term series classification for relative testing and for global quality evaluation, respectively

Number of tests that rejected the homogeneity hypothesis (5% signif.)	**Long-term series classification for relative testing**	**Short-term series global quality classification**
0	Reference	Useful
1	Candidate	Doubtful
2 or more	Exclude	Doubtful

Considering those criteria, the selected reference series were used to test the existence of inhomogeneities in the candidate series. Additionally, the selected reference series were also tested through an iterative procedure in which they were seen consecutively as candidates and references. Long-term series for which two or more absolute tests rejected the homogeneity hypothesis were not tested using the relative approach.

The established criteria seem reasonable since Wijngaard *et al*. (2003) defined a similar classification, although with a different purpose, for the ECA series. The outcomes of four absolute tests (applied to the ECA testing variables) were grouped together and a classification of the series was made depending on the number of tests rejecting the null hypothesis at the 1% level: ‘useful’ if one or zero tests rejected the null hypothesis, ‘doubtful’ if two tests rejected the null hypothesis, and ‘suspect’ if three or four tests rejected the null hypothesis.

In order to select references for a specific candidate series the most common approach was used (e.g. Tayanç *et al.*, 1998; Boissonnade *et al.*, 2002), i.e. Pearson correlation coefficients between the candidate and the eligible reference series were computed and the two highest correlated ones (>0.70) were taken as reference.

Short-term series were not tested through relative procedures, but the absolute tests results were used to provide a global quality evaluation by classifying them as ‘useful’ or ‘doubtful’. Once again, the criteria used relied on the number tests rejecting the homogeneity hypothesis at a 5% significance level (Table 3.1). Furthermore, short-term series classified as 'useful' were used in the homogeneity assessment performed through the proposed geostatistical simulation approach – see Section 3.2.5 for the methodology and Section 3.3.4 for the discussion of the results from this approach.

3.2.4.2 Construction of composite reference series

Long-term composite (ratio) reference series were tested using the Standard normal homogeneity test (SNHT) for a single break (Alexandersson, 1986; Alexandersson and Moberg, 1997), the Pettit test (Pettit, 1979), and the Buishand range test (Buishand, 1982).

The ratios used are defined as (Alexandersson and Moberg, 1997):

$$Q_i = Y_i \Big/ \left\{ \left[\sum_{j=1}^{k} \rho_j{}^2 X_{ji} \overline{Y} / \overline{X}_j \right] \Big/ \sum_{j=1}^{k} \rho_j{}^2 \right\} \quad (3.26)$$

where: Y denotes the candidate series; Y_i is a specific value of the testing variable at year i; X_j denotes one of the reference neighbouring stations (the jth of a total of k); X_{ji} is a specific value from that site; $\overline{Y}$ and $\overline{X}_j$ are mean values of Y and X_j, respectively, and were computed for one common time period for all series (otherwise, Alexandersson and Moberg (1997) verified that the size of non-homogeneities may be underestimated or missed by the test); and, finally, ρ_j is the correlation coefficient between the candidate site and a surrounding station j ($j=1,\ldots, k$) and it was estimated from one common time period for all series.

The values of the denominator in Equation (3.26) (expressed within brackets) are called reference values since they are intended to be reasonable and stable estimates for the candidate series using a set of neighbouring reference stations. If the candidate has no inhomogeneities then the resulting ratio series will oscillate around 1.

Finally, the composite (ratio) reference series are defined as the standardized series:

$$Z_i = \left(Q_i - \overline{Q}\right) \Big/ \sigma_Q \quad (3.27)$$

and were used for homogeneity testing with the Standard normal homogeneity test (SNHT) for a single break (Section 3.2.2.4), the Pettit test (Section 3.2.2.5), and the Buishand range test (Section 3.2.2.6).

3.2.5 Proposed geostatistical simulation approach

Most of the statistical procedures, including nonparametric tests, require serially independent data. When sample data are serially correlated, the presence of serial correlation in time series

will affect the ability of the tests to correctly assess the significance of inhomogeneities detection. However, it is a standard procedure to relax this hypothesis for annual data. The autocorrelation assumption, while acceptable for some annual climate series, is not realistic for daily or monthly series, where there is much empirical evidence of autocorrelation (Reeves *et al*., 2007).

Relative methods using composite reference series are traditional approaches for the homogenization of climate records. Composite reference series are computed as a weighted average of data from neighbouring stations by using some measure of statistical similarity (usually the correlation coefficient or an inverse function of the distance) between them (Peterson *et al.*, 1998). A combined use of those measures was proposed by Romero *et al.* (1998) in order to increase the contribution of the closer stations, both in spatial and correlation terms.

Considering the previous discussion, the motivation to use a geostatistical approach for the inhomogeneities detection in climate series was that these procedures account for the *joint spatial* and *temporal dependence* between observations, and enhance the pre-eminence of the closer stations, both in spatial and correlation terms.

Sequential simulation is a widely used tool for obtaining a set of equiprobable simulated realizations of variables from natural phenomena, honouring their spatial distribution and uncertainty (e.g. Goovaerts 1997, Ch. 8). While sequential indicator simulation (SIS) and sequential Gaussian simulation (SGS) require the transformation of original variables, direct sequential simulation (DSS) has been proposed (Journel, 1994) for simulating directly in the original data space and does not rely on multi-Gaussian assumptions. Journel (1994) showed that for the sequential simulation algorithm to reproduce a specific covariance model it suffices that simulated values are drawn from local distributions centred at the simple kriging estimates with a variance corresponding to the simple kriging estimation variance. This result guarantees that the spatial covariance of the original variable is reproduced but not the histogram. To overcome this limitation, Soares (2001) proposed a DSS algorithm that uses the local simple kriging estimates of the mean and variance, not to define the local cumulative distribution function (cdf) but to sample from the global cdf.

For the detection of inhomogeneities, we propose the DSS algorithm introduced by Soares (2001) to calculate the local probability density function (pdf) at a candidate station's location,

using spatial and temporal observations from nearby reference stations, *without* taking into account the candidate's data. Afterwards, the local pdf from each instant in time (e.g., year) is used to verify the existence of irregularities: a breakpoint is identified whenever the interval of a specified probability p (e.g. 0.95) centred in the local pdf does not contain the observed (real) value of the candidate station. In practice, the local pdfs are provided by the histograms of simulated maps, thus this rule implies that if the observed (real) value lies below or above the pre-defined percentiles of the histogram of a given instant in time then it is not considered as homogeneous. If irregularities are detected in a candidate series, the time series can be adjusted by replacing the inhomogeneous records with the mean, or median, of the pdf(s) calculated at the candidate station's location for the inhomogeneous period(s).

The DSS algorithm was preferred over SGS, not only because it succeeds in reproducing the variogram and histogram of a continuous variable, but also because it does not require any transformation of the original variables. The SGS guarantees the reproduction of the variogram of the Gaussian variable, but it is not assured that the variogram is reproduced after the back-transformation of the original variable, especially when the distribution of the original variable is very asymmetric.

The methodology of the geostatistical simulation approach can be summarized as follows.

Let $\{z(u_\alpha, t_i): \alpha=0, 1,\ldots,n-1; i=1,\ldots,T\}$ be the set of climate data measured at n locations u_α and in t_i time instants (e.g., years). The n monitoring stations do not have to be all informed at the same T time instants (i.e., a number of z-values can be missing). Let $\{z(u_0, t_i): i=1,\ldots,T\}$ denote the candidate time series. The set of climate observations correspond to outcome values (realizations) of a spatiotemporal random variable Z(u, t) that can take a series of values at any location in space u and instant in time t according to a probability distribution.

Using the set of time series corresponding to the reference stations,

$$\{z(u_\alpha, t_i): \alpha=1,\ldots,n-1; i=1,\ldots,T\},$$

the DSS algorithm is applied in order to obtain a set of m equally probable realizations of Z(u,t) at the candidate station's location and all instants in time: $\{z^s(u_0, t_i): s=1,\ldots,m; t=1,\ldots,T\}$. For a given instant in time t_0, the set of simulated values $\{z^s(u_0, t_0): s=1,\ldots,m\}$ defines the local histogram of the candidate station for that instant. The corresponding

empirical cumulative distribution function gives the estimated probability that the variable Z at location u_0 in space and instant t_0 in time is no greater than any given threshold z:

$$F^*(u_0, t_0; z)=Prob^*\{Z(u_0, t_0) \leq z\}.$$

An inhomogeneous record $z(u_0, t_0)$ is identified if the interval of a specified probability p (e.g. 0.95), centred in the estimated local pdf of the candidate station for the instant t_0, does not contain the observed $z(u_0, t_0)$ value:

$$(3.28) \quad Prob^*\{Z(u_0,t_0) \leq z(u_0,t_0)\} < \frac{1-p}{2} \text{ or } Prob^*\{Z(u_0,t_0) \leq z(u_0,t_0)\} > 1-\frac{1-p}{2}$$

3.2.5.1 Implementation issues

The inhomogeneities detection procedures used in this study followed the hybrid approach proposed by Wijngaard *et al.* (2003) for the ECA dataset, and used as the testing variable the annual wet day count with 1 mm as threshold. For illustration purposes, the geostatistical simulation approach was applied to the testing variable data from 4 candidate stations using data from 62 surrounding stations (reference stations, presumed homogeneous) located in the southern region of Portugal. As with other relative homogeneity testing approaches, reference stations' data are used to account for regional climate changes and to isolate the effects of station irregularities (Peterson *et al.*, 1998).

Techniques that use series from surrounding stations, some times run the test once, relying the reference to be homogeneous, or engage in an iterative procedure in which all the stations in the data set are seen consecutively as candidates and references (Aguilar *et al.*, 2003). Following this methodology, the local pdfs of each year of the candidate series, derived from 50 simulated maps in a 1 km x 1 km grid, were computed using data not only from the 62 references but also from the other 3 candidate stations. The analysed period was 1980-2001.

The geostatistical simulations used an isotropic spherical variogram model that was fitted using the testing variable data from the complete set of 66 monitoring stations: the estimated spatial range was equal to 72 km, and the temporal one was equal to 1.8 years.

3.3 Results and discussion

The data analysis for this study was generated through specific programs developed using SAS software macros, SAS/STAT®, SAS/ETS® and SAS/GRAPH® software, of the SAS System[13] for Windows, Version 8. The geostatistical simulation approach, in particular the DSS algorithm, was implemented using geoMS© – Geostatistical Modelling Software[14].

3.3.1 Basic quality control analysis

This section discusses the main results from the quality control procedures used. The emphasis of this stage was on the quality control of precipitation extremes, thus the objective was the identification of errors and suspicious daily precipitation records.

Routine quality control procedures revealed that all precipitation records were non-negative but many series had non-existent dates, which were properly corrected and missing values were assigned to the variable for those dates.

Several robust location and scale estimates were computed for outlier detection by using all records from the daily time series and by computing estimates for each year. However, those estimates were inconclusive, and identical for both approaches, because the daily precipitation distributions are much skewed.

The upper asymmetric pseudo-standard deviation was computed for all 107 daily precipitation series, but it was inconclusive since the median is equal to zero for all series and the third quartile is different from zero for 11 series only, thus the interquartile range is always equal to zero except for those 11 series.

The biweight estimates of the mean and standard deviation could not be computed because the MAD (median-absolute-deviation) is equal to zero for all daily precipitation series and it appears in the weights denominator of those estimates; similarly, Feng *et al.* (2004) applied this procedure for temperature data only.

---

[13] SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. In the USA and other countries ® indicates USA registration.

[14] geoMS© – Geostatistical Modelling Software was developed by Centro de Modelização de Reservatórios Petrolíferos (CMRP). Copyright CMRP-IST 2000.

Robust alternatives to MAD are the S_n and Q_n standard deviations (Rousseeuw and Croux, 1993). The S_n-standard-deviation is equal to zero for all daily precipitation series, and the Q_n-standard-deviation is approximately equal to 0.22 mm for all series, indicating that the centres of the distributions have low variability.

The next set of procedures aimed to identify questionable data by flagging the daily precipitation records with the following classification scheme: (1) 'useful', (2) 'doubtful', (3) 'suspect', and (4) 'erroneous'. The first criterion relied on data outlying pre-fixed thresholds: records greater than the 99th percentile were flagged as (2); records greater than 100 mm were flagged as (3); and all others as (1). Using this criterion, 44% from the whole 107 series under analysis had records flagged as (3), where 25 of them were long-term series and the other 22 were short-term series. Not surprisingly, the long-term series had an average number of records flagged as (3) approximately equal to 5, and for the short-term ones that average was approximately 3. The total number of records flagged as (3) 'suspect' was equal to 188.

The second criterion used was a subjective evaluation of data previously flagged as (3) 'suspect'. If at least two monitoring stations had daily precipitation records greater than 100 mm on the same day, or within a one day range, their flag was set to (2) 'doubtful'. As a result, the number of records flagged as (3) 'suspect' dropped to 52 and the number of series to 22 (16 long-term and 6 short-term).

The third criterion relied on graphical analysis. All 107 series were plotted against time, and when a peak in the graph seemed suspicious, even if that value was previously classified as (1) 'useful' or (2) 'suspect', a closer look was taken by plotting the data against time together with highly correlated stations for the three month period centred in the suspicious day (e.g. Figure 3.7).

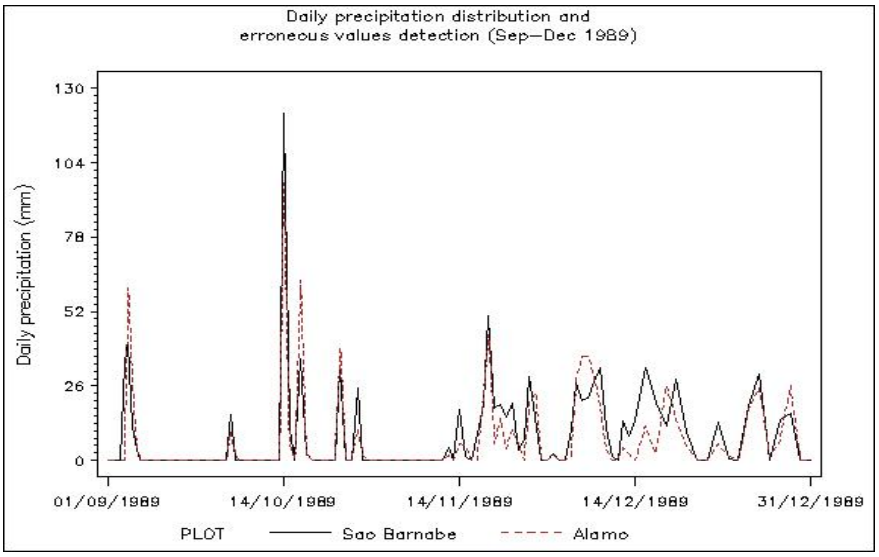

Figure 3.7 – Daily precipitation distribution from September to December 1989 at São Barnabé (29I.01) and Álamo (28K.02)

After a subjective analysis of all the graphs (over 500), almost all the records previously flagged as (3) 'suspect' were set to (2) 'doubtful', and four values previously flagged as (2) were reclassified as (3). Consequently, the number of records flagged as (3) dropped to 7 (corresponding to 3 long-term and 1 short-term series). Conversely, for Alcoutim (29M.01) station the series did not seem to be homogeneous for the period 1954-1959 (Figure 3.8) and therefore all the values of those years were flagged as (3) 'suspect'.

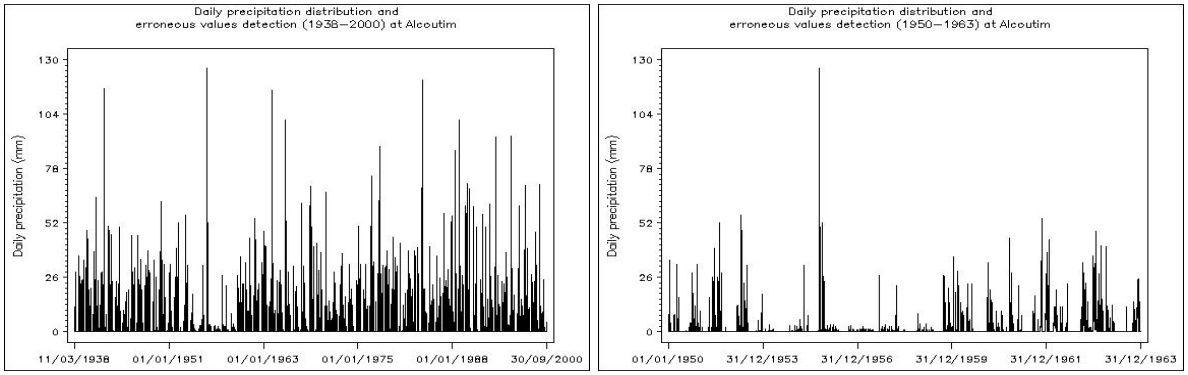

Figure 3.8 – Daily precipitation distribution at Alcoutim (29M.01)

Moreover, four records were reclassified as (4) 'erroneous' (Table 3.2), as they refer to daily precipitation observations greater than 190 mm and the respective peaks in the graphs seemed suspicious (e.g. Figure 3.9 and Figure 3.10). Afterwards, with the purpose of establishing the final classification, the Portuguese Institute for Water (INAG) was contacted in order to clarify if those records were really erroneous. In their reply they explained that the 192.5 mm

precipitation recorded at São Manços (23K.01) in 31/01/1996 referred to the accumulated precipitation from 07/01/1996 to 31/01/1996 (Figure 3.9), and therefore that observation should be set to missing in the daily series. They also informed that the 226 mm recorded at Relíquias (27G.01) in 10/09/1949 was correct, thus it was flagged as (1) 'useful'.

Table 3.2 – Daily precipitation records flagged as (4) 'erroneous' after three criteria (data outlying pre-fixed thresholds, subjective evaluation, and graphical analysis)

Station ID	**Station name**	**Length**	**Period**	**Daily precipitation** (mm)
23K.01	São Manços	Long-term	31/01/1996	192.5
27G.01	Relíquias	Long-term	10/09/1949	226.0
30F.01	Monchique	Long-term	25/10/1997	274.7
29F.01	Cimalhas	Short-term	06/11/1997	200.0

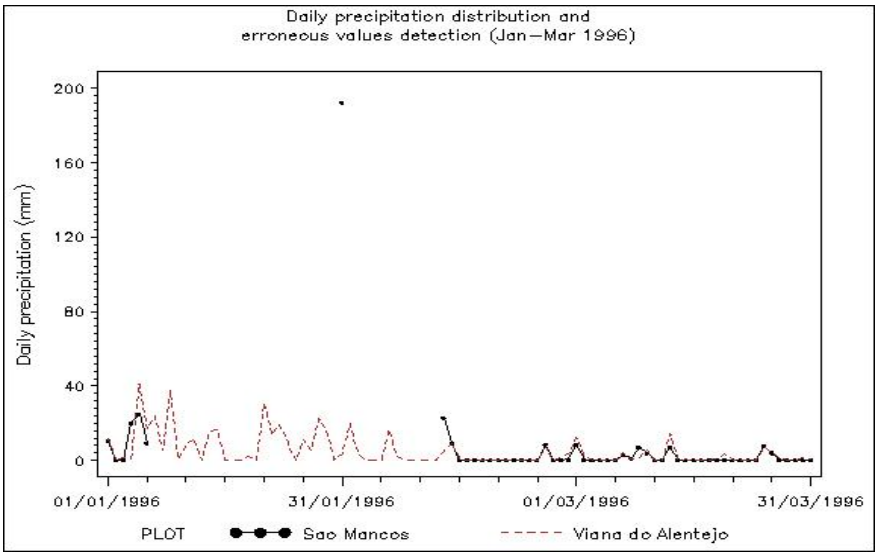

Figure 3.9 – Daily precipitation distribution from January to March 1996 at São Manços (23K.01) and Viana do Alentejo (44I.01)

The record of 274.7 mm in 25/10/1997 at Monchique (30F.01) was flagged as (1) 'useful' (Figure 3.10), since the available metadata provided by INAG indicate that it is a correct record. Moreover, INAG pointed out that Monchique (30F.01) holds the highest national record of 24 hours accumulated precipitation using a recording raingauge, which is equal to 291.6 mm in 26/10/1997 (the observations at Monchique from the SNIRH database refer to accumulative precipitation gauges). Finally, the record of 200 mm in 06/11/1997 at Cimalhas (29F.01) was also flagged as (1) 'useful' since the available metadata provided by INAG indicate that it is a conceivable record.

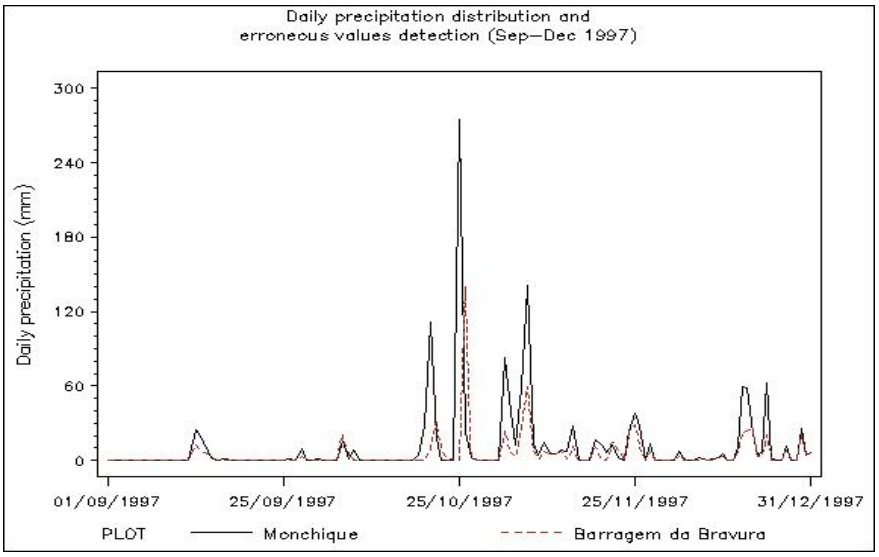

Figure 3.10 – Daily precipitation distribution from September to December 1997 at Monchique (30F.01) and Barragem da Bravura (30E.03)

The last quality control procedure was a 'flat line' check, which identifies data of the same value for at least three (Feng *et al.* (2004) considered seven) consecutive days (not applied to zero precipitation data). For those detected records, the first occurrence was flagged as (0) 'useful', and the following records as (1) 'suspect'. All other records were flagged as (0) 'useful'.

Almost half (49%) of the long-term series and 19% of the short-term ones were flagged with 'suspect' records using this methodology. The average number of runs (blocks of three or four consecutive days having the same value) per station was equal to two. The flagged precipitation values range from 0.1 mm to 5 mm and the most common values were 0.1 mm and 0.2 mm. This seems to indicate that if those flagged values are erroneous they might have been caused by measurement errors (i.e. how precisely very low amounts of precipitation are measured) rather than by other types of errors (e.g., editing).

Since the flagged values are very small and there are in average only four observations flagged per series, they might be considered useful if the aim of the study is to analyse extreme precipitation values, even though there were a significant number of series (32%) with data flagged as 'suspect'.

3.3.2 Absolute testing

The absolute testing stage comprises the application, at all 107 locations, of six statistical tests (see Section 3.2.2) to the testing variable, which is defined as the annual number of days with precipitation amounts above or equal to 1 mm. The results from this homogeneity testing stage are detailed in Appendix II.

Two of the homogeneity tests applied are not distribution free, namely the SNHT and the Buishand range test, and assume that data are independent, identically normally distributed random quantities. Moreover, the remaining nonparametric tests applied also require serially independent data. For those reasons, generalized Durbin-Watson autocorrelation tests and four normality tests were applied to the testing series at all stations.

The Durbin–Watson test (Durbin and Watson, 1950, 1951) is a widely used method of testing for autocorrelation. The generalized Durbin–Watson statistics for 1^{st}, 2^{nd} and 3^{rd} order autocorrelation were computed (for the mathematical description refer to SAS Institute (1999a, p. 354-358)), and conclusions were drawn at the 5% level. The generalized Durbin-Watson tests revealed 1^{st}–order autocorrelation for almost 19% of the series (16 long-term and 4 short-term), and 2^{nd}–order for four series only. None of the testing series had significant 3^{rd}–order autocorrelation.

The four normality tests applied were the Shapiro–Wilk (Shapiro and Wilk, 1965), the Kolmogorov–Smirnov (Kolmogorov, 1933; Smirnov, 1939), the Cramér–von Mises (Cramér, 1928; Von Mises, 1931), and the Anderson–Darling (Anderson and Darling, 1952, 1954) tests. The last three are goodness-of-fit tests based on the empirical distribution function. For details on the statistical computation of the normality tests refer to SAS Institute (1999b, p. 1397-1401). In view of the results from those four normality tests at the 5% level, over 80% of the testing series (36 long-term and 50 short-term) were considered as Gaussian by all of them. On the other hand, the four tests rejected the normality hypothesis for 7.5% of the series (2 long-term and 6 short-term). Taking into consideration these results, we decided to proceed with the homogeneity tests. Furthermore, note that it is a standard procedure to relax those assumptions for annual data.

Approximately 60% of the 107 tested series were considered homogeneous (at the 5% level) by the six statistical tests, and only one of the six tests rejected the null hypothesis for

approximately 22% of the series. Thus, at least two tests rejected the homogeneity hypothesis for the remaining 20 stations.

Taking into consideration the length of the series (Figure 3.11 and Figure 3.12), 17 long-term (stations with at least 30 observation years, 1932-2000) and 47 short-term (stations with at least 10 observation years, 1956-2001) series were considered homogeneous by all tests.

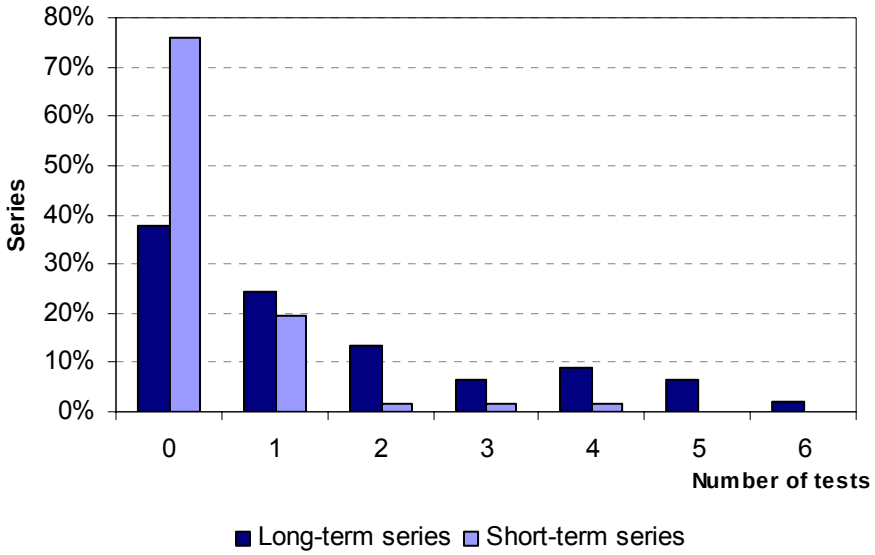

Figure 3.11 – Distribution of long-term and short-term series by number of tests rejecting the homogeneity hypothesis (5% signif). The six tests were applied to the annual number of wet days (threshold 1 mm).

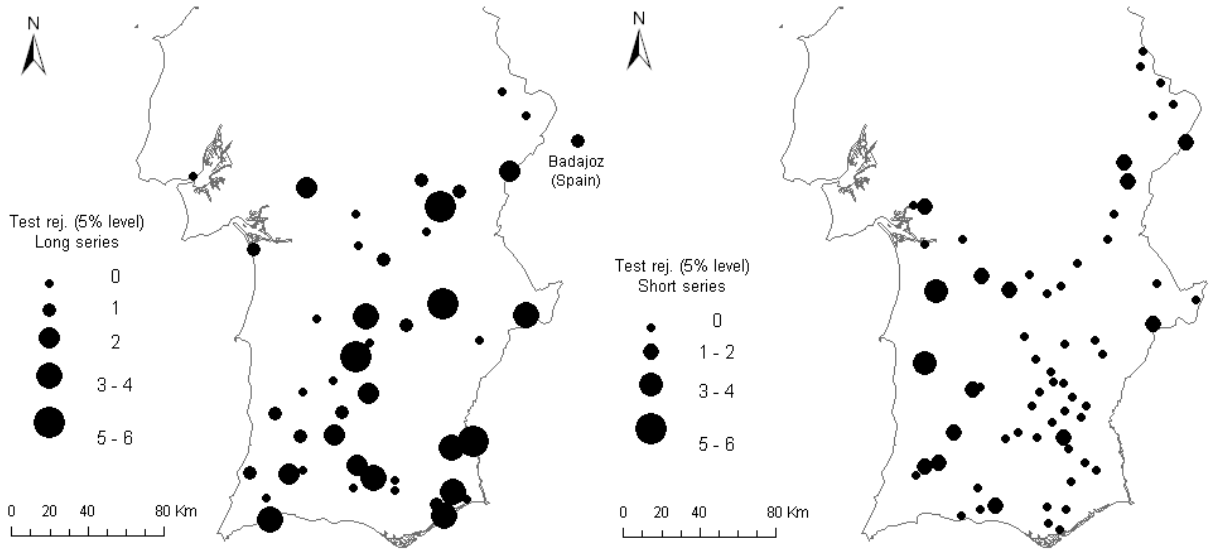

Figure 3.12 – Results from the absolute testing approach. Station dots are scaled with the number of tests that rejected the homogeneity hypothesis at the 5% significance level. The six tests were applied to the annual number of wet days (threshold 1 mm). Left: long-term series. Right: short-term series.

The outcomes from the six tests were then grouped together, and a classification for the long-term (Table 3.3) and short-term (Table 3.4) series was established relying on the number of tests rejecting the homogeneity hypothesis at the 5% significance level (see Section 3.2.4 for methodological details).

Table 3.3 – Distribution of long-term series by number of tests rejecting the homogeneity hypothesis at the 5% significance level, and respective classification for relative testing. The six tests were applied to the annual number of wet days (threshold 1 mm).

Number of tests that rejected the homogeneity hypothesis (5% signif.)	**Number of long-term series**	**Classification for relative testing**
0	17	Reference
1	11	Candidate
2	6	Exclude
3	3	
4	4	
5	3	
6	1	

Approximately 38% of the long-term series were considered appropriate to be selected as reference, and 24% as candidate. Thus, the remaining 38% were excluded from the relative testing analysis.

Table 3.4 – Distribution of short-term series by number of tests rejecting the homogeneity hypothesis at the 5% significance level, and respective global quality evaluation. The six tests were applied to the annual number of wet days (threshold 1 mm).

Number of tests that rejected the homogeneity hypothesis (5% signif.)	**Number of short-term series**	**Global quality classification**
0	47	Useful
1	12	Doubtful
2	1	
3	1	
4	1	
5	0	
6	0	

Not surprisingly, approximately 76% of the short-term series (stations with at least 10 observation years, 1956-2001) were considered homogeneous (at the 5% level) by the six statistical tests, and thus globally evaluated as 'useful'.

Finally, to illustrate the homogenization analysis performed, the results of the six homogeneity tests applied to the annual number of wet days (threshold 1 mm) of two log-term series are discussed. The first one is the series from Beja (666, from the ECA dataset) and the other one is from Aljezur (30E.01, from the SNIRH database), which were both classified as 'candidate' for relative testing purposes.

3.3.2.1 Illustration: results of Beja station

For the testing series from Beja (666), the Wald-Wolfowitz and the Mann-Kendall test considered the series homogeneous, but conversely the Von Neumann ratio test rejected the null hypothesis, at the 5% level, for the period 1941-1990 (Table 3.5).

Table 3.5 – Absolute tests results for Beja (666) at the 5% level for the Wald-Wolfowitz two-tailed test based on large sample approximation, the Mann-Kendall one-tailed test based on large sample approximation, and the Von Neumann ratio test. The three tests were applied to the annual number of wet days (threshold 1 mm).

Period	Wald-Wolfowitz		Mann-Kendall		Von Neumann	
	Standardized* U *statistic	***Result***	***Standardized* S *statistic***	***Result***	**N *statistic***	***Result***
1941–1999	3.75	Homog.	152.92	Homog.		
1941–1990					1.53	Reject
1950–1999					1.73	Homog.
1941–1960					1.00	Reject
1961–1980					2.04	Homog.
1980–1999					1.57	Homog.

Note that the precipitation series from Beja (666) was marked as 'useful' in the ECA dataset (see Section 3.1.1), since none of the four tests considered by the ECA project, including the Von Neumann ratio test, rejected the homogeneity hypothesis at the 1% level.

For Beja (666), the Buishand test statistic was equal to 0.97 for the period 1941-1990, and equal to 1.13 for 1950-1999 (Figure 3.13). The Pettit test statistic was, respectively, equal to 125 and to 233 for the period 1941-1990 and for 1950-1999 (Figure 3.14). Finally, the SNHT

test statistic was equal to 3.27 for the period 1941-1990, and equal to 4.63 for 1950-1999 (Figure 3.15). Hence, these last three tests, which are capable of locating the year where the break is likely, considered the series homogeneous, at the 5% level.

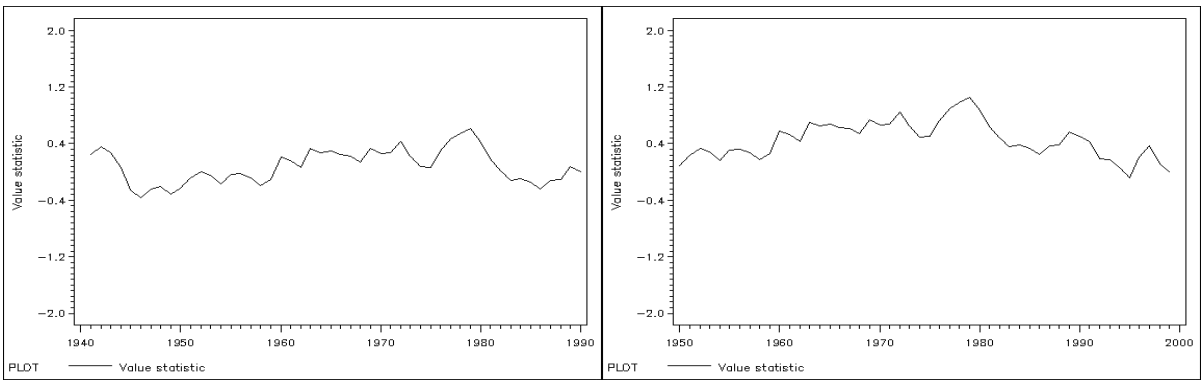

Figure 3.13 – Buishand range test statistic standardized values (left: 1941–1990; right: 1950–1999) for the annual number of wet days (threshold 1 mm) series of Beja (666) station

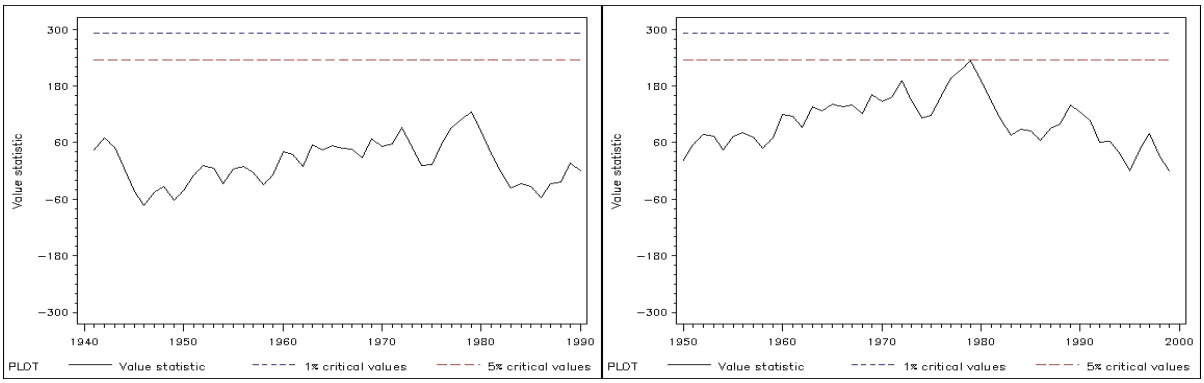

Figure 3.14 – Pettit test results (left: 1941–1990; right: 1950–1999) for the annual number of wet days (threshold 1 mm) series of Beja (666) station

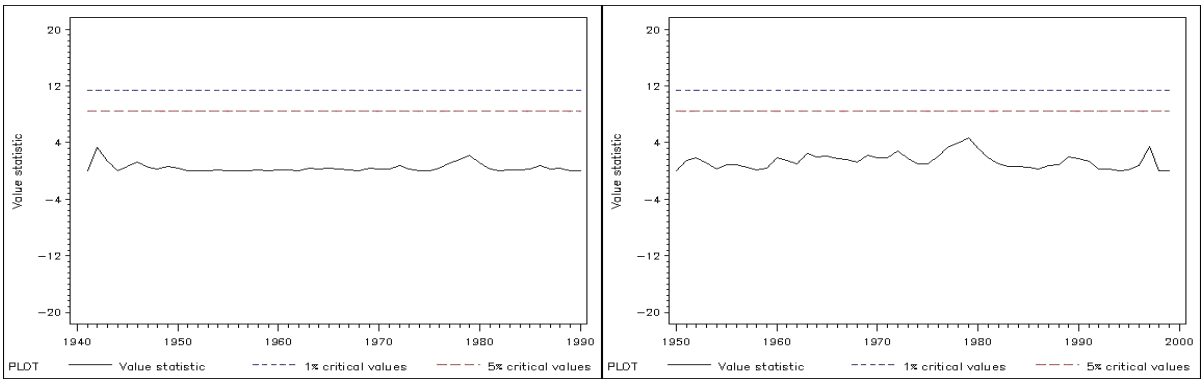

Figure 3.15 – SNHT test results (left: 1941–1990; right: 1950–1999) for the annual number of wet days (threshold 1 mm) series of Beja (666) station

3.3.2.2 Illustration: results of Aljezur station

For the testing series from Aljezur (30E.01) the Wald-Wolfowitz, the Mann-Kendall and the Von Neumann ratio test considered the series homogeneous, at the 5% level (Table 3.6).

Table 3.6 – Absolute tests results for Aljezur (30E.01) at the 5% level for the Wald-Wolfowitz two-tailed test based on large sample approximation, the Mann-Kendall one-tailed test based on large sample approximation, and the Von Neumann ratio test. The three tests were applied to the annual number of wet days (threshold 1 mm).

Period	Wald-Wolfowitz		Mann-Kendall		Von Neumann	
	Standardized* U *statistic	***Result***	***Standardized* S *statistic***	***Result***	**N *statistic***	***Result***
1932–1999	4.09	Homog.	188.91	Homog.		
1932–1981					1.58	Homog.
1950–1999					1.77	Homog.

For Aljezur (30E.01), the Buishand test statistic was equal to 1.552 for the period 1932-1981, and equal to 1.21 for 1950-1999. Hence, during the first fifty years of the testing series, a break was detected in 1942 (Figure 3.16).

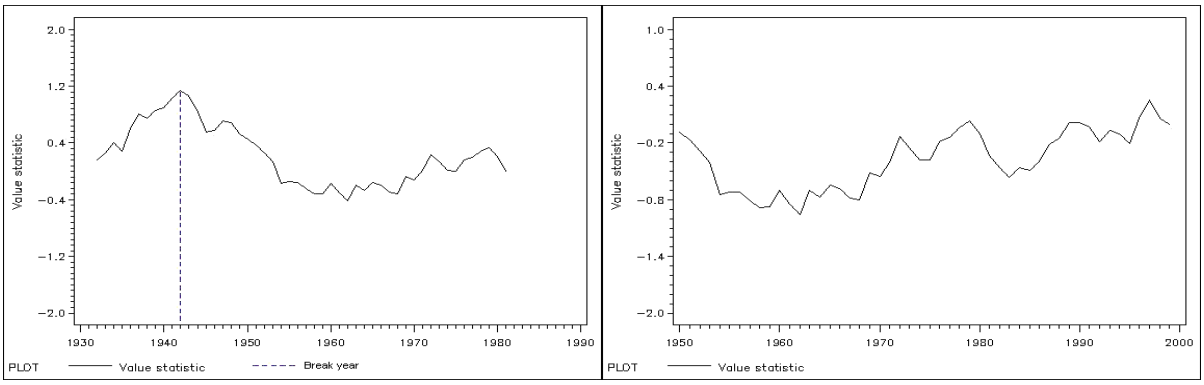

Figure 3.16 – Buishand range test statistic standardized values (left: 1932-1981; right: 1950-1999) for the annual number of wet days (threshold 1 mm) series of Aljezur (30E.01) station

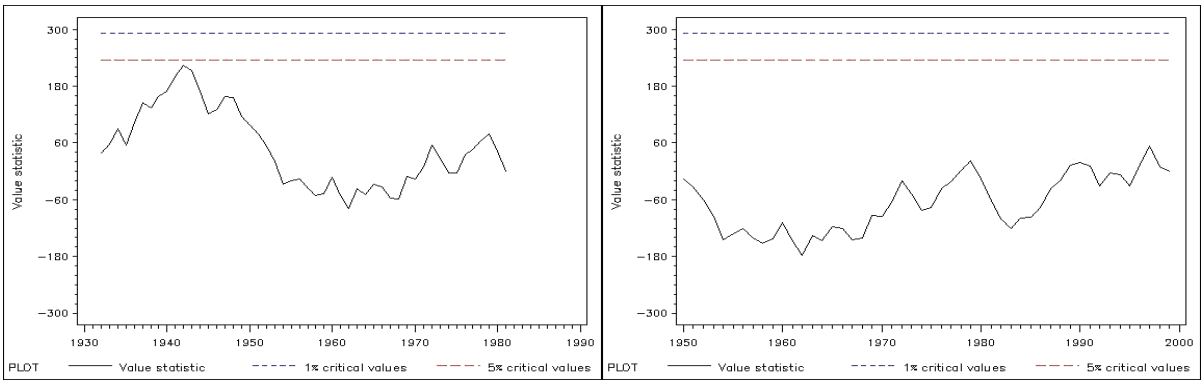

Figure 3.17 – Pettit test results (left: 1932-1981; right: 1950-1999) for the annual number of wet days (threshold 1 mm) series of Aljezur (30E.01) station

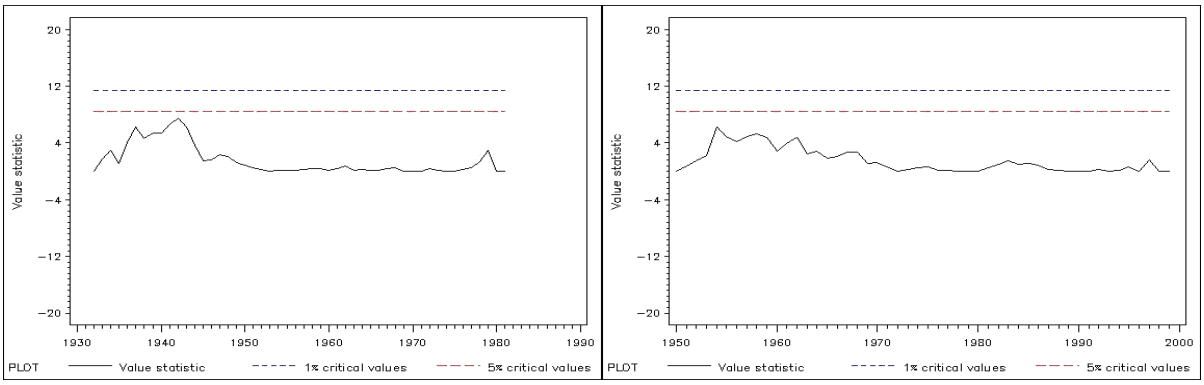

Figure 3.18 – SNHT test results (left: 1932-1981; right: 1950-1999) for the annual number of wet days (threshold 1 mm) series of Aljezur (30E.01) station

The Pettit test statistic was equal to 224 (in 1942, but not significant at 5%) for the period 1932-1981, and equal to (–)178 for 1950-1999 (Figure 3.17). Finally, the SNHT test statistic was equal to 7.54 (in 1942, but not significant at 5%) for the period 1932-1981, and equal to 6.21 for 1950-1999 (Figure 3.18). Thus, these two tests considered as homogeneous the Aljezur (30E.01) testing series, at the 5% level.

3.3.3 Relative testing

The relative testing stage comprises the application of four homogeneity tests, which are capable of locating the year where a break is likely. Three of them were applied to long-term composite (ratio) reference series, namely the SNHT for a single break, the Pettit test, and the Buishand range test. The proposed extension of the Ellipse test (SUR+ Ellipse test) was applied to the testing variable of the long-term series. The methodology and results from this

stage are described by Costa and Soares (2006), and detailed results are presented in Appendix III.

The Buishand, Pettit, and SNHT tests were applied to 28 composite (ratio) reference series and concluded as homogeneous 13 series (5 previously classified as candidates, and 8 as references). For the remaining 15 series, at least one of the tests rejected the homogeneity hypothesis, at the 5% level. Table 3.7 shows the three tests results, including the break years detected and respective relative magnitudes (ratio between the average annual wet day count before and after two consecutive breaks), for the 11 series previously classified as candidates.

Table 3.7 – Buishand, Pettit, and SNHT tests results. Break years detected (and respective relative magnitudes) are presented for the candidate series, at the 5% level. The three tests were applied to composite (ratio) reference series.

Candidate	References	Buishand	Pettit	SNHT
Beja (666)	Azinheira Barros (25G.01) São Manços (23K.01)	Homog.	Homog.	Homog.
Badajoz Talavera (709)	Azinheira Barros (25G.01) São Manços (23K.01)	Homog.	1975 (0.98)	Homog.
Aljezur (30E.01)	Barragem da Bravura (30E.03) Arronches (19N.01)	1968 (1.04)	1968 (1.04)	1968 (1.04)
Picota (30K.02)	Barragem da Bravura (30E.03) Alcaria [Castro Marim] (30L.04)	1988 (0.98)	Homog.	Homog.
Odemira (28F.01)	Azinheira Barros (25G.01) São Manços (23K.01)	Homog.	Homog.	Homog.
Aldeia de Palheiros (28H.01)	Azinheira Barros (25G.01) São Manços (23K.01)	Homog.	Homog.	Homog.
Sabóia (29G.01)	Santiago do Escoural (22H.02) Relíquias (27G.01)	1949 (1.05) 1984 (1.17)	1984 (1.17)	1984 (1.17)
Comporta (23E.01)	Barragem da Bravura (30E.03) Azinheira Barros (25G.01)	Homog.	1986 (1.10)	1986 (1.10)
Viana do Alentejo (24I.01)	Alcáçovas (23I.01) Santiago do Escoural (22H.02)	Homog.	Homog.	Homog.
Azaruja (21K.01)	Santiago do Escoural (22H.02) Lisboa Geofísica (675)	Homog.	Homog.	Homog.
Redondo (22L.01)	São Manços (23K.01) Santiago do Escoural (22H.02)	1963 (1.14)	1963 (1.14)	1963 (1.14)

The SUR+ Ellipse test was applied to the annual number of wet days (threshold 1 mm) at 27 stations. The series from Viana do Alentejo (24I.01) was not tested using this approach

because it was not possible to determine a common period, without too many gaps, for all the series that would be appropriate to model simultaneously (candidates and their respective references).

All the regressors (reference series) parameters of the SUR models are statistically significant at the 5% level, which is why some candidates have only one reference series (Table 3.8). Since each SUR model includes at least two candidate stations' data, some series were tested more than once through different models, depending on the common period of the series included in each model. Consequently, depending on the testing period, different models sometimes provided different results for a specific candidate series (see e.g. the results of Aljezur (30E.01) station in Table 3.8). This problem can be minimized by testing the candidate series through a different model whenever a peak in the graph from the Ellipse test seems suspicious (Figure 3.27, left graph), or by using a combination of statistical tests.

Table 3.8 – Proposed extension of the Ellipse test (SUR+Ellipse) results. Break years detected (and respective relative magnitudes) are presented for the candidate series, for each model tested, at the 5% level. The models were applied to the annual number of wet days (threshold 1 mm).

Candidate	References	Period	Model num.	SUR+Ellipse
Beja (666)	Lisboa Geofísica (675) Relíquias (27G.01) Arronches (19N.01)	1941-1999	5	Homog.
	Relíquias (27G.01) Arronches (19N.01)	1945-1982	13	Homog.
	Lisboa Geofísica (675) Arronches (19N.01)	1956-1997	12	Homog.
Badajoz Talavera (709)	Relíquias (27G.01) Azinheira Barros (25G.01)	1956-1997	6	Homog.
	Relíquias (27G.01) Azinheira Barros (25G.01)	1956-1997	12	Homog.
Aljezur (30E.01)	Arronches (19N.01)	1941-1999	5	Homog.
	Arronches (19N.01)	1956-1997	6	1968 (1.04)
	Arronches (19N.01)	1932-1996	7	Homog.
	Arronches (19N.01)	1932-1994	9	Homog.
	Arronches (19N.01)	1932-1996	10	Homog.
	Arronches (19N.01)	1932-1994	11	Homog.
	Arronches (19N.01)	1956-1997	12	1968 (1.04)
Picota (30K.02)	Barragem da Bravura (30E.03) Alcaria [Castro Marim] (30L.04)	1957-1995	8	Homog.

Candidate	References	Period	Model num.	SUR+Ellipse
Odemira (28F.01)	Santiago do Escoural (22H.02)	1932-1994	9	1952 (0.92)
Aldeia de Palheiros (28H.01)	Santiago do Escoural (22H.02)	1932-1996	7	Homog.
	Santiago do Escoural (22H.02)	1957-1995	8	Homog.
	Santiago do Escoural (22H.02)	1932-1994	9	Homog.
	Santiago do Escoural (22H.02)	1932-1996	10	Homog.
	Santiago do Escoural (22H.02)	1932-1994	11	Homog.
Sabóia (29G.01)	Santiago do Escoural (22H.02)	1932-1994	11	1984 (1.17)
Comporta (23E.01)	Arronches (19N.01)	1956-1997	6	Homog.
	Santiago do Escoural (22H.02)	1957-1995	8	1986 (1.10)
	Arronches (19N.01)	1956-1997	12	Homog.
Azaruja (21K.01)	Santiago do Escoural (22H.02) Lisboa Geofísica (675)	1945-1982	13	Homog.
Redondo (22L.01)	São Manços (23K.01) Relíquias (27G.01)	1945-1982	13	1963 (1.14)

In fact, according to Wijngaard *et al.* (2003), a combination of statistical methods and methods relying on metadata information is considered to be most effective to track down inhomogeneities. Furthermore, that problem might also happen with other testing methods, although it is harder to detect because the tests are usually applied only once. For example, applying the SNHT to the testing variable of Vendas Novas (21G.01) station for the period 1932/92 concludes the series as homogeneous, but testing the period 1938/99 identifies a break in 1943.

The SUR+Ellipse test approach concluded as homogeneous 17 series (5 previously classified as candidates, and 12 as references), at the 5% level. Table 3.8 shows the SUR+Ellipse test results for each model tested, including the break years detected and respective magnitudes, for the 10 series previously classified as candidates.

For a 5% significance level, the SUR+Ellipse test results agree with at least one of the other three tests results for 18 stations (67%). For the remaining 9 stations, the SUR+Ellipse approach considered homogeneous series in 6 stations, where the other methods identified break years, and has identified break years in 3 stations which were considered as homogeneous series by the other methods.

The magnitudes of the breaks detected by the SUR+Ellipse test, but not identified by the other methods, range from –7.6% to 6.88%. Conversely, the magnitudes of the breaks detected by at least one of the other three tests, but not identified by the SUR+Ellipse test range from –14.09% to 10.78%. Hence, there is no apparent connection between the potential breaks magnitudes and the ability of the SUR+Ellipse test to identify them.

Considering the results from the four relative tests, all of them considered as homogeneous 12 (43%) stations' series, whereas at least one of the tests rejected the homogeneity hypothesis for the remaining 16 station series (Figure 3.19), at the 5 % level. Only 4 of the 11 series previously classified as candidates were considered as homogeneous by all the relative tests. Considering the 17 series previously classified as references, 8 of them were considered as homogeneous by all the relative tests.

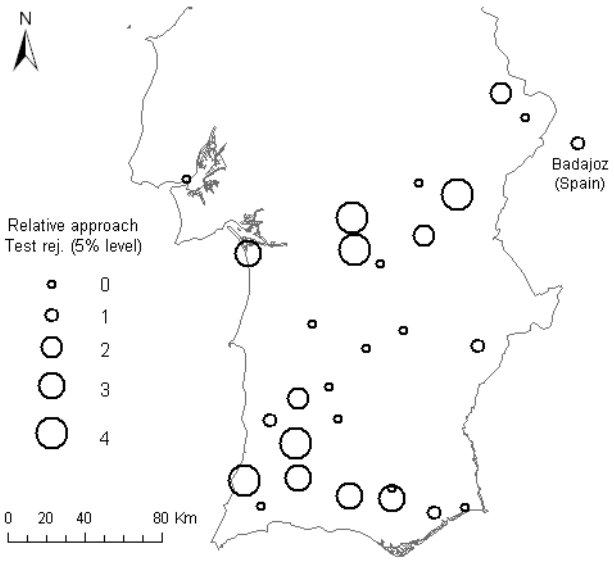

Figure 3.19 – Results from the relative testing approach. Station circles are scaled with the number of tests that rejected the homogeneity hypothesis at the 5% significance level. The Buishand, Pettit, and SNHT tests were applied to 28 composite (ratio) reference series. The proposed SUR+Ellipse test was applied to the annual number of wet days (threshold 1 mm) at 27 stations.

Many of the breaks are detected by all four tests and are mainly located between 1949 and 1954, and around 1986. Thus, there is an apparent trend towards less breaks in recent times (see also Figure 3.20), in contrast to that reported by other homogenization studies (Tuomenvirta, 2001; Wijngaard *et al.*, 2003; Auer *et al.*, 2005; Begert *et al.*, 2005). There were only 20 cases of inhomogeneities detected in the 28 stations' series analysed, which is not surprising since the station selection was based on the absolute testing results. Hence, that

apparent trend may not be true if all the 107 station series were tested through the relative approach. Moreover, the number of inhomogeneities and the breaks magnitudes that can be detected strongly depend on the specifics of the monitoring network and the topography of the different study regions. Furthermore, there are only 4 stations with two break years, whereas the remaining 12 have just one break detected. As expected, for precipitation, with its high variability, few breaks can be detected (Wijngaard *et al.*, 2003).

The average length of the homogeneous subintervals turned out to be 24 years. Taking into consideration the length of the 28 series, the shorter homogeneous period has 3 years (break year detected in 1996 by the SNHT test when applied to the 1956-99 period, at Barranco do Velho (30J.01)), and the longer one has 52 years.

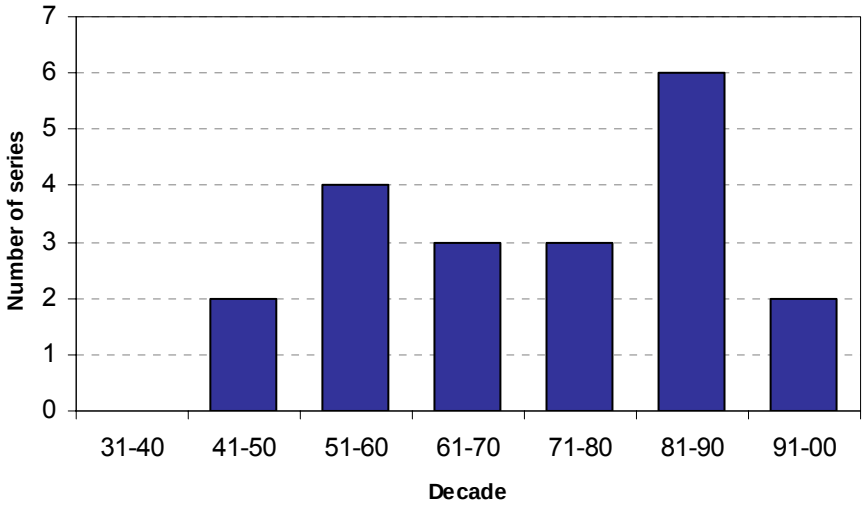

Figure 3.20 – Number of series by breaks detected by the four relative tests (5% level) per decade

The relative magnitudes of the breaks are given by the ratio between the average annual wet day count before and after two consecutive breaks. The breaks magnitudes, of the 20 cases detected, range from –14.1% to 17.3%. The average of the absolute values of the relative magnitudes was 7.3%.

Finally, to illustrate the homogenization analysis performed at this stage, the results of the four homogeneity tests applied are again discussed for the series from Beja (666) and Aljezur (30E.01).

3.3.3.1 Illustration: results of Beja station

Recall from the previous section that, during the absolute testing stage, the Von Neumann ratio test rejected the homogeneity hypothesis for the Beja (666) series, at the 5% significance level. Through relative testing, this series was considered homogeneous by the four tests applied, at the 5% level. Hence, the breaks detected through absolute testing might indicate an abrupt change in the regional climate, or they might be caused by simultaneous changes in the observational network for which relative tests are insensitive.

For Beja (666), the Buishand test statistic was equal to 1.20 for the period 1951-1990, and equal to 1.14 for 1959-1999 (Figure 3.21). The Pettit test statistic was, respectively, equal to 105 and to 85 for the period 1951-1990 and for 1959-1999 (Figure 3.22). Finally, the SNHT test statistic was equal to 4.21 for the period 1951-1990, and equal to 2.0 for 1959-1999 (Figure 3.23). Hence, these last three tests, which are capable of locating the year where the break is likely, considered the series homogeneous, at the 5% level.

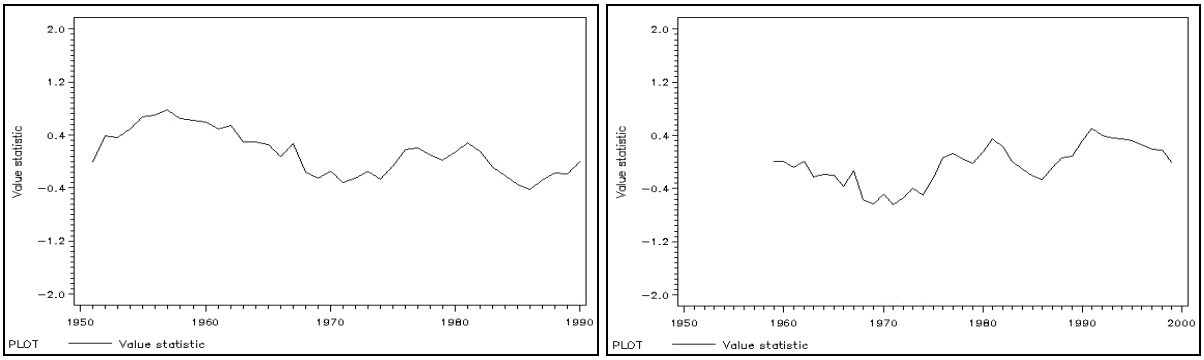

Figure 3.21 – Buishand range test statistic standardized values (left: 1951–1990; right: 1959–1999) for the composite (ratio) reference series of Beja (666) station

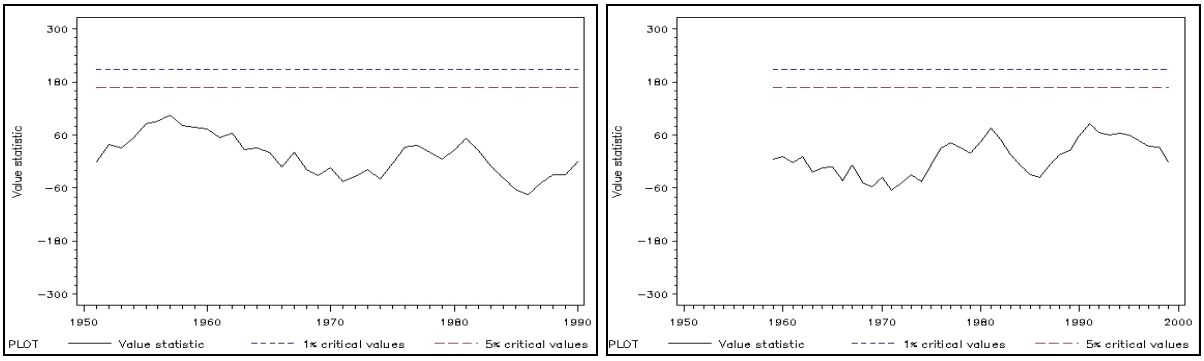

Figure 3.22 – Pettit test results (left: 1951–1990; right: 1959–1999) for the composite (ratio) reference series of Beja (666) station

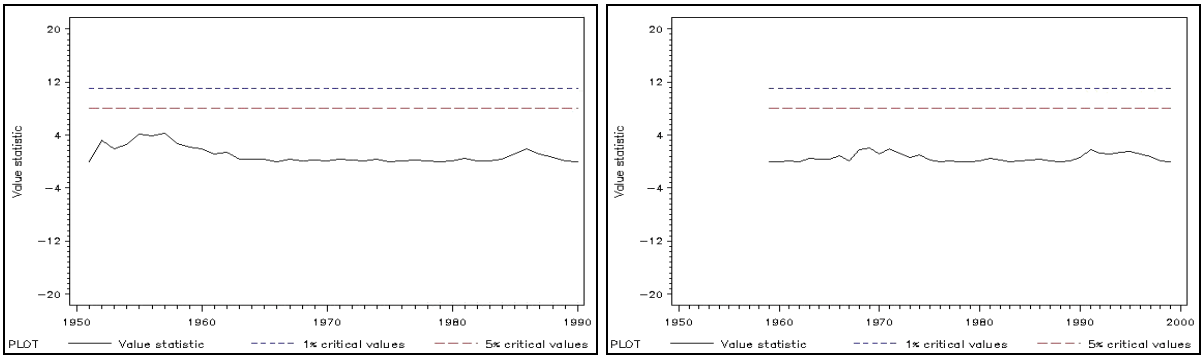

Figure 3.23 – SNHT test results (left: 1951–1990; right: 1959–1999) for the composite (ratio) reference series of Beja (666) station

For the Beja (666) series, the Ellipse test was applied to the residuals from three equations, from the SUR models numbers 5, 12 and 13 (Table 3.8), and all tests considered the series as homogeneous at the 5% level. For illustration purposes, see the SUR+Ellipse test results from models 5 (left graph) and 12 (right graph) in Figure 3.24.

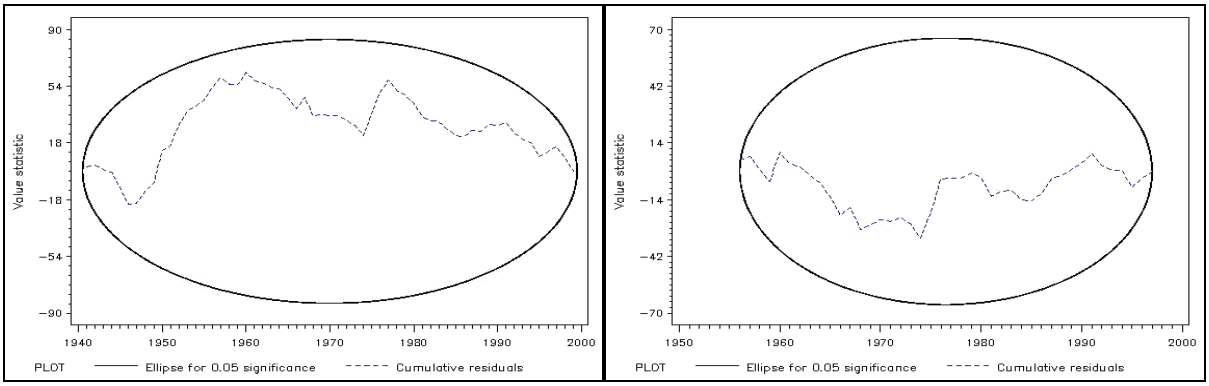

Figure 3.24 – SUR+Ellipse test results (left, SUR model 5: 1941–1999; right, SUR model 12: 1956–1997) for the annual number of wet days (threshold 1 mm) at Beja (666) station

3.3.3.2 Illustration: results of Aljezur station

During the absolute testing stage, the Buishand test detected a break, in 1942, in the Aljezur (30E.01) series, at the 5% significance level. On the other hand, the four relative testing techniques detected a break, in 1968, in the Aljezur (30E.01) series, at the 5% level.

For Aljezur (30E.01), the Buishand test statistic was equal to 1.88 for the period 1956-1995 (Figure 3.25). The Pettit (Figure 3.26, left graph) and the SNHT (Figure 3.26, right graph) tests statistics were equal to (–)255 and to 13.81, respectively, in that period. Hence, these three tests detected a break, in 1968, in the Aljezur (30E.01) series, at the 5% level.

For the Aljezur (30E.01) series, the Ellipse test was applied to the residuals from seven equations, from the SUR models numbers 5, 6, 7, 9, 10, 11 and 12 (Table 3.8). Two Ellipse tests identified breaks in 1968 (models 6 and 12), whereas the other five considered the series as homogeneous, at the 5% level. For illustration purposes, see the SUR+Ellipse test results from models 5 (left graph) and 12 (right graph) in Figure 3.27.

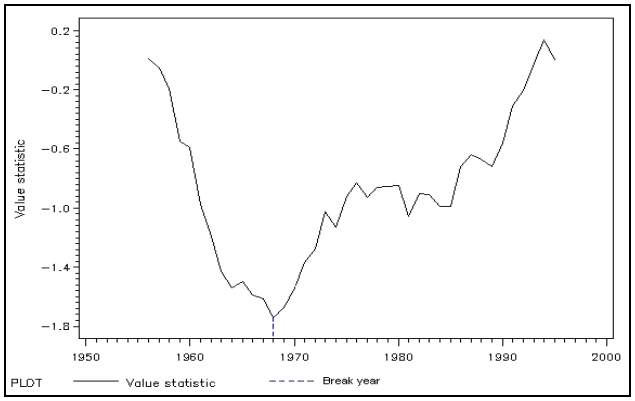

Figure 3.25 – Buishand range test statistic standardized values (1956–1995) for the composite (ratio) reference series of Aljezur (30E.01) station

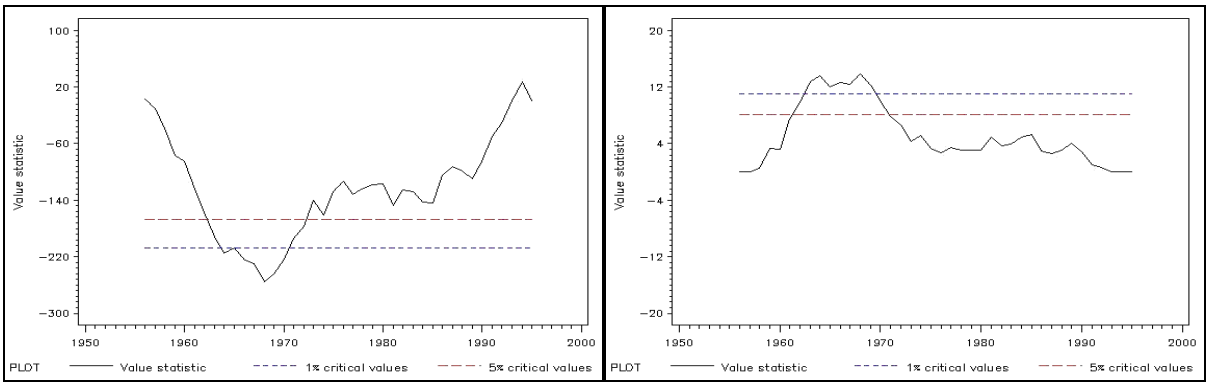

Figure 3.26 – Pettit (left) and SNHT (right) test results (1956–1995) for the composite (ratio) reference series of Aljezur (30E.01) station

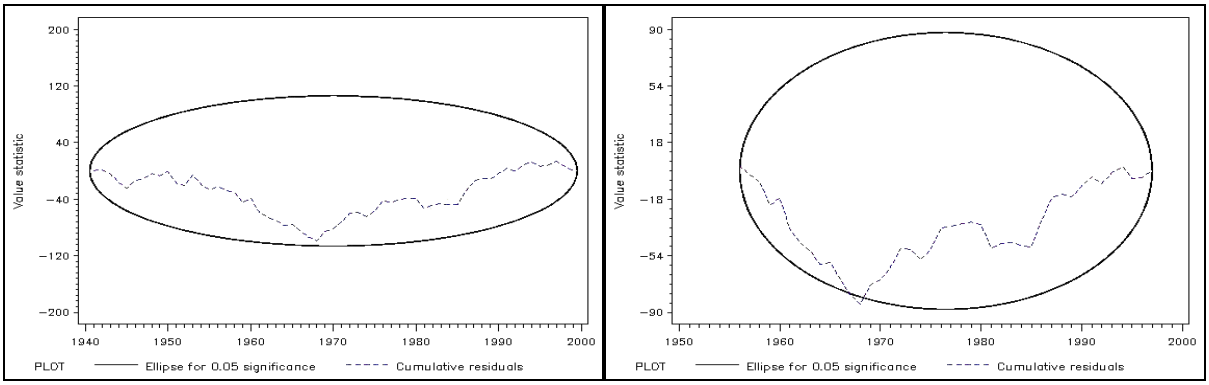

Figure 3.27 – SUR+Ellipse test results (left, SUR model 5: 1941–1999; right, SUR model 12: 1956–1997) for the annual number of wet days (threshold 1 mm) at Aljezur (30E.01) station

As stated before, different models might provide different results for a candidate, depending on the testing period. This problem can be minimized by testing the candidate series through a different model whenever a peak in the graph from the Ellipse test seems suspicious (Figure 3.27, left graph), or by using a combination of statistical tests.

3.3.4 Geostatistical simulation approach

For illustration purposes, the geostatistical simulation approach was applied to the testing variable data, from the period 1980-2001, from 4 candidate stations and 62 reference stations (Figure 3.28). The methodology and results from this approach are described by Costa *et al.* (2008a).

The candidate stations selected are Beja (ECA 666), Aljezur (SNIRH 30E.01), Alferce (SNIRH 30G.01), and Santiago do Escoural (SNIRH 22H.02). The choice of these four stations relied on the results from the relative testing stage, since two of the series were considered homogeneous by all testing procedures within the 1980-2001 period (Beja and Aljezur), whereas for the other two series at least one of the testing procedures identified a break within that period.

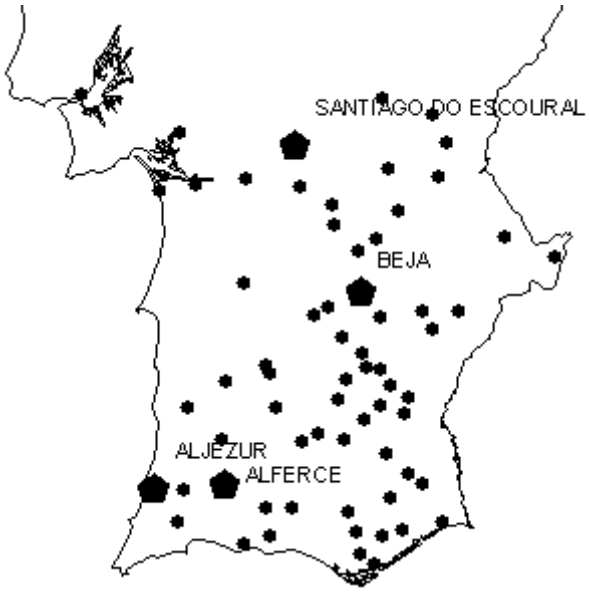

Figure 3.28 – Locations of the 66 monitoring stations. Candidate stations are marked with pentagons.

The 62 reference series were chosen taking into consideration the long-term and short-term series classification provided by the absolute testing results (Section 3.2.4.1). Stations classified as 'candidate', 'reference' and 'useful' (with data within the 1980-2001 period) were selected to illustrate the geostatistical simulation approach (except Badajoz (ECA 709), and a few series from the northeast of the Guadiana basin).

The proposed approach allowed us to identify several inhomogeneities by comparing the observed (real) values of the candidate series, for each year, with the 2.5 and the 97.5th percentiles of the corresponding histograms of 50 simulated maps. The results from this technique were then compared with the results from the relative testing stage in which the SNHT, Pettit and Buishand range tests were applied to composite (ratio) reference series (Section 3.3.3). The geostatistical simulation approach identified not only the same break years (or within one-year range) as the other three testing procedures, but also revealed inhomogeneities in other years that were not detected by any of the three statistical tests, at a 5% significance level.

For Aljezur station, the four approaches considered the series as homogeneous. The series from Beja was considered as homogeneous by the three statistical tests, whereas the geostatistical approach identified a break in 1991 (Figure 3.29).

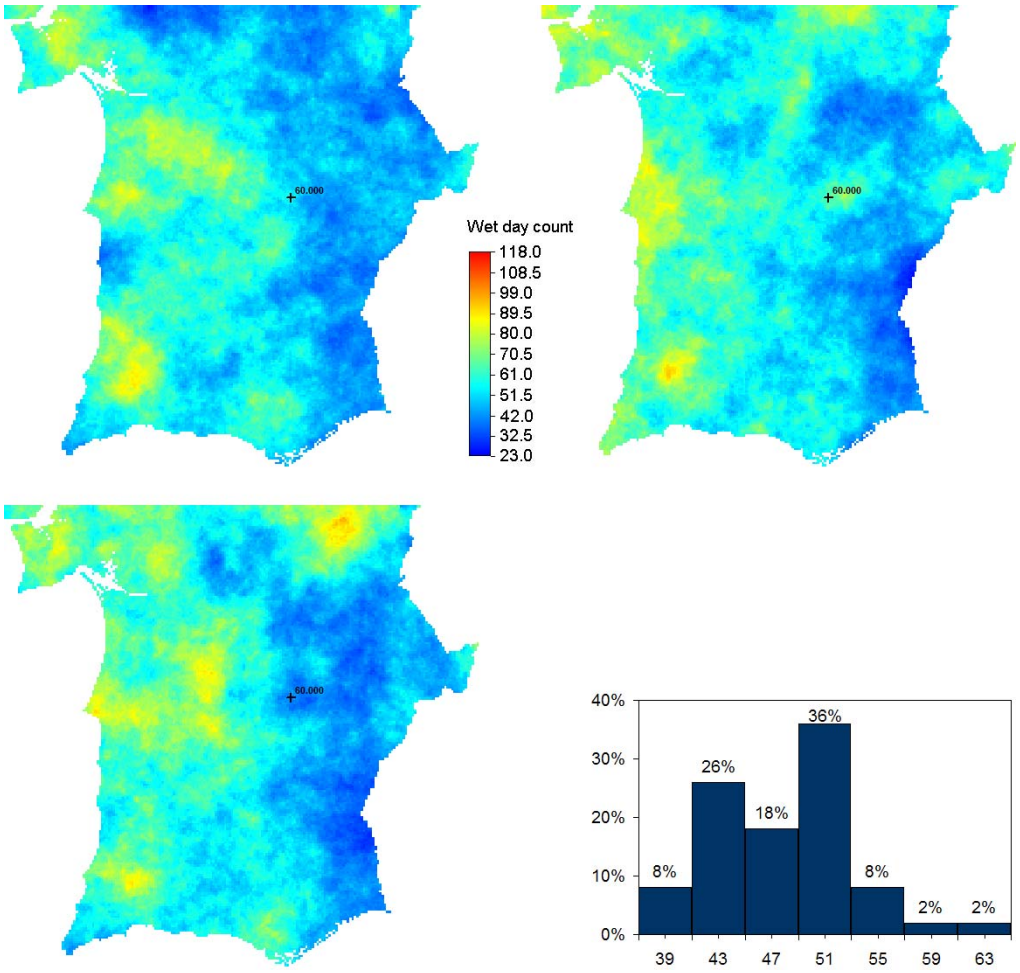

Figure 3.29 – Three simulated realizations of the annual wet day count in 1991, computed without data from Beja, at the nodes of a 1 km x 1 km grid. Histogram of the 50 simulated realizations at Beja location in 1991 (the real value is 60 days)

For Alferce station, the SNHT concluded the series as homogeneous, but the Buishand and Pettit tests detected a break in 1984. Similarly, the geostatistical approach identified a breakpoint in 1983. In fact, at Alferce, the minimum simulated realization of the annual wet day count in 1983 was equal to approximately 33 days, whereas the observed value for this series was 30 days, thus a breakpoint was detected in this year by the proposed technique.

The candidate series from Santiago do Escoural was considered as inhomogeneous by all techniques: the SNHT detected a break in 1989, the Buishand and Pettit tests identified a breakpoint in 1988, and the proposed procedure detected breaks in 1987, 1988 (Figure 3.30, left graph) and 1996 (Figure 3.30, right graph). At Santiago do Escoural, the maximum simulated realization of the annual wet day count in 1987 was equal to approximately 89

days, whereas the observed value was 91 days, so a break was detected in 1987 by the geostatistical simulation approach.

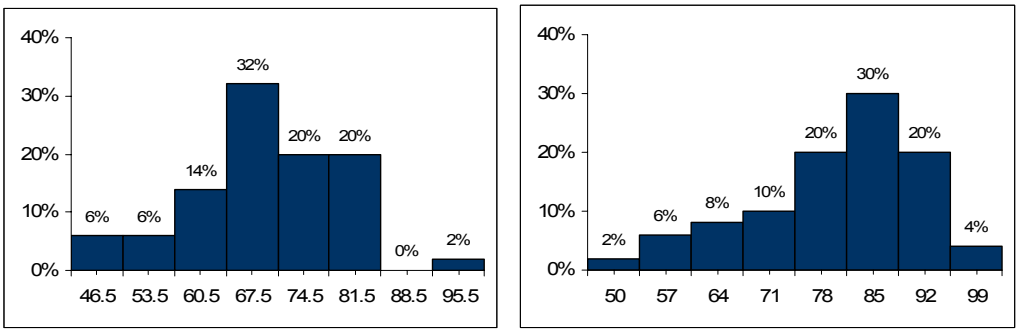

Figure 3.30 – Histograms of the 50 simulated realizations of the annual wet day count at Santiago do Escoural location (left: 1988; right: 1996), computed without data from Santiago do Escoural (the real values are 86 days in 1988, and 96 days in 1996)

These promising results indicate the proposed geostatistical approach as a valuable tool for inhomogeneities detection in climate time series, by accounting for the joint spatial and temporal dependence between observations, and by enhancing the pre-eminence of the closer stations, both in spatial and correlation terms.

All break years identified by the three well-established statistical tests considered were also detected by the proposed technique. Moreover, the geostatistical simulation approach allowed for the identification of breaks near the end of the series that were not detected by the other methods. In fact, this is one of the advantages of the proposed methodology relatively to other testing procedures commonly used which have less power in detecting breakpoints near the start and end of a series (Aguilar *et al.*, 2003).

Furthermore, it accounts for the detection of multiple breaks simultaneously. Another feature of the proposed methodology is that it allows using different sets of neighbouring stations at different years, including shorter and non-complete records.

3.3.5 Overall classification

For the homogenization analysis, a set of 45 long-term and 62 short-term series of daily precipitation were compiled. This section summarises the main results in an overall classification of the daily precipitation time series.

The homogeneity assessment was accomplished through four major stages (e.g. see Figure 3.6 from the *Methodology* section). Following the hybrid approach proposed by Wijngaard *et al.* (2003) for the ECA dataset, we did not attempt to remove non-climatic inhomogeneities from the daily precipitation series, but rather provide a qualitative classification of each station's records. An overall classification of the daily precipitation series was established using four classes: ‘useful’, ‘potentially useful’, ‘potentially suspect’ and ‘suspect’. The criteria used are summarized in Table 3.9, and relied on the results of the homogeneity testing stages.

Table 3.9 – Criteria used to establish the overall classification of the daily series

Classification	Criteria
Useful	All relative approaches considered the series as homogeneous
	Relative break(s) detected might be explained by several months without records
Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
	Absolute break(s) detected might be explained by several months without records
Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Suspect	Relative break(s) detected could not be explained by non-climatic factors

A series was classified as 'useful' when all relative approaches (the four relative statistical tests and the geostatistical approach) considered it as homogeneous. Whenever the daily series had several months without records near a break year, identified by some relative testing procedure, the series was also classified as 'useful', because it is conceivable that the inhomogeneous records were set to missing in the SNIRH database, and the tests rejections were due to them. A series was classified as 'suspect' when at least one of the relative approaches considered it as inhomogeneous and the break(s) detected could not be explained by non-climatic factors.

Considering the series analysed through absolute testing only (both short and long-term), it is difficult to determine if changes or lack of changes result from non-climatic or climatic influences (Peterson *et al*., 1998), since it was not possible to find historic metadata support. Therefore, the intermediate classes, 'potentially useful' and 'potentially suspect', were established. Furthermore, as the short-term series were only analysed through absolute testing,

those series were classified as 'potentially useful' if the six absolute tests considered the series as homogeneous.

Relative approaches that use data from reference stations are usually preferred because they aim to isolate the effects of station irregularities and to account for regional climate changes. However, as mentioned before, a relative approach for the homogenization of all series could not be used, as it would require an iterative procedure in which *all* stations in the data set were seen consecutively as candidates and references, which is out of the scope of this research.

Following those criteria, approximately 13% of the 107 series were classified as 'useful', 55% were classified as 'potentially useful', 19% were classified as 'potentially suspect' and 13% as 'suspect' (Table 3.10).

Table 3.10 – Number of daily precipitation series in the classes 'useful', 'potentially useful', 'potentially suspect' and 'suspect', by series length

Classification	**Long-term series**	**Short-term series**	*Total number of station series*
Useful	14	0	*14*
Potentially useful	8	51	*59*
Potentially suspect	9	11	*20*
Suspect	14	0	*14*
Total num. of station series	*45*	*62*	*107*

Besides the applied criteria and classification of each daily precipitation series, Appendix IV also provides some of the results from the basic quality analysis, namely the number of records flagged through basic quality control procedures. As discussed in Section 3.3.1, records flagged as 'suspect' through the 'flat line' check range from 0.1 mm to 5 mm and the most common values are 0.1 mm and 0.2 mm, thus might be considered useful.

Among the series with records flagged as 'suspect' through subjective flagging procedures, the most problematic ones are Alcoutim (29M.01) and Picota (30K.02), both from the SNIRH database. It might be advisable to set to missing the daily records of the years 1954 to 1959 of Alcoutim, as they were found highly suspicious (this series was classified as 'potentially suspect' using the criteria described in Table 3.9). The daily precipitation records of December 1972 and December 1973 are precisely the same at Picota station, thus it might also be advisable to set them to missing. However, note that even though the responsible

institution is aware of that (Rodrigues, 1990) they were not set to missing in the SNIRH database (this series was classified as 'useful' using the criteria described in Table 3.9).

Figure 3.31 shows the location of the 107 stations according to their series classification. Although defined with different criteria, the qualitative interpretation of the overall classes is similar to the interpretation given for the categories defined for the ECA series (Wijngaard *et al*., 2003). However, it is important to point out that we used the 5% significance level in all statistical tests, whereas those authors used the 1% level. Therefore, our classification is more conservative in the sense that we allowed for the rejection of the homogeneity hypothesis at stations that are considered homogeneous at the 1% significance level.

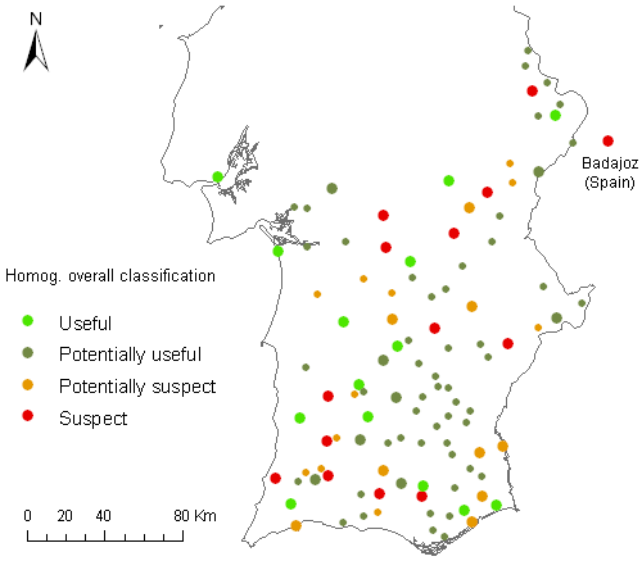

Figure 3.31 – Overall classification of the daily precipitation series by station's location (Large dots: long-term series. Small dots: short-term series)

The series classified as 'useful' seem to be sufficiently homogeneous for trend analysis and variability analysis. The series classified as 'potentially useful' and 'potentially suspect' should be used cautiously, from the perspective of the existence of possible inhomogeneities, as the homogeneity analysis performed might be considered inconclusive – even though all series were considered homogeneous by previous studies (Section 3.1.1). The series classified as 'suspect' should be excluded from trend analysis and variability analysis, as there is strong evidence of inhomogeneities present.

3.3.6 Concluding remarks

The homogeneity assessment of the precipitation time series was accomplished through four major stages, as detailed in Section 3.2 *Methodology*. The results from the procedures implemented and an overall classification of the precipitation series were described and discussed in the preceding sections. However, it is important to mention that other techniques were also investigated during the research work developed for the homogeneity assessment of the precipitation time series.

During the last decades, artificial neural networks (ANNs) have been used in a wide variety of hydrologic contexts ranging from rainfall-runoff modelling and flood forecasting (e.g. Dawson and Wilby, 2001; de Vos and Rientjes, 2005; Kumar *et al.*, 2005), to the prediction of rainfall patterns (e.g. Goswami and Srividya, 1996; Bodri and Čermák, 2000; Luk *et al.*, 2000; Ramírez *et al.*, 2005; Boulanger *et al.*, 2007).

Goswami and Srividya (1996) proposed a generalized structure of a neural network which holds promise for long-range (>2 years) prediction of annual rainfall. Bodri and Čermák (2000) used a feed-forward neural network with back-propagation training algorithm that provided a good fit with the actual monthly data, and showed high feasibility in the prediction of extreme precipitation. The study from Luk *et al.* (2000) focuses on the problems faced in developing a rainfall forecasting model based in ANNs using observed rainfall records in both space and time. These authors identified an optimal set of spatio-temporal inputs for an ANN model to forecast short-term rainfall for an urban catchment. Ramírez *et al.* (2005) generated site-specific quantitative forecasts of daily rainfall, for the region of São Paulo State, Brazil, using a feed-forward neural network with resilient propagation learning algorithm.

Teegavarapu and Chandramouli (2005) tested several methods for the estimation of missing daily precipitation data, including a number of variants of inverse distance weighting methods, ordinary kriging and ANNs. The application and testing of estimation methods developed in their study was carried out in a series of experiments using historical daily rainfall data (1971-2002) available at 20 monitoring stations in the state of Kentucky, USA. Their results suggest that the correlation weighting method, the ANN method (feed-forward network with back-propagation training algorithm) and the kriging approach are the most appropriate to estimate missing daily precipitation data.

According to Bodri and Čermák (2000), the main advantages of the neural network models are the following: (i) no particular knowledge is needed about the system being modelled, unknown effects could be involved through a proper design of the input–output patterns; (ii) relative simplicity of neural network building; and (iii) high ability of the artificial neural networks to reproduce stochastic signals.

As discussed previously in the *Literature review* chapter (Section 2.2.3), further investigation is required to develop procedures for the homogenization of sub-monthly precipitation data (Easterling *et al.*, 1999; Aguilar *et al.*, 2003; Wijngaard *et al.*, 2003; Auer *et al.*, 2005). Therefore, motivated by the preceding discussion, an artificial neural networks approach was investigated for the homogenization of daily precipitation data, particularly multilayer perceptron architectures, which are the most popular types of feed-forward networks. However, this approach failed to succeed due to the lack of data to use in the back-propagation training algorithm. In fact, the major difficulty found was the selection of records from the input data series, because the eligible sets of reference stations did not have enough precipitation observations from common periods. Thus, poor training data inevitably lead to unreliable networks.

Chapter 4: TRENDS IN EXTREME PRECIPITATION

4. Trends in extreme precipitation

Portugal is geographically located in the southwest of the Iberian Peninsula (between 37° and 42°N and 6.5° and 9.5°W). Global circulation and regional climate factors (e.g. latitude, orography, oceanic and continental influences) explain the spatial distribution of rainfall, as well as its intra-annual variability, i.e. seasonal variability (Trigo and DaCamara, 2000; Goodess and Jones, 2002). The precipitation regimes are of a different nature in northern and southern regions of Portugal: in the north, the precipitation regime has an orographic origin, whereas in the south it is mainly associated with vertical motions induced by cyclogenic activity (Trigo and DaCamara, 2000). The inter-annual variability is of a different nature, since the circulation variability is insufficient to explain the observed inter-annual variability of rainfall (Trigo and DaCamara, 2000; Goodess and Jones, 2002; Haylock and Goodess, 2004). In southern Portugal, summer precipitation, almost close to zero during this season, is sometimes associated with local convective activity. These storms can occur with a large degree of independence from the weather circulation type, which characterizes the Iberian circulation for that specific day (Trigo and DaCamara, 2000).

The results obtained by Goodess and Jones (2002) for the Portuguese stations show general agreement with those from Trigo and DaCamara (2000) who considered ten classes of weather circulation types for Portugal. Their results suggest that the cyclonic class is associated with a fairly homogeneous distribution of precipitation over most of the country. Moreover, the ‘rainy’ classes with an Atlantic origin (mainly W and SW; NW to a lesser degree) are to be associated with the observed strong decrease in precipitation from North to South.

Recent studies, based on climate models and past observed records, predict a drier climate in the south of Europe as a result of increased evapotranspiration and a relatively slow decrease of rainfall amounts and precipitation frequency (e.g. Kostopoulou and Jones, 2005; IPCC, 2007; Vicente-Serrano and Cuadrat-Prats, 2007).

In arid and semi-arid regions such as the south of continental Portugal, research on the extent of dryness and temporal trends in heavy rainfall events is an important contribution to evaluate desertification dynamics and to identify areas potentially at risk from land

degradation. However, studies focusing on the role of regional climate change on erosivity and aridity factors are lacking for this region, especially at the local scale (Rosário, 2004b; Pereira *et al*., 2006).

The research on the existence of trends and other temporal patterns in extreme precipitation indices, within the period 1955–1999, at 15 monitoring stations located in southern Portugal is described. This 45-year period was chosen to optimize data availability across the region, taking into consideration the homogenization analysis performed (Chapter 3). Among the numerous indices of precipitation extremes described in the literature (Peterson *et al*., 2001; Frich *et al*., 2002; Kiktev *et al*., 2003; Klein Tank and Können, 2003; Haylock and Goodess, 2004; Kostopoulou and Jones, 2005; Moberg and Jones, 2005), we selected four of them (SDII, R5D, R30 and CDD) and developed two other indices (AII and FDD). Three of the indices (SDII, R5D and R30) provide information on the 'wetness', whereas the other three (CDD, AII and FDD) characterize the 'dryness'. All indicators are based on fixed thresholds and most of them describe moderate climate extremes. The selected indices are appropriate for the purposes of this research, because they might contribute to assess climate dynamics that must be accounted for in impact studies related with water resources management, environmental policies, land use and desertification-related studies for the south of Portugal.

The six daily precipitation indices were subject to a number of diagnosis tests in order to verify the existence of autocorrelation and heteroscedasticity of the regression errors. Depending on the tests' results, trend estimation was performed using three different regression models, namely the simple linear regression model, the autoregressive error model and a heteroscedastic linear model. Moving window statistics (mean and standard deviation) of the precipitation indices were also computed to reduce random fluctuations and provide a clearer view of their underlying behaviour, such as non-linear trends or periods with distinct climatic variability.

This chapter is organized in three sections. The first one (Section 4.1), explains the criteria for station selection from the daily rainfall database that was developed in the previous chapter, and describes the rationale and definitions of the precipitation indices. Section 4.2 presents the methodology used to characterize the dynamic temporal evolution of extreme precipitation indices in the 1955–1999 period. Finally, in Section 4.3, the results are described and discussed, and some conclusions are drawn.

4.1 Study domain and data

4.1.1 Analysis period and data selection

The study domain refers to the south of continental Portugal, and is defined by the Arade, Guadiana, Mira, Ribeiras do Algarve and Sado basins. From the set of 107 stations compiled for homogeneity assessment, one station's data were excluded from the analysis because multiple breakpoints were identified and the homogeneous periods were too short and unreliable; the Badajoz Talavera (709) station, in Spain, was also excluded. The daily rainfall database for the south of Portugal comprises records in the period 1931/2000, but the beginning and ending of each series is highly variable. The selection of stations with quality data for a long common period was developed through several stages.

First, the extreme precipitation indices were computed for the set of 105 stations, regardless of their overall homogeneity classification. However, only the longest homogeneous period was used to build the indices for the series classified as 'suspect'. The extreme precipitation indices are sensitive to the number of missing days, thus the daily records of the selected stations should be as complete as possible. Consequently, for each station, the indices for a specific year were set to missing if there were more than 16% of the days missing for that year (Haylock and Goodess, 2004).

Next, a first set of stations was selected for trend analysis by including all the series classified as 'potentially useful', 'useful' and the longest homogeneous period of the series classified as 'suspect'. In this set, the number of stations with at least 30 years of common observations was very small. Hence, the next stage aimed to select stations classified as 'potentially suspect' with break years near the beginning of the series (identified through absolute testing), so that their longest homogeneous period could also be considered. This allowed us to determine the analysis period 1955/99, which is the longest common period for the final set of 15 series (Figure 4.1). These stations have less than 12% of the days missing in each year, and the data for most stations do not have any missing records.

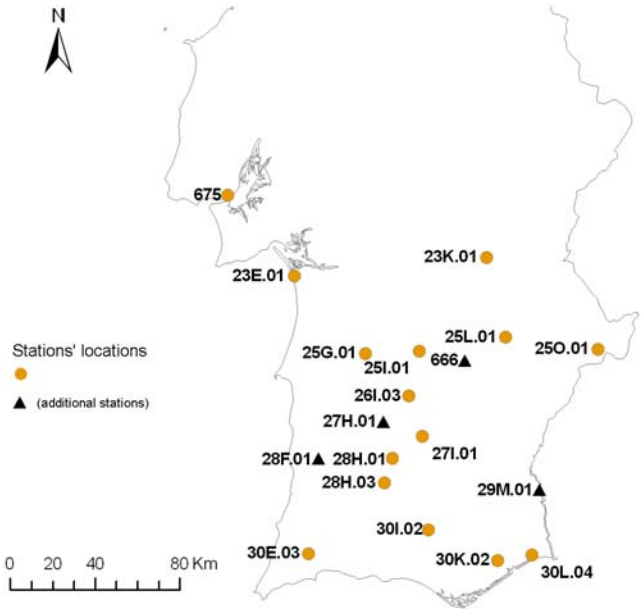

Figure 4.1 – Stations selected for trend analysis. Dots: stations with nearly complete records in 1955-1999. Triangles: additional stations with data within 1940-2000 used to build the regional-average anomaly time series.

4.1.2 Indices of precipitation extremes

There are three main categories of extreme climate indicators: percentile-, threshold- or duration-based indices. The first category of indices is based upon statistical quantities such as percentiles, so the tails of the statistical distribution are examined and days exceeding (not exceeding) a given high (low) percentile are counted. The indices of the second category are based on counts of days crossing a specified fixed value (e.g. the number of days per year with daily precipitation exceeding 20 mm). Duration-based indices allow the characterization of the magnitude of wet/dry spells or heat/cold waves (e.g. the highest consecutive 5-day precipitation total).

Indices based on percentile thresholds have a clear advantage for climate-change detection studies as they compare the changes in the same parts of the precipitation distributions and thus can be used in studies of wide regions containing a broad range of climates (Haylock and Nicholls, 2000; Brunetti *et al.*, 2001; Griffiths *et al.*, 2003; Klein Tank and Können, 2003). On the other hand, indices based on absolute thresholds are beneficial for impact studies as they can be related with extreme events that affect human society and the natural environment (Klein Tank and Können, 2003).

The later set of indices, and indices describing events with short return periods (moderate climate extremes), are suitable for the purposes of this research since they might contribute to assess climate dynamics at the local scale that contribute for land degradation and desertification prone areas of the south of Portugal. Accordingly, we selected four extreme precipitation indices recommended by the joint working group on climate change detection of the World Meteorological Organization – Commission for Climatology (WMO–CCL) and the Research Programme on Climate Variability and Predictability (CLIVAR, Peterson *et al*., 2001; Frich *et al*., 2002; Peterson, 2005), and developed two other indices describing dry conditions.

In the present study only annually specified indices are considered. Their definitions are listed in Table 4.1. Among the indices recommended by the WMO–CCL and the Research Programme on Climate Variability and Predictability, we selected three indices characterizing wet conditions (SDII, R5D, R30), and another one representing dry events (CDD).

Table 4.1 – Acronyms and definitions of the six indices of precipitation extremes

Index	Explanation	Units
SDII	Ratio between the total rain on wet days and the number of wet days (R ≥ 1 mm)	mm
R5D	Highest consecutive 5–day precipitation total	mm
R30	Number of days with daily precipitation totals above or equal to 30 mm	days
CDD	Maximum number of consecutive dry days (R < 1 mm)	days
FDD	Number of dry spells (consecutive period with at least 8 dry days, R < 1 mm)	freq.
AII	Ratio between the total rain on dry days and the number of dry days (R < 10 mm)	mm

The SDII is a simple daily intensity index defined as the average precipitation per wet day. The R5D index is defined as the highest consecutive 5–day precipitation total and can be considered a flood indicator, since it provides a measure of short-term precipitation intensity. The R30 index characterizes the frequency of extremely heavy precipitation events and it is defined as the number of days with daily precipitation totals above or equal to 30 mm. This threshold fits the extreme events regime of the study area as the 30 mm value approximately corresponds to the 95% regional-average percentile of the 1961/90 climate normal. The CDD index corresponds to the maximum number of consecutive dry days, and therefore characterizes the length of the greatest dry spell.

Not only drought but also moderate dry conditions have significant impacts in terms of crop losses, water supply shortages, land degradation and desertification in the south of Portugal. To better understand the pluviometric regime of this region, and the frequency and magnitude of dryness in particular, we developed two indices describing dry events (FDD and AII).

The FDD index is defined as the number of dry spells. For return periods of two years, expected dry-day lengths vary from 60 to 80 days in the study region (Lana *et al*., 2008). The selected indices refer to precipitation events with return periods typically of less than one year, providing relevant information to impact studies. In the FDD definition, a dry spell is a consecutive period with at least 8 dry days. The average length of dry spells in a year ranges from 8 to 12 days in the study region (Lana *et al*., 2008). Therefore, increasing (decreasing) trends of FDD are indicators of a change in the mean frequency of dry events, rather than in the frequency of extremely dry situations. Although a change in the extreme values of the distribution obviously implies a change in the mean.

Regarding the SDII, CDD and FDD indices, a wet day is defined as a day with at least 1 mm of precipitation ($R \geq 1$ mm), thus a dry day has less than 1 mm of precipitation ($R < 1$ mm). Ceballos *et al*. (2004) state that rainfall amounts below this threshold are not absorbed by soils and are evaporated off directly. In fact, Moberg and Jones (2005) agree that, with this definition, a dry day is allowed to have a small amount of precipitation, but generally small enough for the ground not to recover after a long period of dryness. Moreover, thresholds lower than 1 mm can introduce trends in the number of wet days, associated with measurement errors introduced by the observers (Haylock and Nicholls, 2000; Haylock and Goodess, 2004) or by instrument inaccuracies. In fact, taking into consideration the 'flat line' check results (Section 3.3.1), it is prudent to adopt such a threshold for dry days ($R < 1$ mm) because it allows for minimizing the effect of any inaccuracy associated with measurement errors.

In the definition of the AII index, we used the 10 mm threshold to typify a dry day (Ceballos *et al*., 2004; Lana *et al*., 2008). Let *RL10t* be the total rain on days with precipitation amount below 10 mm ($R < 10$ mm), and let *RL10* be the number of days with $R < 10$ mm. Similarly to the SDII, the AII index is defined by *RL10t/RL10* and can be interpreted as a simple aridity index, because it is a numerical indicator of the degree of dryness of the climate at a given location. Below the 10 mm threshold, the rainfall has a small effect on the soil water-content,

since the rainfall evaporates very quickly and hardly drains into the upper soil layer (Ceballos *et al*., 2004). Increasing (decreasing) trends of AII are indicators of change in the normal moisture availability, which is a sensitive issue for desertification susceptible regions.

4.2 Methodology

The six daily precipitation indices were subject to a number of diagnosis tests, at each station's location, in order to verify the existence of autocorrelation and heteroscedasticity of the regression errors. Depending on the tests' results, the trend estimation was performed using three different regression models.

The indices are expressed as annual values Y_t, $t=1,...,T$ with the subscript t referring to the year (also denoted by X_t), and T is the length of the period covered by the station's series. The simple linear regression model, estimated by ordinary least squares (OLS), is

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t, \quad \varepsilon_t \sim IN(0, \sigma^2) \tag{4.1}$$

where ε_t is a disturbance term (error) and the notation $\varepsilon_t \sim \text{IN}(0, \sigma^2)$ indicates that each error is normally and independently distributed with mean *0* and variance σ^2. If the assumption of constant errors variance is violated, the errors are said to be heteroscedastic. If heteroscedasticity is present and OLS regression is computed, the parameter estimates are still unbiased but they are no longer efficient, and inferences from the standard errors are likely to be misleading.

Engle's Lagrange multiplier test for heteroscedasticity (Engle, 1982) allows to test if the errors variance has the form $V(\varepsilon_t) = \sigma_t^2 = \alpha_1 + \alpha_2 X_t$, where α_1 and α_2 are constant parameters. Whenever the null hypothesis of homoscedasticity was rejected, the following *heteroscedastic linear model* was fitted:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \sigma^2(\alpha_1 + \alpha_2 X_t) + \eta_t, \qquad \eta_t \sim IN(0, \sigma_\eta{}^2) \end{aligned} \tag{4.2}$$

where η_t is a disturbance term (error) with constant variance $\sigma_\eta{}^2$. The slope (β_2) of the trend of the heteroscedastic linear model, in Equation (4.2), was calculated by the Yule–Walker

estimation method described in Gallant and Goebel (1976). The Yule–Walker method can be considered as generalized least squares using the OLS residuals to estimate the covariances across observations, thus Griffiths *et al*. (1993) use the term *estimated generalized least squares* (EGLS) for this method.

If the error term is not independent through time, the OLS estimates of the regression coefficients are still unbiased but they are no longer as efficient as they would be if the autocorrelation was taken into account. Furthermore, the statistical tests of the significance of the OLS parameters are not correct because the standard error estimates are invalid. The presence of autocorrelation was investigated using the Durbin–Watson test. Whenever autocorrelation correction was needed, the following *autoregressive error model* was fitted:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + \varepsilon_t, \\ \varepsilon_t &= \rho\varepsilon_{t-1} + \eta_t, \quad \eta_t \sim \mathrm{IN}(0, \sigma^2) \end{aligned} \tag{4.3}$$

where ρ is the autoregressive error model parameter. The slope of the trend of the autoregressive error model, in Equation (4.3), was computed by estimated generalized least squares (EGLS) using the Yule–Walker estimate of ρ.

Whenever the homoscedasticity hypothesis and the independent errors assumption were not rejected by the Lagrange multiplier test and the Durbin–Watson test, respectively, the slope of the trend was estimated by OLS. Many stations' wetness indices had significant non-Gaussian residuals (tested through the Shapiro–Wilk normality test). Other forms of OLS assumptions violations were not investigated (e.g. other forms of heteroscedasticity). Therefore, for all models fitted, the trend significance was assessed through the nonparametric Mann–Kendall test.

The existence of significant trends in anomaly time series was also investigated using the described methodology. In each year, the anomalies of the indices time series were calculated from the base period of 1961–1990 by standardising the individual station's series using the climatological average and standard deviation of the base period. The regional-average anomaly series were computed using the full set of 19 stations' series and the analysis period was set to 1940–2000. As all stations do not contain complete data in this period, the regional anomaly of a year was obtained by weighting the anomalies according to the number of stations available for that year (Frich *et al*., 2002).

Any cyclical components in the variation of time series make it difficult to see the underlying trend. Aiming to improve our understanding of the indices time series, smoothing techniques were used to reduce random fluctuations and provide a clearer view of their underlying behaviour. The temporal variability of the extreme indices was analysed by means of moving window procedures. Moving average smoothing is a smoothing technique used to make the long-term trends and fluctuations of a time series clearer. Moving windows with a time span of 5 and 10 years were used to compute moving average series of the extreme precipitation indices. For the sake of simplicity, these series are denominated the temporal average of extremes (TAE) series. Moving 5 and 10 years standard deviation statistics of the indices were also computed in order to analyse the temporal evolution of their variability. These series are denominated the temporal variability of extremes (TVE) series. The TAE and TVE series were calculated for each station and were then averaged over the 15 stations to obtain the regional-average TAE and TVE series for the period 1955/99.

4.3 Results and discussion

The data analysis for this study was generated through specific programs developed using SAS software macros, SAS/STAT®, SAS/ETS® and SAS/INSIGHT® software, of the SAS System[15] for Windows, Version 8.

A regional correlation analysis, averaging the Spearman rank-order correlation coefficients of the six indices over the 15 stations, revealed that the dryness indices (CDD, FDD, AII) might provide information that is essentially different from the three wetness indices (SDII, R5D, R30), because they are uncorrelated with any of them. Nevertheless, the correlations between AII and the three wetness indices have positive signs, whereas the other two dryness indices show negative signs when correlated with the wetness indices. Interestingly, the correlation between AII and SDII is extremely weak. The correlation between CDD and FDD is negative but weak, which might indicate that an increase (decrease) in the length of the greatest dry spell will not necessarily entail a significant decrease (increase) in the mean frequency of dry events. The three wetness indices are moderately positively correlated with each other.

[15] SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. In the USA and other countries ® indicates USA registration.

4.3.1 Trends in extreme indices

Depending on the diagnoses tests' results, trend estimation was performed using three different regression models, namely the OLS model, the autoregressive error model and the heteroscedastic linear model. Trend significance was tested using the Mann–Kendall test and results are presented at the 5% and 10% significance levels (Table 4.2 and Table 4.3).

Table 4.2 – Trends in precipitation indices estimated with the *OLS model* (O), the *Autoregressive error model* (A) and with the *Heteroscedastic linear model* (H), for the period 1955/99. Significance of trends assessed using the Mann–Kendall test: values in bold face are significant at <5% level (marked with **) and <10% level (marked with *).

Station	Code	CDD	FDD	AII	SDII	R5D	R30
Comporta	23E.01	–0.0094 (O)	0.0066 (O)	–0.0007 (O)	**0.0394**** (A)	**0.8518**** (O)	**0.0566*** (H)
São Manços	23K.01	0.0558 (A)	–0.0086 (O)	**–0.0029**** (O)	0.0156 (O)	**0.5813*** (A)	0.0040 (O)
Azinheira Barros	25G.01	–0.1734 (O)	0.0154 (O)	–0.0011 (O)	0.0066 (H)	0.1577 (O)	0.0194 (O)
Ferreira do Alentejo	25I.01	0.1489 (O)	–0.0067 (O)	**–0.0032*** (H)	**0.0250**** (A)	**0.3814*** (H)	–0.0130 (O)
Pedrogão do Alentejo	25L.01	0.2258 (O)	–0.0061 (O)	**–0.0039**** (A)	–0.0117 (A)	0.1358 (H)	–0.0093 (A)
Santo Aleixo da Restauração	25O.01	–0.0232 (O)	0.0202 (O)	**–0.0040**** (O)	–0.0106 (H)	0.0076 (H)	–0.0144 (A)
Aljustrel	26I.03	**0.5369**** (H)	–0.0121 (O)	**–0.0032**** (H)	–0.0252 (A)	–0.4006 (O)	–0.0066 (A)
Castro Verde	27I.01	–0.2645 (H)	**0.0497**** (A)	–0.0026 (O)	0.0056 (O)	0.1728 (H)	–0.0030 (O)
Aldeia de Palheiros	28H.01	**–0.0944**** (O)	0.0130 (O)	**–0.0034**** (O)	**0.0022**** (H)	0.1886 (O)	**–0.0099**** (H)
Santana da Serra	28H.03	**0.6785**** (O)	–0.0228 (O)	**–0.0039*** (O)	**0.0636**** (H)	**0.7582*** (H)	0.0304 (H)
Barragem da Bravura	30E.03	–0.1025 (H)	–0.0048 (A)	–0.0021 (O)	**0.0306*** (H)	**0.5972**** (O)	**0.0617**** (H)
Sobreira	30I.02	–0.1640 (O)	0.0168 (A)	**–0.0034**** (O)	**0.0588**** (H)	**0.7732*** (A)	0.0377 (H)
Picota	30K.02	–0.3487 (A)	0.0141 (H)	**–0.0023*** (O)	–0.0238 (O)	–0.0183 (H)	–0.0102 (O)
Alcaria (Castro Marim)	30L.04	0.0925 (O)	–0.0194 (O)	–0.0005 (O)	–0.0124 (O)	0.6295 (O)	–0.0014 (O)
Lisboa Geofísica	675	–0.0939 (H)	–0.0080 (O)	–0.0022 (O)	–0.0038 (O)	0.1533 (H)	0.0022 (H)

According with the correlation analysis results, the CDD and FDD indices have oppositely signed trends for most of the stations, although not statistically significant. Hence, these indices do not reflect significant changes neither in the length of dry spells, nor the frequency of dry events. On the other hand, the AII index reflects increases in the magnitude of dryness. A negative station trend in the AII index dominates in the 1955/99 period, which implies a significant increase of aridity over most of the study region (Figure 4.2).

Table 4.3 – Trends in anomaly time series of precipitation indices estimated with the *OLS model* (O), the *Autoregressive error model* (A) and with the *Heteroscedastic linear model* (H), for the period 1955/99. Significance of trends assessed using the Mann–Kendall test: values in bold face are significant at <5% level (marked with **) and <10% level (marked with *).

Station	Code	CDD	FDD	AII	SDII	R5D	R30
Comporta	23E.01	–0.0003 (O)	0.0034 (O)	–0.0064 (O)	**0.0267**** (A)	**0.0263**** (O)	**0.0301*** (H)
São Manços	23K.01	0.0021 (A)	–0.0049 (O)	**–0.0213**** (O)	0.0116 (O)	**0.0278*** (A)	0.0026 (O)
Azinheira Barros	25G.01	–0.0083 (O)	0.0078 (O)	–0.0081 (O)	0.0067 (H)	0.0072 (O)	0.0113 (O)
Ferreira do Alentejo	25I.01	0.0067 (O)	–0.0033 (O)	**–0.0234*** (H)	**0.0208**** (A)	**0.0243*** (H)	–0.0113 (O)
Pedrogão do Alentejo	25L.01	0.0083 (O)	–0.0031 (O)	**–0.0283**** (A)	–0.0053 (A)	0.0057 (H)	–0.0045 (A)
Santo Aleixo da Restauração	25O.01	–0.0009 (O)	0.0095 (O)	**–0.0299**** (O)	–0.0086 (H)	0.0003 (H)	–0.0081 (A)
Aljustrel	26I.03	**0.0235**** (H)	–0.0044 (O)	**–0.0250**** (H)	–0.0117 (A)	–0.0196 (O)	**–0.0041**** (A)
Castro Verde	27I.01	–0.0126 (H)	**0.0273**** (A)	–0.0181 (O)	0.0045 (O)	0.0090 (H)	**–0.0016**** (O)
Aldeia de Palheiros	28H.01	–0.0038 (O)	0.0060 (O)	**–0.0234**** (O)	0.0018 (H)	0.0083 (O)	**–0.0051**** (H)
Santana da Serra	28H.03	**0.0259**** (O)	–0.0123 (O)	**–0.0244*** (O)	**0.0377**** (H)	**0.0254*** (H)	0.0136 (H)
Barragem da Bravura	30E.03	–0.0037 (H)	–0.0024 (A)	**–0.0175*** (O)	0.0144 (H)	0.0198 (O)	0.0182 (H)
Sobreira	30I.02	–0.0059 (O)	**0.0075*** (A)	**–0.0225**** (O)	**0.0254**** (H)	**0.0224*** (A)	0.0104 (H)
Picota	30K.02	**–0.0164*** (A)	0.0103 (H)	**–0.0198**** (O)	**–0.0085*** (O)	–0.0005 (H)	–0.0030 (O)
Alcaria (Castro Marim)	30L.04	0.0034 (O)	–0.0098 (O)	–0.0039 (O)	–0.0041 (O)	0.0140 (O)	–0.0006 (O)
Lisboa Geofísica	675	–0.0010 (H)	–0.0035 (O)	–0.0166 (O)	–0.0027 (O)	0.0055 (H)	**0.0009*** (H)

The SDII monitors precipitation intensity on wet days and presents significant increasing trends in several stations, but without spatial consistency. A few of them also have significant increasing trends in the maximum 5-day precipitation totals (R5D), but this tendency is not significant for the majority of stations in the R30 index which characterizes the frequency of extremely heavy precipitation events. These results agree with other similar studies of the ECA series, corresponding to the same region and baseline period, reported in the literature (Klein Tank and Können, 2003; Miranda *et al*., 2006).

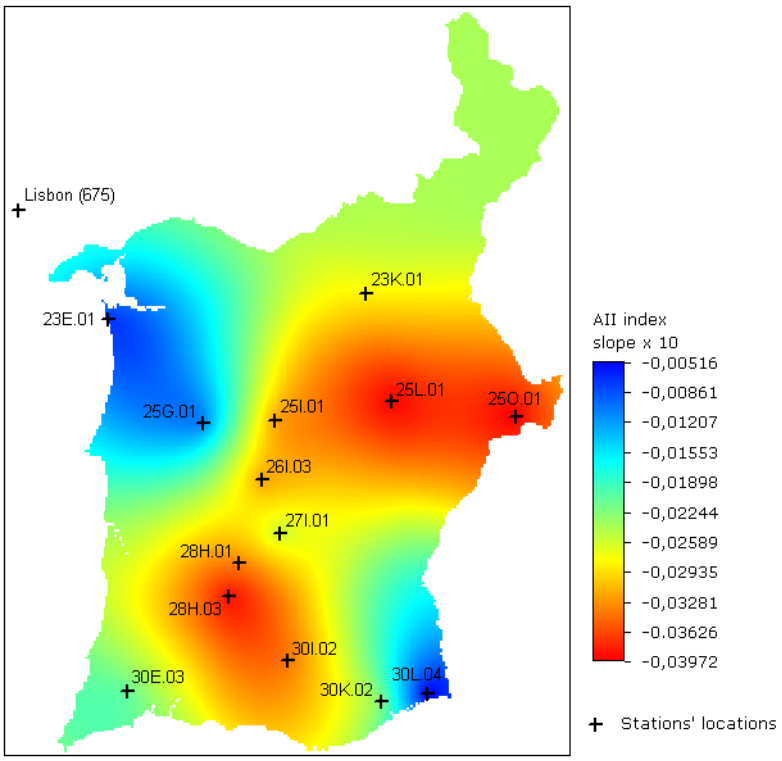

Figure 4.2 – Ordinary kriging interpolation (800m × 800m grid) of the trends per decade in the AII index

The trend results of the anomaly time series (Table 4.3) are exactly the same as the precipitation indices results (Table 4.2) concerning the trend signals, but the trends significance is different for a few stations and indices:

- The anomaly time series of Picota (30K.02) has a significant decreasing trend in the length of dry spells (CDD), while the CDD time series is not significant. On the other hand, the CDD time series of Aldeia de Palheiros (28H.01) is decreasing significantly, but the corresponding anomaly time series is not.

- The anomaly time series of Sobreira (30I.02) has a significant increasing trend in the frequency of dry events (FDD), while the FDD time series is not significant.
- The AII anomaly time series of Barragem da Bravura (30E.03) has a significant decreasing trend, which corresponds to an increase in the magnitude of dryness, while the AII time series is not significant.
- The anomaly time series of Picota (30K.02) has a significant decreasing trend in the precipitation intensity (SDII), while the SDII time series is not significant. On the other hand, the SDII time series of Aldeia de Palheiros (28H.01) and Barragem da Bravura (30E.03) are increasing significantly, while the corresponding anomaly time series are not.
- The R5D time series of Barragem da Bravura (30E.03) is increasing significantly, while the corresponding anomaly time series is not.
- The anomaly time series of Aljustrel (26I.03), Castro Verde (27I.01) and Lisboa (675) have significant trends in the frequency of extreme precipitation (R30), while the R30 time series are not significant for these stations. On the other hand, the R30 time series of Barragem da Bravura (30E.03) is increasing significantly, but the corresponding anomaly time series is not.

Coherent spatial patterns of statistically significant changes emerge in the magnitude of dryness (AII), while the remaining anomaly time series show a lack of spatial consistency. The remaining indicators show mixed patterns of change but significant increases have occurred in the extreme amount derived from short-term precipitation intensity (R5D) in five stations, while for other three stations significant decreases in the number of heavy rainfall events (R30) have occurred.

The existence of significant trends in the regional-average anomaly time series was also investigated (Figure 4.3) using 19 stations' data for the period 1940–2000. As discussed before, absence of spatial consistency and/or significant trends characterizes the majority of the precipitation indices calculated, except for the AII index. Therefore, not surprisingly, this was the only index with a significant decreasing trend (the p-value of the Mann–Kendall test is equal to 0.06) in the regional-average anomaly time series (Figure 4.3c), indicating an increase of dryness over the study region.

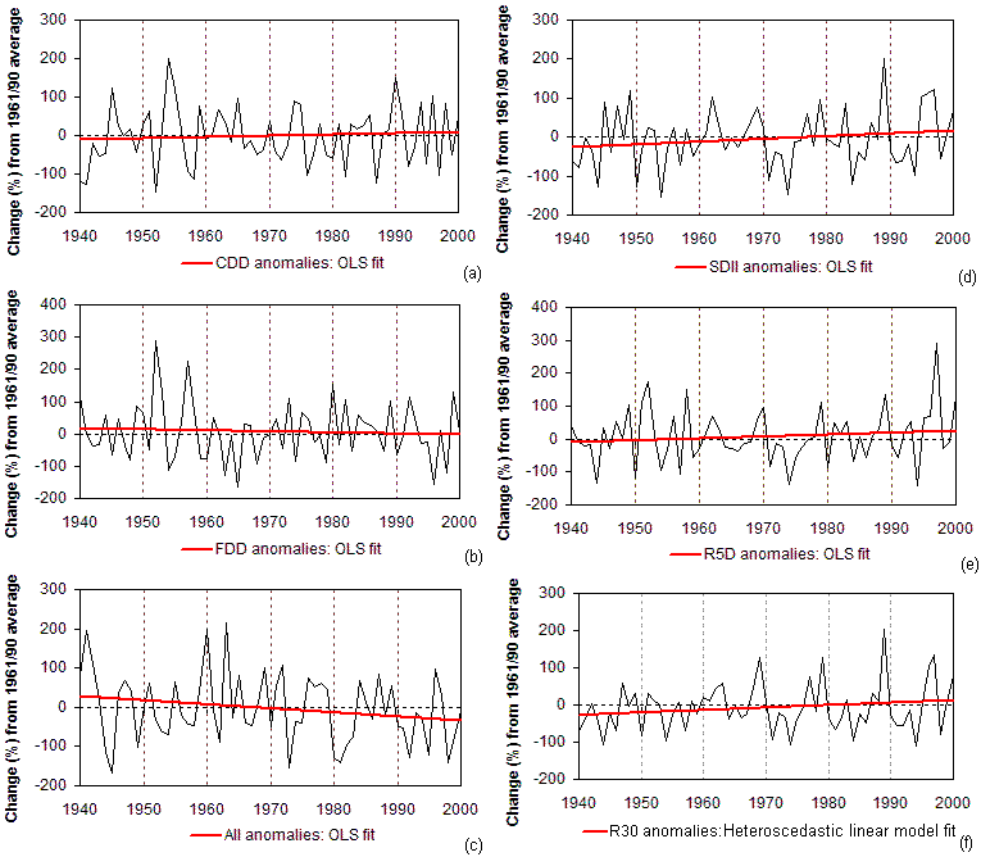

Figure 4.3 – Differences in the average extreme indices' values between 1940 and 2000 from the average 1961/90 value of weighted regional stations. The trend of the AII annual anomalies series (c) is significant at the 6% level.

4.3.2 Dynamic temporal evolution of extremes

In order to reduce random fluctuations and improve our understanding of the indices time series, moving window statistics (mean and standard deviation) with a time span of 5 and 10 years were computed for each station, and then averaged over the 15 stations to obtain a regional-average. This is a very useful approach because non-linear trends in precipitation extremes can be revealed, and periods with distinct climatic variability can be identified. The results obtained using windows with a time span of 5 years are identical to the ones obtained with a time span of 10 years, but a little noisier. For this reason, and for the sake of simplicity, only the later are presented.

The moving average series of the extreme precipitation indices was named TAE series. The temporal dynamics underlying these series were captured through *weighted local polynomial models* (LOWESS smoother proposed by Cleveland, 1979) fitted with a time span of *k* years determined by generalized cross-validation. In order to point out any non-linear trends

underlying the TAE series, simple linear regression models, estimated by OLS, were also fitted.

For all of the precipitation indices considered, the results show non-linear trends in the TAE series within the 1955/99 period at the large majority of stations. These results might explain why the regression models fitted to the indices time series could not significantly capture the trend signal for the majority of the stations' indices. Figure 4.4 shows the results of the regional-average TAE series of the six indices. The non-linear trends and cyclic patterns of the individual stations' TAE series (not shown) are identical to the ones illustrated in Figure 4.4, but much more sharpen (any exceptions are referred in the text).

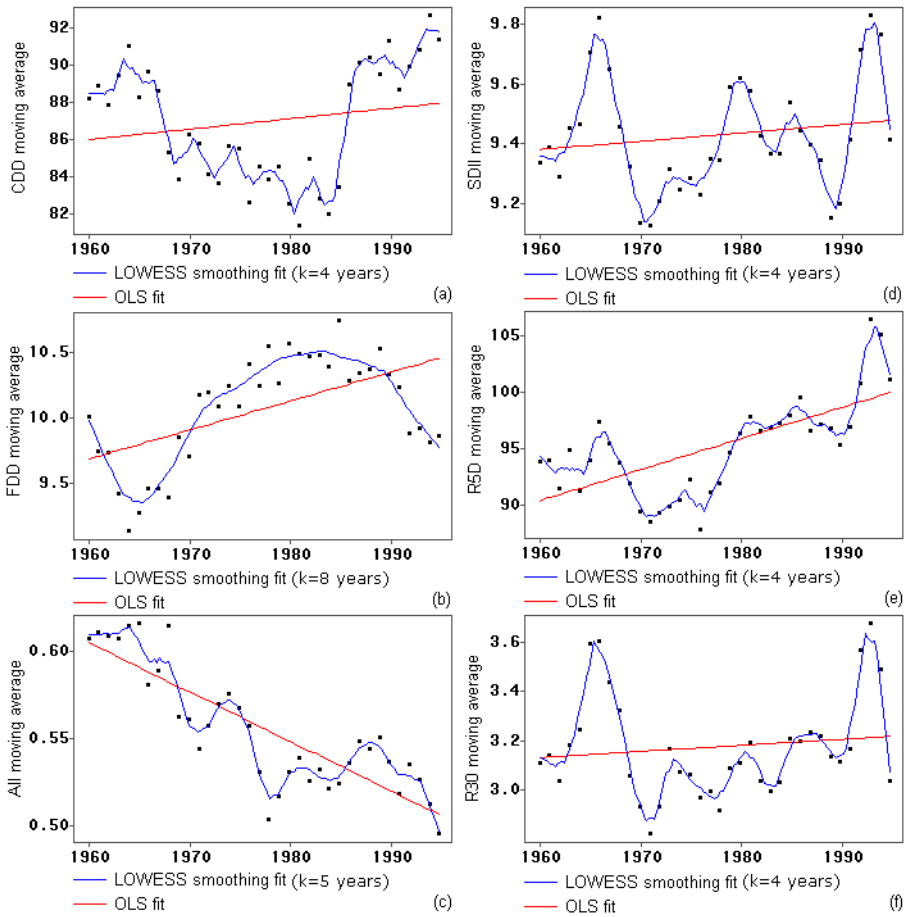

Figure 4.4 – Ordinary least squares fitting (OLS, red line) and weighted local polynomial fitting (LOWESS smoothing, blue line) for each regional-average TAE series (moving average of the extreme index using a time span of 10 years), for the period 1955/99.

As expected from the previous results, the TAE series of AII clearly reflect a strong increase in the magnitude of dryness during the period 1955/99. The maximum length of dry spells, characterized by the TAE series of CDD, has a decreasing trend until the middle of the 1980s

and then suddenly increases, keeping a positive trend through the last decade of the twentieth century. One exception occurs at Pedrogão do Alentejo (25L.01) where the TAE series of CDD has a well-defined cyclic pattern with no trend, and another one occurs at Picota (30K.02) where the duration of dry spells has a decreasing trend during the last three decades of the twentieth century. The TAE series of FDD show a different temporal pattern. The frequency of dry days decreases until the mid 1960s, and then has a parabolic behaviour by increasing until the middle of the 1980s and decreasing afterwards. Those findings led us to perform a regional correlation analysis between the CDD and the FDD indices by decade. The Spearman correlation coefficients between these indices were negative but extremely weak as reported before, but this time with one exception: in the last decade of the twentieth century, the correlation dropped to −0.66. These results seem to indicate that, in recent times, an increase in the length of the greatest dry spell entails a stronger decrease in the mean frequency of dry events. Moreover, those dryness indices remain uncorrelated with any of the other indices within that decade.

The TAE series of R5D clearly reflect a strong increase in the short-term precipitation intensity during the last three decades of the twentieth century, except at Lisbon (675) and Aljustrel (26I.03). The regional-average TAE series of SDII, characterizing the precipitation intensity on wet days, shows a cyclic pattern with a small positive tendency after the 1970s. A closer look at the individual stations' TAE series reveals an opposite behaviour at six stations (codes: 675, 25G.01, 25L.01, 25O.01, 26I.03, 27I.01) located in the centre of the study region. These stations show no trend or a small negative tendency during the last three decades of the twentieth century, whereas all other stations have increasing trends in rainfall amounts on wet days. The temporal pattern in the frequency of extremely heavy precipitation events, characterized by the regional-average TAE series of R30, is similar to the pattern of the regional-average TAE series of SDII until the 1970s. Afterwards, the frequency of extreme rainfall has a cyclic pattern with a positive tendency until the mid 1990s and a sudden decrease at the end of this decade. A few stations' TAE series of R30 (stations' codes: 25I.01, 25O.01, 26I.03, 27I.01) exhibit no trend or a small negative tendency in the last three decades of the twentieth century, whereas all other stations have increasing trends in the frequency of heavy rainfall.

The TVE series (moving 5 and 10 years standard deviation statistics) allow the characterization of the extreme precipitation variability through time, and constitute indicators

of uncertainty associated with the temporal patterns of the extreme precipitation indices in the period 1955/99. Figure 4.5 shows the results of the regional-average TVE series of the six indices.

The TVE series of wetness indices reveal an increase of variability in rainfall frequency and intensity along the 1955/99 period. The variability of the length of dry spells, characterized by the TVE series of the CDD index, has also increased through time, whereas the variability of the frequency of dry spells (FDD) shows a downward pattern with an increase during the 1990s. These results also support the difficulties with capturing trend signals for these indices at the majority of stations. The TVE series of the AII index reflects a decrease in the dryness variability except during the 1990s. These results indicate that extreme precipitation variability and climate uncertainty are greater in recent times.

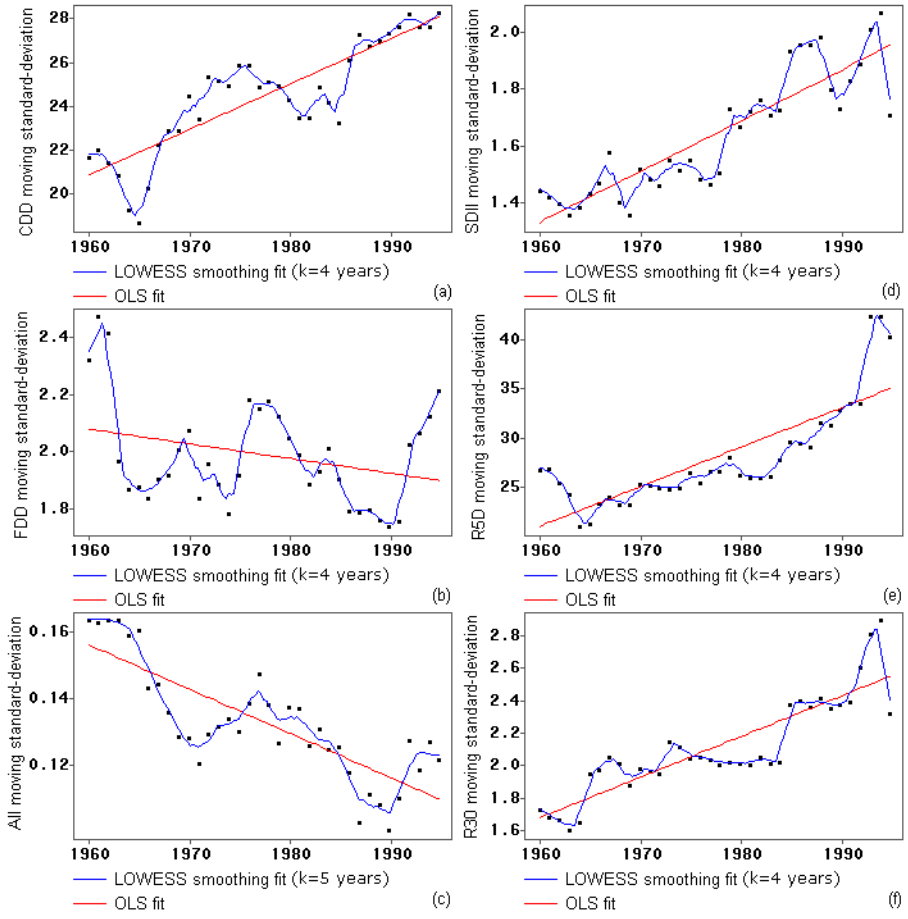

Figure 4.5 – Ordinary least squares fitting (OLS, red line) and weighted local polynomial fitting (LOWESS smoothing, blue line) for each regional-average TVE series (moving standard deviation of the extreme index using a time span of 10 years), for the period 1955/99.

Chapter 5: SPACE-TIME DYNAMICS OF PRECIPITATION EXTREMES

5. Space-time dynamics of precipitation extremes

Information on the spatial variability of extreme precipitation is important for river basins management, flood hazards protection, studies related to climate change, erosion modelling and other applications for hydrological impact modelling. It is long recognized that topography and other geographical factors are responsible for considerable spatial heterogeneity of the precipitation distribution at the sub-regional scale (e.g., Martínez-Cob, 1996; Daly, 2006). A comprehensive review on the complex relationship between precipitation, airflow and physiographic features of mountainous regions is presented by Johansson and Chen (2003), and Smith and Barstad (2004). Accordingly, it is commonly accepted that interpolation techniques that make use of the relationship between existing station data and explanatory physiographic variables (e.g., elevation or distance to the coastline) have the potential to better represent the actual climate spatial patterns, especially in mountainous areas and in regions with complex atmospheric influences (Prudhomme and Reed, 1998; Daly, 2006).

Over the past two decades, efforts have been undertaken by many authors to incorporate elevation, and other physiographic features into the spatial interpolation of rainfall fields. Some examples are multivariate geostatistics such as kriging with external drift or cokriging (Goovaerts, 2000; Nicolau, 2002; Diodato, 2005; Lloyd, 2005), techniques combining distance weighting methods and regression (Faulkner and Prudhomme, 1998; Prudhomme and Reed, 1999; Perry and Hollis, 2005), splines (Hutchinson, 1995; Boer *et al*., 2001), and local regressions (Daly *et al*., 1994; Brunsdon *et al*., 2001). Most studies do not model simultaneously the rainfall space-time patterns, but rather focus on the generation of surfaces of long-term averaged precipitation (e.g., Martínez-Cob, 1996; Nicolau, 2002; Diodato, 2005), or independently derived surfaces for yearly and monthly data (e.g., Lloyd, 2005; Perry and Hollis, 2005).

The number of studies analyzing space-time patterns of extreme precipitation indicators at the regional and local scales is very limited. The literature review of Section 2.3.3 *Indices of precipitation extremes* shows that the large majority of studies only focus on the temporal linear trends of the indices. Although desirable, a spatial analysis is sometimes not feasible

due to the sparse number of monitoring stations over large study regions (Klein Tank and Können, 2003; Moberg *et al*., 2006).

The works of Faulkner and Prudhomme (1998) and Prudhomme and Reed (1998, 1999) focused on an index of extreme rainfall, named RMED – median of the annual maximum rainfall, while Hundecha and Bárdossy (2005) and Perry and Hollis (2005) describe the mapping of several extreme precipitation indices. Faulkner and Prudhomme (1998), Prudhomme and Reed (1999), and Perry and Hollis (2005) used techniques that combine distance weighting methods and regression to produce gridded datasets of extreme precipitation indices, whereas Hundecha and Bárdossy (2005) used kriging with external drift to interpolate daily precipitation observations on a 5 km × 5 km grid and calculated the extreme precipitation indices, afterwards, on derived grids of 5, 10, 25 and 50 km^2.

Interpolation usually leads to a smoothing of the distribution inferred by the observations and thus to a loss of variance. For example, it is well known that kriging is locally accurate in the minimum error variance sense, but does not provide representations of spatial variability given the "smoothing" effect of kriging (Yamamoto, 2005). Moreover, the smoothing depends on the local data configuration, since it is minimal close to the data locations and increases as the location being estimated gets farther away from data locations. Such conditional bias is undesirable considering the modelling of floods or other extreme hydrological processes (Haberlandt, 2007). To overcome this limitation, geostatistical stochastic simulation has become a widely accepted procedure to reproduce the spatial variability and uncertainty of highly variable phenomena in geosciences (e.g., Franco *et al*., 2006; Bourennane *et al*., 2007). Geostatistical simulation methods describe local data variability based on many, equally probable, realizations of the phenomenon, consistent with the data and its statistical characteristics.

The accuracy and uncertainty of gridded data sets is difficult to assess because the field that is being estimated is unknown between data points. Spatial interpolation errors are interdependent functions of the station-network distribution, the efficacy of the interpolation procedure, and the real (but unknown) spatial distribution of the underlying climatic field (Willmott and Matsuura, 2006). Unlike traditional interpolation methods (e.g., cokriging), geostatistical simulation procedures aim at reproducing the spatial uncertainty of the attribute under study. The series of simulated maps can be post-processed and the spatial uncertainty

summarized (Goovaerts, 1997, pp. 431-436). For example, the uncertainty at an unsampled location can be evaluated through spread measures, such as the variance, derived from the corresponding local histogram.

The Algarve and Alentejo regions, in the south of continental Portugal, have large areas with high susceptibility to desertification (Correia, 2004; Rosário, 2004b). As other southern European regions, the rainfall regime is Mediterranean and so highly variable in both the spatial and temporal dimensions. Accordingly, the climate is characterised by a dry and very hot season, and a very irregular distribution of precipitation over the wet season, as well as over the years, with very intense flood peaks and with frequent drought periods. As previously discussed (Section 2.1.1 *Interactions of desertification and climate*), whenever the precipitation variability is associated with extreme phenomena, such as intensive rainfall events or drought situations, it may cause soil degradation and vegetation loss that contribute to the desertification of the most vulnerable regions (Santo *et al*., 2004).

One particularly relevant feature of the rainfall regime in southern Portugal is the occurrence of short but very intensive rainfall events that may lead to significant damages, by causing flash floods that affect small drainage basins (Ramos and Reis, 2002), and high rates of soil erosion (Pimenta, 1998; Ó and Roxo, 2001). In fact, soil degradation by water erosion is one of the major environmental problems to be faced in these regions (Loureiro and Coutinho, 1995). Fragoso and Gomes (2008) concluded that the most southern region (Algarve) is the one where episodes of heavy rainfall are most frequent and exhibits the strongest torrential character. The Alentejo area, north of Algarve, is mainly an agro-silvo-pastoral region and the most affected by desertification and drought (e.g., Roxo *et al*., 1999; Santo *et al*., 2004).

The southern part of the territory, especially the Alentejo region, is a drought prone area that is characterized by scarce precipitation, little runoff and water availability. Several studies characterized the local and regional droughts of these regions (e.g., Paulo *et al*., 2003; Paulo *et al*., 2005; Moreira *et al*., 2006) and, more generally, the drought phenomenon over Portugal (e.g., Santos, 1998; Santo *et al*., 2005). Water deficits are of great ecological and agronomic importance, especially during the dry season, and an irregular precipitation regime highly influences the productivity of rain fed agriculture.

Accordingly, the research on the space-time dynamics of extreme precipitation events, including both dry and wet situations, is an important contribution to evaluate desertification

processes and to identify areas potentially at risk from land degradation in this region, especially at the local scale (Rosário, 2004b; Pereira *et al*., 2006).

In the previous chapter, the existence of trends and other temporal patterns in six extreme precipitation indices were investigated and uncertainty about rainfall patterns evolution was assessed through regression models and smoothing techniques. Now, a subset of three indicators was selected for the characterization of the space-time dynamics of extreme precipitation in southern Portugal in the 1940/99 period: two indices describing wet conditions – R5D (Costa *et al*., 2008c) and R30 – and the proposed AII index, which characterizes dry conditions. The indices definitions and rationale were discussed in Section 4.1.2.

For exploratory purposes, the space-time patterns of the R20 index were also analyzed for the 1970/99 period, and uncertainty was assessed (Costa and Soares, 2007). This indicator is based on the count of days with precipitation above the 20 mm threshold. Additionally, the spatial patterns of the CDD index were also investigated, but this index was discarded due to the spatial inconsistencies found.

For the interpolation and uncertainty assessment of extreme precipitation, we explore the application of direct sequential cosimulation (coDSS), which allows incorporating covariates such as elevation. The choice of cosimulation follows the premises that elevation and precipitation may interact differently not only in space, but also during drier and wetter periods (Goovaerts, 2000; Costa and Soares, 2007). Furthermore, we showed in the previous chapter that there is a significant trend towards a drier climate in southern regions of Portugal, as expected from other studies focusing southern European regions (e.g., Kostopoulou and Jones, 2005; Vicente-Serrano and Cuadrat-Prats, 2007).

Accordingly, the methodology not only accounts for local data variability by using stochastic simulation procedures, but also incorporates space-time models that allow capturing long-term trends of extreme precipitation, and local correlations between elevation and precipitation through time. Elevation was used as secondary information, but other physiographic features were also investigated.

The main objectives of this research can be summarized as follows:

1. to assess the space-time relationships between physiographic features, such as elevation, and extreme precipitation in southern Portugal;
2. whenever considered relevant, to use those relationships to produce space-time scenarios for wet and dry extremes, from 1940 to 1999, on the basis of annual gridded datasets of precipitation indices;
3. to provide an uncertainty evaluation of the produced scenarios;
4. to use those scenarios to produce an additional set of maps of indicators summarizing their underlying space-time dynamics.

The extreme precipitation indices were computed using quality daily precipitation observations measured at 105 monitoring stations with data within the period 1940/99. The direct sequential cosimulation was performed for generating one map per year for the two wetness indices (R5D and R30), using 800 m × 800 m grids and elevation as exhaustive secondary information. For the dryness index (AII), direct sequential simulation was used instead, because no relevant correlations were found with physiographic features.

The produced maps are expected to be useful for the characterization of climate processes that may cause desertification, as well as for regional and local studies related to climate change, land and water resources management, hydrological modelling, and flood mitigation planning.

The study region and data are described in Section 5.1, and the methodology is introduced in Section 5.2. The main results are presented and discussed in Section 5.3. The exploratory study characterizing the space-time patterns of the R20 index (Costa and Soares, 2007) is summarized in Section 5.3.1. Section 5.3.2 describes and discusses the space-time models used, including the relationships between physiographic features and the indices. The space-time patterns of the wetness and dryness scenarios are analysed in Sections 5.3.3 and 5.3.4, respectively, and the uncertainty of the produced maps is assessed.

5.1 Study domain and data

The study domain refers to the south of continental Portugal, and is defined by the Arade, Guadiana, Mira, Ribeiras do Algarve and Sado basins. The domain includes the Algarve region, in the far South, and most of the Alentejo region (limited in the north by the Tejo

River). The study domain has different physiographic characteristics. In the far south, the relief is dominated by the two main Algarve's mountains: Monchique on the west and Caldeirão on the east. In contrast, the Alentejo region is characterized mainly by vast flat to rolling country, the penplain, where the average altitude is approximately 200 m. The São Mamede mountain ridge, the highest in the Alentejo region with an altitude of 1000 meters, lies in the farthest northeast area.

The extreme precipitation indices were computed using daily rainfall series from 105 stations, distributed irregularly over the study region (Figure 5.1), with data within the 1940/99 period (Figure 5.2). Recall from Section 4.1.1 that only the longest homogeneous period was used to build the indices for the series classified as 'suspect' through the homogenization analysis (Section 3.3.5 *Overall classification*). In order to maximize the spatial density of observations, data from stations classified in the intermediate classes 'potentially useful' and 'potentially suspect' were also used. These categories correspond to the results from the absolute testing approach, which might be considered inconclusive without station history information. Furthermore, it was not possible to find historic metadata support for the irregularities identified in stations classified as 'potentially suspect'. However, the tests applied may be very sensitive to changes in the internal time series properties, and the climatic information in the records may be relatively robust even if a negative test result has been obtained (Moberg and Jones, 2005). In this chapter, we place more emphasis on the spatial patterns and the large-scale space-time dynamics than the numeric values of trends. Moreover, the biases potentially introduced in the time series by inhomogeneous data are minimized by using different space-time models for each decade, instead of a single model for the whole analysis period. These arguments should justify the use of a dataset that might contain several inhomogeneous records.

The extreme precipitation indices analysed in this chapter are listed in Table 5.1. Their definitions and rationale were further detailed in Section 4.1.2.

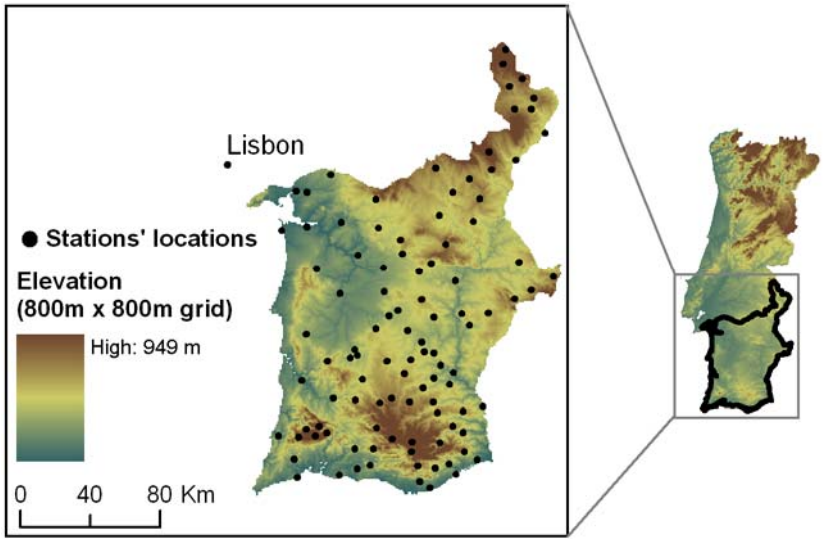

Figure 5.1 – Elevation of the study region in the south of Portugal and stations' locations

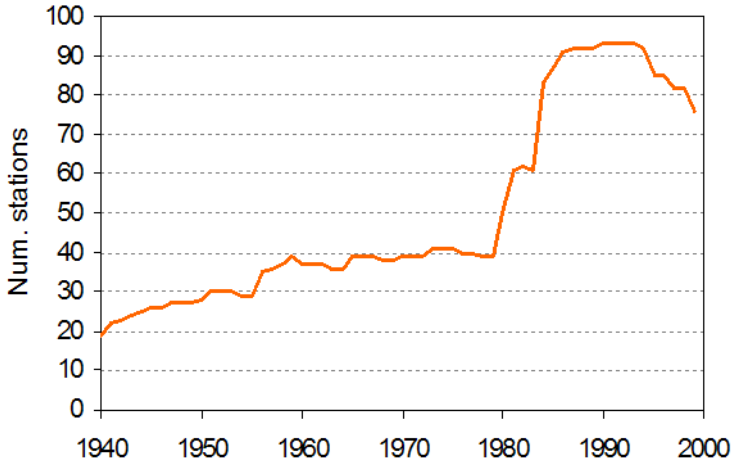

Figure 5.2 – Distribution of the number of available stations by year

Table 5.1 – Acronyms and definitions of the extreme precipitation indices

Index	Explanation	Units
R5D	Highest consecutive 5–day precipitation total	mm
R30	Number of days with daily precipitation totals above or equal to 30 mm	days
AII	Ratio between the total rain on dry days and the number of dry days (R < 10 mm)	mm

Additionally, for exploratory purposes, the space-time patterns of the R20 index were analyzed for the 1970/99 period. This indicator is based on the count of days with precipitation above the 20 mm threshold, thus its rationale is similar to the R30 index. The 20

mm threshold approximately corresponds to the 90% regional-average percentile[16] of the 1961/90 climate normal.

Nicolau (2002) investigated the contribution of several physiographic features for the prediction of precipitation fields in continental Portugal (for further details on this work see Table 2.6 in Section 2.3.2). The 125 explanatory attributes analysed by her include variables[17] related to:

- Elevation
- Slope
- Dominant orientation of the hillsides
- Counting of blockages to the advance of the air masses
- Altimetry platforms reached since the coastline
- Altimetry barriers in the neighbourhood of each cell
- Shortest distance to the coastline
- Distance to the coastline measured according to the W, NW and SW directions.

Nicolau (2002) concluded that elevation was the most important variable to explain the variability of precipitation fields when analyzed on restricted neighbourhoods, whereas the remaining attributes had low correlations with the precipitation variables.

Taking into consideration the thorough study of Nicolau (2002), elevation was considered as the most promising explanatory attribute to be used in the coDSS algorithm. Elevation data were then taken from a digital elevation model (DEM) with a grid resolution of 20 m × 20 m and resampled to an 800 m × 800 m grid mesh. The topographic variable derived is defined as the elevation of the nearest grid point to the meteorological station location, sometimes named smoothed elevation.

The study region's relief is not very complex if compared to the north of continental Portugal or other European study regions (e.g., Prudhomme and Reed, 1999; Perry and Hollis, 2005). On the other hand, Trigo and DaCamara (2000) considered ten classes of weather circulation types for Portugal and verified that the ‘rainy’ classes with an Atlantic origin (mainly W and

[16] Computed using the 15 stations series selected for trend analysis in the previous chapter.

[17] Some of them were evaluated using different radii and directions.

SW) could be associated with the observed strong decrease in precipitation from north to south. Consequently, we decided to investigate the relationship of the indices with the distance from each grid cell to the coastline measured according to the SW direction. This auxiliary attribute was determined for all 800 m × 800 m grid cells using a GIS (Figure 5.3).

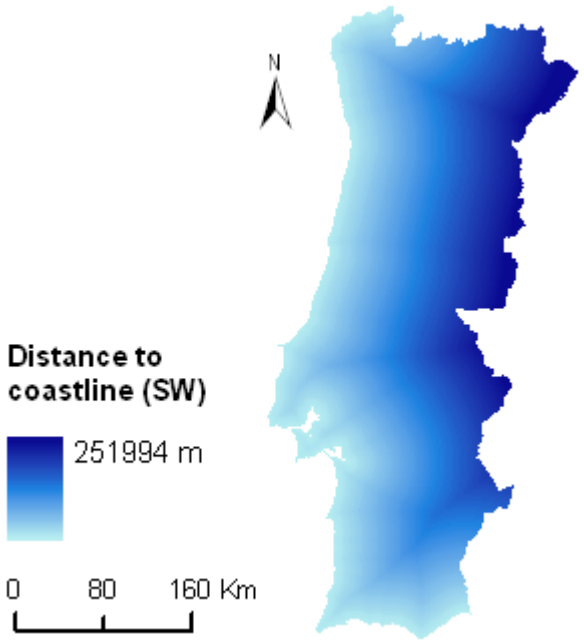

Figure 5.3 – Distance from each grid cell to the coastline according to the SW direction

5.1.1 Spatial inconsistencies of the CDD index

The space-time patterns of the CDD index were also investigated. The CDD index is defined as the maximum number of consecutive dry days per year, and a dry day corresponds to a day with less than 1 mm of precipitation. The spatial inconsistencies found in the CDD index were first revealed by a nugget effect in the decadal experimental semivariograms of the spatial component.

Variograms, which are inverse measures of the correlation for a given vector distance h, describe how the spatial continuity changes as a function of the distance and direction between any pair of points in space and time. In bounded models (e.g., spherical and exponential), variogram values increase with increasing distance of separation until they reach a maximum, named sill, at a distance known as the range. Another parameter that may be added to the variogram model is named nugget. The nugget appears on the variogram as a discontinuity at the origin. The nugget effect results from high variability at short distances

that can be caused by lack of samples or sampling inaccuracy. When a nugget effect is present, the sill is the difference between the limiting value apparently reached by the variogram, when it becomes more or less stable, and the nugget.

For example, the isotropic exponential model fitted to the spatial experimental semivariogram of the 1990s (Figure 5.4) has the following parameters: the estimated nugget is equal to 230 (43% of the variability), the estimated sill is thus equal to 300, and the estimated spatial range is equal to 90 Km. The high variability at short distances in the CDD data revealed by the nugget effect is exemplified in Figure 5.5, which shows stations with high values of the CDD index surrounded by much smaller values, and vice-versa. These are inconsistent spatial features of this index, because it is not reasonable that stations separated by short distances have a very different value for the maximum length of dry spells.

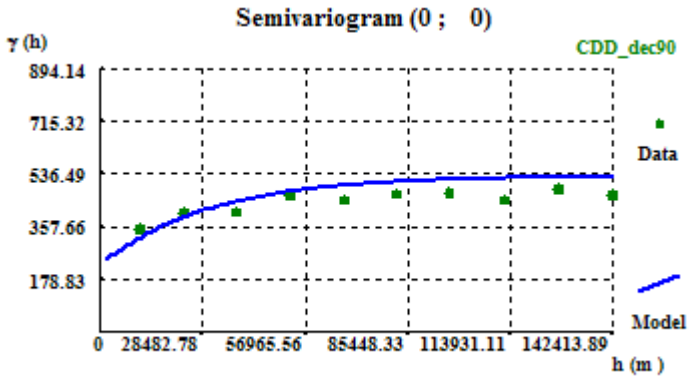

Figure 5.4 – Spatial experimental semivariogram of the 1990s data of the CDD index with the isotropic exponential model fitted

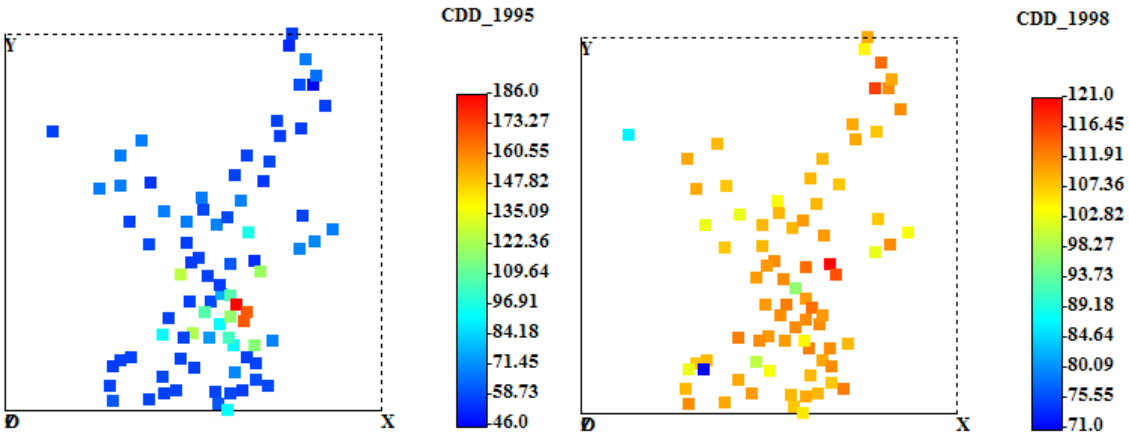

Figure 5.5 – Plot of the CDD data of the years 1995 (left) and 1998 (right)

A thorough space-time analysis of the CDD data and the daily precipitation totals was then carried out by examining the suspicious records and their neighbouring values for each year. We verified that the spatial inconsistencies were not caused by missing daily rainfall records present in the data used to build the index. The spatial inconsistencies were due to the split of a long dry spell into two dry spells because of one, or two, days with small amounts of daily rainfall in some stations.

Let d_1 and d_2 be the first and last days of a dry spell, and d_R be an in-between day with rainfall at some areas/stations of the study region ($d_1<d_R<d_2$). The following situations were found:

- If the length of the period from d_1 to d_R was approximately equal to the period from d_R to d_2, the CDD values were sometimes spatially consistent in that area, even if the precipitation of day d_R was greater than or equal to 1 mm in some stations.
- If the length of d_1–d_R was greater than the length of d_R–d_2, and the precipitation of day d_R was greater than or equal to 1 mm in many stations, three types of situations occurred:
    - At some stations, the CDD values were defined by a dry spell that occurred in another period of the year, other than d_1–d_2. This caused spatial inconsistencies for some stations in several years, while for a few stations (e.g., Lisboa) it did not.
    - Most of the stations had high CDD values defined in the d_1–d_R period, while only one, or a few, had low values defined in the d_R–d_2 period causing spatial inconsistencies.
    - Most of the stations had low CDD values defined in the d_1–d_R period, while only one, or a few, had high values defined in the d_1–d_2 period causing spatial inconsistencies.
- Conversely, if the length of d_1–d_R was smaller than the length of d_R–d_2, and the precipitation of day d_R was greater than or equal to 1 mm in many stations, similar situations have occurred.

An empirical analysis of the daily precipitation totals of days such as d_R indicated that many of them were smaller than 5 mm, a number of them ranged from 5 mm to 10 mm, and a few of them were greater than 10 mm. Therefore, in order to remove these inconsistencies, new 'CDD indices' were defined using different thresholds to characterize a dry day. Let CDD10,

CDD5 and CDD05 denote the indices defined as the maximum number of consecutive dry days per year with a dry day defined as a day with less than 10, 5 and 0.5 mm of precipitation, respectively. The space-time analysis was then repeated for these new 'CDD indices'. For illustration purposes, the parameters of the isotropic exponential models fitted to the spatial experimental semivariograms of the 1990s of these indices are summarized in Table 5.2.

Table 5.2 – Parameters of the exponential models fitted to the spatial experimental semivariograms of the 1990s of CDD indices defined using different thresholds to characterize a dry day

Index	Nugget (% of variability)	Sill	Spatial range (Km)
CDD05 (dry day: R<0.05 mm)	270 (50%)	270	90
CDD5 (dry day: R<5 mm)	190 (32%)	400	90
CDD10 (dry day: R<10 mm)	500 (54%)	430	90

The high variability at short distances remained present in the data even for indices with higher thresholds defining a dry day. Hence, the spatial inconsistencies were not totally removed. Therefore, the CDD index was discarded for further analysis.

5.2 Methodology

This section briefly introduces the kriging techniques and describes the reasoning of the geostatistical simulation algorithms implemented. Interested readers should refer to geostatistical textbooks (e.g., Isaaks and Srivastava, 1989; Goovaerts, 1997) for detailed descriptions of univariate and multivariate geostatistical interpolation methods. For a thorough description of the direct sequential simulation, and cosimulation, algorithms the reader is referred to Soares (2001).

Geostatistical estimators, known as kriging, provide statistically unbiased estimates of surface values from a set of observations at recorded locations, using the estimated spatial (and temporal) covariance model of the observed data.

Consider the two dimensional problem of estimating a primary variable z at an unsampled location u_0. Let $\{z(u_\alpha), \alpha=1, \ldots, n\}$ be the set of primary data measured at n locations u_α. Most of geostatistics is based on the assumption that the set of unknown values is a set of

spatially dependent random variables, hence each measurement $z(u_\alpha)$ is a particular realization of the random variable $Z(u_\alpha)$. Kriging uses a linear combination of neighbouring observations to estimate the unknown value at the unsampled location u_0. This problem can be expressed in terms of random variables as:

$$\hat{Z}(u_0) = \sum_{\alpha=1}^{n} \lambda_\alpha Z(u_\alpha) \tag{5.1}$$

The optimal kriging weights λ_α are determined by solving the kriging equations that result from minimizing the estimation variance while ensuring unbiased estimation of $Z(u_0)$ by $\hat{Z}(u_0)$.

Kriging methods require a stationarity assumption, expressed in two parts. First, the mean of the process is assumed constant and invariant with spatial location (first order stationarity). Second, the variance of the difference between two values is assumed to depend only on the distance h between the two points, and not on their location u (second order stationarity). Stationarity assumptions on kriging are traditionally accounted for by using local search neighbourhoods so that the dependence on stationarity becomes local (Goovaerts, 1997).

When developing the kriging equations the model of spatial covariances, or variogram (inverse function of the spatial covariances), is assumed known. This is a key function of geostatistics and characterizes the variability of the spatial (and temporal) patterns of physical phenomena. Typically, a mathematical variogram model is selected from a small set of authorised ones (e.g. exponential or spherical) and is fitted to experimental semivariogram values calculated from data for given angular and distance classes.

The experimental semivariogram $\hat{\gamma}(h)$ is computed as half the average squared difference between data pairs belonging to a certain angular and distance class:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2 \tag{5.2}$$

where $N(h)$ is the number of pairs of data locations a vector h apart.

In bounded models (e.g., spherical and exponential), variogram functions increase with distance until they reach a maximum, named *sill*, at an approximate distance known as the

range. The range is the distance *h* at which the spatial (or temporal) correlation vanishes, i.e. observations separated by a distance larger than the range are spatially (or temporally) independent observations.

5.2.1 Simple kriging

The simple kriging estimate of $z(u_0)$ is:

$$\hat{z}_{SK}(u_0) - m = \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{SK}(u_0)[z(u_\alpha) - m] \quad (5.3)$$

where m is the stationary mean of $Z(u)$, and $\lambda_\alpha^{SK}(u_0)$ is the weight assigned to datum $z(u_\alpha)$ within a search neighbourhood that comprises $n(u)$ samples.

5.2.2 Collocated cokriging

Consider now the situation where the set of primary data $\{z(u_\alpha), \alpha=1, \ldots, n\}$ is complemented by secondary data available at all estimation grid nodes and denoted by $y(u)$. The collocated cokriging estimate is (Goovaerts, 2000):

$$\hat{z}_{CoK}(u_0) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{CoK}(u_0) z(u_\alpha) + \lambda^{CoK}(u_0)[y(u_0) - m_Y + m_Z] \quad (5.4)$$

where m_Z and m_Y are the global means of the primary and secondary variables, $Z(u)$ and $Y(u)$, respectively. Note that only the secondary datum collocated at the location u_0 being estimated is retained for estimation.

5.2.3 Direct sequential simulation and cosimulation

Let $\{z(u_\alpha, t_i): \alpha=1,\ldots,n; i=1,\ldots,T\}$ be the set of climate data measured at n locations u_α and in t_i time instants (years). The n monitoring stations do not have to be all informed at the same T time instants (i.e., a number of z-values can be missing). The set of climate observations correspond to outcome values (realizations) of a spatiotemporal random variable Z(u, t) that can take a series of values at any location in space u and instant in time t according to a probability distribution.

For each instant in time t_i,, the sequence of the *direct sequential simulation* (DSS) algorithm of a continuous variable can be described as follows (Soares, 2001):

1. Randomly select the spatial location of a node $z(u_0, t_i)$ in a regular grid of nodes to be simulated;
2. Estimate local mean and variance identified with the simple kriging estimator (Equation (5.3)) and kriging variance, respectively. Sample from the global histogram a value $z^s(u_k, t_i)$ centred in the estimated local mean and variance.
3. Return to step (1) until all nodes have been visited by the random path.

Soares (2001) also extended the DSS algorithm for the joint simulation of different variables, thus named *direct sequential cosimulation* (coDSS) algorithm. Instead of simulating all variables simultaneously, it simulates each variable in turn conditioned to the previous simulated variable.

The coDSS algorithm uses collocated simple cokriging to estimate local means and variances, incorporating the secondary information and the relationship between secondary and primary variables. In this study, the collocated cokriging was applied with a Markov-type approximation (Goovaerts, 1997, pp. 237-239) for cross-continuity model. Hence, only the primary variable variogram model and a correlation model between primary and secondary data were required.

The DSS and coDSS algorithms are applied in order to obtain a set of *m* equally probable realizations of Z(u, t) at all grid nodes and all instants in time: $\{z^s(u_\alpha, t_i)$: s=1,…,m; α=1,…,N; i=1,…,T$\}$, where N is the total number of grid nodes to be simulated for each instant in time. For a given instant in time t_0, the set of simulated values $\{z^s(u_0, t_0)$: s=1,…,m$\}$ defines the local histogram at the location (grid node) u_0 for that instant.

To reproduce the spatial distribution and uncertainty of the indices characterizing wet extremes (R5D and R30), *m=100* equiprobable simulated realizations were generated through the coDSS algorithm on 800 m × 800 m grids (*N=74683*), one for each year (*T=60*), using different space-time continuity and correlation models for each decade, as briefly described in the following section and further detailed in Section 5.3.2. Elevation was used as secondary information, but the distance to the coastline measured according to the SW direction was also

investigated (Section 5.3.2.1). For the dryness index AII, direct sequential simulation was used instead, because no relevant correlations were found between this index and physiographic features.

5.2.4 Space-time models

For exploratory purposes, to reproduce the spatial distribution and uncertainty of the R20 index, 100 equiprobable simulated realizations were generated through the coDSS algorithm on 800 m × 800 m grids, one for each year. For each decade, the histogram of simulated values over the study area was allowed to range from zero (minimum observed value) to the maximum observed value of the R20 index (equal to 29, 33 and 37 days for the 1970s, 1980s and 1990s, respectively) plus 10% of each maximum. The coDSS algorithm used a different *correlation value* between the primary and secondary variable (elevation) *for each year*, and a different space-time variogram model of the R20 index for each decade. This approach assumes a constant relationship between elevation and the R20 index over the study region, but allows for changes in the relationship through time.

Simulated images were generated by the coDSS algorithm using for each decade a different space-time variogram model of the primary variable (R5D and R30 indices), and a different *correlation model* between primary and exhaustive secondary data. This strategy allows accounting for possible long-term trends or fluctuations in extreme rainfall, and for *local changes* in the relationship between secondary attributes and extreme precipitation through time.

To allow accounting for changes in correlation across the study area, the relationship between secondary attributes and extreme precipitation, described by the correlation models, was assessed locally. First, for each decade, local correlations were calculated using a search neighbourhood centred at each station's location (further details from this stage are described in Section 5.3.2.1). To reproduce the spatial distribution of the relationship between the secondary attributes and extreme precipitation, the second stage used the DSS algorithm to interpolate the local correlations. In this stage, 50 equiprobable simulated realizations of the local correlations were generated through the DSS algorithm for each decade on 800 m × 800 m grids. The correlation models used later with the coDSS algorithm were determined by

computing the mean of the distribution of the 50 simulated values at each grid node, by decade.

To reproduce the spatial distribution and uncertainty of the dryness index AII, the DSS algorithm used a different space-time variogram model of the primary variable for each decade.

5.2.5 Maps summarizing the space-time dynamics

For a given instant in time t_0, the set of m simulated values $\{z^s(u_0, t_0): s=1,\ldots,m\}$ defines the local histogram at the grid node u_0 for that instant. The space-time scenario for a given year t_0 corresponds to the average of the local histograms that were computed for all grid cells u_α:

$$z^M(u_\alpha, t_0) = \frac{1}{m}\sum_{s=1}^{m} z^s(u_\alpha, t_0) \quad , \alpha = 1,...,N .$$

Similarly, the uncertainty of the space-time scenario for a given year t_0 was evaluated by both the standard deviation and the coefficient of variation of the local histograms.

Using the space-time scenarios of the 1940/99 period (i.e., 60 annual gridded datasets for each index), summary maps of extreme precipitation were produced, namely probability, correlation and trend maps. Let $\{z^M(u_\alpha, t_i): \alpha=1,\ldots,N;\ i=1,\ldots,T\}$ be the set of $T=60$ annual gridded datasets of a given extreme precipitation index I_z.

The probability maps for the wetness indices were computed as follows. At each grid node u_α, the probability of exceeding a given value z_k was evaluated as the proportion of the T estimated values $z^M(u_\alpha, t_i)$ that exceed that threshold. In other words, the probability of I_z to be equal or greater than a fixed threshold z_k was approximated by the corresponding relative frequency computed with the sixty yearly values that were estimated at each grid cell:

$$(5.5) \qquad P[I_Z(u_\alpha) \geq z_k] \approx \frac{1}{T}\sum_{i=1}^{T} w(u_\alpha, t_i) , \ \alpha = 1,...,N$$

where $w(u_\alpha, t_i)$ are indicator data defined as

$$(5.6) \qquad w(u_\alpha, t_i) = \begin{cases} 1 & \text{if } z^M(u_\alpha, t_i) \geq z_k \\ 0 & \text{otherwise} \end{cases} \qquad , \ \alpha = 1,...,N; \ i = 1,...,T$$

Since lower values of the AII index indicate greater dryness, the probability maps for this index correspond to the estimated probability of AII to be less than a fixed threshold z_k:

$$(5.7) \qquad P[I_{All}(u_\alpha) < z_k] \approx \frac{1}{T}\sum_{i=1}^{T} w(u_\alpha, t_i) \qquad , \ \alpha = 1,...,N$$

where $w(u_\alpha, t_i)$ are indicator data defined as

$$(5.8) \qquad w(u_\alpha, t_i) = \begin{cases} 1 & \text{if } z^M(u_\alpha, t_i) < z_k \\ 0 & \text{otherwise} \end{cases} \qquad , \ \alpha = 1,...,N; \ i = 1,...,T$$

Let $\{z^{CV}(u_\alpha, t_i): \alpha=1,\ldots,N; \ i=1,\ldots,T\}$ be the set of $T=60$ annual gridded datasets of the uncertainty evaluation measured by the coefficient of variation for a given extreme precipitation index I_z. Using this set of data, probability maps of the uncertainty of the scenarios to be greater than or equal to fixed thresholds were computed by replacing $z^M(u_\alpha,t_i)$ with $z^{CV}(u_\alpha,t_i)$ and $I_z(u_\alpha)$ with $I_z^{CV}(u_\alpha)$ in equations (5.6) and (5.5), respectively. $I_z^{CV}(u_\alpha)$ denotes the coefficient of variation of the scenarios produced for the index I_z.

The map of local correlations between two extreme precipitation indices, I_Z and I_Y, was produced by computing the Pearson's correlation coefficient at each grid cell u_α using the T estimated values $z^M(u_\alpha, t_i)$ and $y^M(u_\alpha, t_i)$ of the indices I_Z and I_Y, respectively:

$$(5.9) \qquad r(u_\alpha) = \frac{\sum_{i=1}^{T} z^M(u_\alpha, t_i) y^M(u_\alpha, t_i) - T\,\bar{z}^M \bar{y}^M}{(T-1)\, s_{z^M} s_{y^M}} \qquad , \alpha = 1,...,N$$

where $z^M(u_\alpha, t_i)$ are data points measured at the grid cell u_α and time t_i; $\bar{z}^M$ and $\bar{y}^M$ correspond to the average in time of the values $z^M(u_\alpha, t_i)$ and $y^M(u_\alpha, t_i)$, respectively; and, s_{z^M} and s_{y^M} correspond to their respective standard-deviations.

Local trend maps were computed using both parametric and nonparametric estimators of the trend slope magnitude. The parametric yearly trend maps were produced by computing the

ordinary least squares (OLS) estimates of the trend slope magnitude at each grid cell u_α using the indices scenarios for 1940/99:

$$b_{OLS}(u_\alpha) = \frac{\sum_{i=1}^{T} t_i \, z^M(u_\alpha, t_i) - T\,\bar{z}^M\left(\frac{1}{T}\sum_{i=1}^{T} t_i\right)}{\sum_{i=1}^{T} t_i^{\,2} - T\left(\frac{1}{T}\sum_{i=1}^{T} t_i\right)^2} \qquad ,\alpha = 1,\ldots,N \tag{5.10}$$

The nonparametric yearly trend maps are based on the distribution free estimates of the trend slope magnitude, described by Modarres and Silva (2007), computed at each grid cell u_α using the indices scenarios for 1940/99:

$$b_{Free}(u_\alpha) = \text{Median}\left[\frac{z^M(u_\alpha, t_j) - z^M(u_\alpha, t_i)}{(t_j - t_i)}\right] \quad \forall t_i < t_j \qquad ,\alpha = 1,\ldots,N \tag{5.11}$$

where $z^M(u_\alpha, t_j)$ and $z^M(u_\alpha, t_i)$ are data points measured at years t_j and t_i, respectively.

Note that both the DSS and coDSS algorithms provide exact interpolation values at locations with observed data. Consequently, at monitoring stations' locations, the data used to compute the summary maps correspond to the observed values of the indices except for the years with missing data that were thus estimated through geostatistical stochastic simulation.

5.3 Results and discussion

The DSS and coDSS algorithms were implemented using geoMS© – Geostatistical Modelling Software[18]. The pre- and post-processing of data were performed through specific programs developed using SAS software macros and the SAS/STAT® software of the SAS System[19] for Windows, Version 8.

[18] geoMS© – Geostatistical Modelling Software was developed by Centro de Modelização de Reservatórios Petrolíferos (CMRP). Copyright CMRP-IST 2000.

[19] SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. In the USA and other countries ® indicates USA registration.

5.3.1 Exploratory study: R20 index

For the interpolation and uncertainty assessment of the R20 index in the southern region of continental Portugal, for the period 1970/99, we explore the application of direct sequential cosimulation (coDSS), which allows incorporating covariates such as altitude. The R20 index characterizes the frequency of heavy precipitation events. The coDSS algorithm was implemented on 800 m × 800 m grids, one for each year, using elevation as exhaustive secondary information.

The results from this exploratory study are described by Costa and Soares (2007). The relationship between the R20 index and elevation is discussed in Section 5.3.1.1, and the space-time continuity of extreme precipitation is described in Section 5.3.1.2. Finally, the space-time patterns of the extreme precipitation index are analyzed (Section 5.3.1.3), and uncertainty is assessed (Section 5.3.1.4).

5.3.1.1 Relationship between extreme precipitation and elevation

It is recognized that topography and other geographical factors are responsible for considerable spatial heterogeneity of the precipitation distribution at the sub-regional scale. Goovaerts (2000) incorporated a DEM into the spatial interpolation of annual and monthly rainfall observations, averaged over the period of January 1970 to March 1995, and measured in the most southern region of continental Portugal (Algarve). Precipitation generally increases with elevation because of the orographic effect of mountainous terrain. The relief of Algarve, which is within our study region, is dominated by two mountainous regions. Therefore, has expected in that work, the linear correlation coefficient between annual rainfall and elevation was equal to 0.79. Moreover, Goovaerts (2000) found that the correlation coefficient was wicker in dry months (equal to 0.33 in August).

For the implementation of the coDSS procedure, a preliminary analysis revealed that the regional correlation (Pearson's correlation coefficient averaged over the study region) between the extreme precipitation index and elevation ranges from 0.24 to 0.62 within the period 1970/99 (Figure 5.6). The terrain north of the Algarve region does not have high elevations, which might explain those weak to moderate correlations.

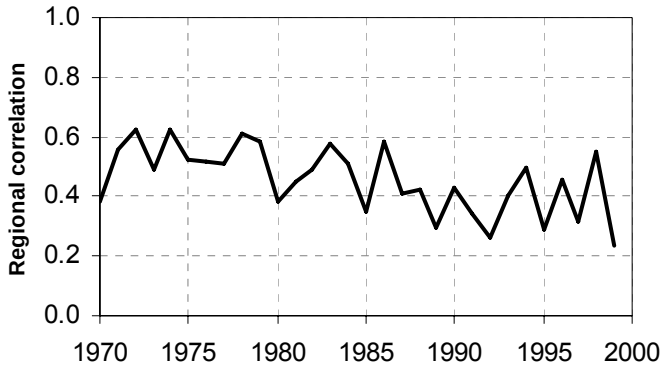

Figure 5.6 – Regional correlation between the R20 index and elevation, by year

Figure 5.6 shows that the relationship between elevation and the R20 index has decreased along the study period. Accordingly, the coDSS algorithm incorporated a different correlation value between the primary and secondary variable for each year.

5.3.1.2 Space-time continuity of extreme precipitation

Variograms were calculated according to Equation (5.2). Experimental space-time semi-variograms were calculated for three decades (1970-79, 1980-89, and 1990-99), and spherical models fitted. The spatial dimension was modeled using an isotropic variogram. The parameters of each variogram model fitted, and the decadal average of the regional correlation between the R20 index and elevation, are summarized in Table 5.3.

Table 5.3 – Average of the regional correlation between the R20 index and elevation, and parameters of the space-time variograms, by decade

Decade	**Regional correlation**	**Parameters of the spherical variogram models**		
		Spatial range (m)	*Temporal range* (years)	*Sill*
1970-79	0.54	46000	4.5	25.09
1980-89	0.45	160000	3	32.71
1990-99	0.38	150000	3.3	26.02

There are no apparent tendencies concerning the temporal component of the semi-variograms. However, the relationship between elevation and the R20 index has decreased along the study period (Figure 5.6), whereas there is evidence of an increase of the spatial continuity of the frequency of extreme precipitation events in the last decades. This evidence is consistent with

the results of Costa *et al*. (2008b) for other extreme precipitation indices computed for the same region within the 1940/99 period.

5.3.1.3 Space-time scenarios

Using the coDSS technique, a set of 100 equiprobable simulated realizations of the extreme precipitation index was computed at each simulated grid node, by year. The space-time inference was performed by means of computing the mean and the median of those distributions. For illustration purposes, the results of 1974 and 1999 are presented since they correspond to the years with the maximum (0.62) and minimum (0.24) regional correlation coefficients between the R20 index and elevation, respectively.

The maps of the median (Figure 5.8 and Figure 5.10) are less smoothed than the maps of the mean (Figure 5.7 and Figure 5.9). Within the study period, the higher values of the R20 index are mostly located in the two mountainous regions of Algarve, though not very evident in the maps of 1974 (Figure 5.7 and Figure 5.8).

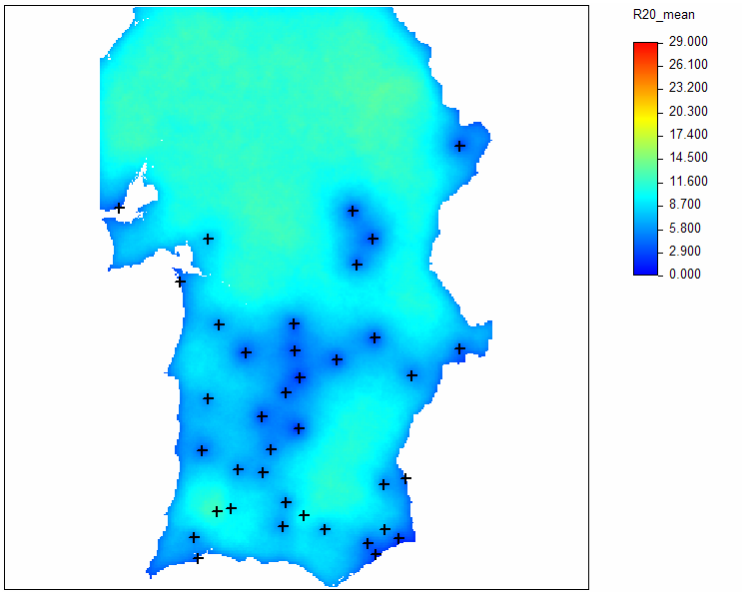

Figure 5.7 – Mean of the distribution of the 100 simulated values of the R20 index for the year 1974 (in days)

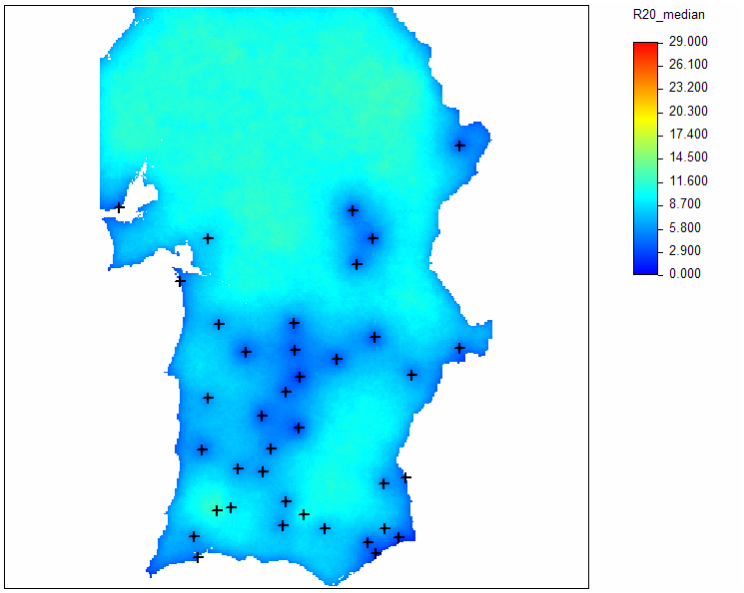

Figure 5.8 – Median of the distribution of the 100 simulated values of the R20 index for the year 1974 (in days)

The spatial continuity of the frequency of extreme precipitation events seems to be lower in the northern part of the maps because of the lack of monitoring station's data in that area. However, it is important to call attention to the fact that most of that area is out of the study region (Figure 5.1).

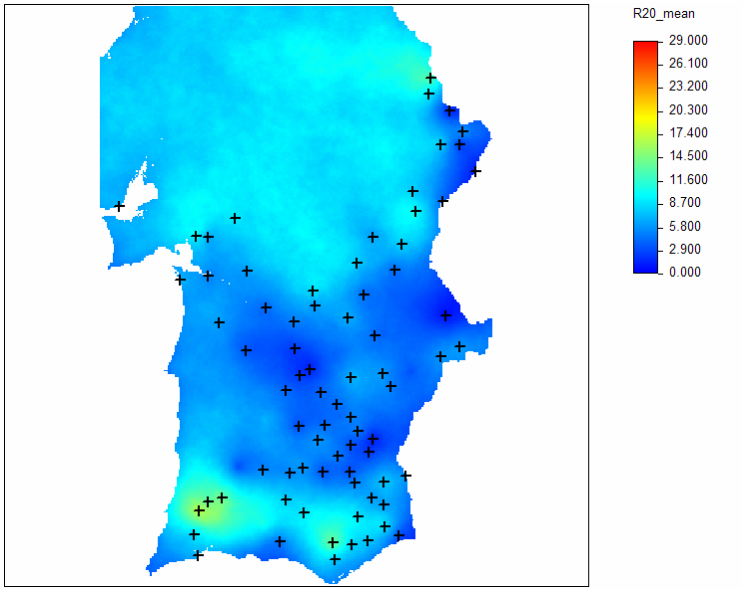

Figure 5.9 – Mean of the distribution of the 100 simulated values of the R20 index for the year 1999 (in days)

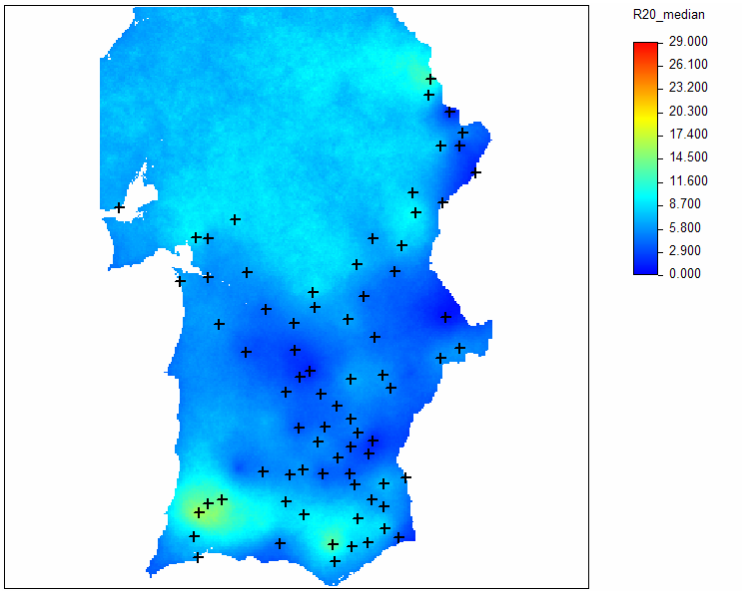

Figure 5.10 – Median of the distribution of the 100 simulated values of the R20 index for the year 1999 (in days)

The contribution of elevation to the prediction of the R20 index values is difficult to assess, especially by just observing the maps of 1974 and 1999. The results of 1974 correspond to the year with the maximum regional correlation coefficient between the R20 index and elevation, thus the mountainous regions of the south were expected to exhibit a more pronounced pattern of heavy precipitation events. However, 1974 was the driest year of the 1970s (Costa *et al*., 2008b), which might explain that difficulty. Furthermore, recall that local variations in the relationship between the frequency of extreme precipitation events and elevation were not accounted for in this exploratory study.

When the maps from the entire study period are analyzed, it is much more evident that elevation contributes more to the prediction of the R20 index when the correlation coefficient is higher, as expected. However, the influence of dry and wet periods in the relationship between elevation and the index is not clear.

5.3.1.4 Uncertainty evaluation

Uncertainty was assessed by means of computing the standard deviation (Figure 5.11 and Figure 5.14), the coefficient of variation (Figure 5.12 and Figure 5.15) and the inter-quartile range (Figure 5.13 and Figure 5.16) of the distribution of the 100 simulated values of the R20 index at each simulated grid node, by year.

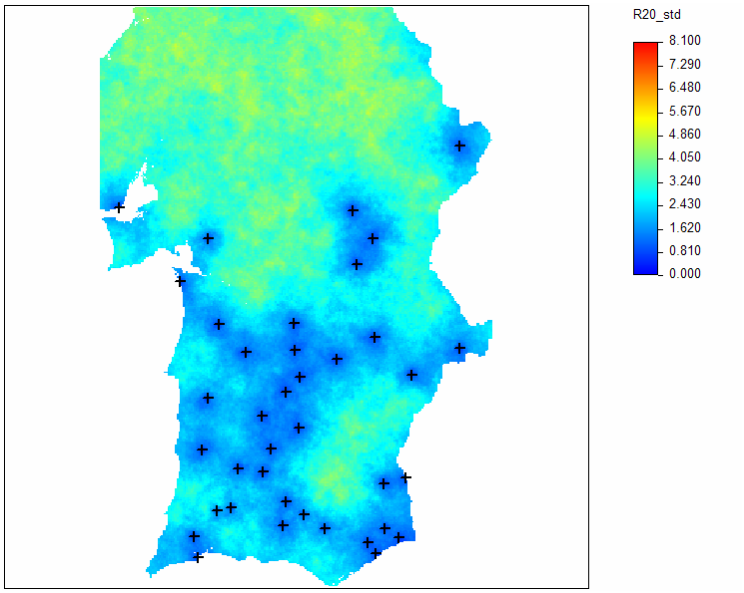

Figure 5.11 – Standard-deviation of the distribution of the 100 simulated values of the R20 index for the year 1974 (in days)

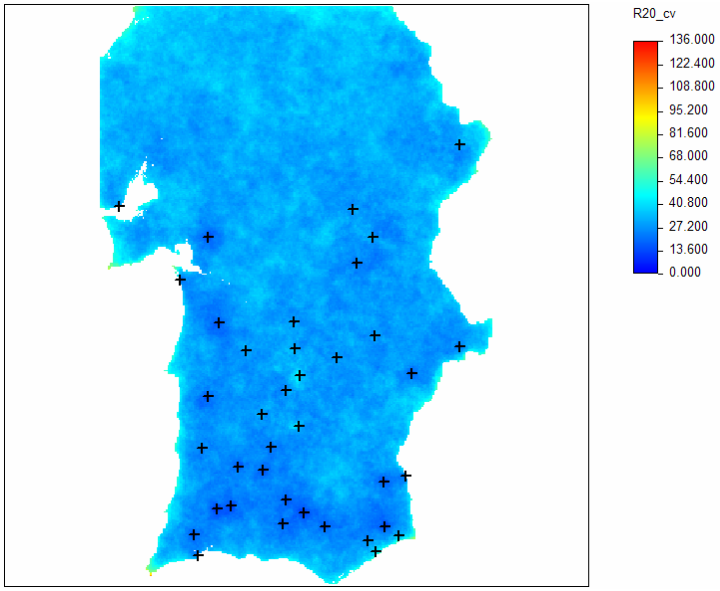

Figure 5.12 – Coefficient of variation of the distribution of the 100 simulated values of the R20 index for the year 1974 (in %)

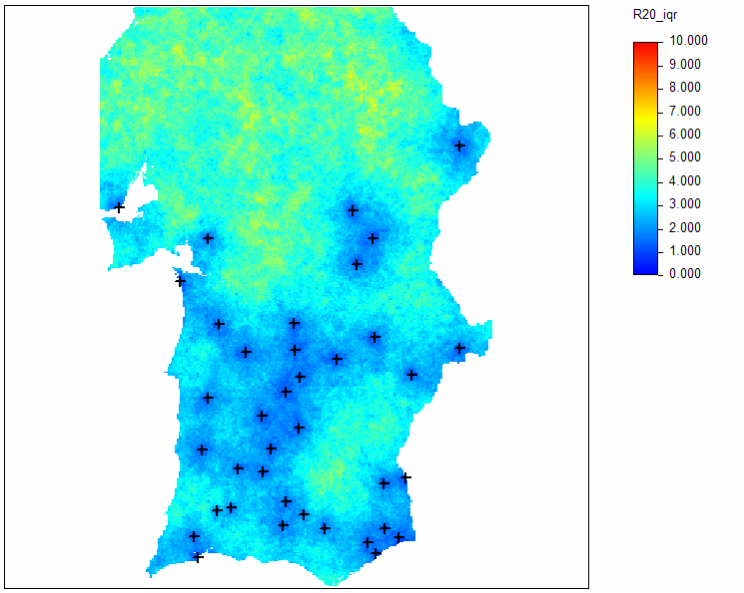

Figure 5.13 – Inter-quartile range of the distribution of the 100 simulated values of the R20 index for the year 1974 (in days)

As expected, the region where the distribution of the R20 index has greater variability, thus more uncertainty, is in the northern part of the maps, corresponding to a region less densely sampled. This is especially evident in the maps of the standard-deviation and inter-quartile range.

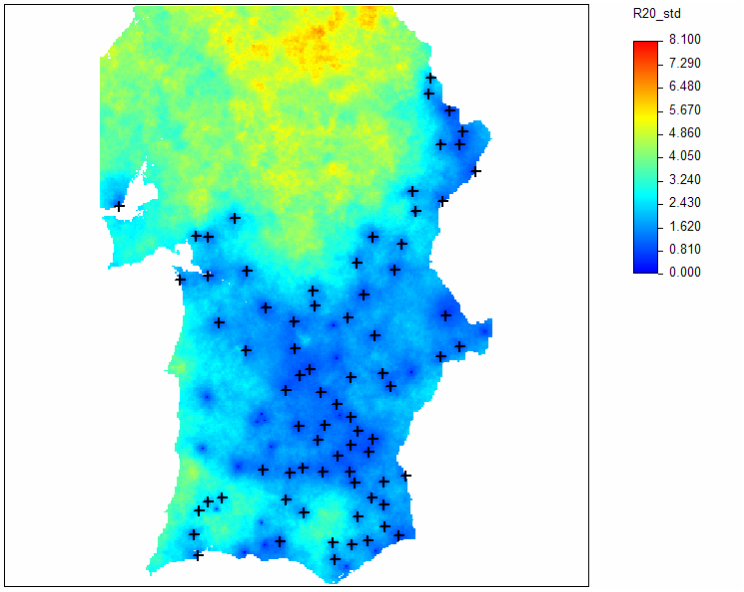

Figure 5.14 – Standard-deviation of the distribution of the 100 simulated values of the R20 index for the year 1999 (in days)

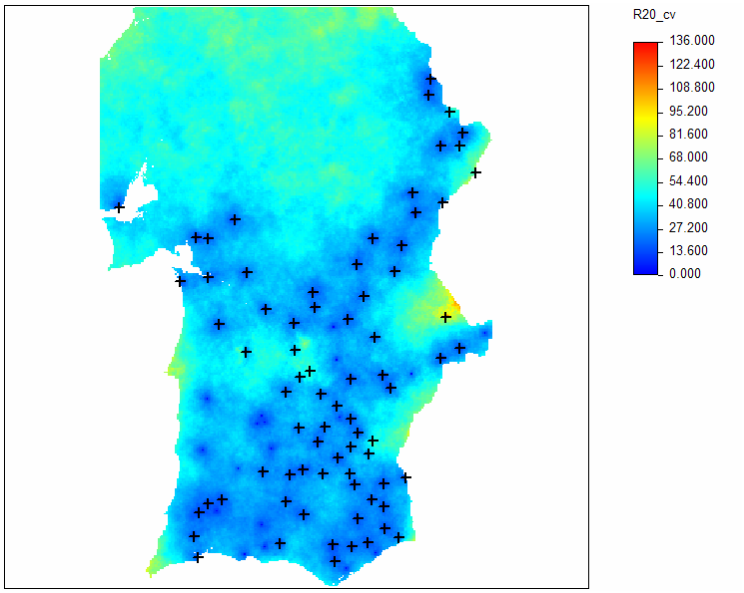

Figure 5.15 – Coefficient of variation of the distribution of the 100 simulated values of the R20 index for the year 1999 (in %)

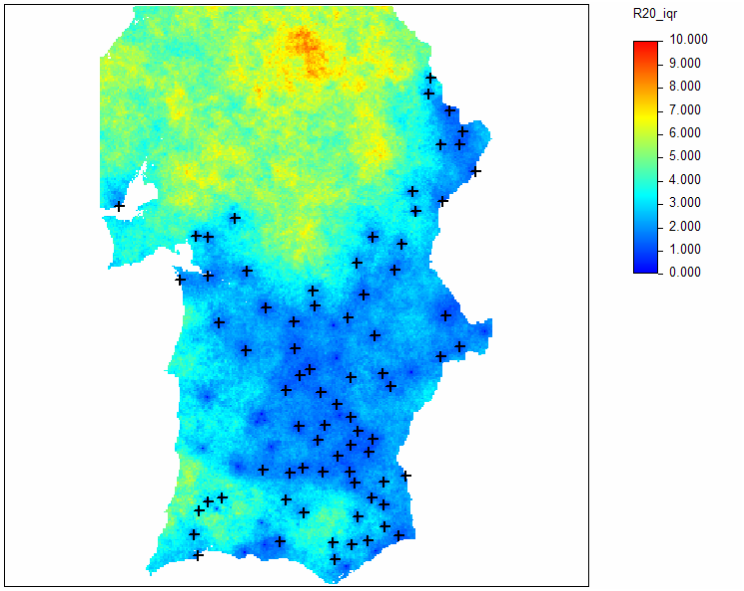

Figure 5.16 – Inter-quartile range of the distribution of the 100 simulated values of the R20 index for the year 1999 (in days)

The coefficient of variation provides a relative measure of the dispersion of variable's values regardless of their units. Hence, the maps of the coefficient of variation are particularly useful because they allow to: compare the uncertainty of the R20 index in different years; compare the uncertainty of R20 with other indices, whenever the same inference methodology is used; and compare the uncertainty of R20 using different inference methodologies.

5.3.2 Space-time models

5.3.2.1 Local correlation models

An exploratory evaluation of the relationships between the extreme precipitation indices and physiographic features (elevation and distance to the coastline) was made by averaging each index at each station, and by computing afterwards the Pearson's correlation coefficient between those values and secondary data (Table 5.4). Elevation was measured by the actual station's altitude and by the station's grid point elevation (smoothed elevation). This regional analysis shows that, in general, the correlation is slightly stronger for the smoothed elevation than the actual one. Other studies on this subject concluded likewise (e.g., Nicolau, 2002; Diodato, 2005).

The relationships between the dryness index AII and the secondary data are very weak, thus the secondary attributes are less useful to improve this index interpolation. Therefore, the computational effort required to incorporate exhaustive secondary information in the coDSS algorithm does not payoff, and DSS was used instead.

Table 5.4 – Linear correlation between the indices and physiographic features

Index	Stations' altitude	Smoothed elevation	Distance to the coastline according to the SW direction
R5D	0.43	0.49	−0.34
R30	0.45	0.52	−0.33
AII	0.10	0.07	−0.01

Although the regional relationships between the wetness indices (R5D and R30) with the distance to the coastline were weak, local correlation models were developed using the DSS algorithm (Appendix V). The local correlation models show extensive patterns of negative correlations and very weak to moderate relationships between the wetness indices with the distance to the coastline. The higher positive correlations occur in mountainous areas of the south. These results confirm that using the distance to the coastline, as an explanatory variable, may be less useful for the interpolation of the wetness indices than using elevation. Figure 5.17 shows regional correlations between the wetness indices with elevation and distance to the coastline by decade. For each index, the relationships evolution is identical along decades.

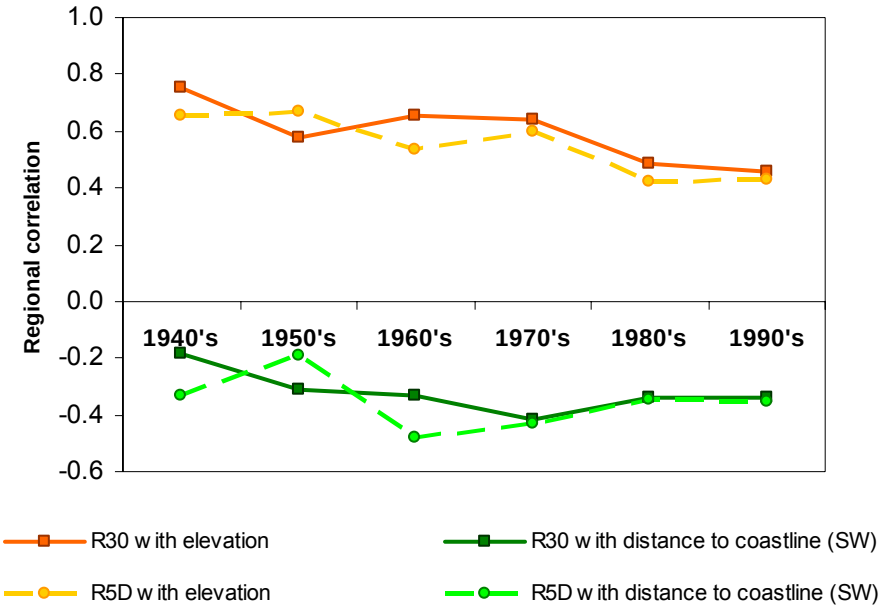

Figure 5.17 – Regional correlations between the wetness indices with elevation and distance to the coastline, by decade

The relationships of the wetness indices (R5D and R30) with smoothed elevation are regionally moderate (Table 5.4) but stronger than with other attributes, as expected from the analysis performed by Nicolau (2002). Hence, these relationships were analysed further and used to produce the gridded datasets of the wetness indices.

Plots of elevation against R5D and R30 values are given in Figure 5.18 and Figure 5.19, respectively. The coefficient of determination, r^2, is small in both cases and so there is little evidence of a (global) linear relationship between elevation and extreme precipitation, as expected (Lloyd, 2005).

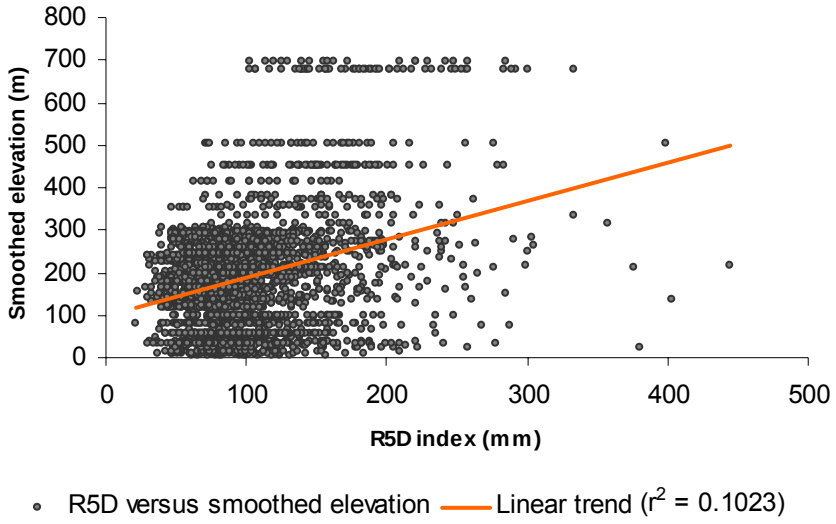

Figure 5.18 – Plot of elevation against R5D values calculated within 1940/99

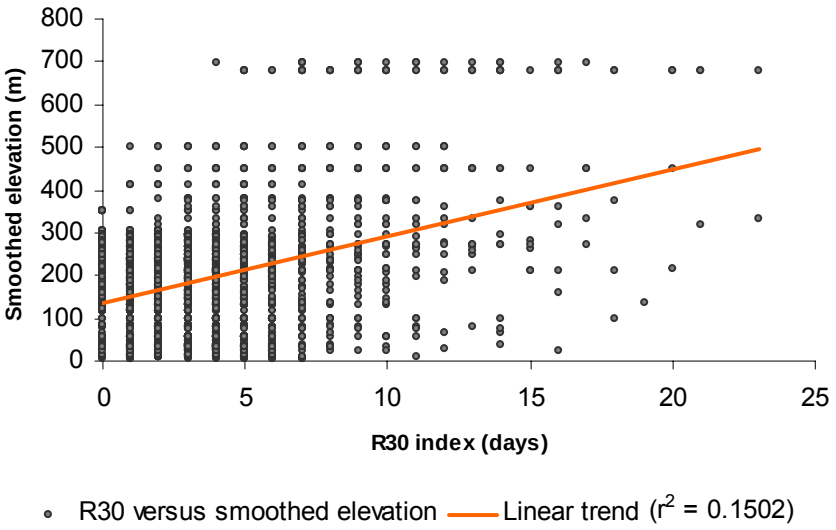

Figure 5.19 – Plot of elevation against R30 values calculated within 1940/99

Moreover, the correlations for elevation against the indices values is not constant through time (Figure 5.20), but rather shows negative trends during the study period, although not statistically significant, as expected from the R20 index analysis (Section 5.3.1). The number of stations used in the computation of these correlations ranges from 19 (in 1940) to 93 (in 1991 and 1992), and is always less than 40 before 1980. Because of the sparse coverage of meteorological stations in some areas, especially until the 1980s (Table 5.5), the local relationships between elevation and extreme precipitation were assessed by decade.

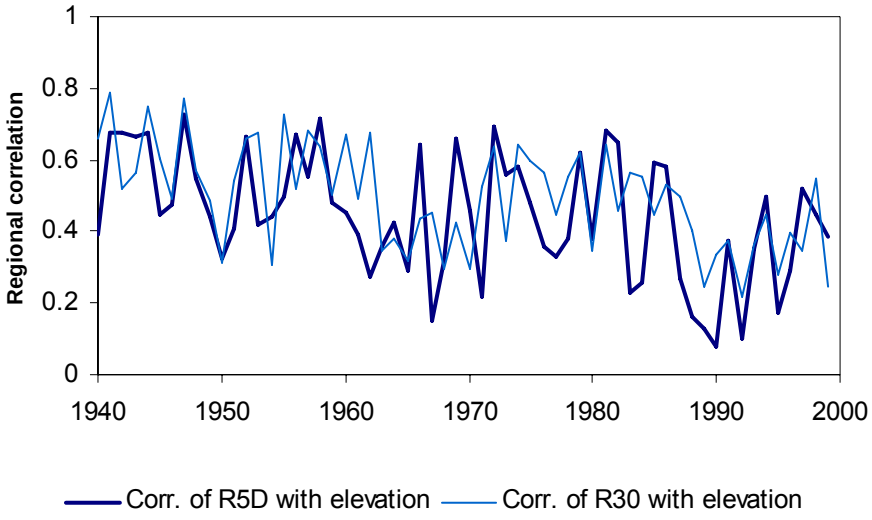

Figure 5.20 – Regional correlations between the wetness indices and elevation, by year

Table 5.5 – Distribution of weather stations (with records within each decade) by elevation classes, and radii of the search neighbourhoods used to calculate the local correlations

Decade	Elevation (m)					Radius (m)
	< 140	140 – 280	281 – 420	421 – 560	561 – 700	
1940/49	9	13	2	2	1	65000
1950/59	16	17	3	2	1	50000
1960/69	16	20	2	2	1	50000
1970/79	17	18	2	2	1	40000
1980/89	31	43	17	1	2	35000
1990/99	31	44	16	1	2	35000

The coDSS algorithm uses a different correlation model between the wetness indices and elevation within each decade. In order to determine these models, first, the relationship between elevation and precipitation was assessed locally by computing, for each decade, Pearson's correlation coefficients using stations' data falling within a circle centred at each station's location. As in earlier decades meteorological stations are scarce, larger radii were used (Table 5.5).

Afterwards, the DSS algorithm was applied to interpolate the local correlations by decade, and 50 simulated maps of local correlations were obtained for each index. These procedures

used space-time spherical variograms of the local correlations, where the spatial dimension was modelled as isotropic. For the R5D index, the estimated range of the spatial dimension was 110000 meters, the range of the temporal dimension was 6 decades, and the estimated sill was 0.053. For the R30 index, the estimated range of the spatial dimension was 130000 meters, the range of the temporal dimension was 6 decades, and the estimated sill was 0.077. Finally, the correlation models were determined by computing the mean of the distribution of 50 simulated values at each grid node, by decade (Figure 5.21 and Figure 5.22).

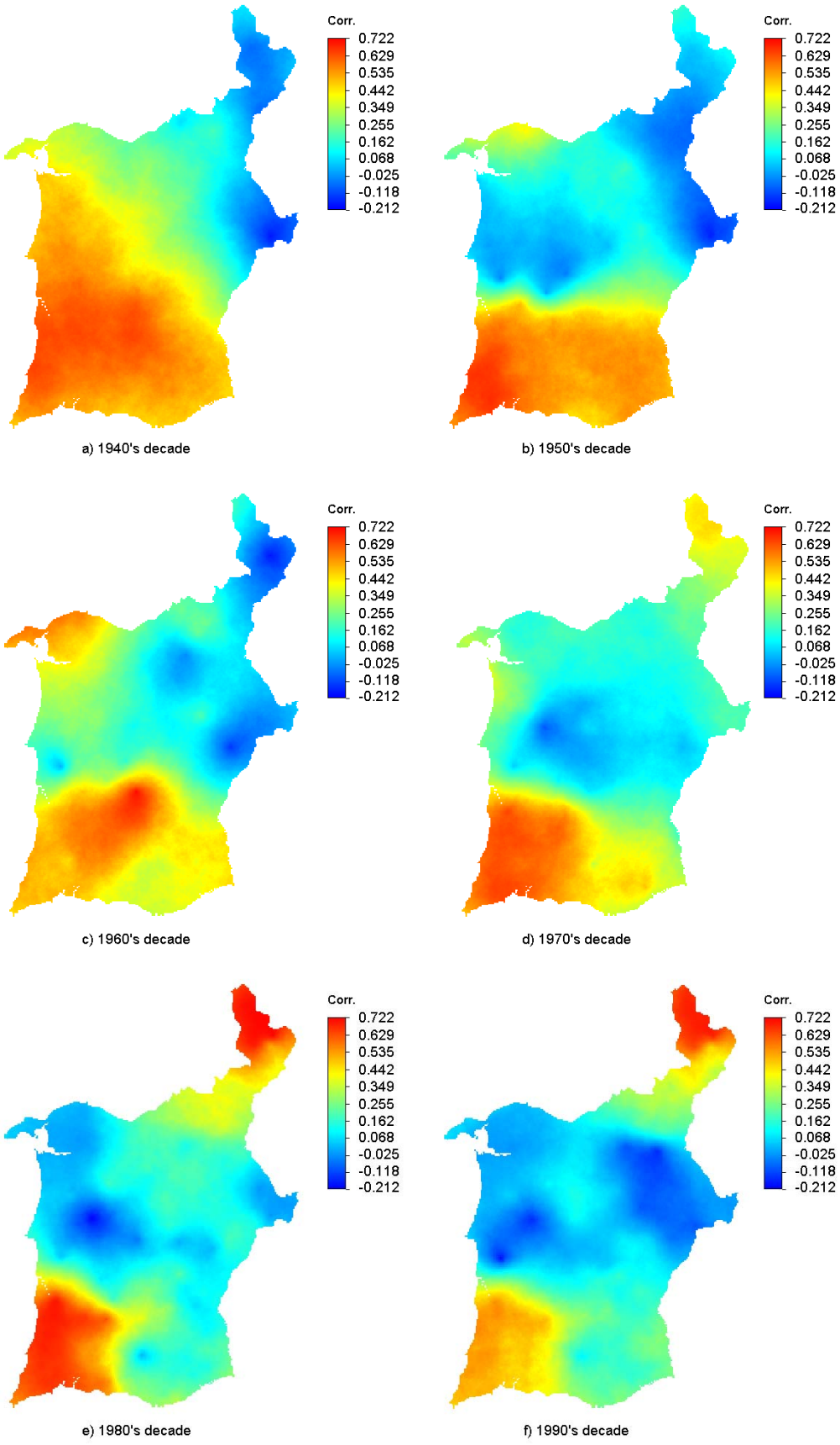

Figure 5.21 – Local correlation models between elevation and R5D values for each decade

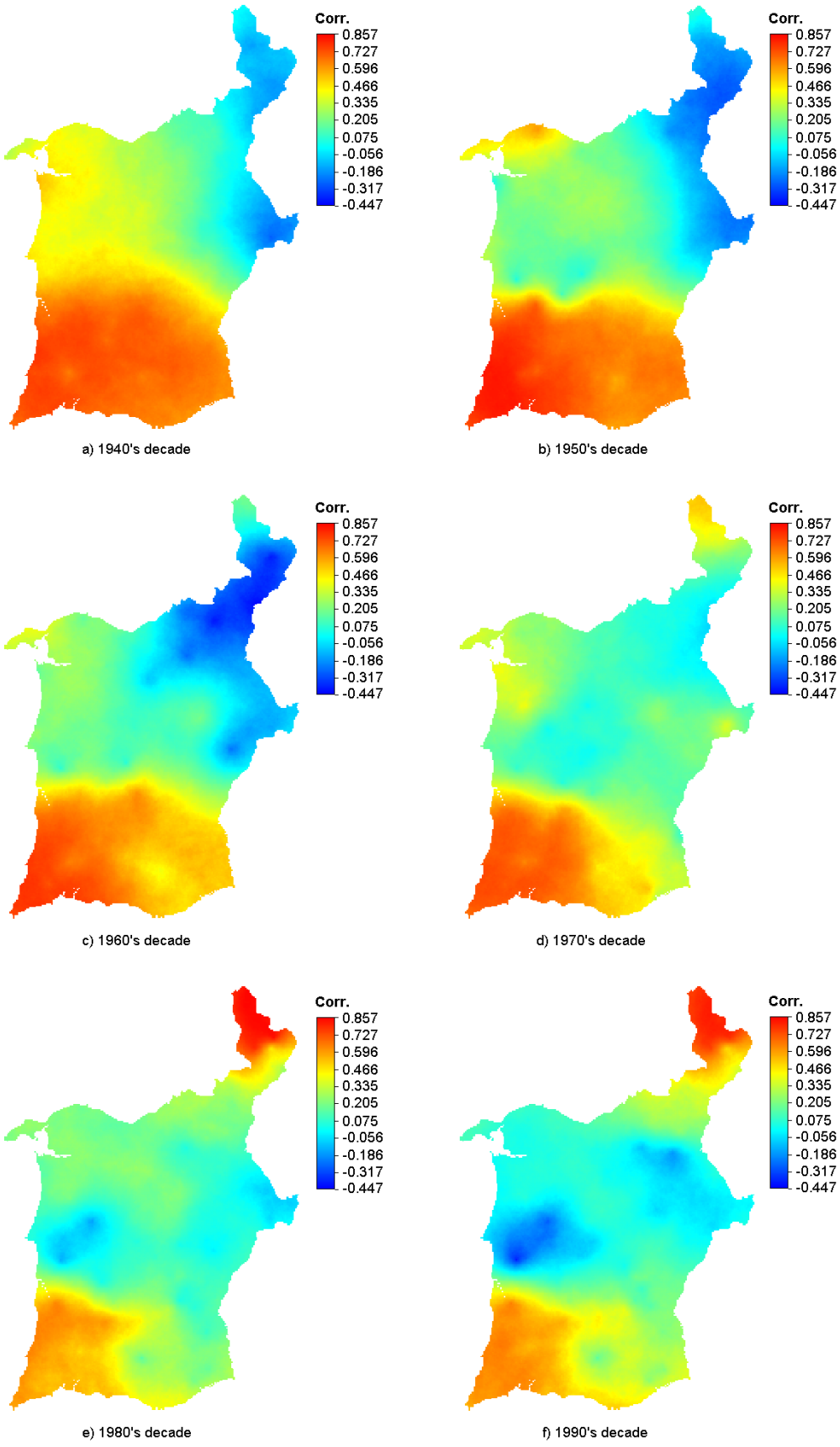

Figure 5.22 – Local correlation models between elevation and R30 values for each decade

The estimated correlations between elevation and the R5D index range from very weak (−0.21) to moderately strong (0.72) across the region and along decades, while for the R30 index the space-time variability is greater since correlations range from −0.45 to 0.86. In the 1940s, the R5D correlations range from −0.18 to 0.65 across the study region, whereas for R30 they range from −0.28 to 0.77. The correlation models of this decade show an unrealistic pattern of correlations on the west of the region, north of Algarve, caused by the scarce number of available stations. Accordingly, the variability associated with estimated correlations over this area is also high. Nevertheless, they correspond to moderate values of approximately 0.50, and so the contribution of elevation to the estimation of the indices will also be moderate. The correlation models of the 1950s exhibit a more realistic pattern, with correlations ranging from −0.17 to 0.67 for R5D, and from −0.31 to 0.83 for R30. The correlation models of the 1960s also show a realistic pattern of correlations, although the west and northeast areas present high variability. Similarly, the 1970s models are realistic but with very high variability within the northeast region. Because of the higher density of stations, the correlation models of the 1980s and 1990s exhibit much less variability than the previous ones.

These results suggest that using elevation as a secondary variable in estimation will increase the accuracy of estimates in some locations, i.e. those mountainous areas where correlations are large (e.g. on the west of Algarve and on the northeast of the study region). In contrast, in places where correlations are small univariate interpolators (e.g., DSS) are likely to provide estimates as accurate as the ones provided by coDSS, but with less computational effort.

5.3.2.2 Space-time continuity of extreme precipitation

Variograms were calculated according to Equation (5.2). In this study, we chose exponential models that capture the major spatial features of the attributes under study within each decade. The spatial variability is assumed identical in all directions (i.e. isotropic) within each decade. The parameters for each exponential variogram of the R5D and R30 indices, used in the coDSS algorithm, are summarized in Table 5.6 and Table 5.7, respectively. For the AII index, the parameters for each exponential variogram used in the DSS algorithm are summarized in Table 5.8.

In what concerns the temporal component, there are no relevant tendencies. However, the range of the models' spatial component shows a strong increase in the spatial continuity of

both the flood indicator (R5D) and the frequency of extremely heavy precipitation events (R30) on the last two decades. These findings are consistent with the preliminary results of the exploratory study for the R20 index (Table 5.3 of Section 5.3.1.2), and of Costa *et al*. (2008b).

Table 5.6 – Parameters of the space-time exponential variograms for the R5D index, by decade (Durão *et al*., 2007)

Decade	Spatial range (m)	Temporal range (years)	Sill
1940-49	85000	4.5	2923.364
1950-59	100000	1.0	2075.247
1960-69	70000	4.0	1263.250
1970-79	70000	5.0	1543.205
1980-89	150000	4.5	2075.301
1990-99	165000	1.3	2803.646

Table 5.7 – Parameters of the space-time exponential variograms for the R30 index, by decade (Durão *et al*., 2007)

Decade	Spatial range (m)	Temporal range (years)	Sill
1940-49	40000	2.5	13.314
1950-59	50000	1.3	8.561
1960-69	65000	1.5	9.981
1970-79	100000	2.5	9.510
1980-89	145000	5.0	13.089
1990-99	160000	4.5	8.984

Table 5.8 – Parameters of the space-time exponential variograms for the AII index, by decade

Decade	Spatial range (m)	Temporal range (years)	Sill
1940-49	90000	3.5	0.043
1950-59	70000	3.5	0.020
1960-69	80000	2	0.031
1970-79	50000	3	0.027
1980-89	70000	5	0.020
1990-99	70000	2	0.022

The analysis of the fitted variogram models shows that their spatial ranges, which express the extent of spatial continuity of the phenomena, are generally increasing over the decades for the wetness indices (Figure 5.23). These results imply that the spatial patterns of extreme precipitation are becoming more homogenous over time, while the levels of local variability are decreasing.

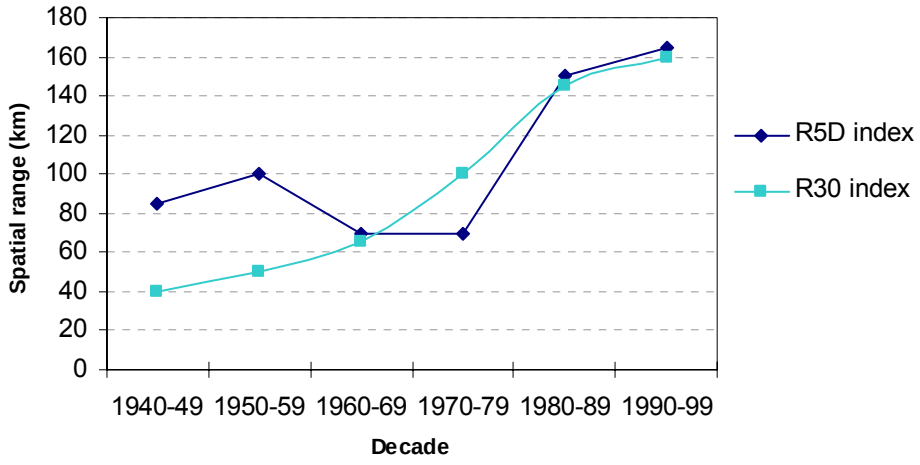

Figure 5.23 – Spatial range of the variogram models fitted for each index

The North Atlantic Oscillation (NAO) is one of the major large-scale atmospheric phenomena influencing the climate of Europe and the North Atlantic. In simple terms, the NAO corresponds to a large-scale meridional oscillation of atmospheric mass between the subtropical anticyclone near the Azores and the subpolar low pressure system near Iceland (Trigo *et al*., 2002). In its positive phase, the NAO corresponds to enhanced westerly flow over the North Atlantic and a northward shift of the mid latitude storm track (Scaife *et al*., 2008). For positive NAO winter months, Central Europe and the Iberian Peninsula experience anomalously anticyclonic circulation and reduced precipitation, which are associated with reduced cloud cover (Trigo *et al*., 2002).

Throughout the last two decades of the twentieth century, the northern centre of the NAO dipole (the Icelandic low) has moved closer to Scandinavia, consequently affecting the precipitation field over the Iberian Peninsula (Goodess and Jones, 2002; Trigo *et al*., 2004). The accumulation of positive modes of the NAO in winter during the last few decades is well documented, being evident since the 1970s and strengthened during the 1980s and 1990s (Trigo *et al*., 2002; Dünkeloh and Jacobeit, 2003; Scaife *et al*., 2008).

Scaife *et al*. (2008) determined the impact of changes in the NAO, such as that observed between the 1960s and 1990s, on the frequency of extreme weather events over Europe. Their conclusions were based on a set of experiments with a general circulation model and on observational data, which were used to compute a set of indices for temperature and precipitation[20] extremes. Scaife *et al*. (2008) verified that changes in the NAO are likely to be responsible for much of the observed change in the frequency of above 90th percentile winter precipitation between the 1960s and 1990s over Europe. Moreover, the results from Goodess and Jones (2002) and Trigo *et al*. (2004) indicate that NAO–rainfall relationships tend to be stronger during the wet seasons of the last decades of the twentieth century in southern Portugal.

Trigo *et al*. (2004) have assessed the impact of the NAO[21] on the mean precipitation and river flow regimes for the three main international Iberian river basins (Douro in the north, Tejo in the centre, and Guadiana in the south). Their study focused on the wet season (Oct. – Apr.) and concluded that the impact of the NAO on precipitation is irregular, presenting a high inter-decadal variability. Generally, the correlation values increase from north to south over the whole 1923/98 period, thus being higher for the Guadiana basin than for the Tejo and Douro basins. Moreover, the most recent sub-period considered (1973/98) shows the highest correlation values (all of them significant at the 5% level) between the NAO index and average precipitation for November to February in the Guadiana basin.

These results are consistent with those from a previous study addressing the large-scale influence of the NAO[22] in the characteristics of Iberian rainfall (Goodess and Jones, 2002). The NAO–rainfall relationships are stronger over central Spain, and in southern Portugal during winter, and weaker along the northern and eastern coasts. Moreover, in winter, the relationships tend to be stronger during the most recent sub-period considered (1978/97).

These changes in the NAO are likely to be responsible for the observed change of spatial continuity of extreme precipitation in southern Portugal, which is especially pronounced during the last two decades of the twentieth century (Figure 5.23).

---

[20] The index used is based on the number of days above the 90th percentile.

[21] This formulation of the NAO index uses Gibraltar as the southern station.

[22] Goodess and Jones (2002) used the NCEP Gibraltar minus Reykjavik index.

5.3.3 Space-time scenarios for wet extremes

Using the coDSS algorithm, a set of 100 equiprobable simulated realizations of the R5D and R30 indices was computed at each simulated grid node, by year. For illustration purposes, two equiprobable realizations of R5D, for 1945 and 1949, are shown in Figure 5.24. The space-time inference was performed by means of computing the mean of those distributions, hereafter referred to as mean-maps or (space-time) scenarios. Uncertainty was assessed by means of computing the standard deviation (STD) and the coefficient of variation (CV) of the distribution of the 100 simulated values at each simulated grid node, by year.

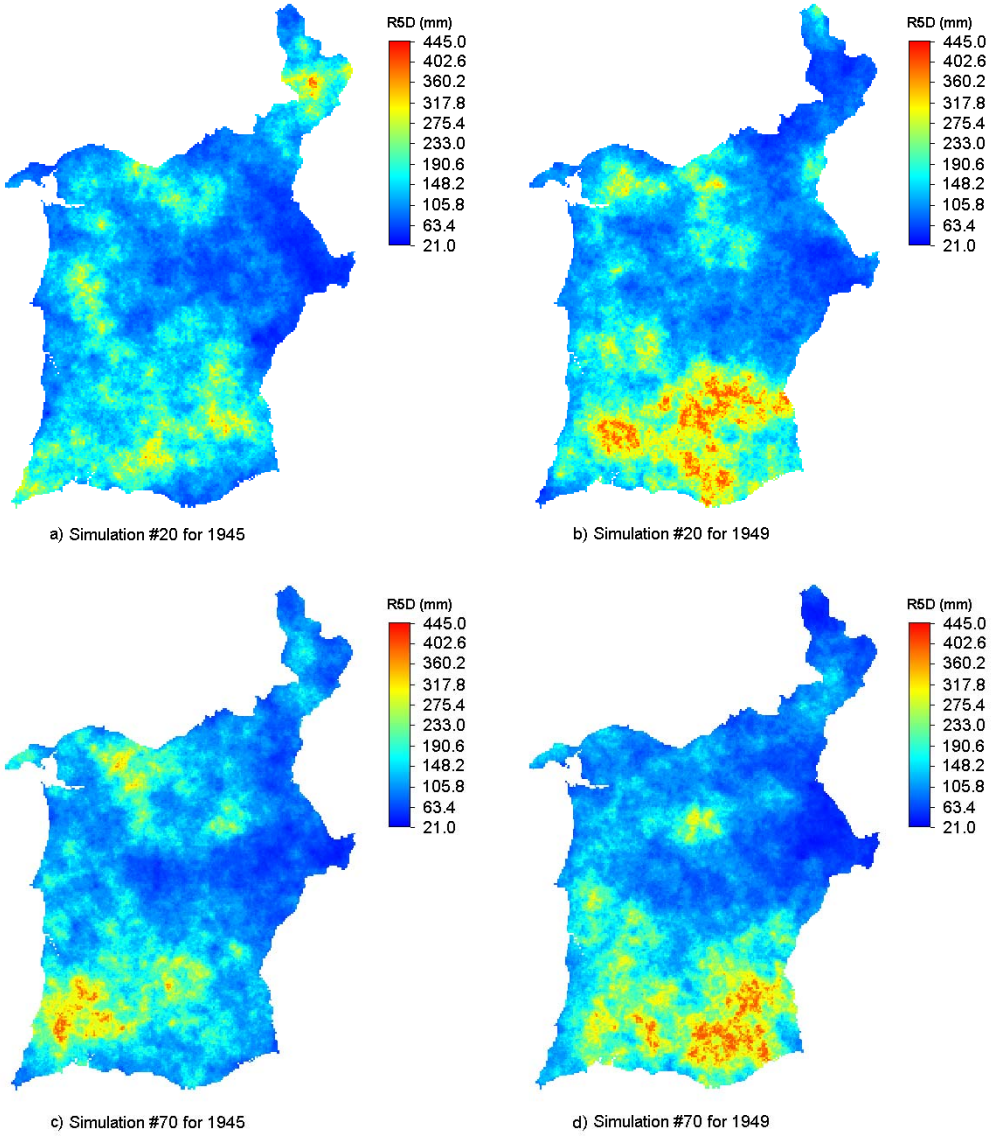

Figure 5.24 – Equiprobable simulated realizations of R5D for 1945 and 1949

For illustration purposes, the flood indicator (R5D) mean-maps of six years are shown in Figure 5.27, while Figure 5.28 and Figure 5.29 show their uncertainty evaluation measured by the standard deviation and the coefficient of variation, respectively. Likewise, six scenarios for the frequency of extreme precipitation events (R30) are shown in Figure 5.30, while Figure 5.31 and Figure 5.32 show their uncertainty evaluation measured by the standard deviation and the coefficient of variation, respectively.

Although the correlation models for the 1940s seemed unrealistic, a visual inspection of all produced scenarios for those years reveals a rather realistic spatial pattern of extreme precipitation. On the other hand, the lack of precipitation data on the northeastern area makes the estimates more uncertain. For that reason, the extreme precipitation values were possibly less accurately estimated over that area in several years. In fact, as expected, one of the regions where the distribution of extreme precipitation shows greater variability, thus more uncertainty, is in the northern part of the maps, corresponding to regions less densely sampled in most years. This is especially evident in the maps of the standard deviation. Moreover, the scenarios for the last two decades of the twentieth century show less uncertainty than the previous ones because the availability of stations over the study region is considerably greater (Figure 5.2).

The spatial uncertainty of the scenarios of both indices can be compared using the corresponding maps of the coefficient of variation. In order to summarize this information, probability maps of spatial uncertainty were computed as described in Section 5.2.5. Figure 5.25 and Figure 5.26 show estimated local probabilities of the coefficient of variation of the scenarios produced for the R5D and R30 indices, respectively, to be greater than or equal to given thresholds. This analysis shows that the scenarios of the R30 index have greater variability over the study region than the scenarios of the R5D index.

Only a few stations are located at medium (>400 m) and high elevations (Table 5.5), thus greater uncertainty would be expected at those regions. However, the uncertainty in the mountainous regions of the south is often small (Figure 5.25 and Figure 5.26), because of the use of elevation as secondary exhaustive information in the spatial interpolation procedure of the wetness indices.

Several common spatial patterns can be observed in scenarios of both indices. In wetter years, the indices exhibit the highest values in mountainous regions of Algarve, especially over the

Monchique mountains, due to the greater influence of altitude there than in the rest of the study domain. For this reason, high values also appear over northeast areas in several years. In drier years, the spatial pattern of extreme precipitation is much smoother as estimates have less variability over the study domain.

As expected from the space-time continuity analysis of the wetness indices (Section 5.3.2.2), the spatial patterns of extreme precipitation are becoming more homogenous over time, while the levels of local variability are decreasing. This is especially noticeable in the maps of the last two decades of the twentieth century.

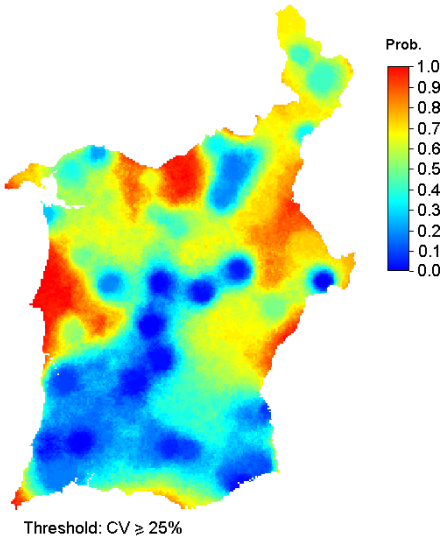

Figure 5.25 – Probability of the uncertainty of the R5D index scenarios, measured by the coefficient of variation, to be greater than or equal to 25%

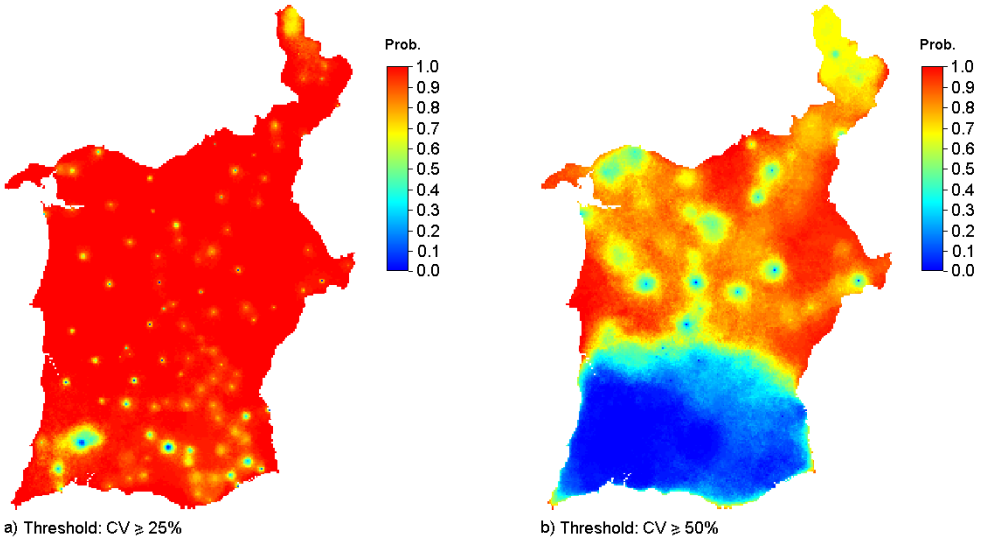

Figure 5.26 – Probability of the uncertainty of the R30 index scenarios, measured by the coefficient of variation, to be greater than or equal to a) 25% and b) 50%

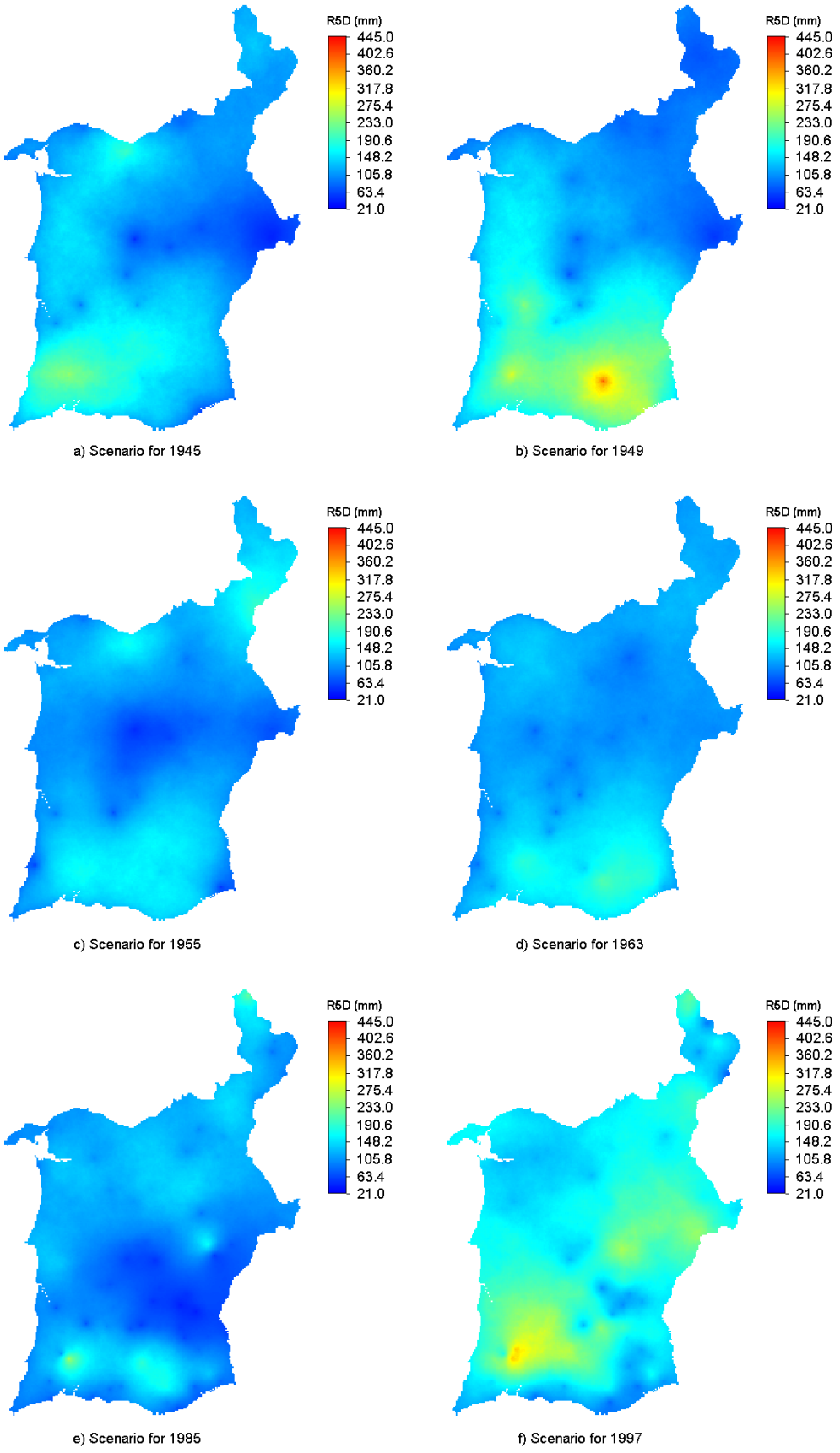

Figure 5.27 – Scenarios for the magnitude of extreme precipitation (R5D index)

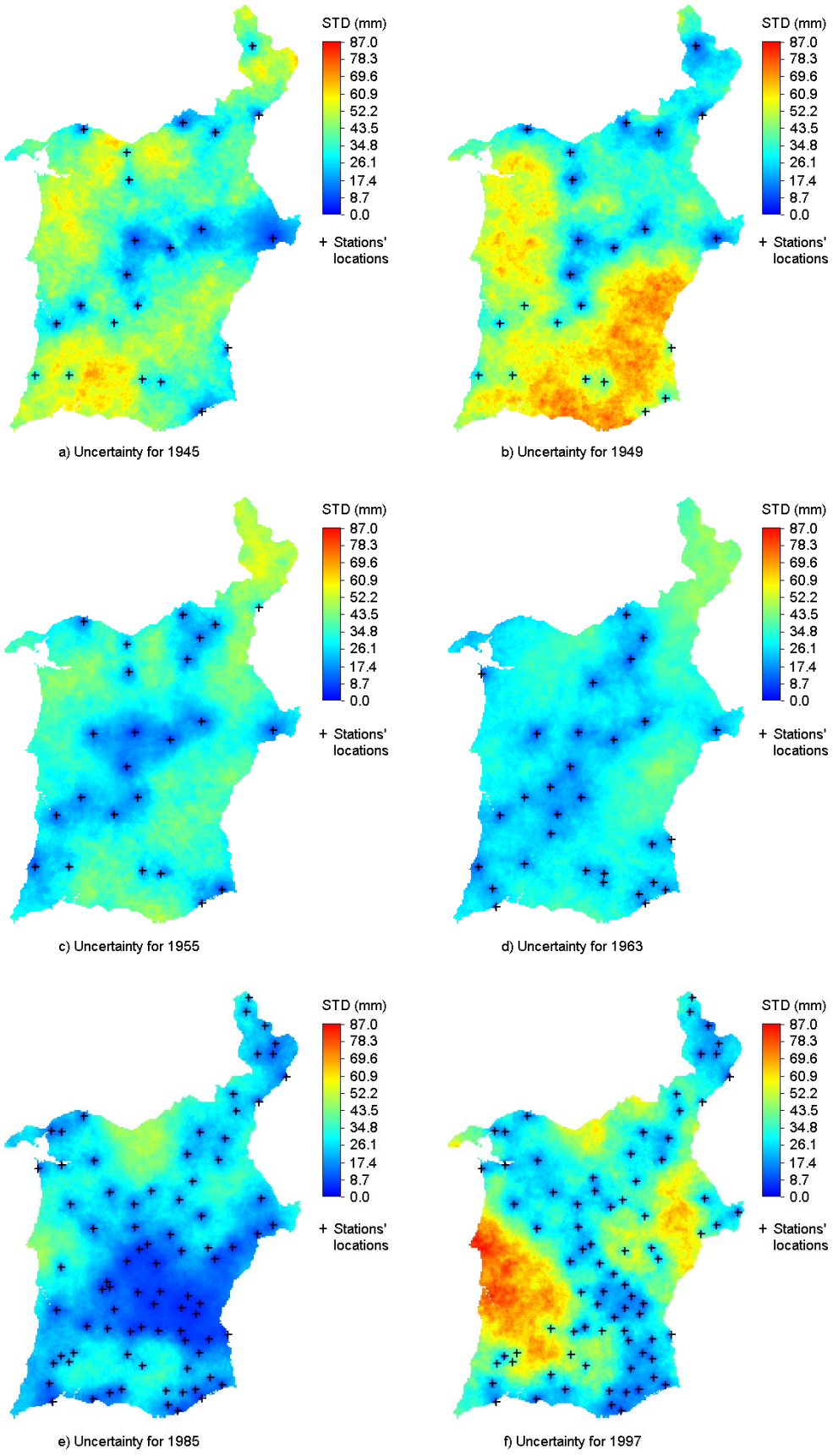

Figure 5.28 – Uncertainty of the scenarios for the magnitude of extreme precipitation (R5D index) measured by the standard-deviation

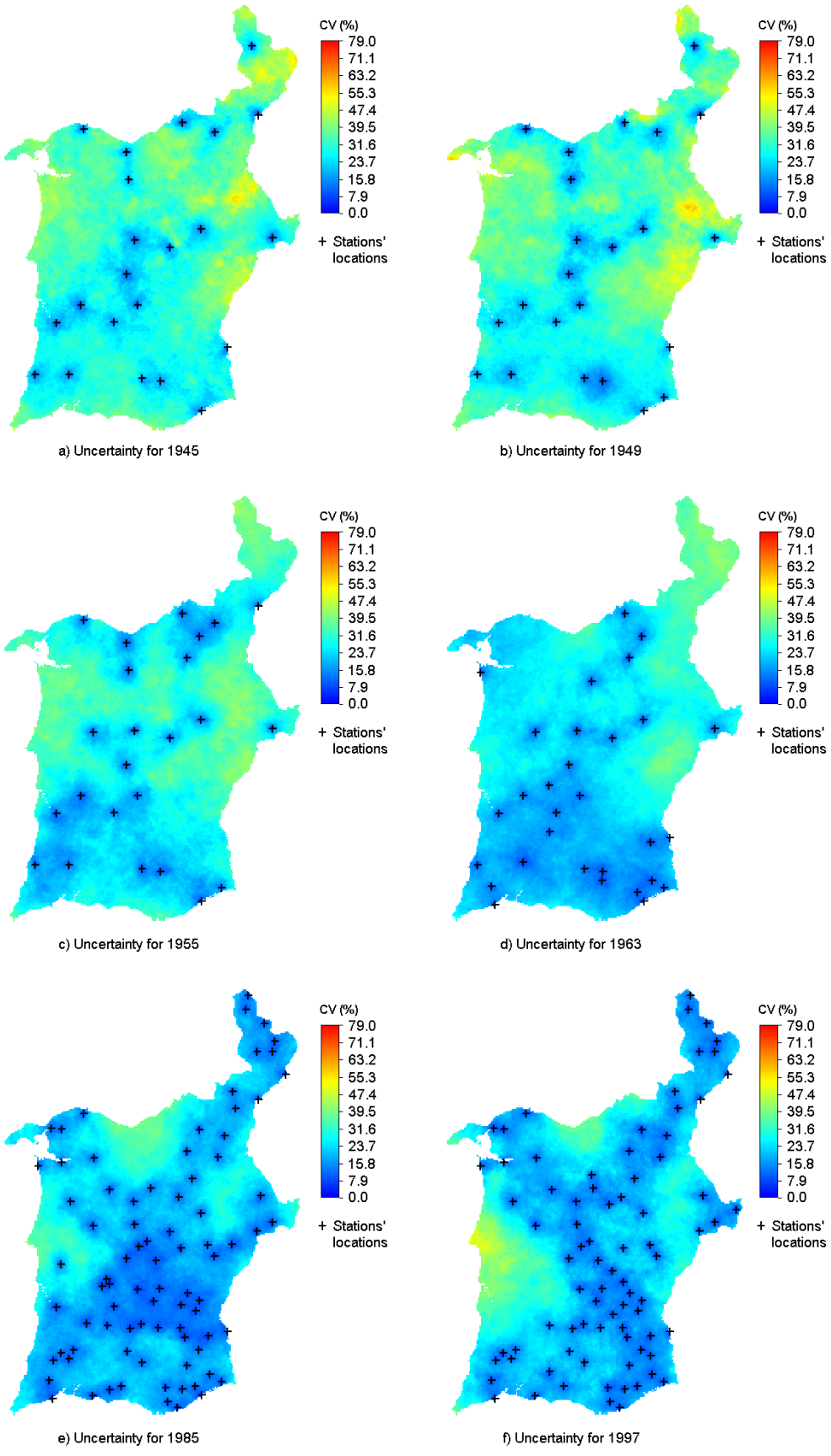

Figure 5.29 – Uncertainty of the scenarios for the magnitude of extreme precipitation (R5D index) measured by the coefficient of variation

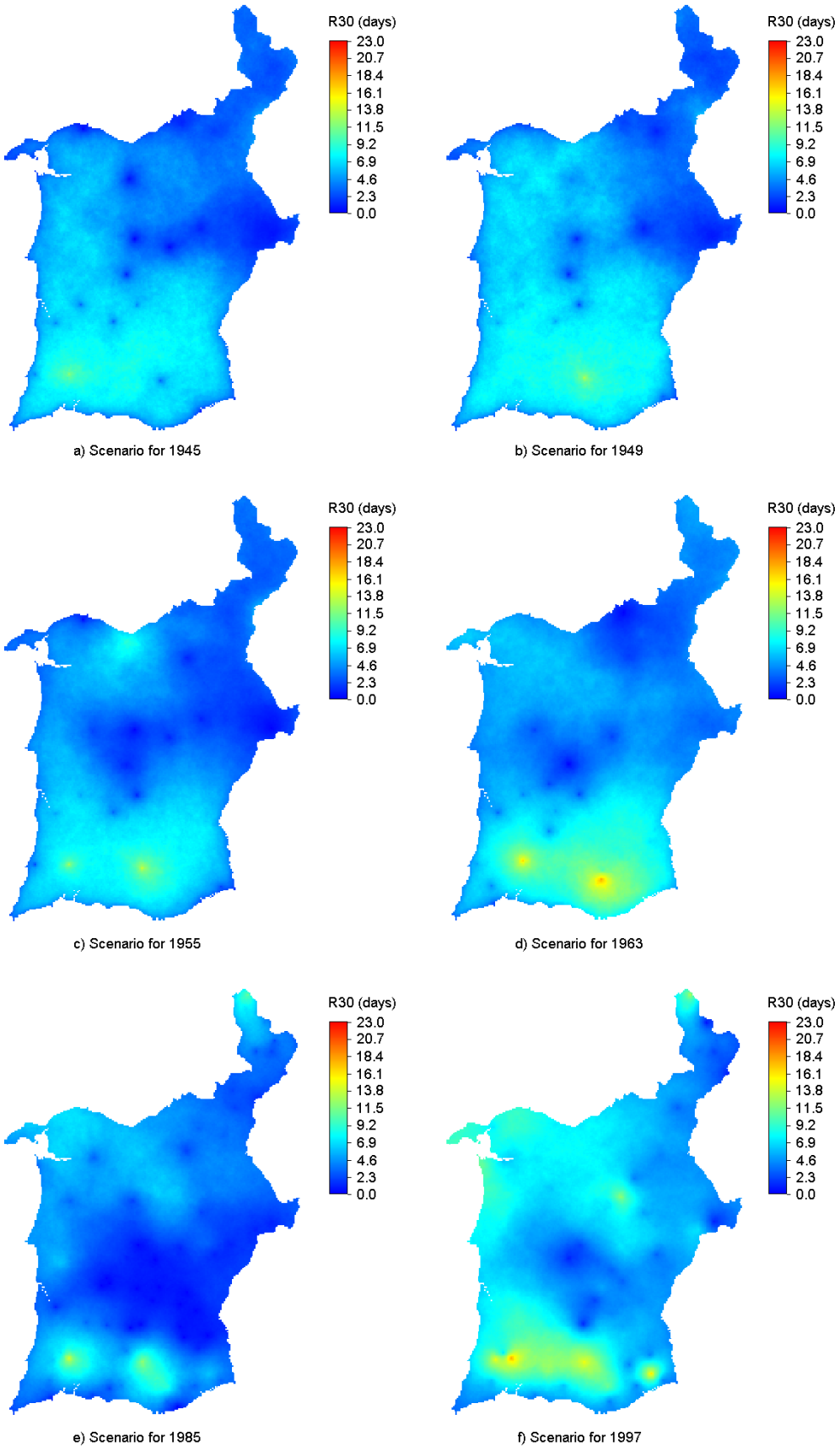

Figure 5.30 – Scenarios for the frequency of extreme precipitation (R30 index)

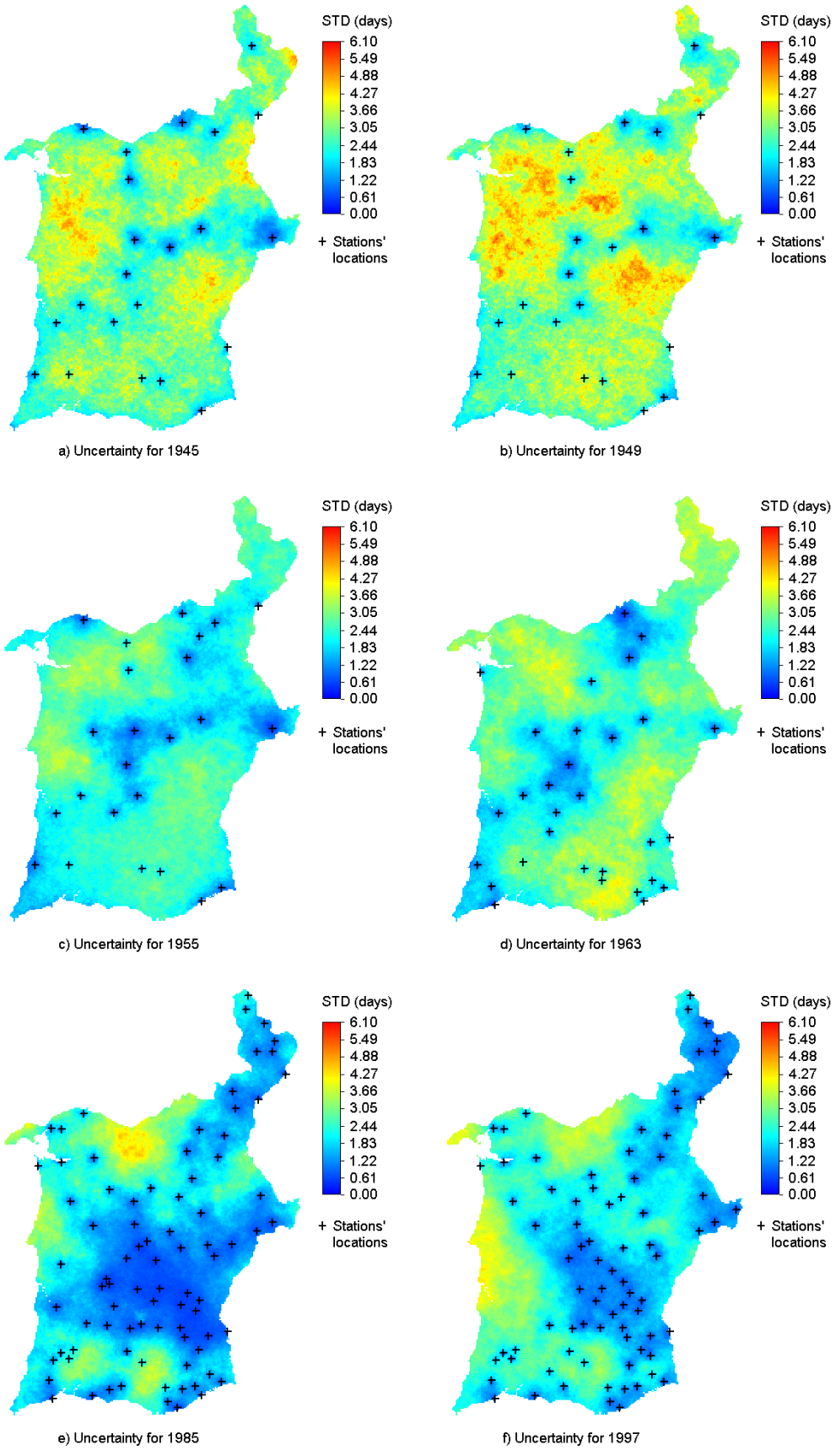

Figure 5.31 – Uncertainty of the scenarios for the frequency of extreme precipitation (R30 index) measured by the standard-deviation

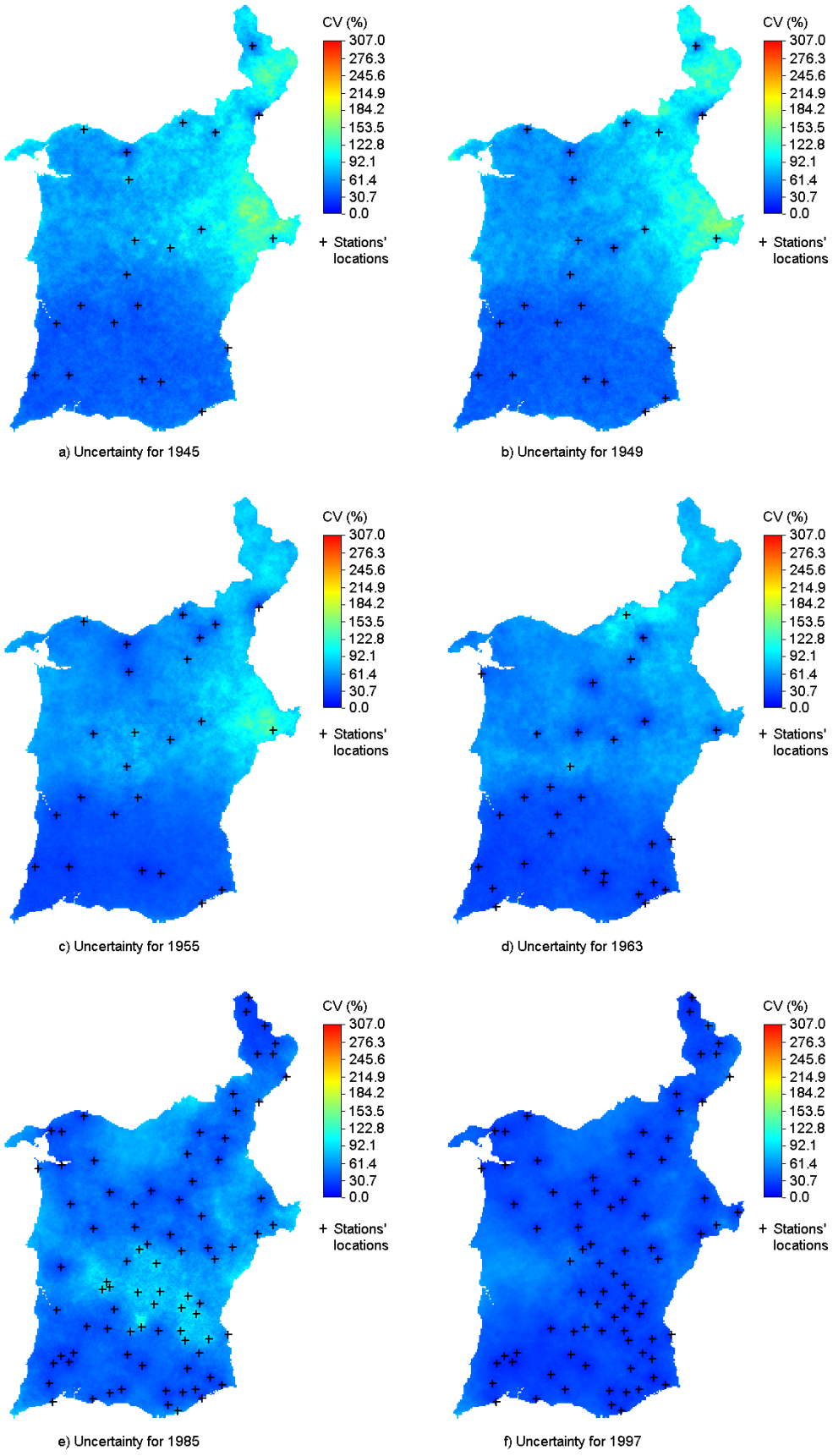

Figure 5.32 – Uncertainty of the scenarios for the frequency of extreme precipitation (R30 index) measured by the coefficient of variation

5.3.3.1 Maps summarizing the space-time dynamics

Using the annual gridded datasets produced, probability maps of extreme precipitation were computed as described in Section 5.2.5. In order to determine appropriate threshold values, the regional histograms of the indices and their basic statistics were calculated using the values from maps corresponding to the climate normal 1961/90 (Table 5.9).

Table 5.9 – Basic statistics of the wetness indices computed from the maps of 1961–1990

Regional statistics	R5D index	R30 index
Mean	108.6	4.3
Standard-deviation	28.3	2.3
Skewness	0.98	1.45
Kurtosis	2.24	4.22
Quantiles		
100% Max	229.4	23.0
99%	193.8	11.4
95%	160.2	8.3
90%	144.2	7.2
75% Q3	123.6	5.4
50% Median	105.4	3.8
25% Q1	89.5	2.6
10%	76.5	1.9
5%	68.6	1.5
1%	56.5	0.9
0% Min	23.2	0.0

The probability maps of both wetness indices show similar spatial patterns of extreme precipitation. Hence, areas that are susceptible to frequent extreme rainfall events are also susceptible to the occurrence of heavy short-term rainfall events. The probability maps corresponding to threshold values equal to the medians of R5D and R30 (Figure 5.33a and Figure 5.34a, respectively) show that the mountainous regions of Algarve, the northeast area, as well as the west coast have high probability of extreme precipitation. On the other hand, the probability maps for the third quartiles of R5D and R30 (Figure 5.33b and Figure 5.34b, respectively) show that the most intense and more frequent rainfall events occur at the Algarve region, especially over the Monchique mountains, as expected.

Probability maps such as these are useful to identify regions at risk of water erosion caused by extreme precipitation events. A probability map could be combined with a vegetation cover map. This would allow the identification of regions at risk of water erosion corresponding to areas with little vegetation cover and high probability of extreme precipitation events. This could be a valuable improvement of the 'Erosion protection' map used to build the 'Vegetation quality index' used by the National Action Programme to Combat Desertification to identify desertification prone areas (Rosário, 2004b).

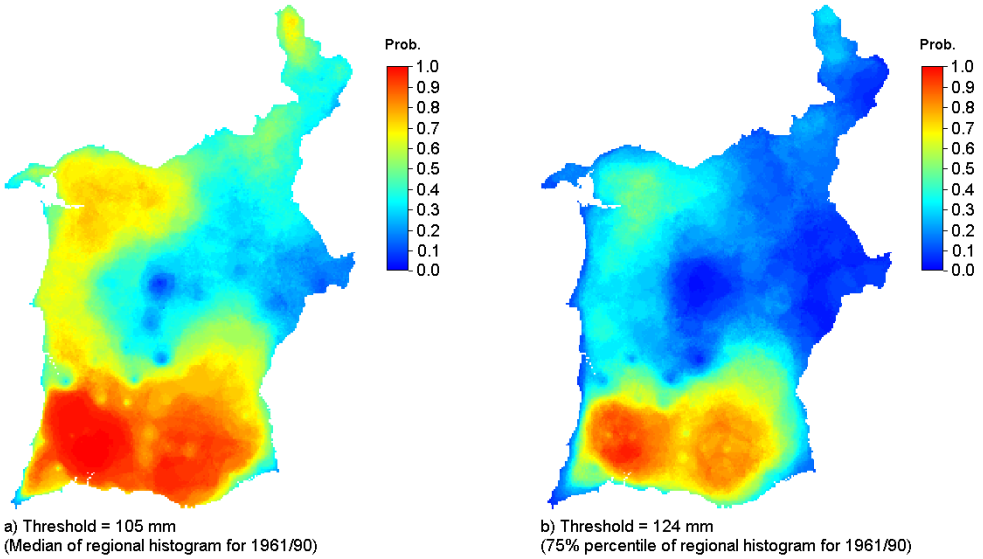

Figure 5.33 – Probability of the magnitude of extreme precipitation (R5D index) to be equal or greater than fixed thresholds

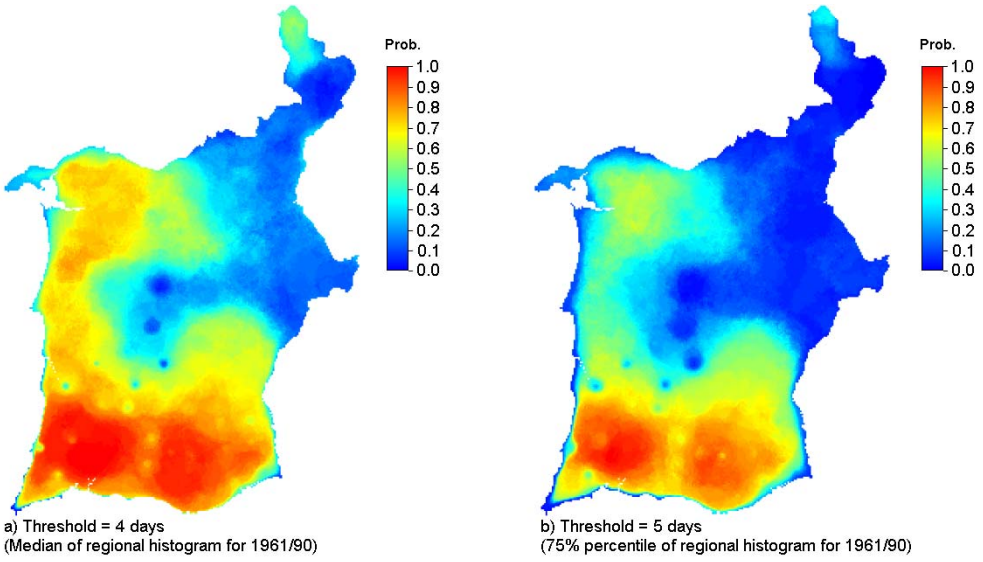

Figure 5.34 – Probability of the frequency of extreme precipitation (R30 index) to be equal or greater than fixed thresholds

The results from Section 4.3 on the regional correlation analysis between the wetness indices, based on 15 stations data, showed that they were moderately positively correlated with each other. Using the 1940/99 scenarios of R5D and R30, a map of local correlations between them was produced by computing the Pearson's correlation coefficient at each grid cell (Figure 5.35). An interesting conclusion from this map is that increasing values of R5D through time entail increasing values of R30 in many areas that have low probabilities of extreme precipitation, and vice-versa. For example, many areas in the mountainous regions of Algarve show weak correlations between the frequency of heavy precipitation and the intensity of short-term rainfall events.

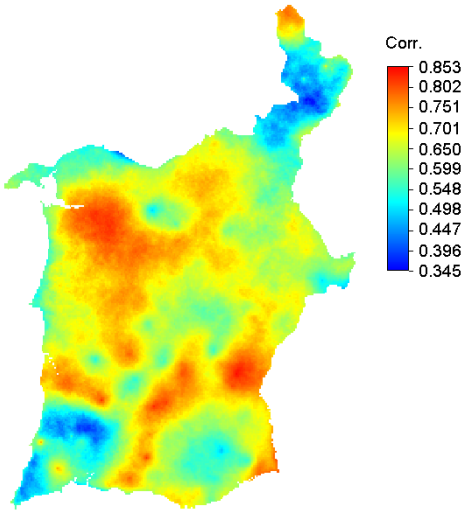

Figure 5.35 – Local correlations between the R5D and R30 indices

Local trend maps were computed using both parametric and nonparametric estimators of the trend slope, in order to compare their resulting spatial patterns. However, to assess the spatial patterns of the trend magnitude, the nonparametric trend map is more appropriate than the parametric one, because of the results from Chapter 4. The yearly estimates of the trends slope were calculated at each grid cell for 1940/99, through OLS and a nonparametric method (Section 5.2.5), using the R5D and R30 scenarios (Figure 5.36 and Figure 5.37, respectively).

Both parametric and nonparametric approaches used to compute the trend slope reveal identical spatial patterns. The trend maps of both wetness indices also show similar spatial patterns. As expected from the trend analysis chapter, there is a pattern of weak, both negative and positive, trend signals of extreme precipitation over the study region. Most of the region

exhibits negative trends of extreme precipitation, and a small area in the northeast has the highest positive trends.

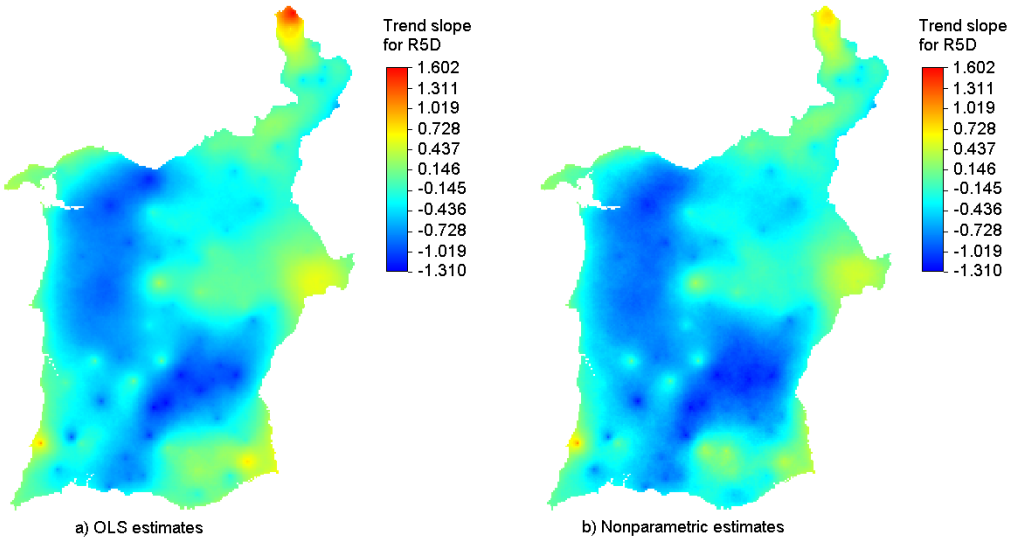

Figure 5.36 – Local trends in the magnitude of extreme precipitation (R5D index)

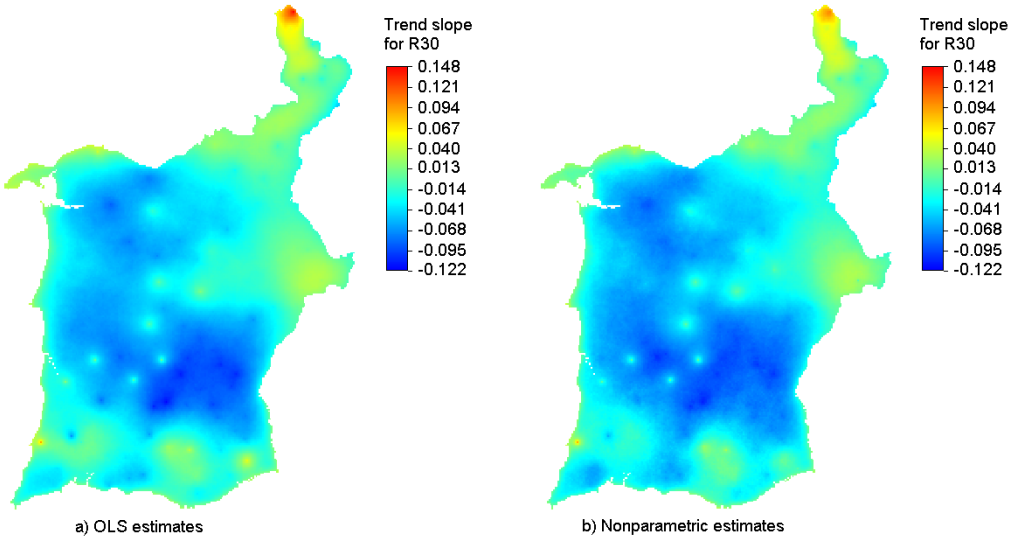

Figure 5.37 – Local trends in the frequency of extreme precipitation (R30 index)

5.3.4 Space-time scenarios for dry extremes

Using the DSS algorithm, a set of 100 equiprobable simulated realizations of the AII index was computed at each simulated grid node, by year. The space-time inference was performed by means of computing the mean of those distributions, hereafter referred to as mean-maps or (space-time) scenarios. Uncertainty was assessed by means of computing the standard

deviation (STD) and the coefficient of variation (CV) of the distribution of the 100 simulated values at each simulated grid node, by year.

Six scenarios for AII are shown in Figure 5.39, while Figure 5.40 and Figure 5.41 show their uncertainty evaluation measured by the standard deviation and the coefficient of variation, respectively. In general, the dryness scenarios of the last two decades of the twentieth century have less uncertainty than earlier years because of the higher availability of data in recent times. Probability maps of spatial uncertainty were computed as described in Section 5.2.5. Figure 5.38 shows the estimated local probabilities of the scenarios' coefficient of variation to be greater than or equal to a given threshold. This analysis shows that the scenarios of the AII index have less variability over the study region than the scenarios of the wetness indices.

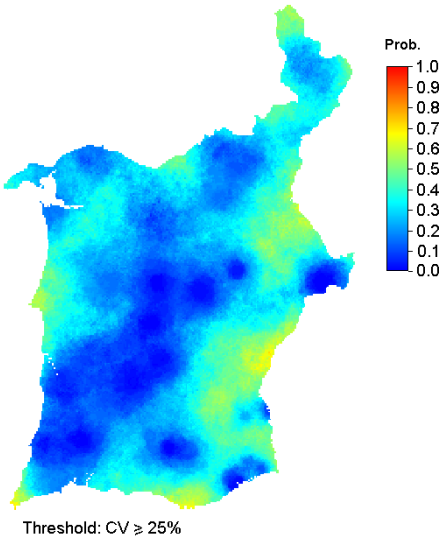

Figure 5.38 – Probability of the uncertainty of the AII index scenarios, measured by the coefficient of variation, to be greater than or equal to 25%

The south of continental Portugal, especially the Alentejo region, is a drought prone area characterized by scarce precipitation, little runoff and water availability. The occurrence of drought spells and scarce precipitation highly influence the productivity of rain fed agriculture in this region, where most of the country's cereals are produced.

The scenarios produced for the AII index clearly show drought and dry situations such as the extreme episodes of 1948/49 and 1980/81 that affected Portugal's territory (Trigo and DaCamara, 2000). The magnitude of aridity (lower values of AII) is higher in interior areas of Alentejo and eastern areas of Algarve in many years.

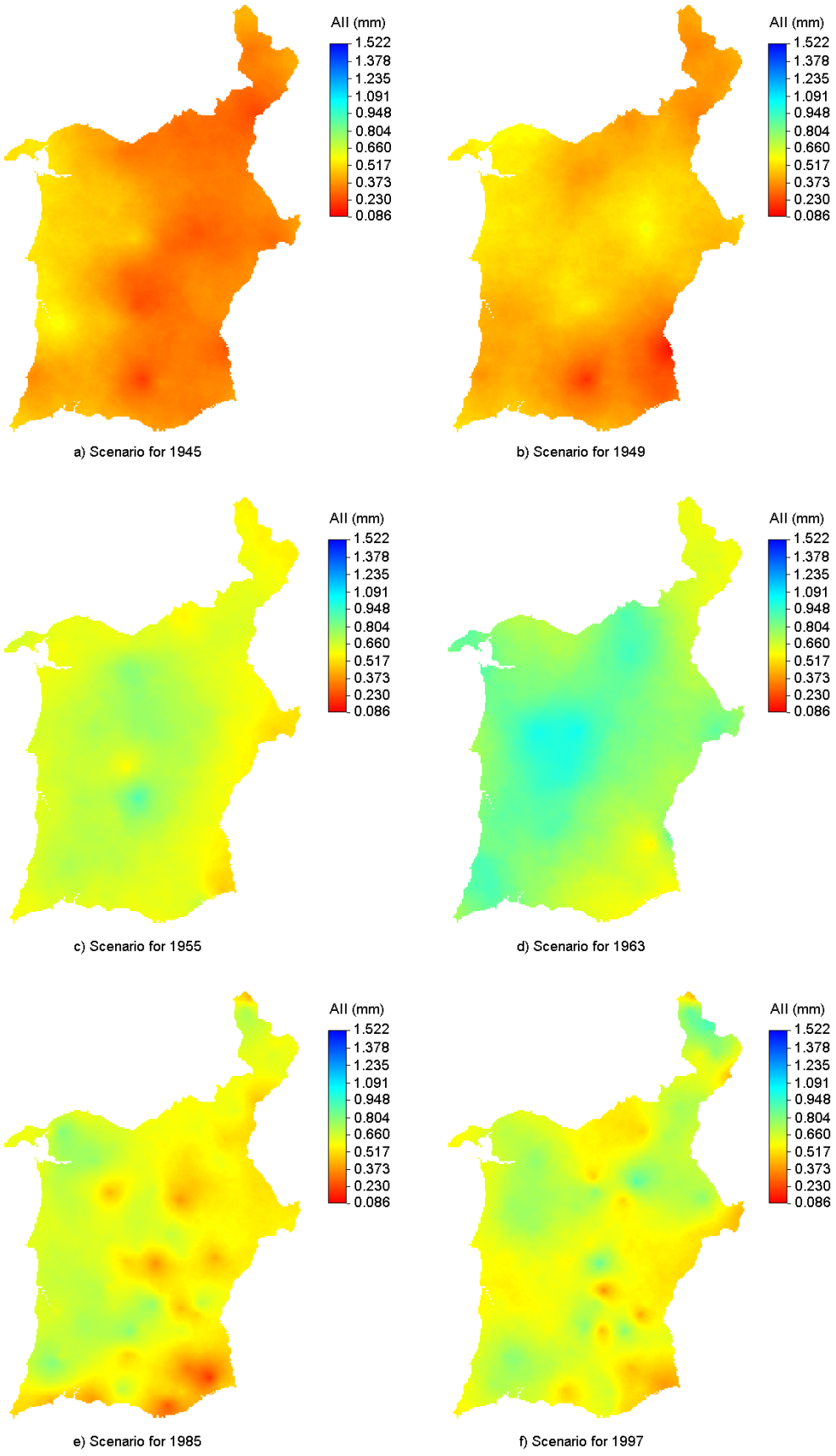

Figure 5.39 – Scenarios for the aridity magnitude (AII index)

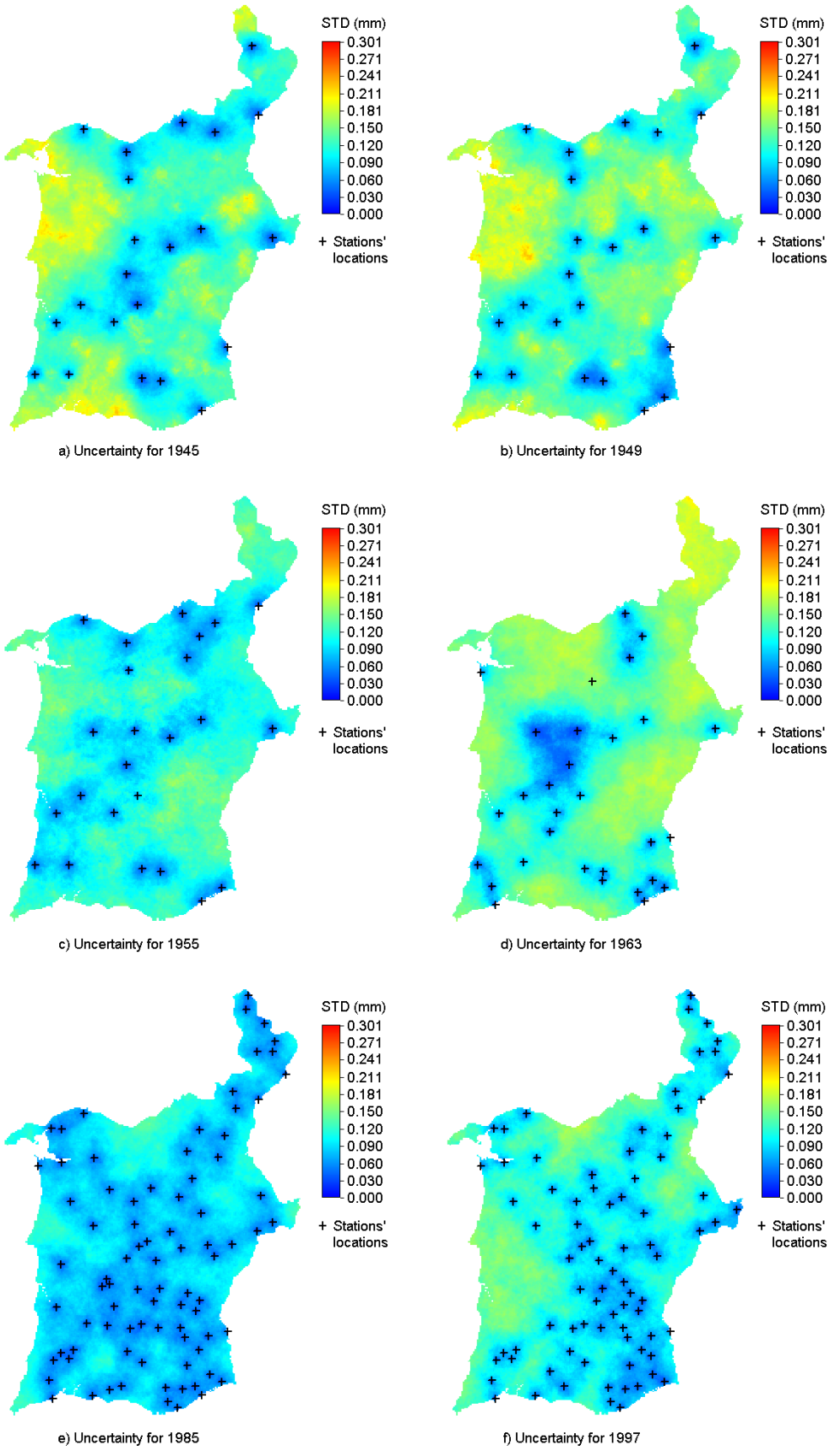

Figure 5.40 – Uncertainty of the scenarios for the aridity magnitude (AII index) measured by the standard-deviation

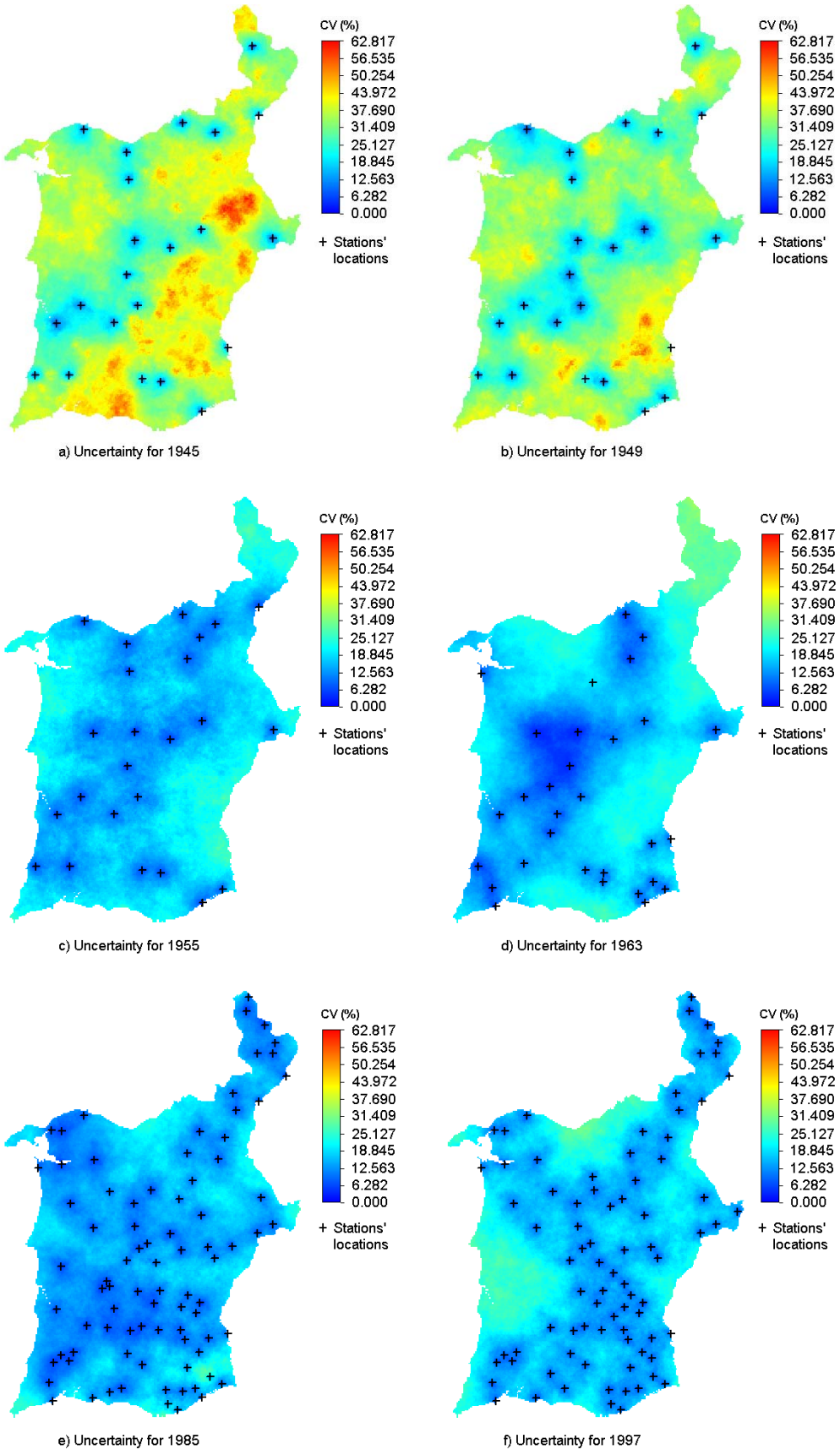

Figure 5.41 – Uncertainty of the scenarios for the aridity magnitude (AII index) measured by the coefficient of variation

5.3.4.1 Maps summarizing the space-time dynamics

Since lower values of the AII index indicate higher dryness, areas with highly negative correlations between AII and the wetness indices correspond to areas that are potentially more susceptible to land degradation caused by water erosion. For this reason, local correlation maps between AII and the wetness indices were produced, even though the correlations were expected to be weak. Figure 5.42 reveals higher correlations of AII with the frequency of extreme precipitation (R30) than with the R5D index. However, in Figure 5.42b, most of the correlation values are approximately equal to 0.3 over the study region, as expected. The strongest negative relationship between AII and extreme precipitation occurs in a small area of the northeast of the study region, for both wetness indices. Since it is a bordering area lacking data in many years, this result is highly uncertain.

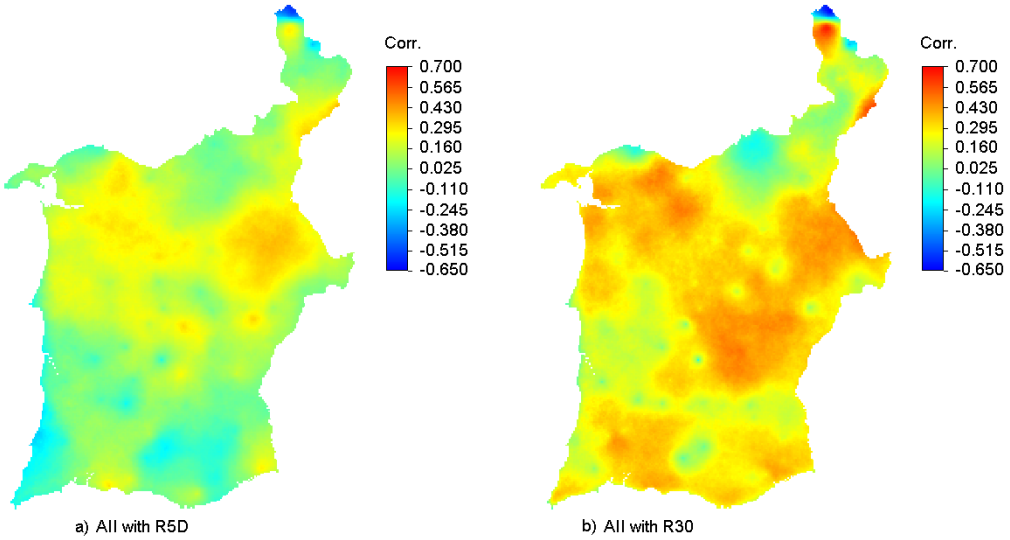

Figure 5.42 – Local correlations between AII and the wetness indices

Using the annual gridded datasets, probability maps of extreme dryness were computed as described in Section 5.2.5. In order to determine appropriate threshold values, regional histograms of the AII index and their basic statistics were calculated using the values from maps corresponding to the climate normal 1961/90 (Table 5.10). The probability maps corresponding to threshold values equal to the median and the first quartile of AII (Figure 5.43) show that the south coast and inland regions of southern Portugal are highly prone to dry conditions, and that the southeast area is extremely susceptible to drought and dryness.

Table 5.10 – Basic statistics of the dryness indices computed from the maps of 1961–1990

Regional statistics	CDD index	AII index
Mean		0.56
Standard-deviation		0.12
Skewness		0.11
Kurtosis		0.22
Quantiles		
100% Max		1.02
99%		0.85
95%		0.75
90%		0.71
75% Q3		0.64
50% Median		0.56
25% Q1		0.48
10%		0.41
5%		0.37
1%		0.29
0% Min		0.09

Figure 5.44 shows the 'Aridity index' map that has been used by the National Action Programme (NAP) to Combat Desertification (Rosário, 2004b). The 'Aridity index' map is one of the four composite indices used to build the susceptibility map of desertification of the NAP, and it should be an indicator of climate quality (Kosmas *et al*., 1999). Currently, the 'Aridity index' map corresponds to the ratio of two maps: the annual average precipitation for the period 1959/60 – 1990/91 mapped through kriging using elevation as external drift, and the annual average potential evapotranspiration for the period 1961–1990 mapped through ordinary kriging. A visual comparison between the NAP's 'Aridity index' map (Figure 5.44) and the probability map corresponding to the *median* of AII (Figure 5.43a) allows to conclude that their spatial patterns are, in general, similar, but extremely dry areas of the southeast (Figure 5.43b) were not captured by the NAP's 'Aridity index'.

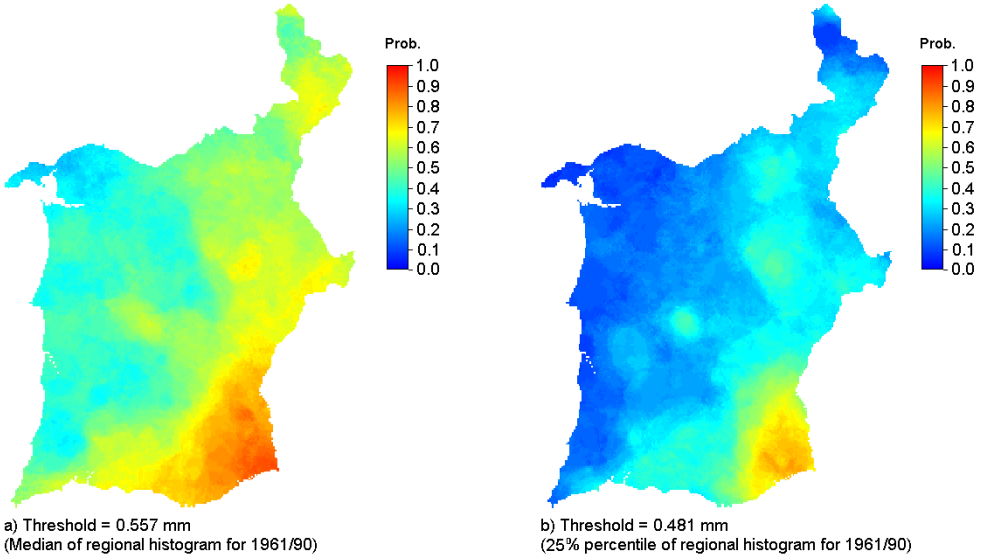

Figure 5.43 – Probability of the aridity magnitude (AII index) to be less than fixed thresholds

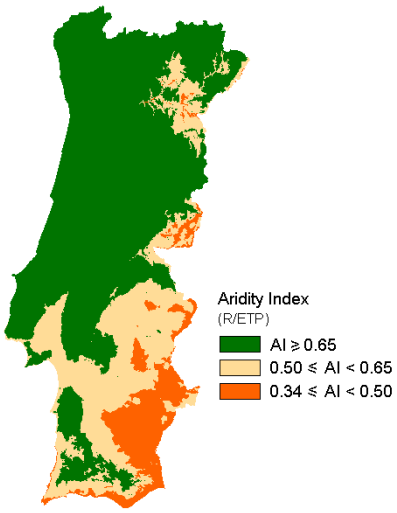

Figure 5.44 – Aridity index used by the National Action Programme to Combat Desertification (Rosário, 2004b). Source data: Programa de Acção Nacional de Combate à Desertificação, http://panda.igeo.pt/pancd/ (retrieved: 14 May 2008).

Yearly estimates of the AII trends slope were calculated at each grid cell for 1940/99, through OLS and a nonparametric method (Section 5.2.5). Both parametric and nonparametric approaches used to compute the trend slope reveal identical spatial patterns (Figure 5.45). As expected from the trend analysis chapter, there is a pattern of negative trend signals for the AII index over most of the study region, which indicates an increasing trend of dry conditions through time, especially in the centre and coastal areas of Algarve.

Note that the gridded trend maps of AII (Figure 5.45) should not be directly compared with the map obtained through ordinary kriging (Figure 4.2, p. 208), or at least they should be cautiously compared, because the analysis period is not the same, neither is the set of stations used in the spatial interpolation process.

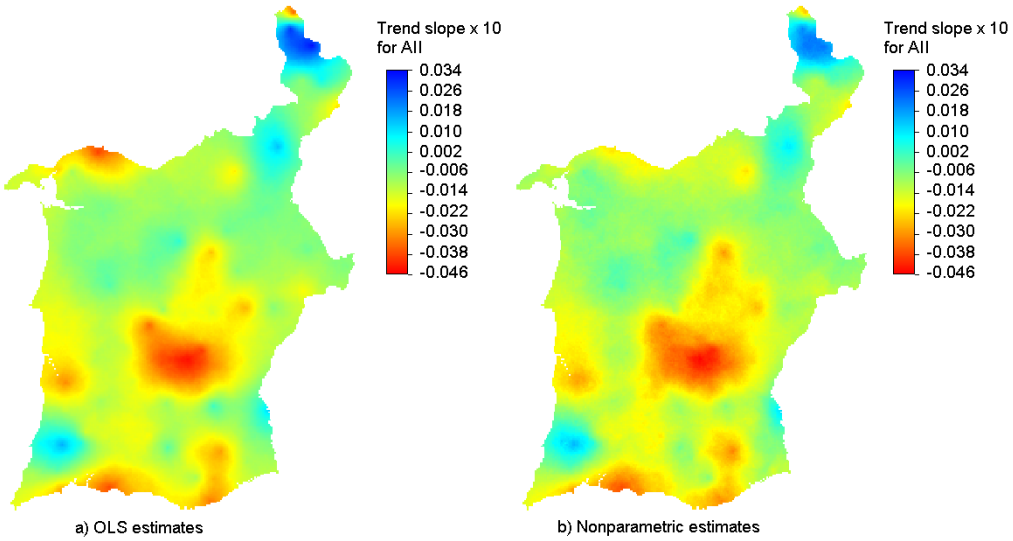

Figure 5.45 – Local decadal trends in the aridity magnitude (AII index)

Chapter 6: CONCLUSION

6. Conclusion

The main purpose of this research work is to characterize the time and space-time dynamics of extreme precipitation indices, including both wet and dry extreme values, at local scales in the south of Portugal, where large areas have high susceptibility to desertification (Correia, 2004; Rosário, 2004b).

A set of 107 series of daily precipitation were compiled for homogeneity assessment (Chapter 3). Wijngaard *et al.* (2003) state that, generally, a combination of statistical methods and methods relying on station history information is considered to be most effective to track down inhomogeneities. However, relative approaches are usually preferred, if the monitoring station network is dense enough, because such techniques account for regional climate changes and isolate the effects of irregularities in a candidate station by using data from reference stations. The methodology used for the homogeneity assessment of the Portuguese precipitation time series comprised both absolute and relative approaches.

During the relative testing stage, we proposed an extension of the Ellipse test (described by Allen *et al.*, 1998), named *SUR+Ellipse test*, which has the advantage of testing simultaneously several candidate series from the same climatic area, taking into account the contemporaneous relationship between them. Moreover, the results show that there is no apparent connection between the potential breaks magnitudes and the ability of the SUR+Ellipse test to identify them. Therefore, this technique is a valuable tool for homogeneity testing climate time series when the station network is dense enough. Like several traditional procedures, a limitation of this approach arises because of missing values and varying availability of stations through time, which makes difficult to select the series for each model.

Additionally, a procedure based on geostatistical stochastic simulation was proposed for the homogenization of climate data (Costa *et al.*, 2008a). The case study results indicate that this approach has a number of potential advantages over traditional ones. Geostatistical techniques allow dealing with the problem of missing values and varying availability of stations through time, by using different sets of neighbouring stations at different time periods (years, months, etc.), and by including shorter and non-complete records. Moreover, the geostatistical

approach avoids the iterative construction of composite reference series because it increases the contribution of records from closer stations, both in spatial and correlation terms, by accounting for the joint spatial and temporal dependence between observations. Multiple breaks can be detected simultaneously, thus this method might be less time consuming than other testing techniques that are used iteratively. Another advantage is that the geostatistical approach seems to be able to identify breakpoints near the start and end of the time series, while traditional approaches have less power in detecting them (Aguilar *et al*., 2003). Accordingly, the promising results from the case study open new research perspectives on the homogenization of climate time series.

This study confirmed that using absolute approaches without metadata information makes it difficult to determine if changes or lack of changes in a station's time series result from inhomogeneities or simply from abrupt changes in the regional climate. The results of the different procedures implemented were used to develop an overall classification of the daily series using four classes: ‘useful’, ‘potentially useful’, ‘potentially suspect’ and ‘suspect’. The intermediate classes were established for stations that were just tested through absolute techniques, since it was not possible to find historic metadata support for the irregularities identified. Therefore, we strongly recommend that further efforts should be made to quality control those series.

From the set of 107 stations compiled for homogeneity assessment, 15 stations with homogeneous daily records in the period 1955/99 were selected for temporal pattern analysis (Chapter 4). Six precipitation indices were then developed to investigate yearly trends and climate dynamics at the local scale in the south of Portugal. The three dryness indices (AII, CDD and FDD) and the three wetness indices (SDII, R5D and R30) describe moderate climate extremes which are relevant for the management of water resources and land use, modelling of erosion, and other applications for ecosystem and hydrological impact modelling. The existence of trends and other temporal patterns in the extreme precipitation indices were investigated and uncertainty about rainfall patterns evolution was assessed through regression models and smoothing techniques.

The proposed AII index is particularly useful to provide information about land vulnerability, especially in agricultural areas such as those located at the south of Portugal. The statistically significant trends of this indicator in many stations and in the moving average series confirms

the desertification-related scenarios that reveal a tendency towards drier climatic conditions in the south of Portugal in the near future. The results of the other two dryness indices provide evidence that, in recent times, an increase in the length of the greatest dry spell entails a decrease in the mean frequency of dry events.

The trend signals of the wetness indices were not statistically significant at the majority of stations. However, the moving window techniques revealed an increase in the short-term precipitation intensity (R5D index) during the last three decades of the twentieth century. Indices characterizing the precipitation intensity on wet days (SDII) and the frequency of extremely heavy precipitation events (R30) have cyclic patterns and different trend signals at the local scale, during the period 1955/99. The results also indicate that extreme precipitation variability and climate uncertainty are greater in recent times.

For the spatial interpolation and uncertainty assessment of two indices of extreme precipitation (R5D and R30), we explored the application of direct sequential cosimulation, which allows incorporating covariates such as elevation. The methodology accounts for local data variability and incorporates space-time models that allow capturing long-term trends of extreme precipitation, and local changes in the relationship between elevation and extreme precipitation through time. For the dryness index (AII), direct sequential simulation was used instead, because no relevant correlations were found with physiographic features.

Annual gridded datasets of the three precipitation indices were produced from 1940 to 1999 using daily precipitation observations measured at 105 monitoring stations (Chapter 5). Uncertainty evaluations of the proposed scenarios were also produced for each year. The spatial resolution of the gridded datasets, 800 m × 800 m grid cells, allows for a detailed assessment of the space patterns of precipitation extremes.

As expected, regions where the distribution of precipitation extremes shows greater variability, thus more uncertainty, correspond to regions less densely sampled. However, although there are few stations located at elevations higher than 420 m, the uncertainty in mountainous regions is noticeably small given that elevation was used as secondary exhaustive information in the spatial interpolation procedure of the wetness indices. The scenarios of the R30 index have greater variability over the study region than the scenarios of the R5D index, whereas the scenarios of the AII index are less uncertain than the wetness indices scenarios.

The results indicate that the relationship between elevation and the wetness indices varies locally and has decreased through time over the study region. The results also indicate that the spatial patterns of precipitation extremes have become more homogenous during the last decades of the twentieth century, which is consistent with the decreasing relationship with elevation. On the other hand, the extreme precipitation variability has increased in recent times in the south of Portugal (Chapter 4). This means that climate uncertainty is becoming greater in the time dimension, but the extreme phenomena are becoming more homogeneous in the spatial dimension.

The probability maps of both wetness indices allow concluding that many areas that are susceptible to frequent extreme rainfall events are also susceptible to the occurrence of heavy short-term rainfall events. The most intense and more frequent rainfall events occur at the Algarve region, especially over the Monchique mountains. Accordingly, many areas of Algarve are at risk of water erosion and floods caused by extreme precipitation events. The gridded trend estimates of R5D and R30 show that most of the study region exhibits weak negative trend signals in extreme precipitation events, while few areas show weak positive signals, particularly the northeast.

Regarding the dryness indices, the spatial inconsistencies found in CDD correspond to high variability at short distances caused by the split of a long dry spell into two dry spells due to one, or two, days with small amounts of daily rainfall in some stations. The proposed AII index is more robust in situations such as these because it is defined as the average precipitation of dry days. Therefore, the proposed AII index seems more appropriate than CDD to characterize drought and dry conditions in arid regions with high climatic variability such as the south of Portugal.

The space-time analysis of AII indicates the southeast region as the most threatened by droughts and extreme dryness. These climatic factors amplify the risks of soil erosion and land degradation in the mountainous areas of this region, which are also prone to extreme rainfall events. Coastal areas of the south and many eastern inland areas are also subject to frequent periods of extreme dryness. Moreover, the gridded trend estimates of AII reveal a tendency towards drier climatic conditions in the coastal areas of the south and in the centre of the study region.

Increased aridity is a robust proximate cause of desertification, both indirectly through greater rainfall variability and directly through prolonged droughts (Geist and Lambin, 2004). Therefore, probability maps such as those computed for the AII are a promising contribution to improve the 'Aridity index' map that has been used by the National Action Programme to Combat Desertification (Rosário, 2004b). In fact, the current 'Aridity index' map is not fully appropriate to map the areas susceptible to desertification because it is based on average computations (Rosário, 2004b; Pereira *et al*., 2006). Therefore, this index does not account for droughts, which are part of the climate driving forces influencing desertification, and it does not capture the dynamic climatic features hidden behind "the means".

Taking into consideration the methodology used and, in particular, the results of the AII index, we believe that our approach is a valuable contribution, not only to improve the knowledge on the time and space-time dynamics of precipitation extremes, but also to compute a new indicator of climate quality, which can be used to develop susceptibility maps of desertification.

The number of studies analyzing space-time patterns of indices of precipitation extremes is very limited, because the large majority of studies only focus on the temporal linear trends of the indices. The direct sequential simulation and cosimulation, and the developed stochastic space-time models in particular, proved to be valuable procedures to improve the knowledge on the space-time dynamics of precipitation extremes and to provide uncertainty assessments of produced scenarios.

The annual gridded datasets allow making available, for each monitoring stations' location, complete time series of the climate indices for the 1940/99 period, because the gaps in the series were infilled. Note that the direct sequential simulation algorithms provide exact interpolation values at locations with observed data. Moreover, an uncertainty evaluation of the infilled records is also available. These datasets of observation-based indices enable a variety of climate and impact studies, including detailed analyses of changes in the occurrence of extremes over the past 60 years.

The results of this work open perspectives for new approaches on the analysis of extreme climate events, particularly in the context of impact studies.

6.1 Limitations

One of the specific aims of this thesis was to assess the homogeneity of the daily precipitation series in order to properly proceed with the research concerning its major objective. However, station history information was not available, at least to us, to support the absolute breaks detected. Further analysis using relative procedures had to be carried out because of this limitation. Procedures based on test iteration, such as the ones used, are computationally intensive and turned out to be time consuming and exacting work. Although a great deal of effort has been made to assess the homogeneity of the daily precipitation series, a significant number of them were not evaluated through relative testing techniques due to time constrains. Nevertheless, in Section 5.1, we discuss a number of arguments that should justify the use of a dataset that might contain several inhomogeneous records.

Geostatistical simulation methods are powerful but complex and computationally intensive. The analysis of the 6000 equally probable maps that were simulated for R5D, R30 and AII, was limited by the computational power available at the time. For this reason, the summary maps were based on the mean-maps (space-time scenarios computed using the mean of the distribution of the simulated values: 100 for each year of 1940/99), rather than directly on the simulated ones. Further insights on the space-time dynamics of the indices could be revealed if the full set of gridded datasets was investigated.

6.2 Recommendations for further research

Three research questions emerge from this thesis, which point out directions for further research:

- *Can the geostatistical simulation approach be used for the homogenization of sub-monthly climate data?*

  Further investigation is required to develop procedures for the homogenization of sub-monthly precipitation data (Aguilar *et al*., 2003; Wijngaard *et al*., 2003; Auer *et al*., 2005). The geostatistical simulation approach seems to be a very promising procedure for this research field, as kriging techniques have proven to succeed in the estimation of missing daily precipitation records (e.g., Kyriakidis *et al*., 2004; Teegavarapu and Chandramouli, 2005; Carrera-Hernández and Gaskin, 2007). Moreover, multivariate

geostatistical simulation algorithms might be used for the homogenization of highly variable elements, such as precipitation, making use of information from explanatory physiographic variables (e.g. elevation).

The inherently high (temporal and spatial) variability of precipitation makes homogenization of precipitation records more difficult to accomplish than other elements (e.g. temperature). Therefore, it is reasonably intuitive that the number of simulated realizations, used by the geostatistical approach to infer the local probability density functions (pdfs), should be higher for precipitation data than for temperature records. However, how many simulated realizations should be used to accurately infer the local pdfs? This and other issues require further investigation and open new perspectives on the homogenization of climate data.

- *What caused the increased spatial homogeneity found in the 'wetness' indices of precipitation extremes in the last decades of the twentieth century? Are there other regions, or countries, experiencing these changes?*

  Changes in the North Atlantic Oscillation (Goodess and Jones, 2002; Haylock and Goodess, 2004; Trigo *et al*., 2004; Rodrigo and Trigo, 2007; Scaife *et al*., 2008) are likely to be responsible for the observed changes in the extreme precipitation indices in southern Portugal, especially during the last two decades of the twentieth century. Nevertheless, other factors such as land cover and land use changes should be considered. Further research on this subject is clearly required.

- *Is the CDD index definition robust enough so that it can be appropriately used in climate change analysis, or impact studies, in other dryland regions?*

  The spatial inconsistencies found in the CDD index indicate that its definition is not robust in highly variable climates such as the one of the south of Portugal. Accordingly, its use may not be appropriate in other Mediterranean regions, or dryland regions, even though that limitation has never been reported in the literature, most likely, because the low density of monitoring stations did not allow a comprehensive spatial analysis for most of the studies.

  Sillmann and Roeckner (2008) verified that the ensemble-mean of the model-based CDD index overestimates the maximum length of dry spells evaluated by most of the

observation-based indices in southern Europe. In these regions, the model overestimates CDD by more than 20 days, except in the 1950s when the simulated CDD is closer to the observed one. Sillmann and Roeckner (2008) argue that the abrupt change in the simulated CDD is not a real effect but can be attributed to a changed masking of missing values in the data used in the observation-based index. Our results indicate that it could also be attributed to the lack of robustness of the CDD in those regions, thus further research on this subject should be pursued.

REFERENCES

References

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J. (2003). *Guidelines on climate metadata and homogenization*. World Meteorological Organization, WMO-TD No. 1186, WCDMP No. 53, Geneva, Switzerland, 55 pp.

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* 111, D05109, doi:10.1029/2005JD006290.

Alexandersson H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology* 6, 661-675.

Alexandersson H, Moberg A. (1997). Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *Int. J. Climatol.* 17(1), 25-34.

Allen RJ, DeGaetano AT. (2000). A method to adjust long-term temperature extreme series for nonclimatic inhomogeneities. *J. Climate* 13, 3680-3695.

Allen RG, Pereira LS, Raes D, Smith M. (1998). "Statistical analysis of weather data sets". In: FAO - Food and Agriculture Organization of the United Nations (ed.), *Crop evapotranspiration - Guidelines for computing crop water requirements*, FAO Irrigation and drainage paper 56, Rome, Annex IV.

Anderson TW, Darling DA. (1952). Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes. *Ann. Math. Stat.* 23, 193-212.

Anderson, T.W. and Darling, D.A. (1954). A test of goodness-of-fit. *J. Am. Stat. Assoc.* 49, 765-769.

Auer I, Böhm R, Jurković A, Orlik A, Potzmann R, Schöner W, Ungersböck M, Brunetti M, Nanni T, Maugeri M, Briffa K, Jones P, Efthymiadis D, Mestre O, Moisselin J-M, Begert M, Brazdil R, Bochnicek O, Cegnar T, Gajić-Čapka M, Zaninović K, Majstorović Ž, Szalai S, Szentimrey T, Mercalli L. (2005). A new instrumental precipitation dataset for the greater alpine region for the period 1800-2002. *Int. J. Climatol.* 25(2), 139-166.

Barnes SL. (1973). *Mesoscale objective map analysis using weighting time-series observations*. NOAA Tech. Memo. ERL NSSL-62., U. S. Dept. of Commerce, 60 pp.

Begert M, Schlegel T, Kirchhofer W. (2005). Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int. J. Climatol.* 25(1), 65-80.

Beguería S, Vicente-Serrano SM. (2006). Mapping the hazard of extreme rainfall by peaks over threshold extreme value analysis and spatial regression techniques. *J. Appl. Meteorol. Clim.* 45(1), 108-124.

Benson GJ, Dambe D, Darnhofer T, Gommes R, Mwongela GN, Pedgley DE, Pérarnaud V. (1997). *Extreme agrometeorological events*. World Meteorological Organization, WMO-TD No. 836, CAgM Report No. 73, Geneva, Switzerland, 182 pp.

Berenson ML, Levine DM. (1996). *Basic business statistics: concepts and applications*. Sixth edition, Prentice Hall International Editions, USA.

Boardman J. (2006). Soil erosion science: Reflections on the limitations of current approaches. *Catena* 68, 73-86.

Bodri L, Čermák V. (2000). Prediction of extreme precipitation using a neural network: application to summer flood occurrence in Moravia. *Adv. Eng. Softw.* 31(5), 311-321.

Boer EPJ, de Beurs KM, Hartkamp AD. (2001). Kriging and thin plate splines for mapping climate variables. *International Journal of Applied Earth Observation and Geoinformation* 3(2), 146-154.

Boissonnade AC, Heitkemper LJ, Whitehead D. (2002). "Weather data: cleaning and enhancement". In: Dischel RS (ed.), *Climate Risk and the Weather Market: Financial Risk Management with Weather Hedges*, Risk Waters, 73-93.

Bonaccorso B, Cancelliere A, Rossi G. (2005). Detecting trends of extreme rainfall series in Sicily. *Advances in Geosciences* 2(1), 7-11.

Boulanger J-P, Martinez F, Penalba O, Segura EC. (2007). Neural network based daily precipitation generator (NNGEN-P). *Clim. Dynam.* 28, 307-324.

Bourennane H, King D, Couturier A, Nicoullaud B, Mary B, Richard G. (2007). Uncertainty assessment of soil water content spatial patterns using geostatistical simulations: An empirical comparison of a simulation accounting for single attribute and a simulation accounting for secondary information. *Ecol. Model.* 205, 323-335.

Brunetti M, Colacino M, Maugeri M, Nanni T. (2001). Trends in the daily intensity of precipitation in Italy from 1951 to 1996. *Int. J. Climatol.* 21(3), 299-316.

Brunsdon C, Mcclatchey J, Unwin DJ. (2001). Spatial variations in the average rainfall-altitude relationship in Great Britain: an approach using geographically weighted regression. *Int. J. Climatol.* 21(4), 455-466.

Buishand TA. (1982). Some methods for testing the homogeneity of rainfall records. *J. Hydrol.* 58, 11-27.

Carrera-Hernández JJ, Gaskin SJ. (2007). Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. *J. Hydrol.* 336, 231-249.

Caussinus H, Lyazrhi F. (1997). Choosing a linear model with a random number of change-points and outliers. *Ann. Inst. Stat. Math.* 49, 761-775.

Caussinus H, Mestre O. (1996). "New mathematical tools and methodologies for relative homogeneity testing". In: *Proceedings of the Seminar for Homogenization of Surface Climatological Data*, Budapest, 6–12 October, 63-82.

Ceballos A, Martínez-Fernández J, Luengo-Ugidos JA. (2004). Analysis of rainfall trends and dry periods on a pluviometric gradient representative of Mediterranean climate in the Duero Basin, Spain. *J. Arid. Environ.* 58(2), 215-233.

Cleveland WS. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74, 829-836.

Correia FN. (2004). *Desertificação em Portugal: incidência no ordenamento do território e no desenvolvimento urbano*. Vol. I: Caracterização dos processos de desertificação e tipologia das zonas afectadas, DGOTDU – Direcção Geral do Ordenamento do território e Desenvolvimento Urbano, Maio de 2004, 185 pp.

Corte-Real J, Qian B, Xu H. (1998). Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation. *Int. J. Climatol.* 18, 619-635.

Costa AC, Soares A. (2006). "Identification of inhomogeneities in precipitation time series using SUR models and the Ellipse test". In: Caetano M, Painho M (eds.), *Proceedings of Accuracy 2006 - 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Instituto Geográfico Português, 419-428.

Costa AC, Soares A. (2007). "Space-time interpolation and uncertainty assessment of an extreme precipitation index using geostatistical cosimulation". In: *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)*, IEEE Computer Society, October 2007, 589-594.

Costa AC, Negreiros J, Soares A. (2008a). "Identification of inhomogeneities in precipitation time series using stochastic simulation". In: Soares A, Pereira MJ, Dimitrakopoulos R (eds.), *geoENV VI – Geostatistics for Environmental Applications*, Springer, 275-282.

Costa AC, Durão R, Soares A, Pereira MJ. (2008b). A geostatistical exploratory analysis of precipitation extremes in southern Portugal. *RevStat: Statistical Journal* 6(1), 21-32.

Costa AC, Durão R, Pereira MJ, Soares A. (2008c). Using stochastic space-time models to map extreme precipitation in southern Portugal. *Natural Hazards and Earth System Sciences* 8(4), 763-773.

Craddock JM. (1979). Methods of comparing annual rainfall records for climatic purposes. *Weather* 34, 332-346.

Cramér H. (1928). On the Composition of Elementary Errors. *Skand. Aktuarietidskr.* 11, 13-74 and 141-180.

Cressman GP. (1959). An operational objective analysis system. *Mon. Weather Rev.* 87(10), 367-374.

Cubasch U, von Storch H, Waszkewitz J, Zorita E. (1996). Estimates of climate change in Southern Europe derived from dynamical climate model output. *Climate Res.* 7, 129-149.

Daly C, Neilson RP, Phillips DL. (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteorol.* 33, 140-157.

Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. (2002). A knowledge-based approach to the statistical mapping of climate. *Climate Res.* 22, 99-113.

Daly C, Gibson WP, Doggett M, Smith J, Taylor G. (2004). "Up-to-date monthly climate maps for the conterminous United States". In: American Meteorological Society (eds.), *Proceedings of 14th AMS Conference on Applied Climatology*, 84th AMS Annual Meeting Combined Preprints, Paper P5.1, Seattle (WA), USA.

Daly C. (2006). Guidelines for assessing the suitability of spatial climate data sets. *Int. J. Climatol.* 26(6), 707-721.

Dawson CW, Wilby RL. (2001). Hydrological modelling using artificial neural networks. *Prog. Phys. Geog.* 25(1), 80-108.

De Luís M, García-Cano MF, Cortina J, Raventós J, González-Hidalgo JC, Sánchez JR. (2001). Climatic trends, disturbances and short-term vegetation dynamics in a Mediterranean shrubland. *Forest Ecol. Manag.* 147(1), 25-37.

Deutsch CV, Journel AG. (1997). *GSLIB: Geostatistical Software Library and User's Guide*. 2nd Edition, Oxford University Press, New York.

de Vos NJ, Rientjes HM. (2005). Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrol. Earth Syst. Sc.* 9, 111-126.

Diodato N. (2005). The influence of topographic co-variables on the spatial variability of precipitation over small regions of complex terrain. *Int. J. Climatol.* 25(3), 351-363.

Diodato N, Ceccarelli M. (2004). Multivariate indicator kriging approach using a GIS to classify soil degradation for Mediterranean agricultural lands. *Ecol. Indic.* 4, 177-187.

Drogue G, Humbert J, Deraisme J, Mahr N, Freslon N. (2002). A statistical-topographic model using an omnidirectional parameterization of the relief for mapping orographic rainfall. *Int. J. Climatol.* 22(5), 599-613.

Ducré-Robitaille J-F, Vincent LA, Boulet G. (2003). Comparison of techniques for detection of discontinuities in temperature series. *Int. J. Climatol.* 23(9), 1087-1101.

Dünkeloh A, Jacobeit J. (2003). Circulation dynamics of Mediterranean precipitation variability 1948–98. *Int. J. Climatol.* 23, 1843-1866.

Durão R, Costa AC, Pereira MJ, Soares A. (2007). "Indices of precipitation extremes in Southern Portugal – a geostatistical approach". In: *Abstracts of the European Geosciences Union 9th Plinius Conference on Mediterranean Storms*, Varenna, Italy, 10-13 September, 2007.

Durbin J, Watson GS. (1950). Testing for serial correlation in least squares regression I. *Biometrika* 37, 409-428.

Durbin J, Watson GS. (1951). Testing for serial correlation in least squares regression II. *Biometrika* 38, 159-178.

Durrans SR, Kirby JT. (2004). Regionalization of extreme precipitation estimates for the Alabama rainfall atlas. *J. Hydrol.* 295, 101-107.

Easterling DR, Peterson TC. (1995a). A new method for detecting and adjusting for undocumented discontinuities in climatological time series. *Int. J. Climatol.* 15, 369-377.

Easterling DR, Peterson TC. (1995b). The effect of artificial discontinuities on recent trends in minimum and maximum temperatures. *Atmos. Res.* 37, 19-26.

Easterling DR, Diaz HF, Douglas AV, Hogg WD, Kunkel KE, Rogers JC, Wilkinson JF. (1999). Long-term observations for monitoring extremes in the Americas. *Climatic Change* 42(1), 285-308.

Eischeid JK, Pasteris PA, Diaz HF, Plantico MS, Lott NJ. (2000). Creating a serially complete national daily time series of temperature and precipitation for the Western United States. *J. Appl. Meteorol.* 39(9), 1580-1591.

Emery X. (2004). Properties and limitations of sequential indicator simulation. *Stoch. Envir. Res. Risk Ass.* 18(6), 414-424.

Engle RF. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987-1007.

Enne G, Zucca C. (2000). *Desertification indicators for the European Mediterranean region: state of the art and possible methodological approaches*. ANPA – Agenzia Nazionale per la Protezione dell' Ambiente, NRD – Nucleo di Ricerca sulla Desertificazione, Università degli Studi di Sassari, Roma, Italy, 261 pp.

Faulkner DS, Prudhomme C. (1998). Mapping an index of extreme rainfall across the UK. *Hydrol. Earth Syst. Sc.* 2, 183-194.

Feng S, Hu Q, Qian W. (2004). Quality control of daily meteorological data in China, 1951-2000: a new dataset. *Int. J. Climatol.* 24(7), 853-870.

Fragoso M, Gomes PT. (2008). Classification of daily abundant rainfall patterns and associated large-scale atmospheric circulation types in Southern Portugal. *Int. J. Climatol.* 28(4), 537-544,.

Franco C, Soares A, Delgado J. (2006). Geostatistical modelling of heavy metal contamination in the topsoil of Guadiamar river margins (S Spain) using a stochastic simulation technique. *Geoderma* 136, 852-864.

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Res.* 19(3), 193-212.

Gallant AR, Goebel JJ. (1976). Nonlinear regression with autoregressive errors. *J. Am. Stat. Assoc.* 71, 961-967.

Gathara ST. (2006). "Overview". In: Gathara ST, Gringof LG, Mersha E, Sinha Ray KC, Spasov P (eds.), *Impacts of desertification and drought and other extreme meteorological*

events, World Meteorological Organization, WMO-TD No. 1343, CAgM Report No. 101, Geneva, Switzerland, 2-11

Geist HJ, Lambin EF. (2004). Dynamic Causal Patterns of Desertification. *Bioscience* 54(9), 817-829.

Goodess CM, Jones PD. (2002). Links between circulation and changes in the characteristics of Iberian rainfall. *Int. J. Climatol.* 22(13), 1593-1615.

Goovaerts P. (1997). *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics Series, Oxford University Press.

Goovaerts P. (1999). Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena* 34, 227-242.

Goovaerts P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* 228, 113-129.

Goswami P, Srividya. (1996). A novel neural network design for long range prediction of rainfall pattern. *Current Science* 70(6), 447-457.

Gouveia C, DaCamara C, Trigo RM. (2008). Droughts monitoring in Portugal using satellite data. *Geophysical Research Abstracts* 10, EGU2008-A-10201.

Greene WH. (2003). *Econometric Analysis*. Fifth edition, Prentice Hall International Edition, New Jersey: Pearson Education.

Griffiths WE, Hill RC, Judge GG. (1993). *Learning and Practicing of Econometrics*. New York: John Wiley & Sons.

Griffiths GM, Salinger MJ, Leleu I. (2003). Trends in extreme daily rainfall across the South Pacific and relationship to the South Pacific Convergence Zone. *Int. J. Climatol.* 23(8), 847-869.

Gringof IG, Mersha E. (2006). "Assessment of desertification, drought and other extreme meteorological events". In: Gathara ST, Gringof LG, Mersha E, Sinha Ray KC, Spasov P (eds.), *Impacts of desertification and drought and other extreme meteorological events*, World Meteorological Organization, WMO-TD No. 1343, CAgM Report No. 101, Geneva, Switzerland, 12-29.

Guan H, Wilson JL, Makhnin O. (2005). Geostatistical mapping of mountain precipitation incorporating autosearched effects of terrain and climatic characteristics. *J. Hydrometeorol.* 6(6), 1018-1031.

Guttman NB. (1998). “Homogeneity, data adjustments, and climatic normals”. In: *7th International Meeting on Statistical Climatology*, 25-29 May 1998, Whistler, British Columbia, Canada.

Haberlandt U. (2007). Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J. Hydrol.* 332, 144-157.

Hasenauer H, Merganicova K, Petritsch R, Pietsch SA, Thornton PE. (2003). Validating daily climate interpolations over complex terrain in Austria. *Agr. Forest. Meteorol.* 119, 87-107.

Haylock MR, Goodess CM. (2004). Interannual variability of European extreme winter rainfall and links with mean large-scale circulation. *Int. J. Climatol.* 24(6), 759-776.

Haylock M, Nicholls N. (2000). Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. *Int. J. Climatol.* 20(13), 1533-1541.

Heim Jr RR. (2000). "Drought indices: A review". In: Wilhite DA (ed.), *Drought: A Global Assessment*, Routledge Hazards and Disasters Series, Vol. 1, Ch. 11, Routledge, London, 159-167.

Heim Jr RR. (2002). A review of Twentieth-Century drought indices used in the United States. *B. Am. Meteorol. Soc.* 83, 1149-1165.

Herrmann SM, Hutchinson CF. (2005). The changing contexts of the desertification debate. *J. Arid. Environ.* 63(3), 538-555.

Hevesi JA, Istok JD, Flint AL. (1992). Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis. *J. Appl. Meteorol.* 31, 661-676.

Hidalgo JC, De Luís M, Raventós J, Sánchez JR. (2003a). Daily rainfall trend in the Valencia region of Spain. *Theor. Appl. Climatol.* 75, 117-130.

Hidalgo JC, Luís M, Vicente SM, Štepánek P, Raventós J, Saz MA, Cuadrat JM, Ferraz J, Creus J. (2003b). "Monthly precipitation data base in Mediterranean climate area of Spain. Reconstruction processes and quality control". In: *Fourth Seminar for Homogenization and Quality Control in Climatological Databases*, Budapest, Hungary, WMO-TD No. 1236, WCDMP No. 56, 105-116.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965-1978.

Hirsch RM, Slack JR. (1984). A nonparametric trend test for seasonal data with serial dependence. *Water Resour. Res.* 20(6), 727-732.

Hirsch RM, Slack JR, Smith RA. (1982). Techniques of trend analysis for monthly water quality data. *Water Resour. Res.* 18(1), 107-121.

Hisdal H, Tallaksen LM (eds.). (2000). *Drought Event Definition*. ARIDE – Assessment of the Regional Impact of Droughts in Europe, ARIDE Technical Report No. 6, Department of Geophysics, University of Oslo, Norway, 45 pp.

Hounam CE, Burgos JJ, Kalik MS, Palmer WC, Rodda J. (1975). *Drought and Agriculture*. Report of the CAgM Working Group on Assessment of Drought, Technical Note No. 138. WMO Publication No. 392, 127 pp.

Hundecha Y, Bárdossy A. (2005). Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century. *Int. J. Climatol.* 25, 1189-1202.

Hutchinson, MF. (1995). Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Syst.* 9, 385-403.

IPCC (2007). "Summary for Policymakers". In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Isaaks EH, Srivastava RM. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press.

Jabbar MT, Chen X. (2006). Land degradation assessment with the aid of geo-information techniques. *Earth Surf. Proc. Land.* 31(6), 777-784.

Johansson B, Chen D. (2003). The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. *Int. J. Climatol.* 23(12), 1523-1535.

Journel AG. (1994). "Modelling uncertainty: some conceptual thoughts". In: Dimitrakopoulos R (ed.), *Geostatistics for the Next Century*, Kluwer Academic Pub., Dordrecht, The Netherlands, 30-43.

Kendall MG. (1975). *Rank correlation methods*. Charles Griffin, London.

Kiktev D, Sexton DMH, Alexander L, Folland CK. (2003). Comparison of modeled and observed trends in indices of daily climate extremes. *J. Climate* 16(22), 3560-3571.

Klein Tank AMG, Können GP. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J. Climate* 16(22), 3665-3680.

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Forland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, López JA, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *Int. J. Climatol.* 22(12), 1441-1453.

Kogan FN. (2000). "Contribution of remote sensing to drought early warning". In: Wilhite DA, Sivakumar MVK, Wood DA (eds.), *Early Warning Systems for Drought Preparedness and Drought Management*, World Meteorological Organization, AGM-2, WMO-TD No. 1037, Geneva, Switzerland, 75-87.

Kohler MA. (1949). Double-mass analysis for testing the consistency of records and for making adjustments. *Bull. Amer. Meteorol. Soc.* 30, 188-189.

Kolmogorov AN. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari* 4, 83-91. Engl. transl. in Selected Works, II. Probability Theory and Mathematical Statistics, Shiryayev AN (ed.), Kluwer, 1992, 139-146.

Kosmas C, Kirkby MJ, Geeson N. (1999). *Manual on key indicators of desertification and mapping environmentally sensitive areas to desertification*. Medalus project – Mediterranean desertification and land use, European Commission Publication EUR 18882, Luxembourg, 87 pp.

Kostopoulou E, Jones PD. (2005). Assessment of climate extremes in the Eastern Mediterranean. *Meteorol. Atmos. Phys.* 89, 69-85.

Krige DG. (1951). A statistical approach to some mine valuations problems at the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa* 52, 119-138.

Kruskal WH. (1952). A nonparametric test for the several sample problem. *Ann. Math. Stat.* 23, 525-540.

Kruskal WH, Wallis WA. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583-621.

Kumar ARS, Sudheer KP, Jain SK, Agarwal PK. (2005). Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrol. Process.* 19(6), 1277-1291.

Květoň V, Žák M. (2003). "Experience regarding homogenisation of temperature time series in the Czech Republic for period 1961-2000". In: *Fourth Seminar for Homogenization and Quality Control in Climatological Databases*, Budapest, Hungary, WMO-TD No. 1236, WCDMP No. 56, 135-142.

Kyriakidis PC, Kim J, Miller NL. (2001). Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics. *J. Appl. Meteorol.* 40(11), 1855-1877.

Kyriakidis PC, Miller NL, Kim J. (2004). A spatial time series framework for simulating daily precipitation at regional scales. *J. Hydrol.* 297, 236-255.

Lana X, Martínez MD, Burgueño A, Serra C, Martín-Vide J, Gómez L. (2008). Spatial and temporal patterns of dry spell lengths in the Iberian Peninsula for the second half of the twentieth century. *Theor. Appl. Climatol.* 91, 99-116.

Lantieri D. (2003). *Potential use of satellite remote sensing for land degradation assessment in drylands. Application to the LADA Project*. LADA – Land Degradation Assessment in Drylands, draft report, 77 pp.

Lanzante JR. (1996). Resistant, robust and nonparametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.* 16, 1197-1226.

Lázaro R, Rodrigo FS, Gutiérrez L, Domingo F, Puigdefábregas J. (2001). Analysis of a 30-year rainfall record (1967–1997) in semi–arid SE Spain for implications on vegetation. *J. Arid. Environ.* 48(3), 373-395.

Lee SH, Maeng SJ. (2003). Frequency analysis of extreme rainfall using L-moment. *Irrig. Drain.* 52(3), 219-230.

Lettenmaier DP. (1988). Multivariate nonparametric tests for trend in water quality. *Water Resour. Bull.* 24(3), 505-512.

Leuangthong O, Deutsch CV. (2003). Stepwise conditional transformation for simulation of multiple variables. *Math. Geol.* 35(2),155-173.

Libiseller C, Grimvall A. (2002). Performance of partial Mann-Kendall test for trend detection in the presence of covariates. *Environmetrics* 13, 71-84.

Llasat M-C, Quintas L. (2004). Stationarity of monthly rainfall series since the middle of the XIXth century. Application to the Case of Peninsular Spain. *Natural Hazards* 31, 613-622.

Lloyd CD. (2005). Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *J. Hydrol.* 308, 128-150.

Lonergan S. (2005). The role of UNEP in desertification research and mitigation. *J. Arid. Environ.* 63, 533-534.

Loureiro NS, Coutinho MA. (1995). Rainfall changes and rainfall erosivity increase in the Algarve (Portugal). *Catena* 24(1), 55-67.

Loureiro NS, Coutinho MA. (2001). A new procedure to estimate the RUSLE EI_{30} index, based on monthly rainfall data and applied to the Algarve region, Portugal. *J. Hydrol.* 250, 12-18.

Luk KC, Ball JE, Sharma A. (2000). A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *J. Hydrol.* 227, 56-65.

Mann HB. (1945). Non-parametric test against trend. *Econometrika* 13, 245-259.

Mann HB, Whitney DR. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18, 50-60.

Marquínez J, Lastra J, García P. (2003). Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *J. Hydrol.* 270, 1-11.

Martínez-Casasnovas JA, Ramos MC, Ribes-Dasi M. (2002). Soil erosion caused by extreme rainfall events: mapping and quantification in agricultural plots from very detailed digital elevation models. *Geoderma* 105, 125-140.

Martínez-Cob A. (1996). Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.* 174, 19-35.

McKee TB, Doesken NJ, Kleist J. (1993). "The relationship of drought frequency and duration to time scales". In: *8th Conference on Applied Climatology*, Anaheim, California, American Meteorological Society, Preprints, 179-184.

Mekis É, Vincent L. (2003). "New developments in the homogenization of precipitation and temperature series in Canada". In: *Fourth Seminar for Homogenization and Quality Control in Climatological Databases*, Budapest, Hungary, WMO-TD No. 1236, WCDMP No. 56, 39-62.

Mendicino G, Versace P. (2007). Integrated drought watch system: a case study in southern Italy. *Water Resour. Manag.* 21, 1409-1428.

Menne MJ, Williams CN. (2005). Detection of undocumented changepoints using multiple test statistics and composite reference series. *J. Climate* 18(20), 4271-4286.

Millán MM, Estrela MJ, Sanz MJ, Mantilla E, Martín M, Pastor F, Salvador R, Vallejo R, Alonso L, Gangoiti G, Ilardia JL, Navazo M, Albizuri A, Artíñano B, Ciccioli P, Kallos G, Carvalho RA, Andrés D, Hoff A, Werhahn J, Seufert G, Versino B. (2005). Climatic feedbacks and desertification: the Mediterranean model. *J. Climate* 18, 684-701.

Miranda PMA, Valente MA, Tomé AR, Trigo R, Coelho MF, Aguiar A, Azevedo EB. (2006). "O clima de Portugal nos séculos XX e XXI". In: Santos FD, Miranda P (eds.), *Alterações climáticas em Portugal. Cenários, impactos e medidas de adaptação*, Projecto SIAM II, Gradiva, Lisboa, 49-113.

Moberg A, Jones PD. (2005). Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. *Int. J. Climatol.* 25(9), 1149-1171.

Moberg A, Jones PD, Lister D, Walther A, Brunet M, Jacobeit J, Alexander LV, Della-Marta PM, Luterbacher J, Yiou P, Chen D, Klein Tank AMG, Saladié O, Sigró J, Aguilar E, Alexandersson H, Almarza C, Auer I, Barriendos M, Begert M, Bergström H, Böhm R, Butler CJ, Caesar J, Drebs A, Founda D, Gerstengarbe F-W, Micela G, Maugeri M, Österle H, Pandzic K, Petrakis M, Srnec L, Tolasz R, Tuomenvirta H, Werner PC, Linderholm H, Philipp A, Wanner H, Xoplaki E. (2006). Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000. *Journal of Geophysical Research* 111, D22106, doi:10.1029/2006JD007103.

Modarres R, Silva VPR. (2007). Rainfall trends in arid and semi-arid regions of Iran. *J. Arid. Environ.* 70(2), 344-355.

Moreira EE, Paulo AA, Pereira LS, Mexia JT. (2006). Analysis of SPI drought class transitions using loglinear models. *J. Hydrol.* 331, 349-359.

Müller-Westermeier G. (2003). "Statistical analysis of results of homogeneity testing and homogenization of long climatological time series in Germany". In: *Fourth Seminar for Homogenization and Quality Control in Climatological Databases*, Budapest, Hungary, WMO-TD No. 1236, WCDMP No. 56, 25-38.

New M, Hulme M, Jones P. (2000). Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate* 13, 2217-2238.

Nicolau R. (1999). *Avaliação da qualidade de séries de registos de precipitação anual com vista à selecção, ajustamento e completamento das mesmas para posterior modelação da variabilidade espacial da pluviosidade*. CNIG – Centro Nacional de Investigação Geográfica, Estudo, 053C/19, Janeiro de 1999.

Nicolau, R. (2002). *Modelação e mapeamento da distribuição espacial da precipitação - uma aplicação a Portugal continental*. Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa. Dissertação apresentada para obtenção do Grau de Doutor, 356 pp.

Nicolau R, Ribeiro L, Rodrigues RR, Pereira HG, Câmara AS. (2002). "Mapping the spatial distribution of rainfall in Portugal". In: Kleingeld WJ, Krige DG (eds.), *Geostatistics 2000*, Cape Town, Vol. 2, Geostatistical Association of Southern Africa, South Africa, 548-558.

Ninyerola M, Pons X, Roure JM. (2007). Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theor. Appl. Climatol.* 89, 195-209.

Ó A, Roxo MJ. (2001): "Driving forces of land use changes in Alentejo and its impact on soil and water". In: Project RICAMARE (ed.), *Workshop on Land Use Changes & Cover and Water Resources in the Mediterranean Region*, Médenine, Tunisia, DG Environment, European Union.

Oliver DS. (2003). Gaussian cosimulation: modelling of the cross-covariance. *Math. Geol.* 35(6), 681-698.

Oz B, Deutsch CV, Tran TT, Xie Y. (2003). DSSIM-HR: A FORTRAN 90 program for direct sequential simulation with histogram reproduction. *Comput. Geosci.* 29, 39-51.

Palmer WC. (1965). *Meteorological drought*. Research Paper No. 45, U.S. Weather Bureau.

PANCD (1999). Portugal, Resolução do Conselho de Ministros nº 69/99, de 9 de Julho de 1999, Aprova o Programa de Acção Nacional de Combate à Desertificação (PANCD) e estabelece procedimentos relativamente à sua concretização, Diário da República Portuguesa, Série I-B, n.º 158/99, p. 4300.

Paulo AA, Pereira LS, Matias PG. (2003). "Analysis of local and regional droughts in southern Portugal using the theory of runs and the Standardised Precipitation Index". In: Rossi G, Cancelliere A, Pereira LS, Oweis T, Shatanawi M, Zairi A (eds.), *Tools for Drought Mitigation in Mediterranean Regions*, Kluwer, Dordrecht, 55-78.

Paulo AA, Ferreira E, Coelho C, Pereira LS. (2005). Drought class transition analysis through Markov and loglinear models, an approach to early warning. *Agr. Water Manage.* 77, 59-81.

Pereira LS, Louro V, Rosário L, Almeida A. (2006). "Desertification, territory and people, a holistic approach in the Portuguese context". In: Kepner WG, Rubio JL, Mouat DA, Pedrazzini F (eds.), *Desertification in the Mediterranean Region: a Security Issue*, NATO Sc.Com., AK/Nato Publishing Unit, Springer-Verlag, Dordrecht, 269-289.

Perry M, and Hollis D. (2005). The generation of monthly gridded datasets for a range of climatic variables over the UK. *Int. J. Climatol.* 25(8), 1041-1054.

Peterson TC. (2005). Climate change indices. *WMO Bulletin* 54(2), 83-86.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Forland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. (1998). Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.* 18(13), 1493-1517.

Peterson TC, Folland C, Gruza G, Hogg W, Mokssit A, Plummer N. (2001). *Report on the activities of the Working Group on Climate Change Detection and Related Rapporteurs 1998–2001*. World Meteorological Organization, WCDMP-No. 47 / WMO-TD No. 1071, Geneva, Switzerland, 143 pp.

Pettit AN. (1979). A non-parametric approach to the change-point detection. *Applied Statistics* 28(2), 126-135.

Piccarreta M, Capolongo D, Boenzi F. (2004). Trend analysis of precipitation and drought in Basilicata from 1923 to 2000 within a southern Italy context. *Int. J. Climatol.* 24(7), 907-922.

Pimenta MT. (1998). *Caracterização da Erodibilidade dos Solos a Sul do Rio Tejo*. Instituto da Água (INAG/DSRH), Lisboa, Julho 1998.

Pimenta MT, Santos MJ, Rodrigues R. (1997). "A proposal of indices to identify desertification prone areas". In: *Jornadas de reflexión sobre el Anexo IV de aplicatión para el Mediterrâneo Norte – Convenio de Lucha contra la Desertificación*, Murcia, Spain, 22-23 May, 1997.

Potter KW. (1981). Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Wea. Rev.* 109, 2040-2045.

Price DT, Mckenney DW, Nalder IA, Hutchinson MF, Kesteven JL. (2000). A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agr. Forest. Meteorol.* 101, 81-94.

Prudhomme C, Reed DW. (1998). Relationships between extreme daily precipitation and topography in a mountainous region: A case study in Scotland. *Int. J. Climatol.* 18(13), 1439-1453.

Prudhomme C, Reed DW. (1999). Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *Int. J. Climatol.* 19(12), 1337-1356.

Puigdefábregas J, Mendizabal T. (1998). Perspectives on desertification: western Mediterranean. *J. Arid. Environ.* 39, 209-224.

Ramírez MCV, Velho HFC, Ferreira NJ. (2005). Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J. Hydrol.* 301, 146-162.

Ramos C, Reis E. (2002). Floods in southern Portugal: their physical and human causes, impacts and human response. *Mitigation and Adaptation Strategies for Global Change* 7(3), 267-284.

Reek T, Doty SR, Owen TW. (1992). A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *B. Am. Meteorol. Soc.* 73(6), 753-762.

Reeves J, Chen J, Wang XL, Lund R, Lu Q. (2007). A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteorol. Clim.* 46, 900-915.

Robertson RK, Mueller UA, Bloom LM. (2006). Direct sequential simulation with histogram reproduction: A comparison of algorithms. *Comput. Geosci.* 32(3), 382-395.

Rodrigo FS, Trigo RM. (2007). Trends in daily rainfall in the Iberian Peninsula from 1951 to 2002. *Int. J. Climatol.* 27, 513-529.

Rodrigues R. (1990). *Caracterização de Episódios Meteorológicos Extremos. O Sotavento Algarvio*. Direcção Geral dos Recursos Naturais, Direcção dos Serviços de Hidrologia, Publicação Nº 10/90, Lisboa, Junho 1990.

Rodriguez R, Llasat M-C, Wheeler D. (1999). Analysis of the Barcelona precipitation series 1850-1991. *Int. J. Climatol.* 19(7), 787-801.

Romero R, Guijarro JA, Ramis C, Alonso S. (1998). A 30-year (1964-1993) daily rainfall data base for the Spanish Mediterranean regions: First exploratory study. *Int. J. Climatol.* 18(5), 541-560.

Rosário L. (2004a). "Sobre os indicadores de desertificação para Portugal continental". In: Louro V (ed.), *Desertificação. Sinais, dinâmicas e sociedade*, Estudos e Documentos, 11, Instituto Piaget, 129-151.

Rosário L. (2004b). *Indicadores de desertificação para Portugal Continental*. Direcção-Geral dos Recursos Florestais (Ed.), Lisboa, Maio 2004, 56 pp.

Rousseeuw PJ, Croux C. (1993). Alternatives to the median absolute deviation. *J. Am. Stat. Assoc.* 88(424), 1273-1283.

Roxo MJ, Mourão JM, Rodrigues L, Casimiro PC. (1999). "The Alentejo region (Mértola municipality, Portugal)". In: Kosmas C, Kirkby MJ, Geeson N (eds.), *Manual on key indicators of desertification and mapping environmentally sensitive areas to desertification*, Medalus project – Mediterranean desertification and land use, European Commission Publication EUR 18882, Luxembourg, 80-87.

Russo A, Trigo R, Soares A. (2008). "Stochastic modelling applied to air quality space-time characterization". In: Soares A, Pereira MJ, Dimitrakopoulos R (eds.), *geoENV VI – Geostatistics for Environmental Applications*, Springer, 83-93.

Ruud PA. (2000). *An Introduction to Classical Econometric Theory*. New York: Oxford University Press.

Santo FE, Pires VC, Silva A. (2004). "Clima de Portugal continental". In: Louro V (ed.), *Desertificação. Sinais, dinâmicas e sociedade*, Estudos e Documentos, 11, Instituto Piaget, 89-101.

Santo FE, Guerreiro R, Pires VC, Pessanha LEV, Gomes IM. (2005). "Monitoring agricultural drought in mainland Portugal". In: Boken VK, Cracknell AP, Heathcote RL (eds.), *Monitoring and Predicting Agricultural Drought: A Global Study*, Oxford University Press, New York, 181-195.

Santos MJ. (1998). *Caracterização e monitorização de secas*. Instituto da Água (INAG/DSRH), Estudo, 052D/68, Lisboa, Dezembro de 1998.

Santos MJ, Henriques R. (1999). *Analysis of the European annual precipitation series*. ARIDE – Assessment of the Regional Impact of Droughts in Europe, Technical Report No. 3, Water Institute, DSRH, Lisbon, Portugal, July 1999.

Santos JA, Corte-Real J, Ulbrich U, Palutikof J. (2007). European winter precipitation extremes and large-scale circulation: a coupled model and its scenarios. *Theor. Appl. Climatol.* 87, 85-102.

SAS Institute Inc. (1999a). *SAS/ETS User's Guide*. Version 8, Cary, NC: SAS Institute Inc., 1546 pp.

SAS Institute Inc. (1999b). *SAS Procedures Guide*. Version 8, Cary, NC: SAS Institute Inc., 1729 pp.

Scaife AA, Folland CK, Alexander LV, Moberg A, Knight JR. (2008). European climate extremes and the North Atlantic Oscillation. *J. Climate* 21, 72-83.

Schreiber K, Harrison J, Sterk G. (2008). "Agrometeorological aspects of desertification". In: *Guide to Agricultural Meteorological Practices (GAMP)*, Draft 3rd Edition, WMO-No.134, Geneva, Switzerland, Ch. 10.

Shapiro SS, Wilk MB. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611.

Sillmann J, Roeckner E. (2008). Indices for extreme events in projections of anthropogenic climate change. *Climatic Change* 86, 83-104.

Sivakumar MVK. (2007). Interactions between climate and desertification. *Agr. Forest. Meteorol.* 142, 143-155.

Smakhtin VU, Hughes DA. (2004). *Review, Automated Estimation and Analyses of Drought Indices in South Asia*. Working Paper 83, International Water Management Institute, Colombo: Sri Lanka, 32 pp.

Smirnov NV. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. (Russian) *Bulletin of Moscow University* 2, 3-16.

Smith RB, Barstad I. (2004). A linear theory of orographic precipitation. *J. Atmos. Sci.* 61(12), 1377-1391.

Soares A. (2001). Direct sequential simulation and cosimulation. *Math. Geol.* 33(8), 911-926.

Solow A. (1987). Testing for climatic change: an application of the two phase regression model. *J. Climate Appl. Meteorol.* 26, 1401-1405.

Steinemann A, Hayes MJ, Cavalcanti L. (2005). "Drought indicators and triggers". In: Wilhite DA (ed.), *Drought and Water Crises: Science, Technology, and Management Issues*, CRC Press, 71-92.

Steinemann A, Cavalcanti L. (2006). Developing multiple indicators and triggers for drought plans. *J. Water Resour. Plng. and Mgmt.* 132(3), 164-174.

Szentimrey T. (1994). “Statistical problems connected with the homogenization of climatic time series”. In: Heino R (ed.), *Climate Variations in Europe*, Proceedings of the European Workshop held in Kirkkonummi (Majvik), Finland, May 1994, Publications of the Academy of Finland 3:94, 330-339.

Szentimrey T. (1995a). "Statistical methods for detection of inhomogeneities". In: *Proceedings of the Regional Workshop on Climate Variability and Climate Change Vulnerability and Adaptation*, Prague, 293-298.

Szentimrey T. (1995b). “General problems of the estimation of inhomogeneities, optimal weighting of the reference stations”. In: *Proceedings of the 6th International Meeting on Statistical Climatology*, Galway, Ireland, 629-631.

Szentimrey T. (1996). “Statistical procedure for joint homogenization of climatic time series”. In: *Proceedings of the Seminar for Homogenization of Surface Climatological Data*, Budapest, Hungary, 47-62.

Szentimrey T. (1999). "Multiple analysis of series for homogenization (MASH)". In: *Proceedings of the Second Seminar for Homogenization of Surface Climatological Data*, Budapest, Hungary, WMO-TD No. 962, WCDMP No. 41, 27-46.

Szentimrey T. (2003). "Multiple analysis of series for homogenization (MASH); Verification procedure for homogenized time series". In: *Fourth Seminar for Homogenization and Quality Control in Climatological Databases*, Budapest, Hungary, WMO-TD No. 1236, WCDMP No. 56, 193-201.

Tarhule A, Woo M-K. (1998). Changes in rainfall characteristics in northern Nigeria. *Int. J. Climatol.* 18(11), 1261-1271.

Tayanç M, Dalfes HN, Karaca M, Yenigün O. (1998). A comparative assessment of different methods for detecting inhomogeneities in Turkish temperature data set. *Int. J. Climatol.* 18(5), 561-578.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. (2006). Going to extremes: an intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change* 79, 185-211.

Teegavarapu RSV, Chandramouli V. (2005). Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *J. Hydrol.* 312, 191-206.

Thiessen AH. (1911). Precipitation averages for large areas. *Monthly Weather Rev.* 39(7), 1082-1084.

Thornton PE, Running SW, White MA. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.* 190, 214-251.

Trigo RM, DaCamara C. (2000). Circulation Weather Types and their impact on the precipitation regime in Portugal. *Int. J. Climatol.* 20(13), 1559-1581.

Trigo RM, Osborn TJ, Corte-Real JM. (2002). The North Atlantic Oscillation influence on Europe climate impacts and associated physical mechanisms. *Climate Res.* 20, 9-17.

Trigo RM, Pozo-Vázquez D, Osborn TJ, Castro-Díez Y, Gámiz-Fortis S, Esteban-Parra MJ. (2004). North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula. *Int. J. Climatol.* 24(8), 925-944.

Tuomenvirta H. (2001). Homogeneity adjustments of temperature and precipitation series - Finnish and Nordic data. *Int. J. Climatol.* 21(4), 495-506.

Türkes M. (1999). Vulnerability of Turkey to desertification with respect to precipitation and aridity conditions. *Tr. J. of Engineering and Environmental Sciences* 23, 363-380.

UNCCD – United Nations Convention to Combat Desertification (1994). *Elaboration of an International Convention to Combat Desertification in countries experiencing serious drought and/or desertification, particularly in Africa*. Final text of the Convention, United Nations General Assembly, A/AC.241/27, 12 September 1994.

UNEP (1992). *World Atlas of Desertification*. United Nations Environmental Program, Edward Arnold, London.

UNEP (1997). *World Atlas of Desertification*. 2 ed., United Nations Environmental Program, Nairobi.

Vargas-Guzmán JA, Dimitrakopoulos R. (2003). Successive nonparametric estimation of conditional distributions. *Math. Geol.* 35(1), 39-52.

Verly G. (1993). "Sequential Gaussian co-simulation: A simulation method integrating several types of information". In: Soares A (ed.), *Geostatistics Troia' 92*, Vol. 1, Kluwer Academic Pub., Dordrecht, The Netherlands, 543-554.

Vincent L. (1998). A technique for the identification of inhomogeneities in Canadian temperature series. *J. Climate* 11, 1094-1104.

Vicente-Serrano SM, Cuadrat-Prats JM. (2007). Trends in drought intensity and variability in the middle Ebro valley (NE of the Iberian peninsula) during the second half of the twentieth century. *Theor. Appl. Climatol.* 88, 247-258.

Vicente-Serrano SM, Saz-Sánchez MA, Cuadrat JM. (2003). Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Res.* 24(2), 161-180.

Von Mises R. (1931). *Wahrscheinliehkeitsrechnung und ihre Anwendung in der Statistik undtheoretischen Physik*. Leipzik, F.Duticke.

Von Neumann J. (1941). Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Stat.* 13, 367-395.

Wald A, Wolfowitz J. (1943). An exact test for randomness in the non-parametric case based on serial correlation. *Ann. Math. Stat.* 14, 378-388.

Weisse AK, Bois P. (2001). Topographic effects on statistical characteristics of heavy rainfall and mapping in the French Alps. *J. Appl. Meteorol.* 40(4), 720-740.

Wijngaard J. (2003). "Homogeneity of daily 'European Climate Assessment and Dataset' series". In: *Proceedings of the Second Seminar for Homogenization of Surface Climatological Data*, Budapest, Hungary, WMO-TD No. 962, WCDMP No. 41, 143-149.

Wijngaard J, Klein Tank AMG, Können GP. (2003). Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.* 23(6), 679-692.

Wilcoxon F. (1945). Individual comparison by ranking methods. *Biometrics* 1, 80-83.

Willmott CJ, Matsuura K. (2006). On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. *Int. J. Geogr. Inf. Sci.* 20(1), 89-102.

Willmott CJ, Robeson SM. (1995). Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.* 15, 221-229.

WMO (2005). *Climate and Land Degradation*. World Meteorological Organization, WMO-No. 989, Geneva, Switzerland, 34 pp.

Xia Y, Winterhalter M, Fabian, P. (1999). A Model to interpolate monthly mean climatological data at Bavarian forest climate stations. *Theor. Appl. Climatol.* 64, 27-38.

Yamamoto JK. (2005). Correcting the smoothing effect of ordinary kriging estimates. *Math. Geol.* 37(1), 69-94.

Yang X, Zhang K, Jia B, Ci L. (2005). Desertification assessment in China: An overview. *J. Arid. Environ.* 63(2), 517-531.

Yue S, Wang CY. (2004). The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resour. Manag.* 18(3), 201-218.

Zellner A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Am. Stat. Assoc.* 57(298), 348-368.

Appendix I DAILY PRECIPITATION SERIES COMPILED

Table 1 – Characterization of the 107 daily precipitation series compiled for the homogenization analysis by location and elevation of the monitoring station, county, and type of network

BASIN / NETWORK	STATION CODE / ID	STATION NAME	LOCATION_X in meters (*IgeoE - coordinate system from the Portuguese Military System*)	LOCATION_Y in meters (*IgeoE - coordinate system from the Portuguese Military System*)	ELEVATION in meters	COUNTY	TYPE OF NETWORK*
ECA&D	666	Beja	223394.5376	116684.4	246	BEJA	#N/A
ECA&D	675	Lisboa Geofísica	111556.7641	194846.2847	77	LISBOA	#N/A
ECA&D	681	Tavira	242933.0733	16869.65469	25	TAVIRA	#N/A
ECA&D	709	Badajoz Talavera	313142.7172	213663.0688	185	(Spain)	#N/A
Arade	29I.01	São Barnabé	197300	43349	250	ALMODÔVAR	Udométrica
Arade	30F.01	Monchique	161964	38903	465	MONCHIQUE	Climatológica
Arade	30G.01	Alferce	168840	40724	328	MONCHIQUE	Udométrica
Arade	30H.04	Santa Margarida	195423	31419	250	LOULÉ	Udométrica
Ribeiras do Algarve	29F.01	Cimalhas	164425	44442	300	MONCHIQUE	Udométrica
Ribeiras do Algarve	29F.02	Foz do Farelo	157033	42626	170	MONCHIQUE	Udométrica
Ribeiras do Algarve	30E.01	Aljezur	141280	39022	48	ALJEZUR	Udométrica
Ribeiras do Algarve	30E.02	Marmelete	152454	38088	375	MONCHIQUE	Udométrica
Ribeiras do Algarve	30E.03	Barragem da Bravura	149795	25802	75	LAGOS	Climatológica
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines	185072	31431	150	SILVES	Udométrica
Ribeiras do Algarve	30H.05	Paderne	193938	22171	80	ALBUFEIRA	Udométrica
Ribeiras do Algarve	30K.02	Picota	239012	22599	174	TAVIRA	Climatológica
Ribeiras do Algarve	30L.03	Faz-Fato	247829	29704	100	TAVIRA	Udométrica
Ribeiras do Algarve	31E.01	Lagos	151952	14907	14	LAGOS	Udométrica
Ribeiras do Algarve	31G.02	Porches	176158	16654	82	LAGOA	Udométrica
Ribeiras do Algarve	31H.02	Algoz	186533	20330	50	SILVES	Climatológica

BASIN / NETWORK	STATION CODE / ID	STATION NAME	LOCATION_X in meters (*IgeoE - coordinate system from the Portuguese Military System*)	LOCATION_Y in meters (*IgeoE - coordinate system from the Portuguese Military System*)	ELEVATION in meters	COUNTY	TYPE OF NETWORK*
Ribeiras do Algarve	31J.01	São Brás de Alportel	220956	21486	325	SÃO BRÁS DE ALPORTEL	Climatológica
Ribeiras do Algarve	31J.04	Estoi	222088	12950	120	FARO	Udométrica
Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)	230950.0732	20372.1926	182	TAVIRA	Udométrica
Ribeiras do Algarve	31K.02	Quelfes	228024.9285	9263.5465	25	OLHÃO	Udométrica
Mira	27G.01	Relíquias	168992	81421	230	ODEMIRA	Udográfica
Mira	28F.01	Odemira	154232	70390	70	ODEMIRA	Udométrica
Mira	28G.01	Barragem de Mira	172573	60474	160	ODEMIRA	Climatológica
Mira	28H.01	Aldeia de Palheiros	189287	70856	210	OURIQUE	Udográfica
Mira	28H.03	Santana da Serra	185391	59177	200	OURIQUE	Udométrica
Mira	29G.01	Sabóia	167726	58240	65	ODEMIRA	Udométrica
Mira	29I.02	Santa Clara-a-Nova	199245	57531	321	ALMODÔVAR	Climatológica
Sado	21G.01	Vendas Novas	170849	188718	150	VENDAS NOVAS	Udométrica
Sado	22E.01	Águas de Moura	151088	179501	10	PALMELA	Udométrica
Sado	22F.03	Moinhola	157331	178972	39	VENDAS NOVAS	Climatológica
Sado	22H.02	Santiago do Escoural	196954	174820	243	MONTEMOR-O-NOVO	Udométrica
Sado	23E.01	Comporta	143060	156704	10	ALCACÉR DO SAL	Climatológica
Sado	23F.01	Montevil	157369	158867	5	ALCACÉR DO SAL	Udométrica
Sado	23G.01	Barragem de Pego do Altar	177106	161408	60	ALCACÉR DO SAL	Climatológica
Sado	23I.01	Alcáçovas	198404	158138	215	VIANA DO ALENTEJO	Udométrica
Sado	24F.01	Grândola	162743	134788	91	GRANDOLA	Climatológica
Sado	24H.02	Barragem do Vale do Gaio	186730	142423	30	ALCACÉR DO SAL	Climatológica
Sado	24I.01	Viana do Alentejo	211058	151269	230	VIANA DO ALENTEJO	Climatológica
Sado	24I.03	Barragem de Odivelas	201320	135012	114	FERREIRA DO ALENTEJO	Climatológica
Sado	24J.02	Alvito	212017	143284	210	ALVITO	Udométrica

BASIN / NETWORK	STATION CODE / ID	STATION NAME	LOCATION_X in meters (*IgeoE - coordinate system from the Portuguese Military System*)	LOCATION_Y in meters (*IgeoE - coordinate system from the Portuguese Military System*)	ELEVATION in meters	COUNTY	TYPE OF NETWORK*
Sado	25G.01	Azinheira Barros	176453	120245	81	SANTIAGO DO CACÉM	Udométrica
Sado	25I.01	Ferreira do Alentejo	201786	121260	123	FERREIRA DO ALENTEJO	Udométrica
Sado	26F.02	Barragem de Campilhas	157018	96675	110	SANTIAGO DO CACÉM	Climatológica
Sado	26I.01	Santa Vitória	209478	110566	153	BEJA	Udométrica
Sado	26I.02	Barragem do Roxo	204402	107509	148	ALJUSTREL	Climatológica
Sado	26I.03	Aljustrel	197000	100109	223	ALJUSTREL	Udométrica
Sado	27G.02	Garvão (Montinho)	182225	83232	110	OURIQUE	Udométrica
Sado	27H.01	Panóias	184805	87851	175	OURIQUE	Udométrica
Sado	27H.02	Barragem do Monte da Rocha	186415	84519	140	OURIQUE	Climatológica
Guadiana	18N.01	São Julião	271329	260536	530	PORTALEGRE	Udométrica
Guadiana	18N.02	Alegrete	270085	252259	458	PORTALEGRE	Udográfica
Guadiana	19N.01	Arronches	273709	239307	300	ARRONCHES	Udométrica
Guadiana	19N.02	Santa Eulália	276621	226474	273	ELVAS	Udométrica
Guadiana	19N.03	Esperança	281064	243698	350	ARRONCHES	Udométrica
Guadiana	19O.02	Barragem do Caia	286079.6492	226417.5134	230	CAMPO MAIOR	Climatológica
Guadiana	19O.03	Degolados	287694	232758	265	CAMPO MAIOR	Udométrica
Guadiana	20O.02	Caia (M. Caldeiras)	294308	212540	170	ELVAS	Udométrica
Guadiana	21K.01	Azaruja	231177	192858	270	ÉVORA	Udométrica
Guadiana	21M.01	Vila Viçosa	261915	202197	370	VILA VIÇOSA	Udométrica
Guadiana	21M.02	Alandroal	263421	191939	350	ALANDROAL	Udométrica
Guadiana	21N.01	Juromenha	277509	197214	206	ALANDROAL	Udométrica
Guadiana	22L.01	Redondo	250838	186792	315	REDONDO	Udométrica
Guadiana	22L.02	Santa Susana	241375	178781	225	REDONDO	Udométrica
Guadiana	22M.01	Santiago Maior	256675	175205	324	ALANDROAL	Udométrica

BASIN / NETWORK	STATION CODE / ID	STATION NAME	LOCATION_X in meters (*IgeoE - coordinate system from the Portuguese Military System*)	LOCATION_Y in meters (*IgeoE - coordinate system from the Portuguese Military System*)	ELEVATION in meters	COUNTY	TYPE OF NETWORK*
Guadiana	23K.01	São Manços	233660	165609	195	ÉVORA	Udométrica
Guadiana	23L.01	Reguengos	252936	161950	210	REGUENGOS MONSARAZ	Udográfica
Guadiana	24J.03	Cuba	221599	133036	160	CUBA	Udométrica
Guadiana	24K.01	Portel	237180	148849	315	PORTEL	Udométrica
Guadiana	24K.02	Vidigueira	228838	137499	190	VIDIGUEIRA	Udométrica
Guadiana	24N.01	Amareleja (D.G.R.N.)	279148	138512	192	MOURA	Udométrica
Guadiana	25L.01	Pedrogão do Alentejo	242611	127877	140	VIDIGUEIRA	Udométrica
Guadiana	25N.01	Sobral da Adiça	276624	117118	200	MOURA	Udométrica
Guadiana	25O.01	Santo Aleixo da Restauração	286353	122242	280	MOURA	Udométrica
Guadiana	25P.01	Barrancos	299228	130067	380	BARRANCOS	Udométrica
Guadiana	26J.04	Albernoa	215041	98920	150	BEJA	Udométrica
Guadiana	26K.01	Salvada	230364	106578	178	BEJA	Udométrica
Guadiana	26L.01	Serpa	246522	108566	190	SERPA	Udométrica
Guadiana	26L.02	Santa Iria	250666	101838	201	SERPA	Udométrica
Guadiana	26M.01	Herdade de Valada	261513	108819	230	SERPA	Climatológica
Guadiana	27I.01	Castro Verde	203510	81240	180	CASTRO VERDE	Udográfica
Guadiana	27J.01	São Marcos da Ataboeira	217105	81966	174	CASTRO VERDE	Udométrica
Guadiana	27J.02	Corte Pequena	224709	86826	167	MÉRTOLA	Udométrica
Guadiana	27J.03	Vale de Camelos	223346.0959	92495.32	135	MÉRTOLA	Climatológica
Guadiana	27K.01	Algodôr	230122	86104	163	MÉRTOLA	Udométrica
Guadiana	27K.02	Corte da Velha	234191	79152	125	MÉRTOLA	Udométrica
Guadiana	28I.01	Almodôvar	205778	60184	270	ALMODÔVAR	Udométrica
Guadiana	28J.01	Alcaria Longa	223936	66105	119	MÉRTOLA	Udográfica
Guadiana	28J.03	Santa Barbara de Padrões	213394	74313	250	CASTRO VERDE	Udométrica

BASIN / NETWORK	STATION CODE / ID	STATION NAME	LOCATION_X in meters (*IgeoE - coordinate system from the Portuguese Military System*)	LOCATION_Y in meters (*IgeoE - coordinate system from the Portuguese Military System*)	ELEVATION in meters	COUNTY	TYPE OF NETWORK*
Guadiana	28K.01	São João dos Caldeireiros	230321	71861	170	MÉRTOLA	Udométrica
Guadiana	28K.02	Álamo	239364	68199	150	MÉRTOLA	Udométrica
Guadiana	28L.01	Mértola	241590	74623	65	MÉRTOLA	Udométrica
Guadiana	29J.05	Guedelhas	216221	58009	290	ALMODÔVAR	Udométrica
Guadiana	29K.01	Martim Longo	232311	52167	290	ALCOUTIM	Climatológica
Guadiana	29K.03	Malfrades	241199	44468	255	ALCOUTIM	Udométrica
Guadiana	29K.04	Penedos	229758	57923	265	MÉRTOLA	Udométrica
Guadiana	29L.01	Pereiro	247058	52823	240	ALCOUTIM	Udométrica
Guadiana	29L.03	Monte dos Fortes	247126	40799	65	CASTRO MARIM	Udométrica
Guadiana	29M.01	Alcoutim	258470	55670	39	ALCOUTIM	Udométrica
Guadiana	30I.02	Sobreira	206386	36970	475	LOULÉ	Udométrica
Guadiana	30J.01	Barranco do Velho	217163	29771	475	LOULÉ	Udométrica
Guadiana	30J.02	Catraia	217595.8648	35135.8264	420	LOULÉ	Udográfica
Guadiana	30K.01	Mercador	233856	34816	330	TAVIRA	Udométrica
Guadiana	30L.04	Alcaria (Castro Marim)	255156	25156	48	CASTRO MARIM	Udométrica

***** '*Udométrica*' means that the station only measures precipitation through accumulative precipitation gauges; '*udográfica*' means that the station only measures precipitation through recording raingauges; and '*climatológica*' means that the station measures precipitation using both instrumentation, and measures other climate variables as well.

Table 2 – Characterization of the 107 daily precipitation series compiled for the homogenization analysis by date of beginning of the precipitation measurements at the monitoring station, available data periods, and series length (years lacking a maximum of 5% of data)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	BEGINNING OF PRECIPITATION MEASUREMENTS	BEGINNING OF RECORDS	ENDING OF RECORDS	NUM. OF DAILY PRECIPITATION RECORDS	SERIES LENGTH (years lacking a maximum of 5% of data)
ECA&D	666	Beja	#N/A	01/01/1941	31/12/1999	21549	59
ECA&D	675	Lisboa Geofísica	#N/A	01/01/1941	31/12/1999	21549	59
ECA&D	681	Tavira	#N/A	01/01/1941	31/12/1994	19723	54
ECA&D	709	Badajoz Talavera	#N/A	01/01/1955	31/08/2001	17045	46
Arade	29I.01	São Barnabé	01/10/1964	01/10/1964	07/11/2002	13917	36
Arade	30F.01	Monchique	04/02/1932	04/02/1932	31/01/1999	24469	55
Arade	30G.01	Alferce	19/08/1958	19/08/1958	30/09/2000	15536	41
Arade	30H.04	Santa Margarida	01/10/1964	01/10/1964	30/09/2000	13242	33
Ribeiras do Algarve	29F.01	Cimalhas	01/10/1980	01/10/1980	31/12/1999	7031	19
Ribeiras do Algarve	29F.02	Foz do Farelo	01/10/1980	01/10/1980	30/09/2000	7641	19
Ribeiras do Algarve	30E.01	Aljezur	08/10/1931	08/10/1931	30/09/2000	25447	68
Ribeiras do Algarve	30E.02	Marmelete	01/10/1959	01/10/1983	30/09/2000	15312	16
Ribeiras do Algarve	30E.03	Barragem da Bravura	01/10/1933	01/10/1955	30/09/2001	24837	45
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines	01/10/1931	01/10/1990	30/09/1999	25416	8
Ribeiras do Algarve	30H.05	Paderne	01/10/1957	01/10/1983	30/09/2000	15707	16
Ribeiras do Algarve	30K.02	Picota	01/10/1933	01/10/1955	30/09/2000	24777	41
Ribeiras do Algarve	30L.03	Faz-Fato	01/10/1945	01/10/1955	30/09/2000	20333	44
Ribeiras do Algarve	31E.01	Lagos	01-10-1864	01/01/1956	31/03/2001	36404	44
Ribeiras do Algarve	31G.02	Porches	01/10/1979	01/10/1979	30/09/1999	7976	19
Ribeiras do Algarve	31H.02	Algoz	01/09/1980	01/09/1980	31/12/1996	8007	16
Ribeiras do Algarve	31J.01	São Brás de Alportel	01/10/1900	01/10/1984	31/08/2001	6178	16

BASIN / NETWORK	STATION CODE / ID	STATION NAME	BEGINNING OF PRECIPITATION MEASUREMENTS	BEGINNING OF RECORDS	ENDING OF RECORDS	NUM. OF DAILY PRECIPITATION RECORDS	SERIES LENGTH (years lacking a maximum of 5% of data)
Ribeiras do Algarve	31J.04	Estoi	01/10/1979	01/10/1983	30/09/2000	7610	16
Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)	01/10/1959	01/10/1983	30/09/2000	15281	16
Ribeiras do Algarve	31K.02	Quelfes	01/08/1980	01/08/1980	30/09/1999	7702	15
Mira	27G.01	Relíquias	01/11/1931	01/11/1931	30/09/2001	25537	67
Mira	28F.01	Odemira	01/11/1931	01/11/1931	30/09/1995	23345	61
Mira	28G.01	Barragem de Mira	01/10/1965	01/10/1969	30/06/2001	11596	23
Mira	28H.01	Aldeia de Palheiros	01/09/1931	01/09/1931	30/09/1997	25204	65
Mira	28H.03	Santana da Serra	09/12/1935	09/12/1935	31/12/2000	23765	62
Mira	29G.01	Sabóia	01/10/1931	01/10/1931	30/09/1995	23376	62
Mira	29I.02	Santa Clara-a-Nova	01/10/1979	05/07/1980	30/11/2000	7454	19
Sado	21G.01	Vendas Novas	01/09/1911	01/09/1911	31/03/2000	32355	46
Sado	22E.01	Águas de Moura	01/10/1931	01/10/1983	27/02/2002	6725	15
Sado	22F.03	Moinhola	01/10/1934	01/10/1972	31/01/2001	10350	27
Sado	22H.02	Santiago do Escoural	21/11/1931	21/11/1931	31/12/2000	25224	67
Sado	23E.01	Comporta	01/04/1934	01/01/1934	01/05/2001	24593	50
Sado	23F.01	Montevil	01/10/1944	01/10/1983	31/12/2000	6302	17
Sado	23G.01	Barragem de Pego do Altar	01/10/1933	01/02/1980	31/12/2000	7640	21
Sado	23I.01	Alcáçovas	14/11/1931	14/11/1931	31/12/2000	25251	66
Sado	24F.01	Grândola	01/10/1931	01/10/1972	31/01/2001	10350	28
Sado	24H.02	Barragem do Vale do Gaio	01/10/1938	01/10/1979	30/06/2001	7944	21
Sado	24I.01	Viana do Alentejo	01/12/1933	01/12/1933	31/01/2001	24534	38
Sado	24I.03	Barragem de Odivelas	06/11/1973	06/11/1973	31/12/2000	9918	27
Sado	24J.02	Alvito	01/10/1939	01/10/1983	31/12/2000	6302	16
Sado	25G.01	Azinheira Barros	09/03/1950	09/03/1950	31/12/2000	18561	50

BASIN / NETWORK	STATION CODE / ID	STATION NAME	BEGINNING OF PRECIPITATION MEASUREMENTS	BEGINNING OF RECORDS	ENDING OF RECORDS	NUM. OF DAILY PRECIPITATION RECORDS	SERIES LENGTH (years lacking a maximum of 5% of data)
Sado	25I.01	Ferreira do Alentejo	08/11/1931	08/11/1931	31/12/2000	25257	68
Sado	26F.02	Barragem de Campilhas	01/10/1952	01/10/1955	31/01/1997	16408	26
Sado	26I.01	Santa Vitória	01/10/1949	01/10/1983	31/12/2000	6302	17
Sado	26I.02	Barragem do Roxo	10/07/1958	10/07/1958	31/12/2000	15516	39
Sado	26I.03	Aljustrel	13/11/1931	13/11/1931	31/12/2000	25252	64
Sado	27G.02	Garvão (Montinho)	01/10/1978	01/04/1979	30/09/1995	6027	15
Sado	27H.01	Panóias	01/10/1933	01/10/1955	30/09/1995	14610	39
Sado	27H.02	Barragem do Monte da Rocha	11/01/1980	11/01/1980	30/09/1995	7540	15
Guadiana	18N.01	São Julião	15/04/1980	15/04/1980	31/03/2000	7291	16
Guadiana	18N.02	Alegrete	10/03/1980	10/03/1980	31/03/2000	7327	19
Guadiana	19N.01	Arronches	01/09/1931	01/09/1931	31/03/2000	25050	68
Guadiana	19N.02	Santa Eulália	01/10/1948	01/10/1982	31/03/2000	6392	17
Guadiana	19N.03	Esperança	12/10/1979	12/10/1979	31/03/2000	7477	16
Guadiana	19O.02	Barragem do Caia	13/01/1963	13/01/1964	31/07/2001	13715	35
Guadiana	19O.03	Degolados	01/10/1980	01/10/1983	31/03/2000	6027	15
Guadiana	20O.02	Caia (M. Caldeiras)	25/10/1979	25/10/1979	31/01/2000	7404	17
Guadiana	21K.01	Azaruja	01/09/1931	01/10/1943	30/09/1983	20607	39
Guadiana	21M.01	Vila Viçosa	01/10/1930	01/10/1980	31/07/2001	7609	20
Guadiana	21M.02	Alandroal	01/09/1931	01/10/1983	31/03/2000	6027	16
Guadiana	21N.01	Juromenha	01/09/1931	01/09/1931	31/03/2000	25050	45
Guadiana	22L.01	Redondo	01/09/1931	01/10/1944	30/09/1983	19967	38
Guadiana	22L.02	Santa Susana	27/03/1949	27/03/1949	31/03/2000	18633	50
Guadiana	22M.01	Santiago Maior	01/04/1949	01/10/1983	31/03/2000	6027	15
Guadiana	23K.01	São Manços	07/04/1942	07/04/1942	31/12/2000	21454	57

BASIN / NETWORK	STATION CODE / ID	STATION NAME	BEGINNING OF PRECIPITATION MEASUREMENTS	BEGINNING OF RECORDS	ENDING OF RECORDS	NUM. OF DAILY PRECIPITATION RECORDS	SERIES LENGTH (years lacking a maximum of 5% of data)
Guadiana	23L.01	Reguengos	01/10/1930	01/10/1984	31/01/2001	5967	15
Guadiana	24J.03	Cuba	01/10/1930	01/10/1985	30/09/1999	5206	13
Guadiana	24K.01	Portel	01/10/1938	01/10/1983	30/09/2000	6210	16
Guadiana	24K.02	Vidigueira	01/10/1948	01/10/1983	31/12/2000	6302	17
Guadiana	24N.01	Amareleja (D.G.R.N.)	01/10/1930	01/10/1983	31/12/2000	6302	17
Guadiana	25L.01	Pedrogão do Alentejo	01/06/1941	01/06/1941	31/12/2000	21764	59
Guadiana	25N.01	Sobral da Adiça	01/08/1980	01/08/1980	31/12/2000	7458	19
Guadiana	25O.01	Santo Aleixo da Restauração	01/10/1930	01/10/1931	31/12/2000	25295	67
Guadiana	25P.01	Barrancos	01/10/1930	01/10/1985	30/09/1999	5206	13
Guadiana	26J.04	Albernoa	01/10/1978	01/10/1983	31/12/2000	6302	17
Guadiana	26K.01	Salvada	01/10/1957	01/10/1983	31/12/2000	6302	17
Guadiana	26L.01	Serpa	01/10/1930	01/10/1984	31/01/2001	5967	16
Guadiana	26L.02	Santa Iria	01/02/1980	01/02/1980	31/12/2000	7640	21
Guadiana	26M.01	Herdade de Valada	01/10/1968	01/10/1968	31/12/2000	11780	30
Guadiana	27I.01	Castro Verde	01/09/1931	01/09/1931	31/01/2001	25356	67
Guadiana	27J.01	São Marcos da Ataboeira	01/10/1956	01/10/1983	31/12/2000	6302	17
Guadiana	27J.02	Corte Pequena	01/10/1979	01/09/1983	30/09/1997	5967	11
Guadiana	27J.03	Vale de Camelos	01/10/1986	01/02/1989	31/12/2000	4352	11
Guadiana	27K.01	Algodôr	01/10/1930	01/10/1986	30/09/2000	5176	13
Guadiana	27K.02	Corte da Velha	01/10/1979	01/10/1979	31/12/2000	7763	21
Guadiana	28I.01	Almodôvar	01/10/1930	01/10/1983	31/12/2000	6302	17
Guadiana	28J.01	Alcaria Longa	01/10/1949	01/10/1985	30/09/2000	5479	14
Guadiana	28J.03	Santa Barbara de Padrões	01/10/1979	01/10/1979	31/12/2000	7763	21
Guadiana	28K.01	São João dos Caldeireiros	01/10/1979	01/10/1983	30/09/2001	6575	17

BASIN / NETWORK	STATION CODE / ID	STATION NAME	BEGINNING OF PRECIPITATION MEASUREMENTS	BEGINNING OF RECORDS	ENDING OF RECORDS	NUM. OF DAILY PRECIPITATION RECORDS	SERIES LENGTH (years lacking a maximum of 5% of data)
Guadiana	28K.02	Álamo	05/03/1980	05/03/1980	31/12/2000	7607	20
Guadiana	28L.01	Mértola	01/10/1938	01/10/1983	31/12/2000	6302	17
Guadiana	29J.05	Guedelhas	01/11/1979	01/11/1979	30/09/2000	7640	19
Guadiana	29K.01	Martim Longo	01/10/1940	01/10/1984	07/08/2001	6155	16
Guadiana	29K.03	Malfrades	01/03/1980	01/03/1980	30/09/2000	7855	19
Guadiana	29K.04	Penedos	01/11/1980	01/11/1980	31/12/2000	7366	19
Guadiana	29L.01	Pereiro	16/07/1957	16/07/1957	30/09/2000	16119	42
Guadiana	29L.03	Monte dos Fortes	01/10/1979	01/10/1983	30/09/2000	6210	15
Guadiana	29M.01	Alcoutim	11/03/1938	11/03/1938	30/09/2000	23155	61
Guadiana	30I.02	Sobreira	11/05/1942	11/05/1942	30/09/2000	21664	56
Guadiana	30J.01	Barranco do Velho	01/10/1934	01/10/1955	30/09/2000	16711	44
Guadiana	30J.02	Catraia	01/09/1931	01/09/1931	31/05/1974	15614	42
Guadiana	30K.01	Mercador	01/10/1959	01/10/1983	30/09/2000	6362	16
Guadiana	30L.04	Alcaria (Castro Marim)	01/04/1946	01/04/1946	30/09/2000	20181	53

Appendix II ABSOLUTE TESTING RESULTS

Table 1 – Summary results from the autocorrelation and normality tests applied to 107 series of annual wet day count (1 mm threshold)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	Durbin-Watson autocorrelation tests (5% level)			NORMALITY TESTS (5% level)			
				1st Order	2nd Order	3rd Order	Shapiro-Wilk	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
ECA&D	666	Beja	1941-1999	Negative	No	No	Normal	Normal	Normal	Normal
ECA&D	675	Lisboa Geofísica	1941-1999	No	No	No	Normal	Normal	Normal	Normal
ECA&D	681	Tavira	1941-1994	No	No	No	Normal	Normal	Normal	Normal
ECA&D	709	Badajoz Talavera	1955-2000	No	Positive	No	Normal	Normal	Normal	Normal
Arade	29I.01	São Barnabé	1965-2000	Negative	No	No	Normal	Normal	Normal	Normal
Arade	30F.01	Monchique	1933-1998	Negative	No	No	Normal	Normal	Normal	Normal
Arade	30G.01	Alferce	1959-1999	No	No	No	Normal	Normal	Normal	Normal
Arade	30H.04	Santa Margarida	1965-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	29F.01	Cimalhas	1981-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	29F.02	Foz do Farelo	1981-1999	Negative	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30E.01	Aljezur	1932-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30E.02	Marmelete	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30E.03	Barragem da Bravura	1956-2000	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines	1991-1998	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30H.05	Paderne	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30K.02	Picota	1957-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	30L.03	Faz-Fato	1956-1999	Negative	-	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	31E.01	Lagos	1956-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	31G.02	Porches	1980-1998	No	No	No	Reject	Reject	Reject	Reject
Ribeiras do Algarve	31H.02	Algoz	1981-1996	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	31J.01	São Brás de Alportel	1985-2000	No	No	No	Normal	Reject	Reject	Reject

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	Durbin-Watson autocorrelation tests (5% level)			NORMALITY TESTS (5% level)			
				1st Order	2nd Order	3rd Order	Shapiro-Wilk	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
Ribeiras do Algarve	31J.04	Estoi	1984-1999	No	No	No	Reject	Reject	Reject	Reject
Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Ribeiras do Algarve	31K.02	Quelfes	1982-1998	No	No	No	Normal	Normal	Normal	Normal
Mira	27G.01	Relíquias	1932-2000	No	Positive	No	Reject	Normal	Normal	Normal
Mira	28F.01	Odemira	1932-1994	No	No	No	Reject	Normal	Normal	Normal
Mira	28G.01	Barragem de Mira	1970-1993	No	No	No	Normal	Normal	Normal	Normal
Mira	28H.01	Aldeia de Palheiros	1932-1996	No	No	No	Normal	Normal	Normal	Normal
Mira	28H.03	Santana da Serra	1936-2000	No	No	No	Normal	Normal	Normal	Normal
Mira	29G.01	Sabóia	1932-1994	Negative	No	No	Normal	Normal	Normal	Normal
Mira	29I.02	Santa Clara-a-Nova	1981-1999	No	No	No	Normal	Normal	Normal	Normal
Sado	21G.01	Vendas Novas	1932-1999	Negative	No	No	Reject	Reject	Reject	Reject
Sado	22E.01	Águas de Moura	1984-2001	No	Positive	No	Normal	Normal	Normal	Normal
Sado	22F.03	Moinhola	1973-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	22H.02	Santiago do Escoural	1932-1999	No	No	No	Reject	Reject	Reject	Reject
Sado	23E.01	Comporta	1934-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	23F.01	Montevil	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	23G.01	Barragem de Pego do Altar	1980-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	23I.01	Alcáçovas	1932-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	24F.01	Grândola	1973-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	24H.02	Barragem do Vale do Gaio	1980-2000	Negative	No	No	Normal	Normal	Normal	Normal
Sado	24I.01	Viana do Alentejo	1934-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	24I.03	Barragem de Odivelas	1974-2000	Negative	No	No	Normal	Normal	Normal	Normal
Sado	24J.02	Alvito	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	25G.01	Azinheira Barros	1951-2000	No	No	No	Normal	Reject	Normal	Normal

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	Durbin-Watson autocorrelation tests (5% level)			NORMALITY TESTS (5% level)			
				1st Order	2nd Order	3rd Order	Shapiro-Wilk	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
Sado	25I.01	Ferreira do Alentejo	1933-2000	Negative	No	No	Normal	Normal	Normal	Normal
Sado	26F.02	Barragem de Campilhas	1956-1994	No	No	No	Normal	Normal	Normal	Normal
Sado	26I.01	Santa Vitória	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	26I.02	Barragem do Roxo	1959-2000	No	No	No	Normal	Normal	Normal	Normal
Sado	26I.03	Aljustrel	1936-2000	Negative	No	No	Reject	Normal	Normal	Normal
Sado	27G.02	Garvão (Montinho)	1980-1994	No	No	No	Normal	Normal	Normal	Normal
Sado	27H.01	Panóias	1956-1994	No	No	No	Normal	Normal	Normal	Normal
Sado	27H.02	Barragem do Monte da Rocha	1980-1994	No	No	No	Normal	Normal	Normal	Normal
Guadiana	18N.01	São Julião	1981-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	18N.02	Alegrete	1981-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	19N.01	Arronches	1932-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	19N.02	Santa Eulália	1983-1999	No	No	No	Normal	Normal	Reject	Reject
Guadiana	19N.03	Esperança	1980-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	19O.02	Barragem do Caia	1965-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	19O.03	Degolados	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	20O.02	Caia (M. Caldeiras)	1980-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	21K.01	Azaruja	1944-1982	No	No	No	Normal	Normal	Normal	Normal
Guadiana	21M.01	Vila Viçosa	1981-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	21M.02	Alandroal	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	21N.01	Juromenha	1932-1999	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	22L.01	Redondo	1945-1982	No	No	No	Reject	Normal	Normal	Normal
Guadiana	22L.02	Santa Susana	1950-1999	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	22M.01	Santiago Maior	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	23K.01	São Manços	1943-2000	No	No	No	Normal	Normal	Normal	Normal

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	Durbin-Watson autocorrelation tests (5% level)			NORMALITY TESTS (5% level)			
				1st Order	2nd Order	3rd Order	Shapiro-Wilk	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
Guadiana	23L.01	Reguengos	1985-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	24J.03	Cuba	1986-1998	No	No	No	Normal	Normal	Normal	Normal
Guadiana	24K.01	Portel	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	24K.02	Vidigueira	1984-2000	No	Positive	No	Normal	Normal	Normal	Normal
Guadiana	24N.01	Amareleja (D.G.R.N.)	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	25L.01	Pedrogão do Alentejo	1942-2000	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	25N.01	Sobral da Adiça	1981-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	25O.01	Santo Aleixo da Restauração	1932-2000	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	25P.01	Barrancos	1986-1998	No	No	No	Normal	Normal	Normal	Normal
Guadiana	26J.04	Albernoa	1984-2000	No	No	No	Reject	Reject	Reject	Reject
Guadiana	26K.01	Salvada	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	26L.01	Serpa	1985-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	26L.02	Santa Iria	1980-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	26M.01	Herdade de Valada	1969-2000	No	No	No	Normal	Normal	Reject	Reject
Guadiana	27I.01	Castro Verde	1932-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	27J.01	São Marcos da Ataboeira	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	27J.02	Corte Pequena	1986-1996	No	No	No	Reject	Reject	Reject	Reject
Guadiana	27J.03	Vale de Camelos	1990-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	27K.01	Algodôr	1987-1999	No	No	No	Normal	Reject	Reject	Reject
Guadiana	27K.02	Corte da Velha	1980-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	28I.01	Almodôvar	1984-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	28J.01	Alcaria Longa	1986-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	28J.03	Santa Barbara de Padrões	1980-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	28K.01	São João dos Caldeireiros	1984-2000	No	No	No	Normal	Normal	Normal	Normal

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	Durbin-Watson autocorrelation tests (5% level)			NORMALITY TESTS (5% level)			
				1st Order	2nd Order	3rd Order	Shapiro-Wilk	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
Guadiana	28K.02	Álamo	1981-2000	No	No	No	Normal	Normal	Normal	Normal
Guadiana	28L.01	Mértola	1984-2000	No	No	No	Reject	Normal	Reject	Reject
Guadiana	29J.05	Guedelhas	1980-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	29K.01	Martim Longo	1985-2000	No	No	No	Normal	Reject	Reject	Reject
Guadiana	29K.03	Malfrades	1981-1999	No	No	No	Reject	Reject	Reject	Reject
Guadiana	29K.04	Penedos	1981-2000	Negative	No	No	Normal	Reject	Reject	Reject
Guadiana	29L.01	Pereiro	1958-1999	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	29L.03	Monte dos Fortes	1984-1999	No	No	No	Reject	Reject	Reject	Reject
Guadiana	29M.01	Alcoutim	1939-1999	Negative	-	No	Normal	Normal	Normal	Normal
Guadiana	30I.02	Sobreira	1943-1999	Negative	No	No	Normal	Reject	Normal	Normal
Guadiana	30J.01	Barranco do Velho	1956-1999	Negative	No	No	Normal	Normal	Normal	Normal
Guadiana	30J.02	Catraia	1932-1973	No	No	No	Normal	Normal	Normal	Normal
Guadiana	30K.01	Mercador	1984-1999	No	No	No	Normal	Normal	Normal	Normal
Guadiana	30L.04	Alcaria (Castro Marim)	1947-1999	No	No	No	Normal	Normal	Normal	Normal

Table 2 – Summary results from the absolute testing approach: tests not capable of locating the year where a break is likely applied to 107 series of annual wet day count (1 mm threshold)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	WALD-WOLFOWITZ (5% level)	MANN-KENDALL (5% level)	VON NEUMANN (5% level)
				Two-tailed test result	One-tailed test result based on large sample approx. (n>7)	
ECA&D	666	Beja	1941-1999	Homogeneous data	Homogeneous data	Break present
ECA&D	675	Lisboa Geofísica	1941-1999	Homogeneous data	Homogeneous data	Homogeneous data
ECA&D	681	Tavira	1941-1994	Homogeneous data	Downward trend present	Homogeneous data
ECA&D	709	Badajoz Talavera	1955-2000	Homogeneous data	Homogeneous data	Break present
Arade	29I.01	São Barnabé	1965-2000	Homogeneous data	Homogeneous data	Break present
Arade	30F.01	Monchique	1933-1998	Trend effec present	Homogeneous data	Break present
Arade	30G.01	Alferce	1959-1999	Homogeneous data	Homogeneous data	Homogeneous data
Arade	30H.04	Santa Margarida	1965-1999	Homogeneous data	Homogeneous data	Homogeneous data
Ribeiras do Algarve	29F.01	Cimalhas	1981-1999	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	29F.02	Foz do Farelo	1981-1999	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	30E.01	Aljezur	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data
Ribeiras do Algarve	30E.02	Marmelete	1984-1999	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	30E.03	Barragem da Bravura	1956-2000	Homogeneous data	Homogeneous data	Homogeneous data
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines	1991-1998	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	30H.05	Paderne	1984-1999	Homogeneous data	Downward trend present	
Ribeiras do Algarve	30K.02	Picota	1957-1999	Homogeneous data	Homogeneous data	Homogeneous data
Ribeiras do Algarve	30L.03	Faz-Fato	1956-1999	Homogeneous data	Homogeneous data	Break present
Ribeiras do Algarve	31E.01	Lagos	1956-1999	Homogeneous data	Downward trend present	Homogeneous data
Ribeiras do Algarve	31G.02	Porches	1980-1998	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	31H.02	Algoz	1981-1996	Homogeneous data	Homogeneous data	

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	WALD-WOLFOWITZ (5% level)	MANN-KENDALL (5% level)	VON NEUMANN (5% level)
				Two-tailed test result	One-tailed test result based on large sample approx. (n>7)	
Ribeiras do Algarve	31J.01	São Brás de Alportel	1985-2000	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	31J.04	Estoi	1984-1999	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)	1984-1999	Homogeneous data	Homogeneous data	
Ribeiras do Algarve	31K.02	Quelfes	1982-1998	Homogeneous data	Homogeneous data	
Mira	27G.01	Relíquias	1932-2000	Homogeneous data	Homogeneous data	Homogeneous data
Mira	28F.01	Odemira	1932-1994	Homogeneous data	Downward trend present	Homogeneous data
Mira	28G.01	Barragem de Mira	1970-1993	Homogeneous data	Homogeneous data	Break present
Mira	28H.01	Aldeia de Palheiros	1932-1996	Homogeneous data	Homogeneous data	Homogeneous data
Mira	28H.03	Santana da Serra	1936-2000	Homogeneous data	Downward trend present	Homogeneous data
Mira	29G.01	Sabóia	1932-1994	Homogeneous data	Homogeneous data	Break present
Mira	29I.02	Santa Clara-a-Nova	1981-1999	Homogeneous data	Homogeneous data	
Sado	21G.01	Vendas Novas	1932-1999	Homogeneous data	Homogeneous data	Break present
Sado	22E.01	Águas de Moura	1984-2001	Homogeneous data	Homogeneous data	
Sado	22F.03	Moinhola	1973-2000	Homogeneous data	Homogeneous data	Break present
Sado	22H.02	Santiago do Escoural	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data
Sado	23E.01	Comporta	1934-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	23F.01	Montevil	1984-2000	Homogeneous data	Homogeneous data	
Sado	23G.01	Barragem de Pego do Altar	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	23I.01	Alcáçovas	1932-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	24F.01	Grândola	1973-2000	Homogeneous data	Upward trend present	Break present
Sado	24H.02	Barragem do Vale do Gaio	1980-2000	Homogeneous data	Homogeneous data	Break present
Sado	24I.01	Viana do Alentejo	1934-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	24I.03	Barragem de Odivelas	1974-2000	Homogeneous data	Homogeneous data	Break present
Sado	24J.02	Alvito	1984-2000	Homogeneous data	Homogeneous data	

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	WALD-WOLFOWITZ (5% level)	MANN-KENDALL (5% level)	VON NEUMANN (5% level)
				Two-tailed test result	One-tailed test result based on large sample approx. (n>7)	
Sado	25G.01	Azinheira Barros	1951-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	25I.01	Ferreira do Alentejo	1933-2000	Homogeneous data	Homogeneous data	Break present
Sado	26F.02	Barragem de Campilhas	1956-1994	Homogeneous data	Homogeneous data	Break present
Sado	26I.01	Santa Vitória	1984-2000	Homogeneous data	Homogeneous data	
Sado	26I.02	Barragem do Roxo	1959-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	26I.03	Aljustrel	1936-2000	Trend effec present	Downward trend present	Break present
Sado	27G.02	Garvão (Montinho)	1980-1994	Homogeneous data	Homogeneous data	
Sado	27H.01	Panóias	1956-1994	Homogeneous data	Homogeneous data	Homogeneous data
Sado	27H.02	Barragem do Monte da Rocha	1980-1994	Homogeneous data	Homogeneous data	
Guadiana	18N.01	São Julião	1981-1999	Homogeneous data	Homogeneous data	
Guadiana	18N.02	Alegrete	1981-1999	Homogeneous data	Homogeneous data	
Guadiana	19N.01	Arronches	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	19N.02	Santa Eulália	1983-1999	Homogeneous data	Homogeneous data	
Guadiana	19N.03	Esperança	1980-1999	Homogeneous data	Homogeneous data	
Guadiana	19O.02	Barragem do Caia	1965-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	19O.03	Degolados	1984-1999	Homogeneous data	Homogeneous data	
Guadiana	20O.02	Caia (M. Caldeiras)	1980-1999	Homogeneous data	Homogeneous data	
Guadiana	21K.01	Azaruja	1944-1982	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	21M.01	Vila Viçosa	1981-2000	Homogeneous data	Upward trend present	Homogeneous data
Guadiana	21M.02	Alandroal	1984-1999	Homogeneous data	Homogeneous data	
Guadiana	21N.01	Juromenha	1932-1999	Homogeneous data	Downward trend present	Break present
Guadiana	22L.01	Redondo	1945-1982	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	22L.02	Santa Susana	1950-1999	Homogeneous data	Downward trend present	Break present
Guadiana	22M.01	Santiago Maior	1984-1999	Homogeneous data	Homogeneous data	

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	WALD-WOLFOWITZ (5% level)	MANN-KENDALL (5% level)	VON NEUMANN (5% level)
				Two-tailed test result	One-tailed test result based on large sample approx. (n>7)	
Guadiana	23K.01	São Manços	1943-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	23L.01	Reguengos	1985-1999	Homogeneous data	Homogeneous data	
Guadiana	24J.03	Cuba	1986-1998	Homogeneous data	Homogeneous data	
Guadiana	24K.01	Portel	1984-1999	Homogeneous data	Homogeneous data	
Guadiana	24K.02	Vidigueira	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	24N.01	Amareleja (D.G.R.N.)	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	25L.01	Pedrogão do Alentejo	1942-2000	Trend effec present	Homogeneous data	Break present
Guadiana	25N.01	Sobral da Adiça	1981-2000	Homogeneous data	Homogeneous data	
Guadiana	25O.01	Santo Aleixo da Restauração	1932-2000	Trend effec present	Homogeneous data	Break present
Guadiana	25P.01	Barrancos	1986-1998	Homogeneous data	Homogeneous data	
Guadiana	26J.04	Albernoa	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	26K.01	Salvada	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	26L.01	Serpa	1985-2000	Homogeneous data	Homogeneous data	
Guadiana	26L.02	Santa Iria	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	26M.01	Herdade de Valada	1969-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	27I.01	Castro Verde	1932-2000	Homogeneous data	Downward trend present	Homogeneous data
Guadiana	27J.01	São Marcos da Ataboeira	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	27J.02	Corte Pequena	1986-1996	Homogeneous data	Homogeneous data	
Guadiana	27J.03	Vale de Camelos	1990-2000	Homogeneous data	Homogeneous data	
Guadiana	27K.01	Algodôr	1987-1999	Homogeneous data	Homogeneous data	
Guadiana	27K.02	Corte da Velha	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	28I.01	Almodôvar	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	28J.01	Alcaria Longa	1986-1999	Homogeneous data	Homogeneous data	
Guadiana	28J.03	Santa Barbara de Padrões	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	WALD-WOLFOWITZ (5% level)	MANN-KENDALL (5% level)	VON NEUMANN (5% level)
				Two-tailed test result	One-tailed test result based on large sample approx. (n>7)	
Guadiana	28K.01	São João dos Caldeireiros	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	28K.02	Álamo	1981-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	28L.01	Mértola	1984-2000	Homogeneous data	Homogeneous data	
Guadiana	29J.05	Guedelhas	1980-1999	Homogeneous data	Homogeneous data	
Guadiana	29K.01	Martim Longo	1985-2000	Homogeneous data	Homogeneous data	
Guadiana	29K.03	Malfrades	1981-1999	Homogeneous data	Homogeneous data	
Guadiana	29K.04	Penedos	1981-2000	Homogeneous data	Upward trend present	
Guadiana	29L.01	Pereiro	1958-1999	Homogeneous data	Homogeneous data	Break present
Guadiana	29L.03	Monte dos Fortes	1984-1999	Homogeneous data	Homogeneous data	
Guadiana	29M.01	Alcoutim	1939-1999	Trend effec present	Homogeneous data	Break present
Guadiana	30I.02	Sobreira	1943-1999	Trend effec present	Homogeneous data	Break present
Guadiana	30J.01	Barranco do Velho	1956-1999	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	30J.02	Catraia	1932-1973	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	30K.01	Mercador	1984-1999	Homogeneous data	Homogeneous data	
Guadiana	30L.04	Alcaria (Castro Marim)	1947-1999	Homogeneous data	Homogeneous data	Homogeneous data

Table 3 – Summary results from the absolute testing approach: tests capable of locating the year where a break is likely applied to 107 series of annual wet day count (1 mm threshold), and series classification

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	BUISHAND (5% level)	PETTIT (5% level)	SNHT (5% level)	SERIES CLASSIFICATION
ECA&D	666	Beja	1941-1999	Homogeneous data	Homogeneous data	Homogeneous data	Candidate
ECA&D	675	Lisboa Geofísica	1941-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
ECA&D	681	Tavira	1941-1994	Homogeneous data	Break: 1979	Break: 1978	Reject
ECA&D	709	Badajoz Talavera	1955-2000	Homogeneous data	Homogeneous data	Homogeneous data	Candidate
Arade	29I.01	São Barnabé	1965-2000	Homogeneous data	Break: 1972	Homogeneous data	Reject
Arade	30F.01	Monchique	1933-1998	Homogeneous data	Homogeneous data	Homogeneous data	Reject
Arade	30G.01	Alferce	1959-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Arade	30H.04	Santa Margarida	1965-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Ribeiras do Algarve	29F.01	Cimalhas	1981-1999			Break: 1982	Doubtful
Ribeiras do Algarve	29F.02	Foz do Farelo	1981-1999			Break: 1983	Doubtful
Ribeiras do Algarve	30E.01	Aljezur	1932-1999	Break: 1942	Homogeneous data	Homogeneous data	Candidate
Ribeiras do Algarve	30E.02	Marmelete	1984-1999			Homogeneous data	Useful
Ribeiras do Algarve	30E.03	Barragem da Bravura	1956-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines	1991-1998				Useful
Ribeiras do Algarve	30H.05	Paderne	1984-1999			Homogeneous data	Doubtful
Ribeiras do Algarve	30K.02	Picota	1957-1999	Homogeneous data	Break: 1972	Homogeneous data	Candidate
Ribeiras do Algarve	30L.03	Faz-Fato	1956-1999	Break: 1986	Break: 1986	Homogeneous data	Reject
Ribeiras do Algarve	31E.01	Lagos	1956-1999	Break: 1972	Break: 1979	Break: 1972	Reject
Ribeiras do Algarve	31G.02	Porches	1980-1998			Homogeneous data	Useful
Ribeiras do Algarve	31H.02	Algoz	1981-1996			Homogeneous data	Useful
Ribeiras do Algarve	31J.01	São Brás de Alportel	1985-2000			Homogeneous data	Useful
Ribeiras do Algarve	31J.04	Estoi	1984-1999			Homogeneous data	Useful

Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)	1984-1999			Homogeneous data	Useful
Ribeiras do Algarve	31K.02	Quelfes	1982-1998			Homogeneous data	Useful
Mira	27G.01	Relíquias	1932-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Mira	28F.01	Odemira	1932-1994	Homogeneous data	Homogeneous data	Homogeneous data	Candidate
Mira	28G.01	Barragem de Mira	1970-1993	Homogeneous data	Homogeneous data	Homogeneous data	Doubtful
Mira	28H.01	Aldeia de Palheiros	1932-1996	Homogeneous data	Break: 1979	Homogeneous data	Candidate
Mira	28H.03	Santana da Serra	1936-2000	Homogeneous data	Break: 1979	Homogeneous data	Reject
Mira	29G.01	Sabóia	1932-1994	Homogeneous data	Homogeneous data	Homogeneous data	Candidate
Mira	29I.02	Santa Clara-a-Nova	1981-1999			Homogeneous data	Useful
Sado	21G.01	Vendas Novas	1932-1999	Homogeneous data	Homogeneous data	Break: 1943	Reject
Sado	22E.01	Águas de Moura	1984-2001			Homogeneous data	Useful
Sado	22F.03	Moinhola	1973-2000	Homogeneous data	Homogeneous data	Homogeneous data	Doubtful
Sado	22H.02	Santiago do Escoural	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Sado	23E.01	Comporta	1934-2000	Homogeneous data	Homogeneous data	Break: 1935	Candidate
Sado	23F.01	Montevil	1984-2000			Homogeneous data	Useful
Sado	23G.01	Barragem de Pego do Altar	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data	Useful
Sado	23I.01	Alcáçovas	1932-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Sado	24F.01	Grândola	1973-2000	Homogeneous data	Break: 1979	Break: 1982	Doubtful
Sado	24H.02	Barragem do Vale do Gaio	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data	Doubtful
Sado	24I.01	Viana do Alentejo	1934-2000	Homogeneous data	Break: 1989	Homogeneous data	Candidate
Sado	24I.03	Barragem de Odivelas	1974-2000	Homogeneous data	Homogeneous data	Homogeneous data	Doubtful
Sado	24J.02	Alvito	1984-2000			Homogeneous data	Useful
Sado	25G.01	Azinheira Barros	1951-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Sado	25I.01	Ferreira do Alentejo	1933-2000	Break: 1958	Break: 1958	Break: 1958	Reject
Sado	26F.02	Barragem de Campilhas	1956-1994	Homogeneous data	Breaks: 1979, 1989	Break: 1979	Doubtful
Sado	26I.01	Santa Vitória	1984-2000			Homogeneous data	Useful

Sado	26I.02	Barragem do Roxo	1959-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Sado	26I.03	Aljustrel	1936-2000	Break: 1972	Break: 1972	Break: 1972	Reject
Sado	27G.02	Garvão (Montinho)	1980-1994			Break: 1983	Doubtful
Sado	27H.01	Panóias	1956-1994	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Sado	27H.02	Barragem do Monte da Rocha	1980-1994			Homogeneous data	Useful
Guadiana	18N.01	São Julião	1981-1999			Homogeneous data	Useful
Guadiana	18N.02	Alegrete	1981-1999			Homogeneous data	Useful
Guadiana	19N.01	Arronches	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	19N.02	Santa Eulália	1983-1999			Homogeneous data	Useful
Guadiana	19N.03	Esperança	1980-1999			Homogeneous data	Useful
Guadiana	19O.02	Barragem do Caia	1965-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	19O.03	Degolados	1984-1999			Homogeneous data	Useful
Guadiana	20O.02	Caia (M. Caldeiras)	1980-1999			Break: 1989	Doubtful
Guadiana	21K.01	Azaruja	1944-1982	Homogeneous data	Break: 1952	Homogeneous data	Candidate
Guadiana	21M.01	Vila Viçosa	1981-2000	Homogeneous data	Homogeneous data	Homogeneous data	Doubtful
Guadiana	21M.02	Alandroal	1984-1999			Break: 1989	Doubtful
Guadiana	21N.01	Juromenha	1932-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reject
Guadiana	22L.01	Redondo	1945-1982	Break: 1962	Homogeneous data	Homogeneous data	Candidate
Guadiana	22L.02	Santa Susana	1950-1999	Break: 1979	Break: 1979	Break: 1979	Reject
Guadiana	22M.01	Santiago Maior	1984-1999			Homogeneous data	Useful
Guadiana	23K.01	São Manços	1943-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	23L.01	Reguengos	1985-1999			Homogeneous data	Useful
Guadiana	24J.03	Cuba	1986-1998			Homogeneous data	Useful
Guadiana	24K.01	Portel	1984-1999			Homogeneous data	Useful
Guadiana	24K.02	Vidigueira	1984-2000			Homogeneous data	Useful
Guadiana	24N.01	Amareleja (D.G.R.N.)	1984-2000			Homogeneous data	Useful

Guadiana	25L.01	Pedrogão do Alentejo	1942-2000	Break: 1954	Break: 1954, 1972	Breaks: 1946, 1954, 1965	Reject
Guadiana	25N.01	Sobral da Adiça	1981-2000			Break: 1983	Doubtful
Guadiana	25O.01	Santo Aleixo da Restauração	1932-2000	Break: 1958	Break: 1958	Homogeneous data	Reject
Guadiana	25P.01	Barrancos	1986-1998			Homogeneous data	Useful
Guadiana	26J.04	Albernoa	1984-2000			Homogeneous data	Useful
Guadiana	26K.01	Salvada	1984-2000			Homogeneous data	Useful
Guadiana	26L.01	Serpa	1985-2000			Homogeneous data	Useful
Guadiana	26L.02	Santa Iria	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data	Useful
Guadiana	26M.01	Herdade de Valada	1969-2000	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	27I.01	Castro Verde	1932-2000	Homogeneous data	Break: 1979	Homogeneous data	Reject
Guadiana	27J.01	São Marcos da Ataboeira	1984-2000			Homogeneous data	Useful
Guadiana	27J.02	Corte Pequena	1986-1996			Homogeneous data	Useful
Guadiana	27J.03	Vale de Camelos	1990-2000			Homogeneous data	Useful
Guadiana	27K.01	Algodôr	1987-1999			Homogeneous data	Useful
Guadiana	27K.02	Corte da Velha	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data	Useful
Guadiana	28I.01	Almodôvar	1984-2000			Homogeneous data	Useful
Guadiana	28J.01	Alcaria Longa	1986-1999			Homogeneous data	Useful
Guadiana	28J.03	Santa Barbara de Padrões	1980-2000	Homogeneous data	Homogeneous data	Homogeneous data	Useful
Guadiana	28K.01	São João dos Caldeireiros	1984-2000			Homogeneous data	Useful
Guadiana	28K.02	Álamo	1981-2000	Homogeneous data	Homogeneous data	Homogeneous data	Useful
Guadiana	28L.01	Mértola	1984-2000			Homogeneous data	Useful
Guadiana	29J.05	Guedelhas	1980-1999			Homogeneous data	Useful
Guadiana	29K.01	Martim Longo	1985-2000			Homogeneous data	Useful
Guadiana	29K.03	Malfrades	1981-1999			Homogeneous data	Useful
Guadiana	29K.04	Penedos	1981-2000			Break: 1995	Doubtful
Guadiana	29L.01	Pereiro	1958-1999	Break: 1983	Homogeneous data	Break: 1995	Reject

Guadiana	29L.03	Monte dos Fortes	1984-1999			Homogeneous data	Useful
Guadiana	29M.01	Alcoutim	1939-1999	Break: 1959	Break: 1959	Break: 1959	Reject
Guadiana	30I.02	Sobreira	1943-1999	Break: 1954	Homogeneous data	Break: 1949	Reject
Guadiana	30J.01	Barranco do Velho	1956-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	30J.02	Catraia	1932-1973	Homogeneous data	Homogeneous data	Homogeneous data	Reference
Guadiana	30K.01	Mercador	1984-1999			Homogeneous data	Useful
Guadiana	30L.04	Alcaria (Castro Marim)	1947-1999	Homogeneous data	Homogeneous data	Homogeneous data	Reference

Appendix III RELATIVE TESTING RESULTS

Table 1 – Summary results from the relative testing approach: Buishand, Pettit, and SNHT tests applied to composite (ratio) reference series of the annual wet day count (1 mm threshold)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES	PERIOD	BUISHAND (5% level)	PETTIT (5% level)	SNHT (5% level)
ECA&D	666	Beja	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	1951-1999	Homogeneous data	Homogeneous data	Homogeneous data
ECA&D	709	Badajoz Talavera	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	1955-2000	Homogeneous data	Break: 1975	Homogeneous data
Ribeiras do Algarve	30E.01	Aljezur	Barragem da Bravura (Rib. Algarve 30E.03) Arronches (Guadiana 19N.01)	1956-1999	Break: 1968	Break: 1968	Break: 1968
Ribeiras do Algarve	30K.02	Picota	Barragem da Bravura (Rib. Algarve 30E.03) Alcaria [Castro Marim] (Guadiana 30L.04)	1957-1999	Break: 1988	Homogeneous data	Homogeneous data
Mira	28F.01	Odemira	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	1952-1994	Homogeneous data	Homogeneous data	Homogeneous data
Mira	28H.01	Aldeia de Palheiros	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	1951-1995	Homogeneous data	Homogeneous data	Homogeneous data
Mira	29G.01	Sabóia	Santiago do Escoural (Sado 22H.02) Relíquias (Mira 27G.01)	1932-1994	Breaks: 1949; 1984	Break: 1984	Break: 1985
Sado	23E.01	Comporta	Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	1956-2000	Homogeneous data	Break: 1986	Break: 1986
Sado	24I.01	Viana do Alentejo	Alcáçovas (Sado 23I.01) Santiago do Escoural (Sado 22H.02)	1934-1999	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	21K.01	Azaruja	Santiago do Escoural (Sado 22H.02) Lisboa Geofísica (ECA 675)	1944-1982	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	22L.01	Redondo	São Manços (Guadiana 23K.01) Santiago do Escoural (Sado 22H.02)	1945-1982	Break: 1963	Break: 1963	Break: 1963
TESTS FOR REFERENCE SERIES							
ECA&D	675	Lisboa Geofísica	Santiago do Escoural (Sado 22H.02) Alcáçovas (Sado 23I.01)	1941-1999	Homogeneous data	Homogeneous data	Homogeneous data
Arade	30G.01	Alferce	Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	1959-1999	Break: 1984	Break: 1984	Homogeneous data
Arade	30H.04	Santa Margarida	Barragem da Bravura (Rib. Algarve 30E.03) Relíquias (Mira 27G.01)	1965-1999	Break: 1978	Break: 1978	Break: 1978

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES	PERIOD	BUISHAND (5% level)	PETTIT (5% level)	SNHT (5% level)
Ribeiras do Algarve	30E.03	Barragem da Bravura	São Manços (Guadiana 23K.01) Azinheira Barros (Sado 25G.01)	1956-2000	Homogeneous data	Homogeneous data	Homogeneous data
Mira	27G.01	Relíquias	Alcáçovas (Sado 23I.01) Santiago do Escoural (Sado 22H.02)	1932-1999	Break: 1969	Break: 1969	Homogeneous data
Sado	22H.02	Santiago do Escoural	Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	1932-1999	Break: 1988	Break: 1988	Break: 1989
Sado	23I.01	Alcáçovas	Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	1932-1999	Break: 1960	Break: 1960	Break: 1960
Sado	25G.01	Azinheira Barros	São Manços (Guadiana 23K.01) Lisboa Geofísica (ECA 675)	1951-1999	Homogeneous data	Homogeneous data	Homogeneous data
Sado	26I.02	Barragem do Roxo	Azinheira Barros (Sado 25G.01) Barragem da Bravura (Rib. Algarve 30E.03)	1959-2000	Homogeneous data	Homogeneous data	Homogeneous data
Sado	27H.01	Panóias	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	1956-1994	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	19N.01	Arronches	Santiago do Escoural (Sado 22H.02) Relíquias (Mira 27G.01)	1932-1999	Homogeneous data	Homogeneous data	Break: 1954
Guadiana	19O.02	Barragem do Caia	Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	1965-2000	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	23K.01	São Manços	Alcáçovas (Sado 23I.01) Lisboa Geofísica (ECA 675)	1943-1999	Homogeneous data	Homogeneous data	Break: 1950
Guadiana	26M.01	Herdade de Valada	Barragem do Roxo (Sado 26I.02) Azinheira Barros (Sado 25G.01)	1969-2000	Homogeneous data	Homogeneous data	Break: 1995
Guadiana	30J.01	Barranco do Velho	Alcáçovas (Sado 23I.01) Barragem da Bravura (Rib. Algarve 30E.03)	1956-1999	Break: 1976	Break: 1976	Breaks: 1975; 1996
Guadiana	30J.02	Catraia	Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	1932-1973	Homogeneous data	Homogeneous data	Homogeneous data
Guadiana	30L.04	Alcaria (Castro Marim)	Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	1947-1999	Homogeneous data	Homogeneous data	Homogeneous data

Table 2 – Summary results from the relative testing approach: SUR+Ellipse test applied to the annual wet day count (1 mm threshold)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES (In light gray: independent variables, non significant at the 5% level, that were removed from the models)	MODEL ('m' indicates average of references as independent variable)	PERIOD	SUR + Ellipse (5% level)
ECA&D	666	Beja	Lisboa Geofísica (ECA 675) Relíquias (Mira 27G.01) Arronches (Guadiana 19N.01)	5	1941-1999	Homogeneous data
			Lisboa Geofísica (ECA 675) Relíquias (Mira 27G.01) Arronches (Guadiana 19N.01)	13	1945-1982	Homogeneous data
			Lisboa Geofísica (ECA 675) Relíquias (Mira 27G.01) Arronches (Guadiana 19N.01)	12	1956-1997	Homogeneous data
ECA&D	709	Badajoz Talavera	Relíquias (Mira 27G.01) Azinheira Barros (Sado 25G.01)	6	1956-1997	Homogeneous data
			Relíquias (Mira 27G.01) Azinheira Barros (Sado 25G.01)	12	1956-1997	Homogeneous data
Ribeiras do Algarve	30E.01	Aljezur	Arronches (Guadiana 19N.01)	5	1941-1999	Homogeneous data
			Arronches (Guadiana 19N.01)	6	1956-1997	Break: 1968
			Arronches (Guadiana 19N.01)	7	1932-1996	Homogeneous data
			Arronches (Guadiana 19N.01)	9	1932-1994	Homogeneous data
			Arronches (Guadiana 19N.01)	10	1932-1996	Homogeneous data
			Arronches (Guadiana 19N.01)	11	1932-1994	Homogeneous data
			Arronches (Guadiana 19N.01)	12	1956-1997	Break: 1968
Ribeiras do Algarve	30K.02	Picota	Barragem da Bravura (Rib. Algarve 30E.03) Alcaria [Castro Marim] (Guadiana 30L.04) Azinheira Barros (Sado 25G.01) Lisboa Geofísica (ECA 675)	8	1957-1995	Homogeneous data
Mira	28F.01	Odemira	Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	9	1932-1994	Break: 1952

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES (In light gray: independent variables, non significant at the 5% level, that were removed from the models)	MODEL ('m' indicates average of references as independent variable)	PERIOD	SUR + Ellipse (5% level)
Mira	28H.01	Aldeia de Palheiros	Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	7	1932-1996	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	8	1957-1995	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	9	1932-1994	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	10	1932-1996	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	11	1932-1994	Homogeneous data
Mira	29G.01	Sabóia	Santiago do Escoural (Sado 22H.02)	11	1932-1994	Break: 1984
Sado	23E.01	Comporta	Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	6	1956-1997	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	8	1957-1995	Break: 1986
			Santiago do Escoural (Sado 22H.02) Arronches (Guadiana 19N.01)	12	1956-1997	Homogeneous data
Sado	24I.01	Viana do Alentejo	*Not possible to determine a common period, without too many gaps, for all the series that would be appropriate to include in a model*			
Guadiana	21K.01	Azaruja	Santiago do Escoural (Sado 22H.02) Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01) São Manços (Guadiana 23K.01)	13	1945-1982	Homogeneous data
Guadiana	22L.01	Redondo	São Manços (Guadiana 23K.01) Lisboa Geofísica (ECA 675) Relíquias (Mira 27G.01) Santiago do Escoural (Sado 22H.02)	13	1945-1982	Break: 1963
TESTS FOR REFERENCE SERIES						
ECA&D	675	Lisboa Geofísica	Santiago do Escoural (Sado 22H.02) Alcáçovas (Sado 23I.01)	14 32m	1941-1997	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Alcáçovas (Sado 23I.01)	17	1956-1995	Homogeneous data

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES (In light gray: independent variables, non significant at the 5% level, that were removed from the models)	MODEL ('m' indicates average of references as independent variable)	PERIOD	SUR + Ellipse (5% level)
Arade	30G.01	Alferce	Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	15	1959-1997	Homogeneous data
			Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	21	1959-1999	Break: 1984
Arade	30H.04	Santa Margarida	Barragem da Bravura (Rib. Algarve 30E.03) Relíquias (Mira 27G.01)	16 28m	1965-1999	Homogeneous data
Ribeiras do Algarve	30E.03	Barragem da Bravura	São Manços (Guadiana 23K.01) Azinheira Barros (Sado 25G.01)	17	1956-1995	Homogeneous data
			São Manços (Guadiana 23K.01) Azinheira Barros (Sado 25G.01)	22	1956-1994	Homogeneous data
Mira	27G.01	Relíquias	Santiago do Escoural (Sado 22H.02)	14 32m	1941-1997	Homogeneous data
			Alcáçovas (Sado 23I.01)	18	1940-1997	Homogeneous data
			Santiago do Escoural (Sado 22H.02)	19	1940-1997	Homogeneous data
Sado	22H.02	Santiago do Escoural	Arronches (Guadiana 19N.01)	18	1940-1997	Homogeneous data
			Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	22 29m	1956-1994	Homogeneous data Break: 1960
			Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	27	1932-1973	Homogeneous data
Sado	23I.01	Alcáçovas	Arronches (Guadiana 19N.01)	19	1940-1997	Homogeneous data
			Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	20	1951-1999	Homogeneous data
			Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	24	1943-1999	Break: 1960
			Arronches (Guadiana 19N.01) Relíquias (Mira 27G.01)	26 30m	1956-1999	Homogeneous data
Sado	25G.01	Azinheira Barros	São Manços (Guadiana 23K.01) Lisboa Geofísica (ECA 675)	20	1951-1999	Homogeneous data

BASIN / NETWORK	STATION CODE / ID	STATION NAME	REFERENCE SERIES (In light gray: independent variables, non significant at the 5% level, that were removed from the models)	MODEL ('m' indicates average of references as independent variable)	PERIOD	SUR + Ellipse (5% level)
Sado	26I.02	Barragem do Roxo	Azinheira Barros (Sado 25G.01) Barragem da Bravura (Rib. Algarve 30E.03)	21	1959-1999	Homogeneous data
			Azinheira Barros (Sado 25G.01) Barragem da Bravura (Rib. Algarve 30E.03)	23	1966-2000	Homogeneous data
Sado	27H.01	Panóias	Azinheira Barros (Sado 25G.01) São Manços (Guadiana 23K.01)	22 29m	1956-1994	Homogeneous data
Guadiana	19N.01	Arronches	Santiago do Escoural (Sado 22H.02) Relíquias (Mira 27G.01)	15	1959-1997	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Relíquias (Mira 27G.01)	17	1956-1995	Homogeneous data
			Santiago do Escoural (Sado 22H.02) Relíquias (Mira 27G.01)	23	1966-2000	Break: 1988
Guadiana	19O.02	Barragem do Caia	Barragem da Bravura (Rib. Algarve 30E.03) Azinheira Barros (Sado 25G.01)	23	1966-2000	Homogeneous data
Guadiana	23K.01	São Manços	Lisboa Geofísica (ECA 675)	24	1943-1999	Break: 1950
Guadiana	26M.01	Herdade de Valada	Barragem do Roxo (Sado 26I.02) Azinheira Barros (Sado 25G.01)	25	1969-1999	Homogeneous data
Guadiana	30J.01	Barranco do Velho	Barragem da Bravura (Rib. Algarve 30E.03)	26 30m	1956-1999	Homogeneous data
			Barragem da Bravura (Rib. Algarve 30E.03)	31m	1956-1999	Homogeneous data
Guadiana	30J.02	Catraia	Arronches (Guadiana 19N.01)	27	1932-1973	Homogeneous data
Guadiana	30L.04	Alcaria (Castro Marim)	Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	16 28m	1965-1999	Homogeneous data
			Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	21	1959-1999	Homogeneous data
			Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	25	1969-1999	Homogeneous data
			Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	26 30m	1956-1999	Homogeneous data
			Lisboa Geofísica (ECA 675) Arronches (Guadiana 19N.01)	31m	1956-1999	Homogeneous data

Table 3 – Summary results from the relative testing approach: break years identified by at least one of the relative procedures, relative magnitudes (%) of the breaks and length of homogeneous periods

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	BREAK YEARS	MAGNITUDE OF 1st BREAK	MAGNITUDE OF 2nd BREAK	LENGTH OF 1st HOMOG. PERIOD	LENGTH OF 2nd HOMOG. PERIOD	LENGTH OF 3rd HOMOG. PERIOD
ECA&D	666	Beja	1941-1999						
ECA&D	709	Badajoz Talavera	1955-2000	1975	-1.91%		20	25	
Ribeiras do Algarve	30E.01	Aljezur	1932-1999	1968	3.67%		36	31	
Ribeiras do Algarve	30K.02	Picota	1957-1999	1988	-2.41%		31	11	
Mira	28F.01	Odemira	1932-1994	1952	-7.60%		20	42	
Mira	28H.01	Aldeia de Palheiros	1932-1996						
Mira	29G.01	Sabóia	1932-1994	1949 + 1984	4.71%	17.26%	17	35	10
Sado	23E.01	Comporta	1934-2000	1986	10.23%		52	14	
Sado	24I.01	Viana do Alentejo	1934-2000						
Guadiana	21K.01	Azaruja	1944-1982						
Guadiana	22L.01	Redondo	1945-1982	1963	13.71%		18	19	
TESTS FOR REFERENCE SERIES									
ECA&D	675	Lisboa Geofísica	1941-1999						
Arade	30G.01	Alferce	1959-1999	1984	8.06%		25	15	
Arade	30H.04	Santa Margarida	1965-1999	1978	-12.10%		13	21	
Ribeiras do Algarve	30E.03	Barragem da Bravura	1956-2000						
Mira	27G.01	Relíquias	1932-2000	1969	3.48%		37	31	
Sado	22H.02	Santiago do Escoural	1932-1999	1960 + 1988	6.88%	-8.94%	28	28	11
Sado	23I.01	Alcáçovas	1932-2000	1960	-3.96%		28	40	
Sado	25G.01	Azinheira Barros	1951-2000						
Sado	26I.02	Barragem do Roxo	1959-2000						
Sado	27H.01	Panóias	1956-1994						
Guadiana	19N.01	Arronches	1932-1999	1954 + 1988	5.03%	-0.57%	22	34	11

BASIN / NETWORK	STATION CODE / ID	STATION NAME	PERIOD	BREAK YEARS	MAGNITUDE OF 1st BREAK	MAGNITUDE OF 2nd BREAK	LENGTH OF 1st HOMOG. PERIOD	LENGTH OF 2nd HOMOG. PERIOD	LENGTH OF 3rd HOMOG. PERIOD
Guadiana	19O.02	Barragem do Caia	1965-2000						
Guadiana	23K.01	São Manços	1943-2000	1950	-0.52%		7	50	
Guadiana	26M.01	Herdade de Valada	1969-2000	1995	-10.11%		26	5	
Guadiana	30J.01	Barranco do Velho	1956-1999	1976 + 1996	10.78%	-14.09%	20	20	3
Guadiana	30J.02	Catraia	1932-1973						
Guadiana	30L.04	Alcaria (Castro Marim)	1947-1999						

Appendix IV OVERALL CLASSIFICATION OF THE DAILY PRECIPITATION SERIES

Table 1 – Number of records flagged through basic quality control procedures and overall classification of the daily precipitation series

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
ECA&D	666	Beja		11	Suspect	Relative break(s) detected could not be explained by non-climatic factors
ECA&D	675	Lisboa Geofísica		2	Useful	All relative approaches considered the series as homogeneous
ECA&D	681	Tavira		2	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
ECA&D	709	Badajoz Talavera			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Arade	29I.01	São Barnabé		15	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Arade	30F.01	Monchique		6	Potentially useful	Absolute break(s) detected might be explained by several months without records
Arade	30G.01	Alferce			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Arade	30H.04	Santa Margarida			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	29F.01	Cimalhas			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	29F.02	Foz do Farelo			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	30E.01	Aljezur	2		Suspect	Relative break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	30E.02	Marmelete			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	30E.03	Barragem da Bravura		2	Useful	All relative approaches considered the series as homogeneous

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Ribeiras do Algarve	30H.03	São Bartolomeu de Messines			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	30H.05	Paderne			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	30K.02	Picota	62	4	Useful	Relative break(s) detected might be explained by several months without records **Note**: it is advisable to set to missing the records of Dec. 1972 and Dec. 1973
Ribeiras do Algarve	30L.03	Faz-Fato			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	31E.01	Lagos		2	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Ribeiras do Algarve	31G.02	Porches		2	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	31H.02	Algoz		2	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	31J.01	São Brás de Alportel		4	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	31J.04	Estoi			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	31K.01	Santa Catarina (Tavira)			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Ribeiras do Algarve	31K.02	Quelfes			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Mira	27G.01	Relíquias			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Mira	28F.01	Odemira		2	Useful	Relative break(s) detected might be explained by several months without records
Mira	28G.01	Barragem de Mira		2	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Mira	28H.01	Aldeia de Palheiros			Useful	All relative approaches considered the series as homogeneous
Mira	28H.03	Santana da Serra			Potentially useful	Absolute break(s) detected might be explained by several months without records
Mira	29G.01	Sabóia		4	Suspect	Relative break(s) detected could not be explained by non-climatic factors
Mira	29I.02	Santa Clara-a-Nova			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	21G.01	Vendas Novas		4	Potentially useful	Absolute break(s) detected might be explained by several months without records
Sado	22E.01	Águas de Moura			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	22F.03	Moinhola			Potentially useful	Absolute break(s) detected might be explained by several months without records
Sado	22H.02	Santiago do Escoural			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Sado	23E.01	Comporta		2	Useful	Relative break(s) detected might be explained by several months without records
Sado	23F.01	Montevil		12	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	23G.01	Barragem de Pego do Altar			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	23I.01	Alcáçovas	1		Suspect	Relative break(s) detected could not be explained by non-climatic factors
Sado	24F.01	Grândola			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Sado	24H.02	Barragem do Vale do Gaio			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Sado	24I.01	Viana do Alentejo			Useful	All relative approaches considered the series as homogeneous

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Sado	24I.03	Barragem de Odivelas		5	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Sado	24J.02	Alvito			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	25G.01	Azinheira Barros		4	Useful	All relative approaches considered the series as homogeneous
Sado	25I.01	Ferreira do Alentejo			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Sado	26F.02	Barragem de Campilhas		2	Potentially useful	Absolute break(s) detected might be explained by several months without records
Sado	26I.01	Santa Vitória			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Sado	26I.02	Barragem do Roxo			Useful	All relative approaches considered the series as homogeneous
Sado	26I.03	Aljustrel		4	Potentially useful	Absolute break(s) detected might be explained by several months without records
Sado	27G.02	Garvão (Montinho)			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Sado	27H.01	Panóias			Useful	All relative approaches considered the series as homogeneous
Sado	27H.02	Barragem do Monte da Rocha			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	18N.01	São Julião			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	18N.02	Alegrete			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	19N.01	Arronches			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Guadiana	19N.02	Santa Eulália			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Guadiana	19N.03	Esperança			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	19O.02	Barragem do Caia		2	Useful	All relative approaches considered the series as homogeneous
Guadiana	19O.03	Degolados			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	20O.02	Caia (M. Caldeiras)		2	Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	21K.01	Azaruja			Useful	All relative approaches considered the series as homogeneous
Guadiana	21M.01	Vila Viçosa			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	21M.02	Alandroal		2	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	21N.01	Juromenha			Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	22L.01	Redondo			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Guadiana	22L.02	Santa Susana		2	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	22M.01	Santiago Maior			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	23K.01	São Manços			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Guadiana	23L.01	Reguengos			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	24J.03	Cuba			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	24K.01	Portel			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Guadiana	24K.02	Vidigueira			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	24N.01	Amareleja (D.G.R.N.)			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	25L.01	Pedrogão do Alentejo			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	25N.01	Sobral da Adiça			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	25O.01	Santo Aleixo da Restauração		5	Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	25P.01	Barrancos			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	26J.04	Albernoa			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	26K.01	Salvada			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	26L.01	Serpa			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	26L.02	Santa Iria		2	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	26M.01	Herdade de Valada		5	Suspect	Relative break(s) detected could not be explained by non-climatic factors
Guadiana	27I.01	Castro Verde	4	13	Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	27J.01	São Marcos da Ataboeira			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	27J.02	Corte Pequena			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	27J.03	Vale de Camelos			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Guadiana	27K.01	Algodôr			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	27K.02	Corte da Velha			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28I.01	Almodôvar	1		Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28J.01	Alcaria Longa			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28J.03	Santa Barbara de Padrões			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28K.01	São João dos Caldeireiros			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28K.02	Álamo			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	28L.01	Mértola			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	29J.05	Guedelhas			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	29K.01	Martim Longo		7	Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	29K.03	Malfrades			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	29K.04	Penedos		2	Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	29L.01	Pereiro			Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors
Guadiana	29L.03	Monte dos Fortes			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	29M.01	Alcoutim	2191	12	Potentially suspect	Absolute break(s) detected could not be explained by non-climatic factors. **Note**: it is advisable to set to missing the records of 1954-59

BASIN / NETWORK	STATION CODE / ID	STATION NAME	SUBJECTIVE FLAGGING	'FLAT LINE' CHECK FLAGGING	SERIES OVERALL CLASSIFICATION	CRITERION / COMMENTS
			Num. Records flagged as (3) 'suspect'	Num. Records flagged as (1) 'suspect'		
Guadiana	30I.02	Sobreira			Potentially useful	Absolute break(s) detected might be explained by several months without records
Guadiana	30J.01	Barranco do Velho			Suspect	Relative break(s) detected could not be explained by non-climatic factors
Guadiana	30J.02	Catraia		2	Useful	All relative approaches considered the series as homogeneous
Guadiana	30K.01	Mercador			Potentially useful	Short-term series previously classified as 'useful' (the 6 absolute tests considered the series as homogeneous)
Guadiana	30L.04	Alcaria (Castro Marim)		2	Useful	All relative approaches considered the series as homogeneous

Appendix V LOCAL CORRELATION MODELS FOR THE WETNESS INDICES AND DISTANCE TO THE COASTLINE

Local correlation models for the wetness indices and distance to the coastline

In order to determine local correlation models for the wetness indices (R5D and R30) and distance to the coastline according to the SW direction, first, the relationships were assessed locally by computing, for each decade, Pearson's correlation coefficients using stations' data falling within a circle centred at each station's location. As in earlier decades meteorological stations are scarce, larger radii were used (Table 1). Second, the DSS algorithm was applied to interpolate the local correlations by decade, and 50 simulated maps of local correlations were obtained for each index.

Table 1 – Radii of the search neighbourhoods used to calculate the local correlations

Decade	Radius (m)
1940/49	65000
1950/59	50000
1960/69	50000
1970/79	40000
1980/89	35000
1990/99	35000

These procedures used space-time spherical variograms of the local correlations, where the spatial dimension was modelled as isotropic. For the R5D index, the estimated range of the spatial dimension was 75000 meters, the range of the temporal dimension was 5 decades, and the estimated sill was 0.063. For the R30 index, the estimated range of the spatial dimension was 70000 meters, the range of the temporal dimension was 4.5 decades, and the estimated sill was 0.069.

The correlation models were then determined by computing the mean of the distribution of 50 simulated values at each grid node, by decade (Figure 1 and Figure 2).

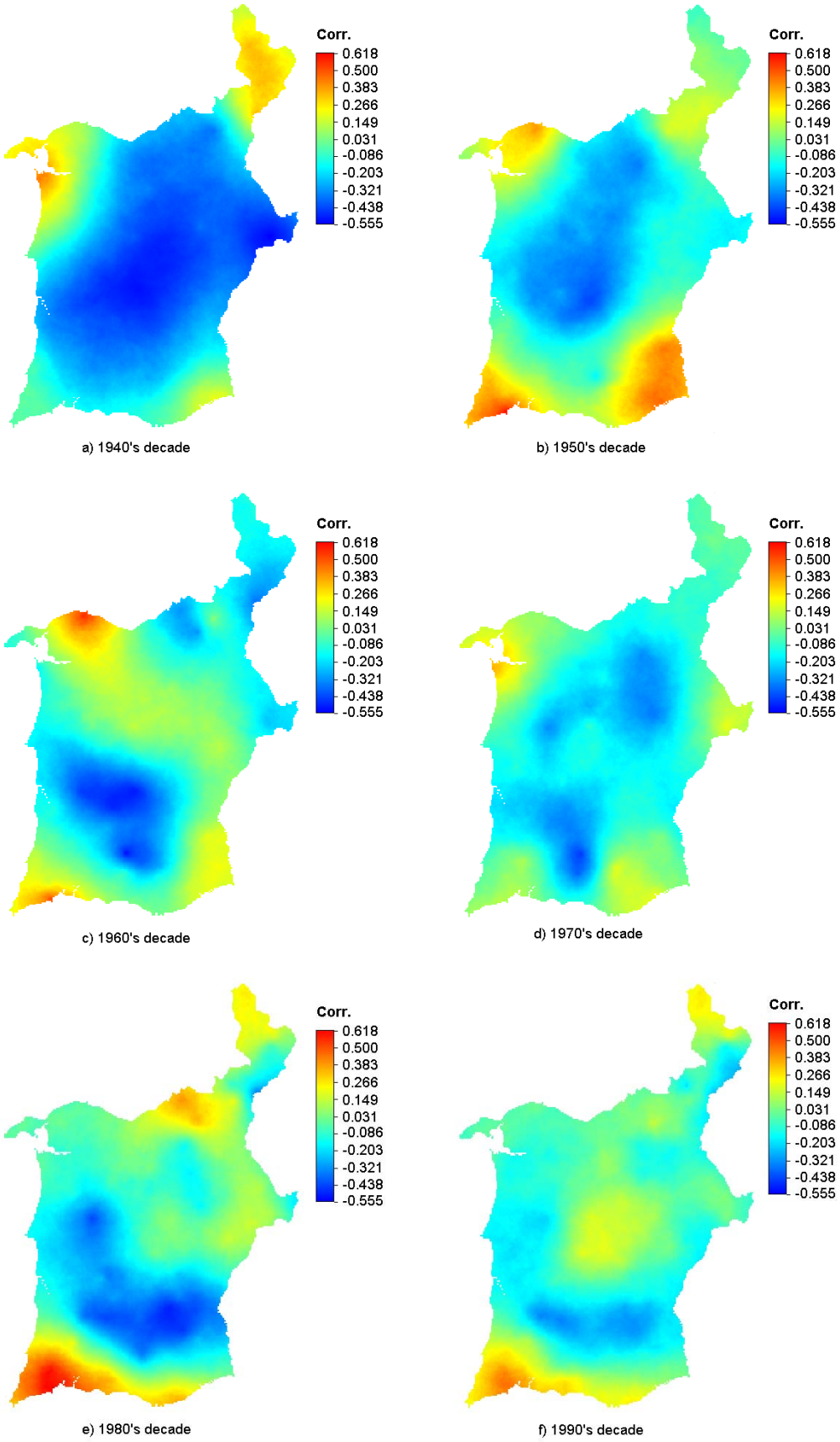

Figure 1 – Local correlation models between R5D values and distance to the coastline (SW) for each decade

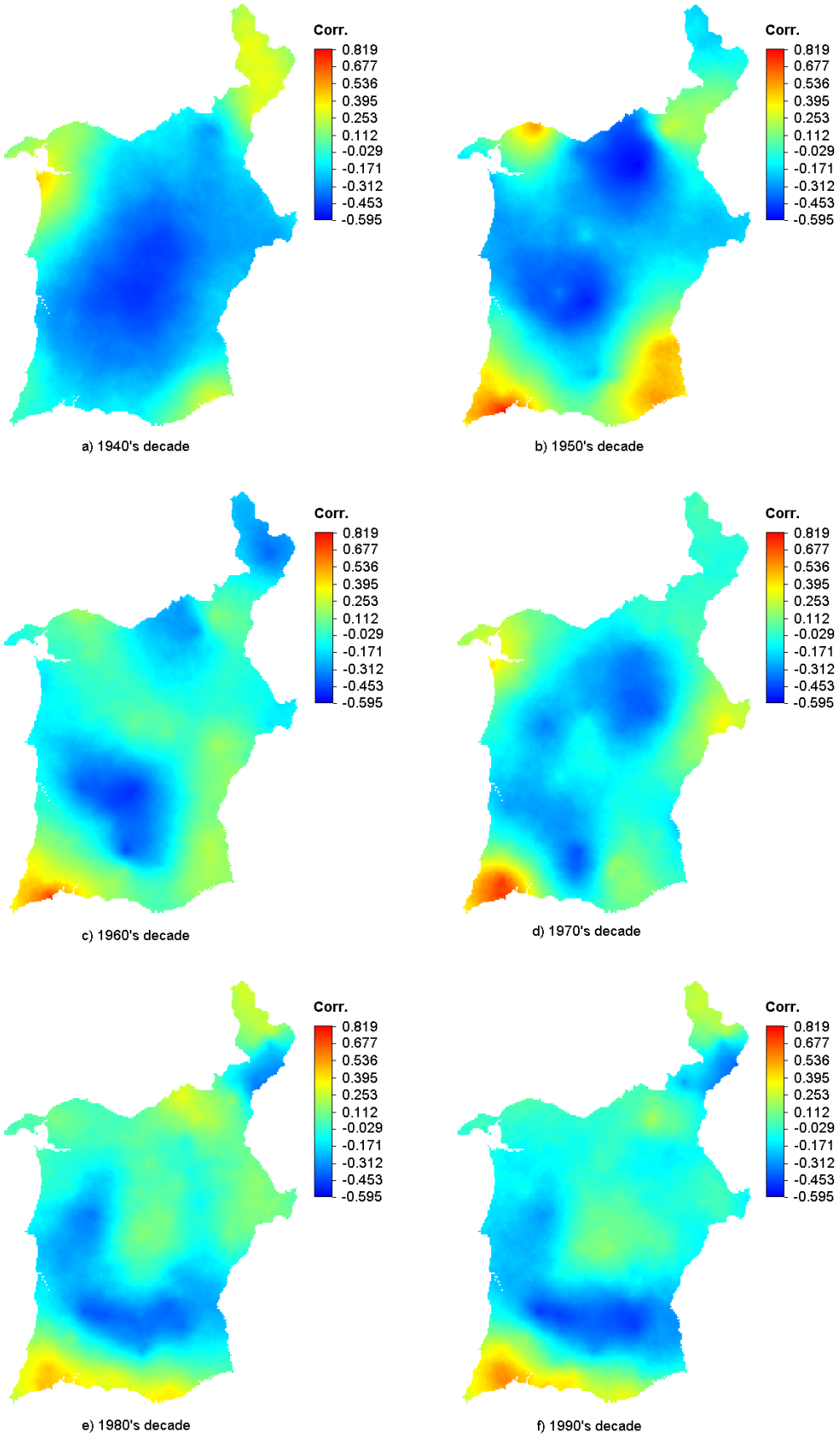

Figure 2 – Local correlation models between R30 values and distance to the coastline (SW) for each decade